DIVERSITY AND ECOLOGY OF THE *ROSEOBACTER* CLADE AND OTHER MARINE MICROBES AS REVEALED BY METAGENOMIC AND METATRANSCRIPTOMIC APPROACHES

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen
Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen

im Promotionsprogramm Biologie
der Georg-August University School of Science (GAUSS)

vorgelegt von
Bernd Wemheuer
aus Northeim
Göttingen, 2013

Betreuungsausschuss

Prof. Dr. Rolf Daniel, Genomische und Angewandte Mikrobiologie und Göttinger Genomlabor, Institut für Mikrobiologie und Genetik

Prof. Dr. Stefanie Pöggeler, Genetik eukaryotischer Mikroorganismen, Institut für Mikrobiologie und Genetik

Mitglieder der Prüfungskommission

Referent:	Prof. Dr. Rolf Daniel, Genomische und Angewandte Mikrobiologie und Göttinger Genomlabor, Institut für Mikrobiologie und Genetik
Korreferentin:	Prof. Dr. Stefanie Pöggeler, Genetik eukaryotischer Mikroorganismen, Institut für Mikrobiologie und Genetik

Weitere Mitglieder der Prüfungskommission:

Jun.-Prof. Dr. Kai Heimel, Mikrobielle Zellbiologie, Institut für Mikrobiologie und Genetik

Prof. Dr. Burkhard Morgenstern, Bioinformatik, Institut für Mikrobiologie und Genetik

PD Dr. Christoph Scherber, Agrarökologie, Department für Nutzpflanzenwissenschaften

Prof. Dr. Stefan Vidal, Agrarentomologie, Department für Nutzpflanzenwissenschaften

Tag der mündlichen Prüfung: 21. 01. 2014

“We have to remember that what we observe is not nature itself, but nature exposed to our method of questioning”

Werner Heisenberg

(German physicist and Noble Price laureate)

TABLE OF CONTENTS

ABBREVIATIONS

AAP	aerobic anoxygenic photosynthesis
ACE	ACE richness estimator
ANOVA	Analysis of Variance
a.s.l.	above sea line
BLAST	Basic Local Alignment Search Tool
BLASTN	BLAST search using a nucleotide query
bp	base pairs
°C	degree Celsius
ca.	circa
cm	centimetre
CO	carbon monoxide
CTD	conductivity, temperature, depth
D.	*Dactylis*
DEPC	diethylpyrocarbonate
DGGE	denaturing gradient gel electrophoresis
DMSP	dimethylsulfoniopropionate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DOC	dissolved organic carbon
DOM	dissolved organic matter
E	East
EDTA	Ethylenediaminetetraacetic acid
e.g.	exempli gratia, for example
EU	European Union
EU-27	Member of the European Union
F.	*Festuca*
Fig.	figure
Figs.	figures
g	gram
GOS	Global Ocean Sampling
GrassMan	Grassland Management Experiment

h	hour(s)
ha	hectare
i.e.	id est, that is
k	kilo
L	litre
L.	*Lolium*
m	meter
µ	micro
min	minute
mm	millimetre
mRNA	messenger RNA
N	North
NAC	North Atlantic cluster
NCBI	National Centre for Biotechnology Information
NGS	next generation sequencing
$n_{max.}$	maximal OUT number
NPK	Fertilizer containing Nitrogen, Phosphorous, and Potassium
nt	nucleotides
O.	*Octadecabacter*
ORF	open reading frame
OTU	operational taxonomic unit
P.	*Planktomarina*
PCR	polymerase chain reaction
PFLA	Phospholipid Fatty Acid
pH	power of hydrogen
POC	particulate organic carbon
POM	particulate organic matter
QIIME	Quantitive Insights Into Microbial Ecology
R.	*Roseobacter*
RCA	*Roseobacter* clade affiliated
rDNA	DNA coding for ribosomal DNA
rRNA	ribosomal RNA

PCoA	Principal Coordinate Analysis
PGPR	Plant Growth Promoting Rhizobacteria
RNA	ribonucleic acid
rpm	rotations per minute
rRNA	ribosomal RNA
sec	second
SSU	Small Subunit of the Ribosome
Tab.	table
TAE	tris-acetate-EDTA
Taq	*Thermus aquaticus*
TE	Tris EDTA
Tr.	Treatment
Tris	tris(hydroxymethyl)aminomeethane
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
USA	United States of America
UV	ultraviolet
V	volt
vs.	versus
yr	year

Abbreviations for nucleotides

A	adenine
C	cytosine
G	guanine
R	purines (adenine or guanine)
T	thymine (5-methyluracil)
Y	pyrimidines (cytosine, thymine or uracil)

LIST OF RELEVANT PUBLICATIONS

1) **Wemheuer B**, Wemheuer F, and Daniel R. RNA-Based Assessment of Diversity and Composition of Active Archaeal Communities in the German Bight. *Archaea*,vol. 2012, Article ID 695826, 8 pages, 2012. doi:10.1155/2012/695826.

2) Wemheuer B, Güllert S, Billerbeck S, Giebel H-A, Voget S, Simon M, and Daniel R (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. FEMS Microb Ecol 87: 378–389. doi: 10.1111/1574-6941.12230

3) **Wemheuer B**, Meier D, Klempert P, Billerbeck S, Giebel HA, Scherber C, Simon M, and Daniel R. From Germany to Norway: simultaneous assessment of total and active bacterial community structures in the North Sea along a latitudinal gradient. *(in preparation)*

4) Kanukollu S, **Wemheuer B**, Herber J, Billerbeck S, Lucas J, Daniel R, Simon M, Cypionka H, and Engelen B: Community analysis of free-living and attached *Roseobacter*-affiliated bacteria indicate an overlap in diversty between matured particles and the sediment surface. *(in preparation)*

5) Voget S, **Wemheuer B**, Brinkhoff T, Vollmers J, Dietrich S, Giebel HA, Beardsley C, Bakenhus I, Billerbeck S, Daniel R, and Simon M**:** Adaptation of an abundant Roseobacter RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. ISME J 2014 AUG 1. doi:10.1038/ismej.2014.134

6) Simon M, Billerbeck S, Brinkhoff T, Dogs M, Müllenmeister S, Seibt M, Seidel M, Smits M, Wagner-Döbler I, Wang H, **Wemheuer B**, and Wurst M (2012) Composition and activity of the bacterioplankton communities in the Drake Passage and Antarctic Peninsula region with a special emphasis on the Ro*seobacter* clade and dissolved organic matter. In: Lucassen M (ed) The expedition of the Research Vessel "Polarstern" to the Antarctic in 2012 (ANT-XXVIII/4) Reports on Polar and Marine Research 653: 49-54.

7) Sahm K**,** John P, Nacke H, **Wemheuer B**, Grote R, Daniel R, and Antranikian G (2013) High abundance of heterotrophic prokaryotes in hydrothermal springs of the Azores as revealed by a network of 16S rRNA gene-based methods. Extremophiles 17: 649-662.

8) Krohn-Molt I, **Wemheuer B**, Alawi M, Poehlein A, Güllert S, Schmeisser C, Pommerening-Röser A, Grundhoff A, Daniel R, Hanelt D, and Streit WR. Metagenome survey of a multispecies and algae-associated biofilm reveals key elements of bacterial-algae interactions in photobioreactors. Appl Environ Microbiol 79 (20) 6196-6206. doi: 10.1128/AEM.01641-13

9) Wemheuer F, Kretzschmar D, **Wemheuer B**, Daniel R, and Vidal S. Impact of grassland management regimes on bacterial endophyte diversity differs with grass species and season. *(in preparation)*

CHAPTER A

SUMMARY

Marine *Bacteria* and *Archaea* represent the major part of the total biodiversity on Earth. They are main drivers of biogeochemical cycles in marine environments and play an integral role in ecosystem structuring. Therefore, the understanding how marine prokaryotic communities are structured and how these communities react towards altered environmental conditions is of fundamental importance. In this thesis, the diversity and ecology of *Bacteria* and *Archaea* in the pelagic realm of the North Sea was studied using different culture-independent approaches. In particular, the abundance and ecological role of the *Roseobacter* clade was analyzed. Moreover, the response of the investigated prokaryotic communities to changing environmental conditions was examined.

One objective was to analyze active archaeal and bacterial community structures in the southern North Sea. Moreover, the impact of a phytoplankton bloom on prokaryotic community composition and diversity was elucidated. For this purpose, 14 marine water samples were collected in May 2010 in the German Bight in and outside of a phytoplankton bloom. Community structures were assessed by pyrotag sequencing of 16S rRNA transcripts. A total of 62,045 and 211,769 16S rRNA sequences were used to examine archaeal and bacterial community composition and diversity, respectively. Different marine lineages of the *Gamma-* and *Alphaproteobacteria* including the SAR92 clade and the *Roseobacter* clade affiliated (RCA) cluster dominated the bacterial community. The archaeal community mainly composed of *Halobacteria* as well as minor proportions of *Thermoplasmata* and members of the Marine Group I.

Both archaeal and bacterial community structures were influenced by the presence of the examined phytoplankton bloom. In addition to structural changes, the bacterial richness was reduced in bloom presence indicating that only certain bacteria could thrive during the bloom. To exploit the impact of the phytoplankton bloom on gene expression, environmental mRNA from three samples was extracted, enriched, and sequenced resulting in a total of 988,022 sequences for all three samples. Changes in various metabolic pathways including photosynthesis and protein metabolism were recorded as response to the bloom. The observed changes were induced either directly by the bloom or indirectly by altered environmental parameters as most of these parameters were significantly correlated to bloom presence.

Another objective was to assess the structure and diversity of total and active bacterial communities along a latitudinal gradient ranging from the German Bight to the Norwegian coast. For this purpose, surface water samples were taken in July 2011 on board of the research vessel Heincke. Community structures were examined by pyrotag sequencing of 16S rRNA genes and transcripts. A total of 382,507 16s rRNA gene sequences were used to assess bacterial community structures. These communities were dominated by different groups affiliated to the *Bacteroidetes* and *Alphaproteobacteria* including the marine groups NS5, NS7, and NS9, different subclusters of the *Roseobacter* clade, and members of the SAR116 clade. Moreover, *Cyanobacteria* were found in minor abundance with *Synechococcus* as the dominant genus.

Total and active bacterial community structures were significantly altered along the latitudinal gradient whereas the bacterial richness was unaffected. The abundance of certain marine groups such as the *Roseobacter* CHAB-I-5 and NAC11-7 clusters either decreased or increased with rising latitude. The recorded changes in bacterial community structures along the studied gradient might be explained by changing environmental conditions which are probably caused by the different water currents found along the gradient. Additionally, the composition of the *Rhodobacteraceae* in the free-living and in the particle-associated fraction in different depths as well as in the sediment along the latitudinal gradient was investigated. The composition of this family significantly varied between the different fractions.

In addition to the overall bacterial community structure, the ecological contribution of an abundant member of the RCA cluster, *Planktomarina temperata*, to nutrient cycling in the pelagic realm of the North Sea was analyzed. This cluster exhibited more than 20% relative abundance in samples derived from the first sampling. Consequently, the largest mRNA dataset of this cruise and a metagenome obtained by direct sequencing of gDNA from the same sample was mapped on its genome. Around one million pyrosequencing reads were mapped in total. Large parts of the genome were retrieved by the mapping approach with approximately 86% of the genome transcribed. Active features included the photosynthetic operon, the two CO dehydrogenase systems, and the assimilatory sulfate reduction pathway.

In addition to marine ecosystems, the diversity of other prokaryotic communities in three non-marine habitats was examined. In the first survey, we studied the prokaryotic community in two hyperthermal springs on the Azores. Environmental DNA was extracted from two sediment samples and further analyzed using different metagenomic methods. The first spring was dominated by heterotrophic bacterial genera including *Caldicellulosiruptor*, *Dictyoglomus*, and *Fervidobacterium* as well as by the chemolithoautotrophic genus *Sulfurihydrogenibium*. The archaeal community comprised *Crenarchaeota* with *Thermoproteaceae* and *Desulfurococcaceae* as the dominant families. The bacterial community in the other spring was mainly composed of the heterotrophic genus *Acidicaldus* and the chemolithoautotrophic genus *Acidithiobacillus*.

In the second study, the diversity and genomic potential of a bacterial biofilm associated with two freshwater algae, *Chlorella vulgaris* and *Scenedesmus obliquus*, in a photobioreactor was investigated. DNA was extracted from collected biofilm samples and analyzed using different metagenomic approaches. The recorded community size was rather limited with approximately 30 bacterial species. The majority of the observed *Bacteria* were affiliated to *Alphaproteobacteria*, *Betaproteobacteria*, and *Bacteroidetes*. Analysis of the genomic potential revealed a high metabolic diversity with respect to the utilization of polymers, aromatic, and non-aromatic compounds.

In the third study, the bacterial endophytic community structures in three agricultural important grasses, *Dactylis glomerata*, *Festuca rubra*, and *Lolium perenne*, was examined. Moreover, the response of these communities to fertilizer application, mowing frequency, and seasonal changes over two consecutive years was investigated. Endophytic community structures were studied by DGGE-based analysis of 16S rRNA genes amplified from plant DNA. The communities were dominated by different members of the *Firmicutes* and *Proteobacteria* with *Bacillus* and *Pseudomonas* as the predominate genera. The studied management regimes significantly altered the endophytic communities in *L. perenne* and *F. rubra* but not in *D. glomerata*. On the other hand, season significantly affected the community structures in all three grasses. Moreover, as community structures were subjected to seasonal variations, the recorded impact of management regimes differed between the two investigated years.

CHAPTER B

GENERAL INTRODUCTION

"Ocean: A body of water occupying about two-thirds of a world made for man - who has no gills."

Ambrose Bierce (American journalist and writer)

1 DIVERSITY AND ECOLOGY OF MARINE MICROORGANISMS

1.1 OCEANIC DIVISIONS

The world ocean covers approximately 71% of the Earth's surface and contains more than 95% of the global water. Moreover, it is the planet's largest ecosystem. In general, the ocean can be divided into the pelagic realm (water column) and the benthic realm (sea floor) (Fig. 1). With respect to depth and light penetration, the pelagic realm is vertically divided into a photic and an aphotic zone. The photic zone is the zone where light is abundant, photosynthesis can occur, and thus the major part of marine primary production happens. As a consequence, life in the aphotic zone relies on other energy sources such as material sinking from the surface to the ground (marine snow) or hydrothermal vents. In addition to its vertical separation, the pelagic realm is horizontally divided into a neritic (coastal regions) and an oceanic zone (open ocean). The neritic zone exhibits a different nutrient composition and availability than the oceanic zone due to upwelling ground water or the inflow of rivers. The benthic realm is subdivided into different zones on the basis of depth, which extend from high tide to the continental shelf and deep-sea zones.

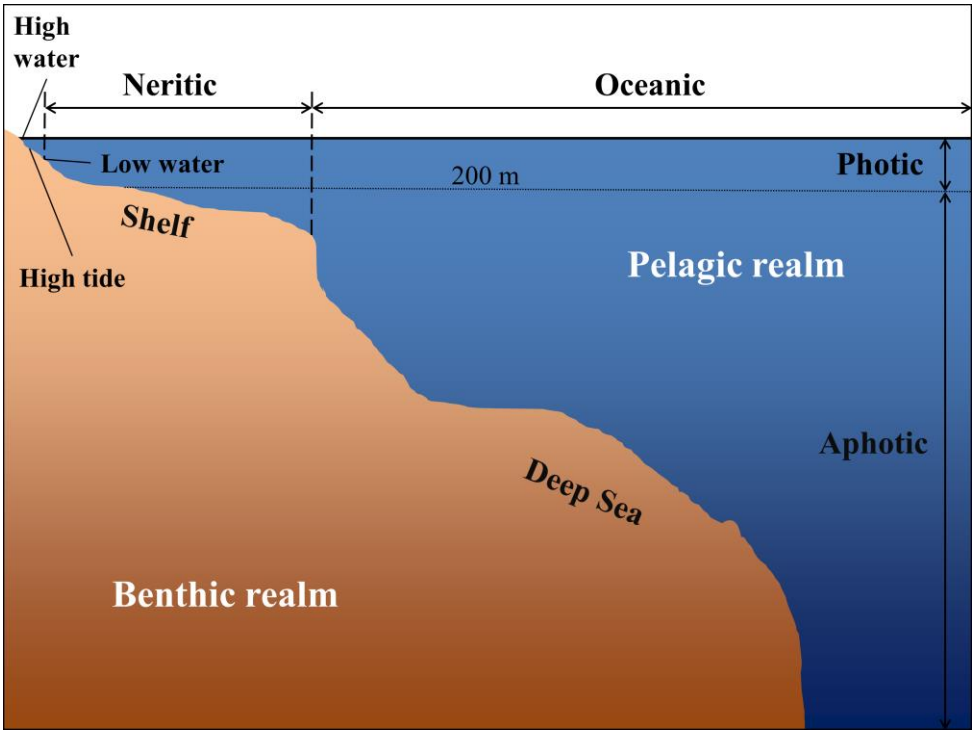


Figure 1: Oceanic divisions (© Bernd Wemheuer).

1.2 STUDY AREA

The studies presented in this thesis were mainly focused on the analysis of microbial communities in the pelagic realm of the North Sea. The North Sea is a marginal of the Atlantic Ocean on the European continental shelf. Coastal shelf areas of the temperate zone are highly productive because of the continuous nutrient supply by rivers. The counter-clockwise circulation of the water masses is caused by the main currents inflowing from the English Channel and the Norwegian trench (Fig. 2). During the last 40 years, the North Sea, in particular its southern region underwent high nutrient loading and warming (McQuatters-Gollop *et al.*, 2007; Wiltshire *et al.*, 2010).

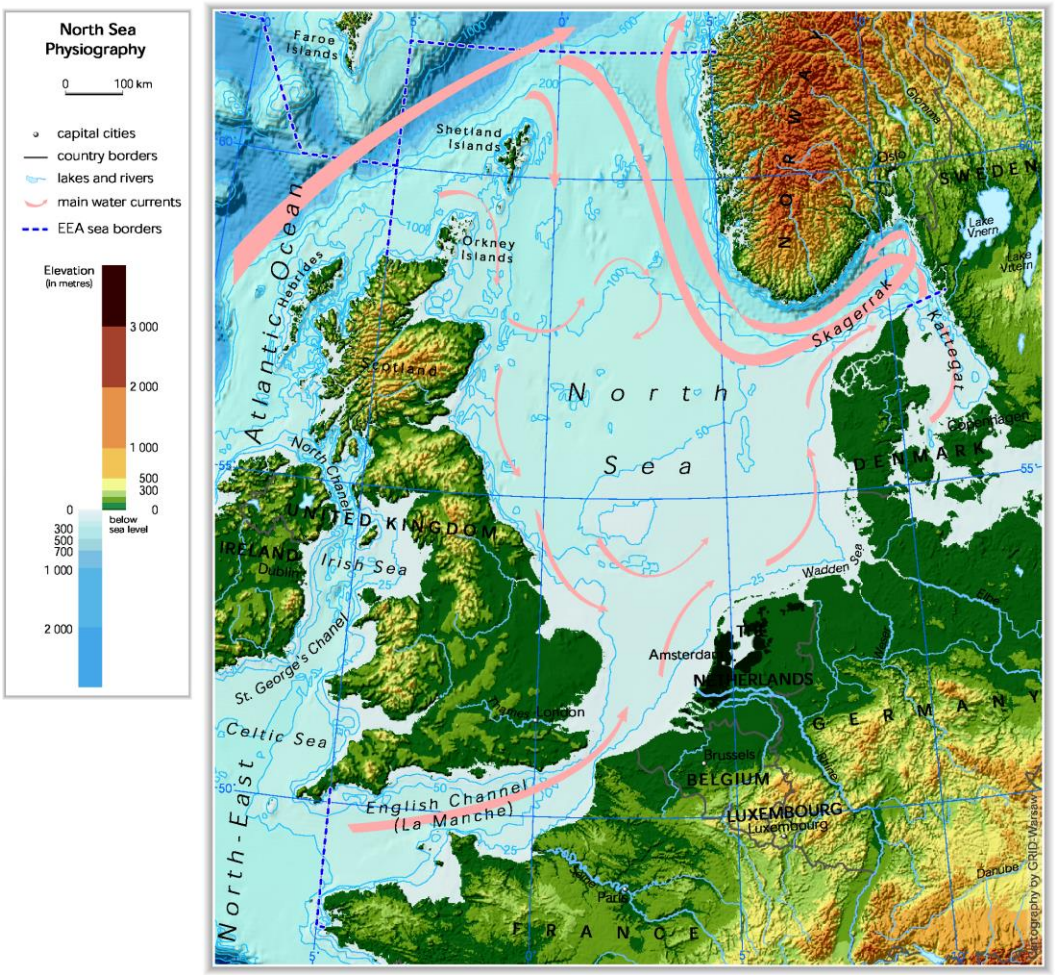


Figure 2: Physiography of the North Sea (© European Environment Agency).

1.3 MARINE MICROBIAL ECOLOGY

According to the National Oceanic and Atmospheric Administration (USA), the world ocean harbors an enormous biodiversity and promotes the life of nearly 50 percent of all species on Earth (http://www.noaa.gov/ocean.html). Although all marine ecosystems are generally characterized by high salinities (around 3.5%) and low nutrient availability, various ecological niches are formed. All of these niches are colonized by marine microbes. The biology in all marine ecosystems is strongly influenced by different factors such as temperature and salinity. Moreover, anthropogenic factors such as marine pollution or overfishing changed the structures of these ecosystems during the last century. In addition large parts, especially the deep sea, remain poorly examined.

Microbes (organisms from 0.2 to 100µm) account for 50% of the global biomass. These bacteria, archaea, protists, and unicellular fungi are responsible for 98% of the primary production (Whitman *et al.*, 1998) and mediate all biogeochemical cycles in the ocean (Atlas & Bartha, 1993). One example is the microbial contribution to the carbon cycle. A large fraction of the primary products formed is dissolved (dissolved organic matter; DOM) by various mechanisms in the food web (Fig. 3). This part of the primary production is almost exclusively accessible to heterotrophic *Bacteria* and *Archaea* (Azam & Malfatti, 2007). Therefore, marine microbes are important for global nutrient cycling. During the last years, there has been an increasing interest in the ecology of these important organisms and their active contribution to ecosystem functioning.

The overall role of marine microorganism in different marine ecosystems has been frequently addressed over the last years. However, the active contribution of single microbial clades or lineages is still poorly understood mainly due to the non-cultivability of these organisms. Currently available isolates hardly represent marine microbial diversity. Most marine microorganisms are uncharacterized both genetically and biochemically (Rusch *et al.*, 2007). As a consequence, it is difficult to link individual organisms, clusters, or phyla to certain biogeochemical processes and ecosystem services (Heidelberg *et al.*, 2010).

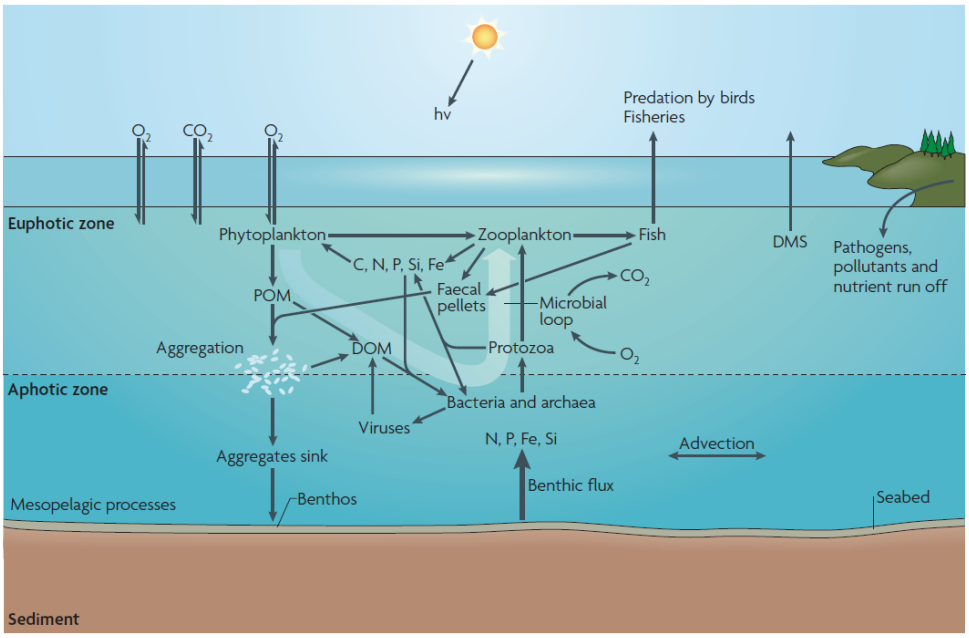


Figure 3: Microbial structuring of a marine ecosystem (taken from (Azam & Malfatti, 2007). (DMS = dimethyle sulfide; DOM = dissolved organic matter; POM particulate organic matter)

1.4 CULTURE-INDEPENDENT ANALYSIS OF MARINE MICROBIAL COMMUNITIES

It has been calculated that one milliliter of oceanic sea water contains up to one million microorganisms (Curtis *et al.*, 2002). Most of these organisms cannot be studied as they are reluctant to common culturing techniques. As a consequence, cultivation-independent approaches are the only possibility to assess and exploit both the diversity and the ecology of marine microbial communities. Studies using these approaches have greatly advanced our understanding of marine microbial diversity and ecology (e.g., Venter *et al.*, 2004, Giovannoni & Stingl, 2005). Early studies used traditional techniques such as Sanger sequencing-based analysis of 16S rRNA gene libraries to examine bacterial community structures (e.g., Giovannoni *et al.*, 1990, Britschgi & Giovannoni, 1991). However, these studies failed to investigate the full extent of microbial diversity due the size and complexity of marine microbial communities. The advent of next-generation sequencing revolutionized metagenomic research and

has been applied for in-depth investigation of bacterial communities in almost all marine ecosystems (e.g., Bolhuis & Stal, 2011, Lee *et al.*, 2011, Thompson *et al.*, 2011)). However, most studies on marine diversity and ecology applied DNA-based approaches (e.g., Schmidt *et al.*, 1991, Selje *et al.*, 2004, Giebel *et al.*, 2009). Therefore, these surveys assessed the diversity of the total microbial community but were not able to distinguish between active and inactive fraction. The application of culture-independent approaches using environmental RNA and proteins as starting material (Fig. 4) and thus targeting the active fraction of the microbial community provided novel and unique insights into diversity and ecological contribution of marine microbes (e.g., Teeling *et al.*, 2012, Wemheuer *et al.*, 2012, Wemheuer *et al.*, 2013).

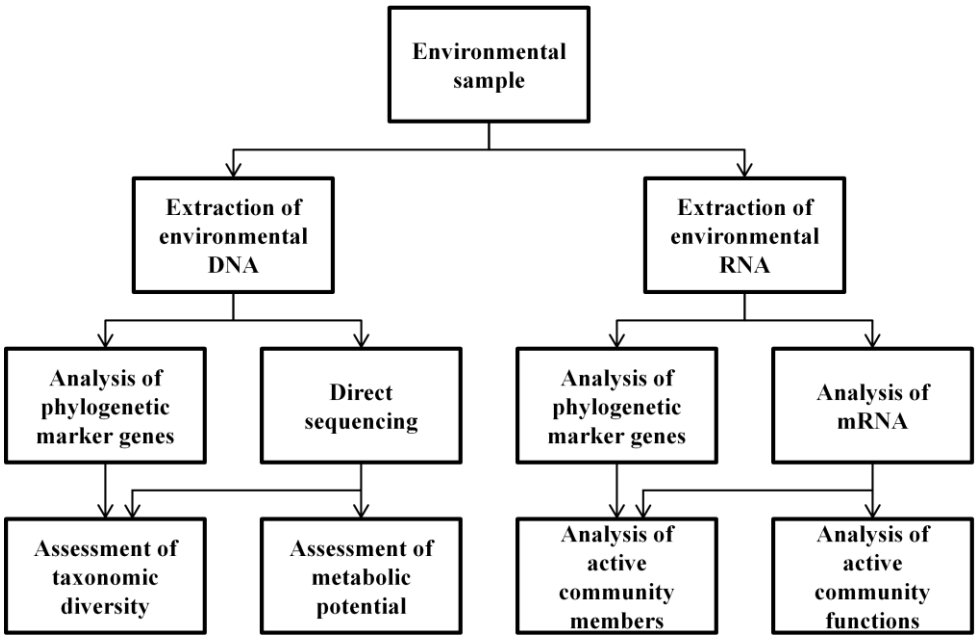


Figure 4: Metagenomic and metatranscriptomic ways to investigate environmental samples (modified from Simon & Daniel, 2011).

1.5 STRUCTURE OF MARINE MICROBIAL COMMUNITIES WITH FOCUS ON THE ROSEOBACTER CLADE

Certain marine clades are ubiquitous and abundant in almost all marine environments (Fig. 5). One of these groups is the *Roseobacter* clade, a monophyletic group within the *Rhodobacteraceae* (*Alphaproteobacteria*). Members of this clade have been found in almost every marine environment including polar regions (Brinkmeyer *et al.*, 2003). They constitute up to 25% of all marine prokaryotes (Buchan *et al.*, 2005). Members of the *Roseobacter* clade have been identified as biofilm-forming, free-living, or as symbionts of phytoplankton and sponges (as reviewed in Brinkhoff *et al.*, 2008).

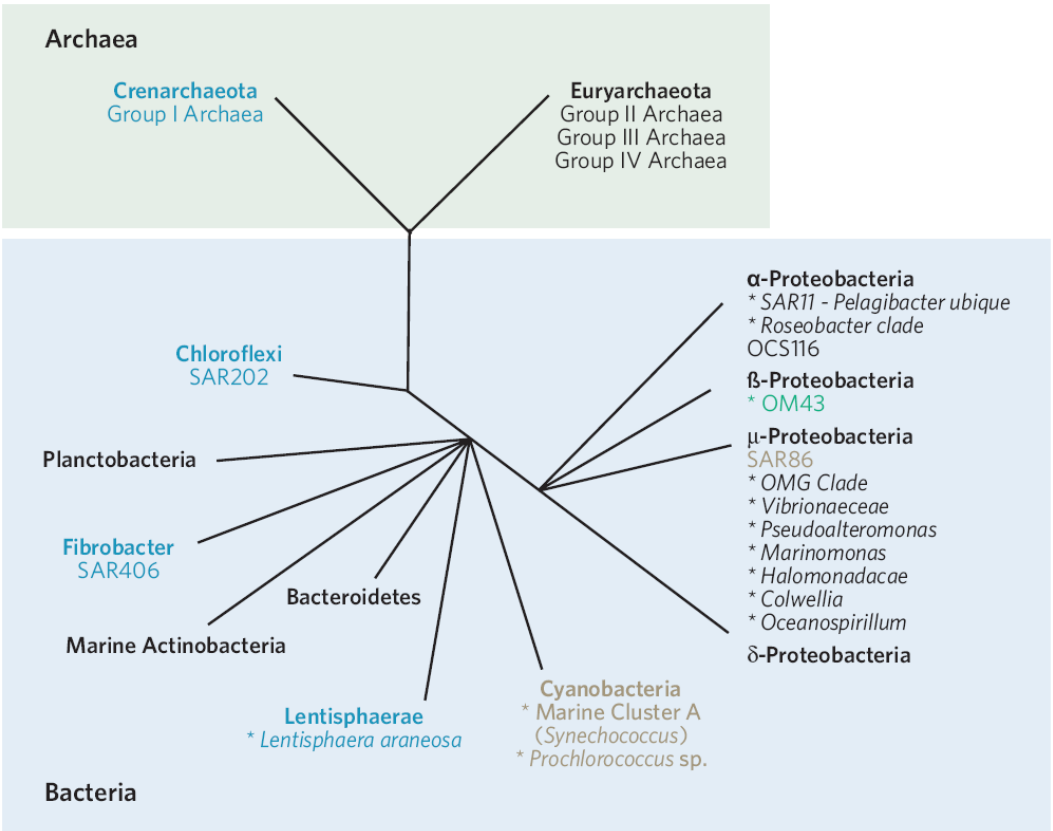


Figure 5: Main phyla of marine pelagic prokaryotes. Ubiquitous groups are written in black (Giovannoni & Stingl, 2005).

The use of a multitude of organic compounds, the production of secondary metabolites, and other metabolic pathways such as aerobic anoxygenic photosynthesis (AAnP) or carbon monoxide (CO) oxidation are thought to contribute to the success of the *Roseobacter* clade in marine environments. This knowledge is mainly based on the high numbers of available isolates and the analysis of the corresponding genomes. Several of these genomes have been investigated during the past years (Wagner-Döbler *et al.*, 2010, Kalhoefer *et al.*, 2011, Vollmers *et al.*, 2013). These studies found several novel traits which might also contribute to the success of the *Roseobacter* clade in marine ecosystems including several potential heavy metal resistance genes and xanthorhodopsins (Kalhoefer *et al.*, 2011, Vollmers *et al.*, 2013). However, most of the investigated isolates do not represent the predominate subclusters of the *Roseobacter* clade. What is responsible for the success of this clade in marine environments is still not fully understood.

2 DIVERSITY AND ECOLOGY OF MICROORGANISMS IN OTHER ECOSYSTEMS

In addition to marine ecosystems, the diversity and ecology of microbes in different natural and artificial habitats was studied, i.e., two geothermal springs derived from the Azores, a photobioreactor, and different agricultural important grass species

2.1 GEOTHERMAL SPRINGS

Sites of volcanic activity can be found all over the world. With regard to physicochemical conditions, they provide a variety of different ecological niches for microorganisms. These niches exhibit extreme environmental conditions, such as high temperature, low or high pH, and the presence of toxic ions. One well studied example of such a niche are hydrothermal springs (Hugenholtz *et al.*, 1998, MEYER-DOMBARD *et al.*, 2005, Kvist *et al.*, 2007, Wemheuer *et al.*, 2013). Their extreme features are expected to lead to limited biodiversity, thus, making hydrothermal habitats ideal model systems to study principles of community structure and function (Sahm *et al.*, 2013). Moreover, extremophilic microorganisms inhabiting hydrothermal springs are considered to be the closest living descendants of the earliest life forms on Earth (Olsen *et al.*, 1994).

2.2 PHOTOBIOREACTORS

Rising carbon dioxide levels in the atmosphere and the global energy crisis resulted in an increasing demand for alternative energy sources. Algal biofuel produced by microalgae is a promising candidate to supply this demand, but the process is still some way from being carbon neutral or commercially viable (Scott *et al.*, 2010). Microalgae comprise a large and phylogenetically very heterogeneous group of microorganisms. They perform oxygenic photosynthesis and thereby convert CO_2 to high amounts of cellular lipids as potential source of biofuels (Abomohra *et al.*, Chisti, 2007, Gouveia & Oliveira, 2009, Williams &

Laurens, 2010, Wiley *et al.*, 2011). Since photobioreactors allow the continuous cultivation of microalgae under relatively controlled conditions, they have been in the focus of many research projects often linked to large-scale biofuel production (Chisti, 2007, Ugwu *et al.*, 2008, Scott *et al.*, 2010, Williams & Laurens, 2010). However, one fundamental problem of biofuel production is that there is still little background knowledge established (Scott *et al.*, 2010). This knowledge might enable us to solve frequent problems in algal farming such as biofouling and the growth of microbial biofilms in the reactor. Moreover, it can help us to increase yield and quality of the growing algae.

2.3 ENDOPHYTIC BACTERIA

Almost all plant species are colonized by microorganisms including endophytic bacteria (Senthilkumar *et al.*, 2011). Endophytic bacteria are defined as bacteria that can be isolated from within plants or from surface-disinfested plant tissue, and that have no apparent harmful effects on the plant (Hallmann *et al.*, 1997). Endophytic bacteria have been reported to promote plant growth, plant yield, and the overall plant health by a number of mechanisms. This includes the production of phytohormones and antibiotics (Bacon & Hinton, 2006, Compant *et al.*, 2010) as well as enhanced nutrient availability and nitrogen fixation (Stoltzfus *et al.*, 1997, Rosenblueth & Martinez-Romero, 2006). Moreover, plants infected with endophytic bacteria have a higher resistance to plant pathogens (e.g., (Hallmann *et al.*, 1998, Krechel *et al.*, 2002, Siddiqui & Shaukat, 2003, Compant *et al.*, 2005) and environmental stresses (e.g., (Sturz & Nowak, 2000, Bacon & Hinton, 2006, Bacon & Hinton, 2011). Many biotic factors including plant age or tissue as well as abiotic factors such as soil conditions, temperature, or crop rotation influence the bacterial endophytic community (Hallmann *et al.*, 1997, Fuentes-Ramírez *et al.*, 1999, Sessitsch *et al.*, 2002, Seghers *et al.*, 2004, Hardoim *et al.*, 2012). Although their important role in agricultural cropping systems is frequently appreciated (e.g., (Hallmann *et al.*, 1997, Kobayashi & Palumbo, 2000, Bacon & Hinton, 2006, Maksimov *et al.*, 2011, Senthilkumar *et al.*, 2011), the effect of

different management regimes and seasonal changes on the overall bacterial endophytic community in different grass species has not been studied yet.

3 FOCUS OF THIS THESIS

This main objective of this thesis was to investigate the diversity and ecology of microbial communities in marine ecosystems (Fig. 7). In addition, microbes in other habitats were studied (Fig. 7). To gain insights into the diversity and ecology of the examined microbial communities, different culture-independent approaches were employed. Metagenomic (DNA-based) methods were used to assess phylogenetic structures and to exploit the genomic potential of the studied communities. In addition, metatranscriptomic (RNA-based) approaches were used to investigate the active microbial community fraction. With regard to the high diversity of the studied ecosystems, next generation sequencing techniques were employed as these approaches provide a more accurate picture of microbial community structure and potential.

The first two studies were focused on the analysis of samples taken on board of the RV Heincke in May 2010. The aim of this cruise was to examine the diversity of archaeal and bacterial communities in the southern North Sea. Moreover, the influence of a phytoplankton bloom on the ambient microbial community was studied. The next two surveys analyzed samples taken along a latitudinal gradient in the North Sea on board of the RV Heincke in July 2011. Bacterial communities were analyzed with respect to changes of environmental conditions and geographical parameters. In addition, the composition of the *Roseobacter* clade and other *Rhodobacterales* in different consortia (free-living, particle-associated, sediment) and different depths (from surface to sediment) was investigated. The fifth study investigated the abundance and ecological function of an abundant *Roseobacter* clade isolate, *Planktomarina temperata*. The sixth study is article about the sampling on board of the RV Polarstern in the Southern Ocean. The last three surveys were focused on the analysis of other, non-marine samples. In the first one, the diversity and ecology of *Archaea* and *Bacteria* in two hyperthermal springs on the Azores were studied using different metagenomic approaches. In the second one, structure and genomic potential of a bacterial community in a photobioreactor were investigated to better understand the dynamics and function of this artificial habitat. The last study aimed at assessing the bacterial endophytic community in three agricultural important grass species

and to evaluate the impact of different management regimes and season changes on these communities.

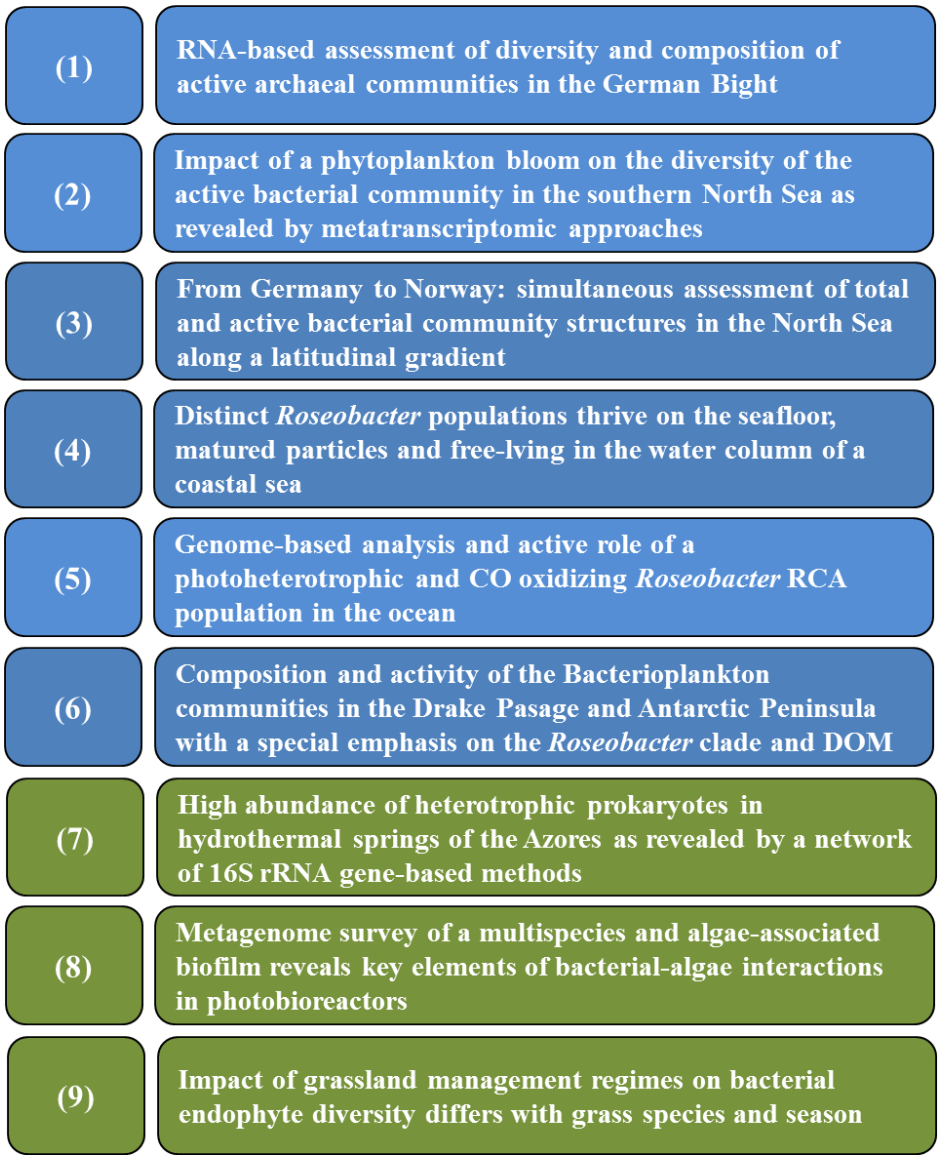


Figure 7: Overview about the studies presented in this thesis. All marine studies are highlighted in blue, non-marine studies in green.

CHAPTER C

PUBLICATIONS

STUDY 1:

RNA-BASED ASSESSMENT OF DIVERSITY AND COMPOSITION OF ACTIVE ARCHAEAL COMMUNITIES IN THE GERMAN BIGHT

WEMHEUER B[1], WEMHEUER F[2], AND DANIEL R[1]

ARCHAEA, VOL. 2012, ARTICLE ID 695826
(doi:10.1155/2012/695826)

[1]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY; [2]DEPARTMENT FOR CROP SCIENCES, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 6, D-37077 GÖTTINGEN, GERMANY

Author contributions to the work:

Performed the experiments: BW

Analyzed data: BW

Contributed data on water properties and analysis of these data: BW

Wrote the publication: BW, FW, RD

Conceived and designed the experiments: BW, RD

Hindawi Publishing Corporation
Archaea
Volume 2012, Article ID 695826, 8 pages
doi:10.1155/2012/695826

Research Article

RNA-Based Assessment of Diversity and Composition of Active Archaeal Communities in the German Bight

Bernd Wemheuer,[1] Franziska Wemheuer,[2] and Rolf Daniel[1]

[1] *Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany*
[2] *Section of Agricultural Entomology, Department for Crop Sciences, Georg-August University of Göttingen, Grisebachstraße 6, 37077 Göttingen, Germany*

Correspondence should be addressed to Rolf Daniel, rdaniel@gwdg.de

Received 6 September 2012; Accepted 2 October 2012

Academic Editor: Michael Hoppert

Archaea play an important role in various biogeochemical cycles. They are known extremophiles inhabiting environments such as thermal springs or hydrothermal vents. Recent studies have revealed a significant abundance of *Archaea* in moderate environments, for example, temperate sea water. Nevertheless, the composition and ecosystem function of these marine archaeal communities is largely unknown. To assess diversity and composition of active archaeal communities in the German Bight, seven marine water samples were taken and studied by RNA-based analysis of ribosomal 16S rRNA. For this purpose, total RNA was extracted from the samples and converted to cDNA. Archaeal community structures were investigated by pyrosequencing-based analysis of 16S rRNA amplicons generated from cDNA. To our knowledge, this is the first study combining next-generation sequencing and metatranscriptomics to study archaeal communities in marine habitats. The pyrosequencing-derived dataset comprised 62,045 archaeal 16S rRNA sequences. We identified *Halobacteria* as the predominant archaeal group across all samples with increased abundance in algal blooms. *Thermoplasmatales* (*Euryarchaeota*) and the Marine Group I (*Thaumarchaeota*) were identified in minor abundances. It is indicated that archaeal community patterns were influenced by environmental conditions.

1. Introduction

It has been calculated that one mL of oceanic sea water contains up to 10^6 different microorganisms [1]. These archaea, bacteria, protists, and unicellular fungi contribute 98% to the primary biomass production and are involved in almost all biogeochemical cycles [2]. It has been estimated that the global ocean harbors approximately 1.3×10^{28} archaeal cells and 1.3×10^{28} bacterial cells, which together constitute 63% to 90% of the entire marine picoplankton [3]. In addition, high numbers of *Archaea* have been found in marine sediments [4].

In contrast to their relatives living in extreme environments, little is known on marine *Archaea.* This is partly due to the unavailability of pure cultures. Marine *Archaea* might be involved in the oceanic nitrogen cycle as some marine *Crenarchaeota* are capable of nitrification [5]. However, our knowledge of the archaeal role in oceanic ecology is rudimentary and their influence on global biogeochemical cycles is largely unexplored [6].

Culture-independent approaches have greatly advanced our knowledge of the diversity and ecology of marine microbial communities [7–9]. Next-generation sequencing (NGS) contributed to this advancement. For example, many different ecosystems such as soil [10, 11] or sea water [12] have been studied by DNA-based high throughput sequencing of 16S rRNA gene fragments and analysis of the obtained sequences. The main drawback of DNA-based metagenomic approaches is the inability to distinguish between active and inactive community members.

Active members and functions of microbial communities are accessible by employing RNA-based metatranscriptomic approaches. For example, Urich et al. [13] analyzed the composition and metabolic potential of active soil microbial

communities by sequencing of reverse transcribed total RNA. Other studies analyzed gene expression in ocean surface waters [8] or in a deep-sea hydrothermal plume [14]. However, mainly bacterial communities and their capabilities were analyzed in these studies.

In this paper, we investigated the composition of active archaeal communities in surface water derived from the southeastern part of the North Sea, the German Bight. The northwest of the German Bight is separated from the remaining North Sea by the Doggerbank, a large sandbank. Large coastal parts of the bight are shallow with water depths of approximately 2 to 12 meters. In our investigation, we collected seven water samples at different locations and depths in these shallow offshore areas.

The aim of our study was to assess the active archaeal community structures in the southern North Sea employing next-generation sequencing of 16S rRNA amplicons generated by reverse transcription polymerase chain reaction (RT-PCR). To our knowledge, this is the first study using this combined approach to study marine archaeal communities.

2. Material and Methods

2.1. Sampling and Sample Preparation. Seven marine water samples were taken for archaeal community analysis. Approximately 50 liters of sea water per sampling site were collected on board of the research vessel Heincke in May 2010 employing a conductivity, temperature, and depth (CTD) profiler. All sites were located in the German Bight. Sea water samples were prefiltered through a 10 μm-mesh-size nylon net and a filter sandwich consisting of a precombusted (4 h at 450°C) 47 mm-diameter glass fiber filter (Whatman GF/D; Whatman, Maidstone, UK) and a 47 mm-diameter (pore size 3.0 μm) polycarbonate filter (Nuclepore, Whatman). Bacterioplankton was harvested by filtration of 1 L prefiltered sea water through a filter sandwich consisting of a glass fiber filter (Whatman GF/F) and a 47 mm-diameter (pore size 0.2 μm) polycarbonate filter (Nuclepore, Whatman).

Additionally, marine phytoplankton samples were collected by employing a plankton net (pore size 55 μm). The composition of the algal community was determined by microscopy of the collected samples.

2.2. RNA Extraction and Purification. Total RNA was extracted as described by Weinbauer et al. [15]. One 47 mm-diameter filter (pore size 0.2 μm) was used per sample. Subsequently, RNA was purified employing the RNeasy Mini Kit as recommended by the manufacturer (Qiagen, Hilden, Germany).

To remove residual DNA from RNA samples, Ambions TURBO DNase (Invitrogen, Carlsbad, USA) was used according to the instructions of the manufacturer with one modification: subsequent to a standard reaction, 0.5 μL of TURBO DNase per 10 μg of RNA was added to the mixture, and incubation was performed at 37°C for 15 min. Phenol/Chloroform/Isoamyl alcohol (25 : 24 : 1) was used to inactivate the DNase.

The presence of remaining DNA was tested by PCR using the 16S rRNA gene as a target gene for amplification. The following two primer sets were employed: 8F/518R (5′- AGAGTTTGATCCTGGCTCAG-3′ [16] and 5′-ATTACCGCGGCTGCTGG-3′ [17]) and 1055F/1378R (5′-ATGGCTGTCGTCAGCT-3′ [18] and 5′-CGGTGTGTACA-AGGCCCGGGAACG-3′ [19]).

The PCR reaction mixture (25 μL) for amplification of the target gene contained 2.5 μL of 10-fold Mg-free *Taq* polymerase buffer (Fermentas, St. Leon-Rot, Germany), 200 μM of each of the four desoxynucleoside triphosphates, 1.75 mM $MgCl_2$, 0.4 μM of each primer, 1 U of *Taq* DNA polymerase (Fermentas), and approximately 100 ng of purified RNA sample as template. The following thermal cycling scheme was used: initial denaturation at 94°C for 2 min, 28 cycles of denaturation at 94°C for 1.5 min, annealing at 55°C for 1 min, followed by extension at 72°C for 40 s. The final extension was carried out at 72°C for 10 min.

2.3. Synthesis of cDNA from Total RNA. cDNA was synthesized from total RNA by employing the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen) with modifications of the first strand synthesis protocol: 10 μL of total RNA (up to 5 μg) were mixed with 1 μL of random hexamer primers (Roche, Mannheim, Germany) and 1 μL dNTP mixture containing 10 mM of each of the four desoxynucleoside triphosphates. The mixture was incubated for 10 min at 70°C and chilled on ice. Four μL 5x first-strand buffer, 1 μL of 0.1 M DTT, and 1 μL RNA protect (Fermentas) were added, and the reaction mixture was incubated for 2 min at 25°C. Subsequently, 1 μL of SuperScript II reverse transcriptase was added. The reaction was incubated for 10 min at 25°C and then for 1 h at 45°C. The generated cDNA was subjected to 16S rRNA PCR.

2.4. Amplification of 16S rRNA and Pyrosequencing. To analyze archaeal diversity, the V3–V5 region of the archaeal 16S rRNA was amplified by PCR. The PCR reaction (25 μL) contained 5 μL of 5-fold Phusion GC buffer (Finnzymes, Vantaa, Finland), 200 μM of each of the four desoxynucleoside triphosphates, 1.5 mM $MgCl_2$, 4 μM of each primer (Table 1), 2.5% DMSO, 1 U of Phusion High Fidelity Hot Start DNA polymerase (Finnzymes), and approximately 50 ng of cDNA. The following thermal cycling scheme was used: initial denaturation at 98°C for 5 min, 25 cycles of denaturation at 98°C for 45 s, annealing at 68°C for 45 s, followed by extension at 72°C for 30 s. The final extension was carried out at 72°C for 5 min. Negative controls were performed by using the reaction mixture without template. Primer sequences for amplification of the V3–V5 region [20] as well as 454 adaptors with the unique MIDs for each sample are listed in Table 1. The resulting PCR products were checked for appropriate size and then purified by using the peqGOLD Gel Extraction Kit (Peqlab, Erlangen, Germany) as recommended by the manufacturer. Three independent PCR reactions were performed per sample, purified by gel extraction, and pooled in equal amounts. Quantification of the PCR products was performed using the Quant-iT dsDNA

Table 1: Primers used for amplification of the V3–V5 region of the archaeal 16S rRNA [20].

Sample	Primer	Sequence (5′-3′)			
		454-Adaptor (Lip-A Kit)	Key	Unique MID	Archaeal 16S rRNA specific
655	ARC344F	CGTATCGCCTCCCTCGCGCCA	TCAG	ACTGTACAGT	ACGGGGYGCAGCAGGCGCGA
658	ARC344F	CGTATCGCCTCCCTCGCGCCA	TCAG	AGACTATACT	ACGGGGYGCAGCAGGCGCGA
659	ARC344F	CGTATCGCCTCCCTCGCGCCA	TCAG	AGCGTCGTCT	ACGGGGYGCAGCAGGCGCGA
660	ARC344F	CGTATCGCCTCCCTCGCGCCA	TCAG	AGTACGCTAT	ACGGGGYGCAGCAGGCGCGA
664	ARC344F	CGTATCGCCTCCCTCGCGCCA	TCAG	ATAGAGTACT	ACGGGGYGCAGCAGGCGCGA
670	ARC344F	CGTATCGCCTCCCTCGCGCCA	TCAG	CACGCTACGT	ACGGGGYGCAGCAGGCGCGA
671	ARC344F	CGTATCGCCTCCCTCGCGCCA	TCAG	CAGTAGACGT	ACGGGGYGCAGCAGGCGCGA
All	ARC915R	CTATGCGCCTTGCCAGCCCGC	TCAG	ACAGTATATA	GTGCTCCCCCGCCAATTCCT

BR Assay Kit and a Qubit fluorometer (Invitrogen) as recommended by the manufacturer. The Göttingen Genomics Laboratory determined the sequences of the 16S rRNA by using a Roche GS-FLX 454 pyrosequencer with Titanium chemistry (Roche, Mannheim, Germany).

2.5. Processing and Analysis of Pyrosequencing Derived Data Sets. Sequence data were deposited in the sequence read archive of the National Center for Biotechnology Information under accession number SRA056839. Generated 16S rRNA datasets were processed and analyzed employing the QIIME 1.4 software package and other tools [21]. The sequences were initially processed according to the denoising of 454 datasets workflow. Sequences shorter than 300 bp, with an average quality value below 25, or possessing homopolymers longer than 8 bp were removed. Afterwards, the sequences were denoised. Cutadapt was used to truncate remaining primer sequences [22]. Chimeric sequences were removed using UCHIME and the Green Genes Gold dataset as reference database [23–25].

Remaining sequences were clustered employing the UCLUST algorithm [23] and the following QIIME scripts: pick_otus.py and pick_rep_set.py. The sequences were clustered in operational taxonomic units (OTUs) at 3% and 1% genetic dissimilarity. Phylogenetic composition was determined using the QIIME assign_taxonmy.py script. A BLAST alignment [26] against the most recent Silva ARB database [27] was thereby performed. Sequences were classified with respect to the taxonomy of their best hit in the ARB database. Finally, OTU tables were generated.

2.6. Rarefaction Analysis and Diversity Analysis. Rarefaction curves, Shannon indices [28], and Chao1 indices [29] were calculated employing QIIME scripts. In addition, the maximal number of OTUs (n_{max}) was estimated for each sample in R (version 2.15) [30] using the data derived from the QIIME rarefaction analysis and a nonlinear regression model based on Michaelis-Menten kinetics [31].

To compare archaeal community structures across all samples based on phylogenetic or count-based distance metrics, a principal coordinate analysis (PCoA) was performed using QIIME. The following scripts were successively used to generate a phylogenetic tree at 1% genetic distance prior to PCoA calculation: align_seqs.py (PyNAST algorithm), filter_alignment.py, and make_phylogeny.py. The tree and the respective OTU table were used to generate PCoAs employing the "beta_diversity_through_plots.py" script.

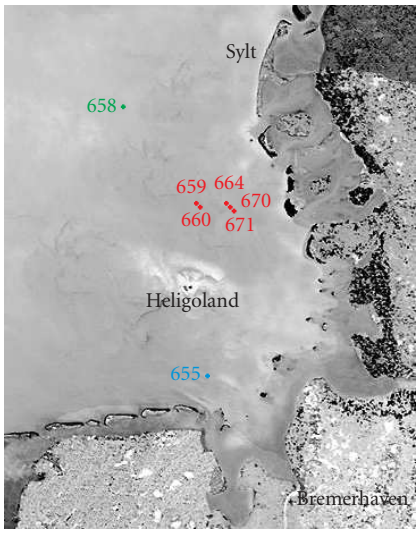


Figure 1: Satellite image of the German Bight showing the locations of the seven sampling sites (Image: ESA/NASA - SOHO/LASCO). Samples taken during an algal bloom (samples 659, 660, 664, 670, and 671) are shown in red. Sample 655 taken at a river outfall and sample 658 originating from outside the algal bloom in blue and green, respectively.

3. Results

3.1. Environmental Parameters. Marine water samples for archaeal community analysis were randomly collected at seven different locations in the German Bight (Figure 1, Table 2). Five samples (sites 659, 660, 664, 670, and 671) were taken in presence of an algal bloom. The other two samples derived from a river outfall (655) and from a site outside the algal bloom (658). The algal blooms observed during the sampling were mainly dominated by the genus *Phaeocystis*. Diatoms of the genus *Rhizosolenia* and some dinoflagellates were also identified but only in minor abundances.

Table 2: Parameters of sampling sites analyzed in this study.

	Site	Latitude °N	Longitude °E	Depth (m)	T (°C)	Salinity (psu)	Fluorescence (mg/m^3)	Transmission (%)
655	River outfall	53°53.729	8°02.979	2	11.09	30.24	1.21	57.2
658	No bloom	54°45.754	7°26.780	2	9.73	32.71	0.49	81.23
659	Bloom	54°27.450	7°59.360	9	10.80	30.64	2.76	60.14
660	Bloom	54°27.250	8°00.110	2	10.83	30.65	1.89	72.28
664	Bloom	54°28.400	8°11.830	2	10.90	30.76	1.14	87.28
670	Bloom	54°27.570	8°12.420	2	11.43	30.83	—*	75.72
671	Bloom	54°26.940	8°12.970	2	11.70	31.04	—*	76.59

* Fluorescence was not measured due to a malfunction of the profiler.

Environmental factors at all seven sampling sites were monitored employing a CTD profiler (Table 2). Temperatures and salinities ranged from 9.73 to 11.70°C and from 30.24 to 32.71 psu, respectively. The lowest temperature and highest salinity were measured at site 658. All other sites showed similar conditions. Fluorescence was higher at bloom sites due to a higher chlorophyll concentration, whereas transmission was reduced due to a higher turbidity in the water.

3.2. Archaeal Community Structure Revealed by 16S rRNA-Based Analysis. To assess archaeal community structures, total RNA was extracted from the samples. Approximately 5 μg of total RNA per filter were extracted from each sample. After removal of contaminating DNA and small RNAs, 0.25 to 1.5 μg of RNA were used as template for cDNA synthesis. The V3–V5 region of the 16S rRNA was amplified from the generated cDNA. The resulting PCR products were subjected to pyrosequencing. Sequence processing including quality filtering, denoising, and removal of potential chimeric sequences resulted in recovery of 62,090 high quality sequences with a read length of ≥300 bp across all 7 samples. The average read length was 506 bp. The number of sequences per sample ranged from 4,301 to 23,070. We were able to assign 62,045 sequences to the domain *Archaea* and to classify all of these sequences below the domain level. The classified sequences were affiliated to three archaeal phyla with twelve archaeal classes or similar phylogenetic groups. *Euryarchaeota* was the most abundant archaeal phylum (99.25%) and *Halobacteria* the predominant class across all samples (>98.1%) (Figure 2). Most of the sequences affiliated to the *Halobacteria* (97.81%) were affiliated to uncultured members of the Deep Sea Hydrothermal Vent Group 6 (DHVEG-6) [32]. Interestingly, *Halobacteria* were more abundant in bloom samples than in other samples (Figure 2). Other archaeal groups present in all samples were the Marine Group I (*Thaumarchaeota*) [33] and the *Thermoplasamata* (*Euryarchaeota*). Sequences affiliated to the latter archaeal group belonged to the uncultured members of the CCA47 [34] group and the Marine Group II [33].

3.3. Diversity and Species Richness of Archaeal Communities. To determine the archaeal diversity and richness, rarefaction analyses were performed with QIIME [21]. Alpha diversity analysis was performed at the same level of surveying effort (3100 randomly selected sequences per sample). The observed OTU number in the archaeal picoplankton ranged from 252 to 454 OTUs (1% genetic distance) and from 250 to 417 OTUs (3% genetic distance) (Table 3). The maximal expectable number of clusters for every sample was determined by nonlinear regression based on the Michalis-Menten equitation. The average OTU coverages were 62.3% and 62.6% at 1% and 3% genetic distance, respectively. Shannon indices ranged from 3.74 to 7.74 (1% genetic distance) and from 3.63 to 7.62 (3% genetic distance).

Comparison of the rarefaction analyses with the number of OTUs determined by Chao1 richness estimator revealed that at 1%and 3% genetic distances the rarefaction curves (Figure 3) were not saturated and the richness estimators indicated that 41.34% to 73.41% of the estimated richness, respectively, were recovered by the sequencing effort (Table 3). Thus, we did not survey the full extent of taxonomic diversity at these genetic distances, but a substantial fraction of the archaeal diversity within individual samples was assessed at genetic divergence of 3%.

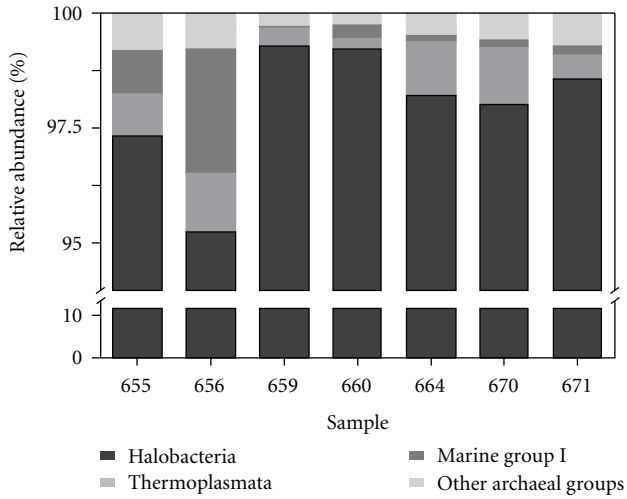


Figure 2: Relative sequence abundances of different archaeal phyla and classes. *Euryarchaeota*, especially *Halobacteria* (98.14%), were highly abundant. *Thermoplasmata* (0.75%) and the Marine Group 1 (0.58%) were found to some extent. All archaeal classes and groups (abundance < 0.5%) are depicted together.

Table 3: Archaeal diversity and richness values at 1% and 3% genetic distance. Numbers of observed OTUs as well as Shannon and Chao1 values were calculated with QIIME [16]. The maximal OTU number (n_{max}) in each sample was calculated by nonlinear modeling. Coverage was determined based on observed OTUs and n_{max}. To compare community structures, 3100 randomly selected sequences form every sample were used.

Sample	Observed OTUs		Max. OTUs (n_{max})		Coverage (%)		Shannon index (H′)		Chao1	
	1%	3%	1%	3%	1%	3%	1%	3%	1%	3%
655	293	268	510	468	57.45	57.26	4.37	4.02	530	470
658	451	428	555	524	81.26	81.68	7.74	7.62	636	583
659	252	250	446	441	56.50	56.69	3.74	3.63	498	470
660	281	269	516	496	54.46	54.23	3.95	3.74	551	509
664	346	327	516	488	67.05	67.01	4.81	4.65	569	486
670	454	417	782	717	58.06	58.16	5.21	5.07	785	674
671	399	370	649	586	61.48	63.14	5.09	4.96	1227	895

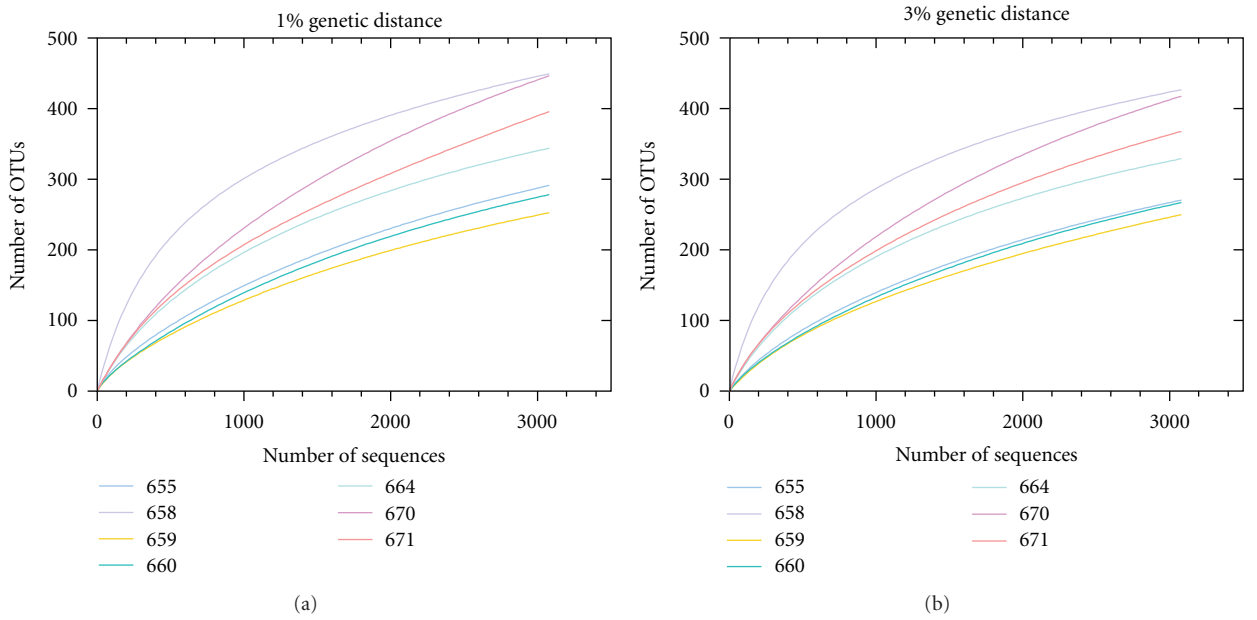


Figure 3: Rarefaction curves for all seven sampling sites. Curves were calculated at 1% (a) and 3% (b) genetic distance level employing QIIME [21]. Description of samplings sites is shown in Table 2.

3.4. Beta Diversity of the Bacterioplankton Community. Changes of the active bacterial community in response to different environmental conditions were examined by principal coordinate analysis (PCoA) (Figure 4). Surveying effort had no or little effect on diversity and community structure. However, the PCoA analysis revealed that all samples exhibiting similar environmental parameters such as temperature and salinity were assigned to one site of the plot. In addition, all bloom samples tend to cluster together. Sample 658 taken outside the algal bloom was completely separated from all other samples.

4. Discussion

Marine environments contain a high microbial biodiversity, and marine microbial communities play major roles in many biogeochemical cycles. Studies using culture-independent approaches have greatly contributed to our understanding of the extent of microbial diversity [35]. Most of these studies focused on marine bacteria, whereas very little is known on the diversity and ecology of marine *Archaea*. Recent metagenomic studies provided evidence for ammonium-oxidizing *Archaea* being capable of nitrification [36]. Some marine crenarchaeal lineages are thought to be important nitrifiers in planktonic marine systems [37]. These results indicate that *Archaea* are important players in the global nitrogen cycle. However, detailed comparative ecological studies to understand archaeal community patterns and environmental drivers that shape these communities are missing [37].

This study focused on assessing the active archaeal community structure and richness in picoplankton samples derived from the German Bight by metatranscriptomic approaches. To our knowledge, this is the first study using an RNA-based approach combined with NGS to analyze archaeal community compositions. In addition, the obtained average read length (506 bp) is higher than in most other

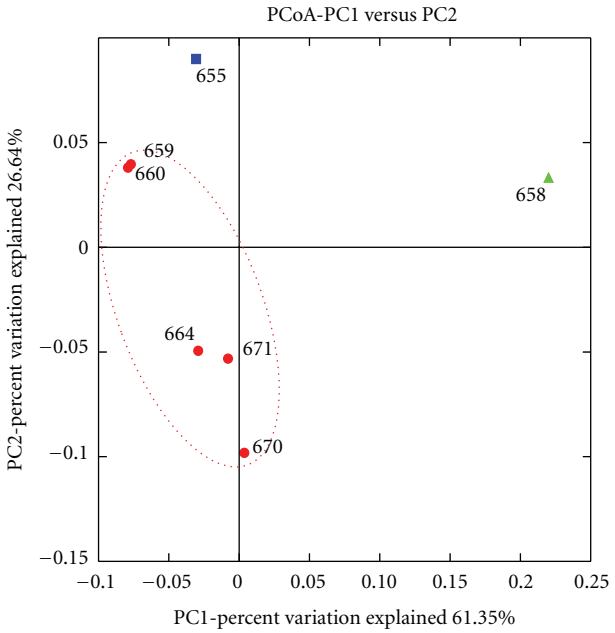


Figure 4: Weighted UniFrac 2D Principal Coordinate Analysis plot for beta diversity analysis. Samples taken during an algal bloom (samples 659, 660, 664, 670, and 671) are shown in red. Sample 655 taken at a river outfall and sample 658 originating from outside the algal bloom in blue and green, respectively.

studies employing NGS sequencing of 16S rRNA gene amplicons [38, 39]. The majority of sequences obtained was affiliated to the *Euryarchaeota*. Sapp et al. [40] studied marine sediments derived from the Oyster Ground (North Sea) and found high abundances of members of this phylum in their samples. We identified *Halobacteria* as the most abundant archaeal group. Members of this group can grow aerobically as well as anaerobically. Large halobacterial blooms appear reddish due to production of retinal-containing rhodopsins. Rhodopsins are photoactive membrane proteins with a highly conserved tertiary structure [41] and may serve as an additional possibility to conserve energy. This is advantageous in marine environments, as the concentration of dissolved organic matter and other nutrients is usually low [42]. Most of the halobacterial sequences analyzed in this study were affiliated to the Deep Sea Hydrothermal Vent Euryarchaeotal Group 6 (DHVE-6). This group was originally described as a hydrothermal vent lineage [43]. It was later renamed Miscellaneous Euryarchaeotic Group, as members of this group were also found in marine sediment [44] and in soil [45]. Another archaeal group found in all samples was the Marine Group I (MG-I). It was originally identified by sequencing of environmental 16S rRNA genes derived from sea water [46, 47]. Members of MG-I account for large fractions of marine prokaryotic picoplankton and prokaryotic communities in deep sea water (below 3000 m). *Thermoplasmata* were the third most abundant archaeal class in the investigated samples. Most sequences were affiliated to the CCA47 group. This group was originally identified by 16S rRNA gene analysis of oxygen-depleted marine environments [48]. Later, Ferrer et al. [34] found members of this group in anoxic subsaline sediments. A few sequences assigned to Thermoplasmata were also affiliated to the Marine Group II. DeLong [33] suggested that members of Marine Group II (*Euryarchaeota*) are more abundant in temperate sea water than Marine Group I (*Crenarchaeota*) members. We found the opposite, as we recorded a higher abundance of Marine Group I members in the studied samples. Marine Group II members were almost absent in the investigated samples. One reason for this discrepancy might be that large parts of the German Bight are strongly influenced by tidal currents. Thus, these currents might whirl up archaeal cells from the sediment to the surface water, as most of the identified groups were originally described as inhabitants of marine sediments. Nonetheless, the number of studies targeting archaeal communities in the water column is substantially lower than that on marine sediments. Due to this knowledge gap, the habitat preference of these archaeal groups cannot be deduced definitely.

The impact of environmental conditions onto archaeal community composition and richness has been rarely studied. Auguet et al. [37] performed a general analytical approach to find community patterns of uncultured *Archaea* along environmental gradients or habitat types. Their results indicate that habitat types have a greater effect on archaeal community structures than other environmental conditions. All samples investigated in our study originated from almost the same habitat type, except for samples 655 and 658, which were collected at a river outfall region and outside of the algal bloom, respectively. Accordingly, all samples derived from the bloom showed an almost identical community composition. In addition, sample 655 showed a consimilar community structure. This indicates that similar environmental factors, such as temperature, salinity, and high nutrient availability during algal blooms or at river outfalls, have a similar impact onto composition of active archaeal communities.

Herfort et al. [49] studied archaeal communities in the southwestern North Sea via Denaturing Gradient Gel Electrophoresis (DGGE) and showed a positive correlation between the abundance of *Euryarchaeota* and chlorophyll concentrations, whereas the abundance of *Crenarchaeota* was negatively correlated with the chlorophyll concentration. Teeling et al. [50] investigated bacterial communities near Helgoland. They demonstrated that bacterial community structures were highly influenced by the presence of an algal bloom. In our study, we investigated the influence of algal blooms on archaeal diversity by PCoA. Sample taken in presence of a bloom shared a more similar community structure. This indicates that marine archaeal communities are also influenced by algal blooms or by environmental parameters correlated with bloom presence. We observed an increased number of *Halobacteria* in bloom samples. This might be correlated with the high amounts of organic matter in blooms. *Halobacteria* are the most active organisms with respect to organic matter degradation in hypersaline environments [37]. Thus the higher abundance of *Halobacteria* in algal bloom samples might indicate an involvement in marine organic matter degradation under high nutrient conditions found during algal blooms.

Due to the lack of pure cultures and large comparative investigations, robust conclusions on contributions of marine archaeal communities to biogeochemical cycles cannot be drawn. In this study, we found highly diverse and active archaeal communities in the surface water of the German Bight. Their ecological role is unknown, and further research including analyses of expressed functional genes needs to be performed to unravel the role of marine *Archaea*.

Acknowledgments

The authors thank the crew of the research vessel Heincke for their valuable support and the people from Oldenburg, especially Meinhard Simon and Helge-Ansgar Giebel, for the hydrographic data and valuable help prior and during the sampling. This work was funded by the Deutsche Forschungsgemeinschaft (DFG) as part of the collaborative research center TRR 51.

References

[1] T. P. Curtis, W. T. Sloan, and J. W. Scannell, "Estimating prokaryotic diversity and its limits," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 16, pp. 10494–10499, 2002.

[2] M. L. Sogin, H. G. Morrison, J. A. Huber et al., "Microbial diversity in the deep sea and the underexplored "rare biosphere"," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 32, pp. 12115–12120, 2006.

[3] M. B. Karner, E. F. Delong, and D. M. Karl, "Archaeal dominance in the mesopelagic zone of the Pacific Ocean," *Nature*, vol. 409, no. 6819, pp. 507–510, 2001.

[4] H. Agogué, M. Brink, J. Dinasquet, and G. J. Herndl, "Major gradients in putatively nitrifying and non-nitrifying Archaea in the deep North Atlantic," *Nature*, vol. 456, no. 7223, pp. 788–792, 2008.

[5] E. F. DeLong and D. M. Karl, "Genomic perspectives in microbial oceanography," *Nature*, vol. 437, no. 7057, pp. 336–342, 2005.

[6] S. J. Giovannoni and U. Stingl, "Molecular diversity and ecology of microbial plankton," *Nature*, vol. 437, no. 7057, pp. 343–348, 2005.

[7] T. Brinkhoff, D. Fischer, J. Vollmers et al., "Biogeography and phylogenetic diversity of a cluster of exclusively marine myxobacteria," *International Society for Microbial Ecology Journal*, vol. 6, no. 6, pp. 1260–1272, 2012.

[8] J. Frias-Lopez, Y. Shi, G. W. Tyson et al., "Microbial community gene expression in ocean surface waters," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 10, pp. 3805–3810, 2008.

[9] J. C. Venter, K. Remington, J. F. Heidelberg et al., "Environmental genome shotgun sequencing of the sargasso sea," *Science*, vol. 304, no. 5667, pp. 66–74, 2004.

[10] H. Nacke, A. Thürmer, A. Wollherr et al., "Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils," *PLoS ONE*, vol. 6, no. 2, Article ID e17000, 2011.

[11] C. Will, A. Thürmer, A. Wollherr et al., "Horizon-specific bacterial community composition of german grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes," *Applied and Environmental Microbiology*, vol. 76, no. 20, pp. 6751–6759, 2010.

[12] M. Vila-Costa, J. M. Gasol, S. Sharma, and M. A. Moran, "Community analysis of high- and low-nucleic acid-containing bacteria in NW Mediterranean coastal waters using 16S rDNA pyrosequencing," *Environmental Microbiology*, vol. 14, no. 6, pp. 1390–1402, 2012.

[13] T. Urich, A. Lanzén, J. Qi, D. H. Huson, C. Schleper, and S. C. Schuster, "Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome," *PLoS ONE*, vol. 3, no. 6, Article ID e2527, 2008.

[14] R. A. Lesniewski, S. Jain, K. Anantharaman, P. D. Schloss, and G. J. Dick, "The metatranscriptome of a deep-sea hydrothermal plume is dominated by water column methanotrophs and lithotrophs," *International Society for Microbial Ecology Journal*. In press.

[15] M. G. Weinbauer, I. Fritz, D. F. Wenderoth, and M. G. Höfle, "Simultaneous extraction from bacterioplankton of total RNA and DNA suitable for quantitative structure and function analyses," *Applied and Environmental Microbiology*, vol. 68, no. 3, pp. 1082–1087, 2002.

[16] V. I. Miteva, P. P. Sheridan, and J. E. Brenchley, "Phylogenetic and physiological diversity of microorganisms isolated from a deep greenland glacier ice core," *Applied and Environmental Microbiology*, vol. 70, no. 1, pp. 202–213, 2004.

[17] G. Muyzer, E. C. De Waal, and A. G. Uitterlinden, "Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA," *Applied and Environmental Microbiology*, vol. 59, no. 3, pp. 695–700, 1993.

[18] M. J. Ferris, A. Masztal, and D. H. Martin, "Use of species-directed 16S rRNA gene PCR primers for detection of Atopobium vaginae in patients with bacterial vaginosis," *Journal of Clinical Microbiology*, vol. 42, no. 12, pp. 5892–5894, 2004.

[19] M. Hartmann and F. Widmer, "Reliability for detecting composition and changes of microbial communities by T-RFLP genetic profiling," *FEMS Microbiology Ecology*, vol. 63, no. 2, pp. 249–260, 2008.

[20] Z. Yu, R. García-González, F. L. Schanbacher, and M. Morrison, "Evaluations of different hypervariable regions of archaeal 16S rRNA genes in profiling of methanogens by Archaea-specific PCR and denaturing gradient gel electrophoresis," *Applied and Environmental Microbiology*, vol. 74, no. 3, pp. 889–893, 2008.

[21] J. G. Caporaso, J. Kuczynski, J. Stombaugh et al., "QIIME allows analysis of high-throughput community sequencing data," *Nature Methods*, vol. 7, no. 5, pp. 335–336, 2010.

[22] M. Martin, "Cutadapt removes adapter sequences from high-throughput sequencing reads," *EMBnet.Journal*, vol. 17, no. 1, pp. 10–12, 2011.

[23] R. C. Edgar, B. J. Haas, J. C. Clemente, C. Quince, and R. Knight, "UCHIME improves sensitivity and speed of chimera detection," *Bioinformatics*, vol. 27, no. 16, Article ID btr381, pp. 2194–2200, 2011.

[24] P. D. Schloss, D. Gevers, and S. L. Westcott, "Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies," *PloS ONE*, vol. 6, no. 12, Article ID e27310, 2011.

[25] T. Z. DeSantis, P. Hugenholtz, N. Larsen et al., "Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB," *Applied and Environmental Microbiology*, vol. 72, no. 7, pp. 5069–5072, 2006.

[26] C. Camacho, G. Coulouris, V. Avagyan et al., "BLAST+: architecture and applications," *BMC Bioinformatics*, vol. 10, article 421, 2009.

[27] E. Pruesse, C. Quast, K. Knittel et al., "SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB," *Nucleic Acids Research*, vol. 35, no. 21, pp. 7188–7196, 2007.

[28] C. E. Shannon, "A mathematical theory of communication," *SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[29] A. Chao and J. Bunge, "Estimating the number of species in a stochastic abundance model," *Biometrics*, vol. 58, no. 3, pp. 531–539, 2002.

[30] RDevelopmentCoreTeam, *R: A Language and Environment For Statistical Computing*, 2012.

[31] L. Michaelis and M. L. Menten, "Die kinetik der invertinwirkung," *Biochemistry Zeitung*, vol. 49, no. 333–369, p. 352, 1913.

[32] J. Kan, S. Clingenpeel, R. E. Macur et al., "Archaea in Yellowstone Lake," *International Society for Microbial Ecology Journal*, vol. 5, no. 11, pp. 1784–1795, 2011.

[33] E. F. DeLong, "Oceans of archaea," *ASM News*, vol. 69, no. 10, p. 9, 2003.

[34] M. Ferrer, M. E. Guazzaroni, M. Richter et al., "Taxonomic and functional metagenomic profiling of the microbial community in the anoxic sediment of a sub-saline shallow lake (Laguna de Carrizo, Central Spain)," *Microbial Ecology*, vol. 62, no. 4, pp. 824–837, 2011.

[35] C. Simon and R. Daniel, "Metagenomic analyses: past and future trends," *Applied and Environmental Microbiology*, vol. 77, no. 4, pp. 1153–1161, 2011.

[36] R. Cavicchioli, M. Z. DeMaere, and T. Thomas, "Metagenomic studies reveal the critical and wide-ranging ecological importance of uncultivated archaea: the role of ammonia oxidizers," *BioEssays*, vol. 29, no. 1, pp. 11–14, 2007.

[37] J. C. Auguet, A. Barberan, and E. O. Casamayor, "Global ecological patterns in uncultured Archaea," *International Society for Microbial Ecology Journal*, vol. 4, no. 2, pp. 182–190, 2010.

[38] O. O. Lee, Y. Wang, J. Yang, F. F. Lafi, A. Al-Suwailem, and P. Y. Qian, "Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea," *International Society for Microbial Ecology Journal*, vol. 5, no. 4, pp. 650–664, 2011.

[39] J. A. Huber, D. B. Mark Welch, H. G. Morrison et al., "Microbial population structures in the deep marine biosphere," *Science*, vol. 318, no. 5847, pp. 97–100, 2007.

[40] M. Sapp, E. R. Parker, L. R. Teal, and M. Schratzberger, "Advancing the understanding of biogeography-diversity relationships of benthic microorganisms in the North Sea," *FEMS Microbiology Ecology*, vol. 74, no. 2, pp. 410–429, 2010.

[41] J. L. Spudich, C. S. Yang, K. H. Jung, and E. N. Spudich, "Retinylidene proteins: structures and functions from archaea to humans," *Annual Review of Cell and Developmental Biology*, vol. 16, pp. 365–392, 2000.

[42] M. T. Madigan and T. D. Brock, *Brock Biology of Microorganisms*, Pearson/Benjamin Cummings, San Francisco, Calif, USA, 2010.

[43] K. Takai and K. Horikoshi, "Genetic diversity of archaea in deep-sea hydrothermal vent environments," *Genetics*, vol. 152, no. 4, pp. 1285–1297, 1999.

[44] K. B. SØRensen, A. Lauer, and A. Teske, "Archaeal phylotypes in a metal-rich and low-activity deep subsurface sediment of the Peru Basin, ODP Leg 201, Site 1231," *Geobiology*, vol. 2, no. 3, pp. 151–161, 2004.

[45] K. Takai, D. P. Moser, M. DeFlaun, T. C. Onstott, and J. K. Fredrickson, "Archaeal diversity in waters from deep South African gold mines," *Applied and Environmental Microbiology*, vol. 67, no. 12, pp. 5750–5760, 2001.

[46] E. F. DeLong, "Archaea in coastal marine environments," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 12, pp. 5685–5689, 1992.

[47] J. A. Fuhrman, K. McCallum, and A. A. Davis, "Novel major archaebacterial group from marine plankton," *Nature*, vol. 356, no. 6365, pp. 148–149, 1992.

[48] T. Stoeck and S. Epstein, "Novel eukaryotic lineages inferred from small-subunit rRNA analyses of oxygen-depleted marine environments," *Applied and Environmental Microbiology*, vol. 69, no. 5, pp. 2657–2663, 2003.

[49] L. Herfort, S. Schouten, B. Abbas et al., "Variations in spatial and temporal distribution of Archaea in the North Sea in relation to environmental variables," *FEMS Microbiology Ecology*, vol. 62, no. 3, pp. 242–257, 2007.

[50] H. Teeling, B. M. Fuchs, D. Becher et al., "Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom," *Science*, vol. 336, no. 6081, pp. 608–611, 2012.

STUDY 2:

IMPACT OF A PHYTOPLANKTON BLOOM ON THE DIVERSITY OF THE ACTIVE BACTERIAL COMMUNITY IN THE SOUTHERN NORTH SEA AS REVEALED BY METATRANSCRIPTOMIC APPROACHES

WEMHEUER B[1], GÜLLERT S[1,3], BILLERBECK S[2], GIEBEL HA[2], VOGET S[1], SIMON M[2], AND DANIEL R[1]

FEMS MICROBIOLOGY ECOLOGY: 87: 378–389
(doi: 10.1111/1574-6941.12230)

[1]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY; [2]INSTITUTE FOR CHEMISTRY AND BIOLOGY OF THE MARINE ENVIRONMENT (ICBM), CARL-VON-OSSIETZKY-UNIVERSITY OF OLDENBURG, CARL-VON-OSSIETZKY-STR. 9-11, D-26111 OLDENBURG, GERMANY; [3]PRESENT ADDRESS: BIOZENTRUM KLEIN FLOTTBEK, MIKROBIOLOGIE, OHNHORSTSTR. 18, D-22609 HAMBURG, GERMANY

Author contributions to the work:

Performed the experiments: BW, SG

Analyzed data: BW, SG, SV

Contributed data on water properties and analysis of these data: SB, HAG, MS

Wrote the publication: BW, MS, RD

Conceived and designed the experiments: BW,RD

FEMS MICROBIOLOGY ECOLOGY

RESEARCH ARTICLE

Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches

Bernd Wemheuer[1], Simon Güllert[1], Sara Billerbeck[2], Helge-Ansgar Giebel[2], Sonja Voget[1], Meinhard Simon[2] & Rolf Daniel[1]

[1]Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Göttingen, Germany; and [2]Biology of Geological Processes – Aquatic Microbial Ecology, Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl-von-Ossietzky-University Oldenburg, Oldenburg, Germany

Correspondence: Rolf Daniel, Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany. Tel.: 0049 551 393827; fax: 0049 551 3912181; e-mail: rdaniel@gwdg.de

Present address: Simon Güllert, Microbiology and Biotechnology, Biocenter Klein Flottbek, Ohnhorststr. 18, D-22609 Hamburg, Germany

Received 21 June 2013; revised 18 September 2013; accepted 19 September 2013. Final version published online 25 October 2013.

DOI: 10.1111/1574-6941.12230

Editor: Riks Laanbroek

Keywords
bacterial diversity; bacterioplankton; pyrotag sequencing; German Bight; metatranscriptome.

Abstract

Despite their importance for ecosystem functioning, little is known about the composition of active marine bacterioplankton communities. Hence, this study was focused on assessing the diversity of these communities in the southern North Sea and examining the impact of a phytoplankton spring bloom on the ambient bacterioplankton community. Community composition in and outside the bloom was assessed in 14 samples by pyrosequencing-based analysis of 16S rRNA gene amplicons generated from environmental RNA. The data set comprised of 211 769 16S rRNA gene sequences. *Proteobacteria* were the predominant phylogenetic group with *Alphaproteobacteria* and *Gammaproteobacteria* as the most abundant classes. *Actinobacteria* and *Bacteroidetes* were identified in minor abundances. Active bacterial communities were dominated by few lineages such as the *Roseobacter* RCA cluster and the SAR92 clade. Community structures of three selected samples were also assessed by direct sequencing of cDNA generated from rRNA-depleted environmental RNA. Generated data sets comprised of 988 202 sequences. Taxonomic assignment of the reads confirmed the predominance of *Proteobacteria*. The examined phytoplankton spring bloom affected the bacterioplankton community structures significantly. Bacterial richness was reduced in the bloom area, and the abundance of certain bacterial groups was affected by bloom presence. The SAR92 clade and the *Roseobacter* RCA cluster were significantly more abundant and active in the bloom. Functions affected by the bloom include photosynthesis, protein metabolism, and DNA metabolism.

Introduction

Cultivation-independent approaches have greatly advanced our understanding of the ecology and diversity of marine microbial communities (e.g. Venter *et al.*, 2004; Giovannoni & Stingl, 2005). In early DNA-based studies, traditional techniques such as Sanger sequencing-based analysis of 16S rRNA gene libraries or fingerprinting methods were used to analyze microbial community structures at different marine locations (e.g. Giovannoni *et al.*, 1990; Schmidt *et al.*, 1991; Muyzer *et al.*, 1993). These approaches were often limited to the analysis of relatively small numbers of clones or samples. Taking into account the high diversity and community size of marine microbial communities, only a small fraction of the bacterial diversity was unraveled by these studies. Recent studies analyzing 16S rRNA gene fragments by next-generation sequencing technologies have been applied for in-depth investigation of bacterial communities in diverse ecosystems such as soil (e.g. Will *et al.*, 2010; Nacke *et al.*, 2011) or different marine ecosystems (e.g. Kirchman *et al.*, 2010; Vila-Costa *et al.*, 2012). These studies provided detailed insights into the composition of the ambient microbial communities and revealed the

existence of many taxa not known from the previous less sensitive approaches.

Coastal shelf seas of the temperate zone are highly productive because of the continuous nutrient supply by rivers. A typical shelf sea with these characteristics is the North Sea and in particular its southern region, the German Bight, which underwent high nutrient loading and warming during the last 40 years (McQuatters-Gollop *et al.*, 2007; Wiltshire *et al.*, 2010). Nutrients and plankton communities have been studied intensely in this dynamic coastal sea. Studies concerning the bacterioplankton community in the German Bight show that distinct lineages of *Alphaproteobacteria* and *Gammaproteobacteria* as well as *Flavobacteria* constitute the majority of this community (Alderkamp *et al.*, 2006; Rink *et al.*, 2011; Teeling *et al.*, 2012; Sintes *et al.*, 2013). *Archaea* are also present in the prokaryotic picoplankton with dominance of *Halobacteria* and minor proportions of *Thermoplasmata* and members of the marine group I (Wemheuer *et al.*, 2012). Among *Alphaproteobacteria*, the SAR11 and *Roseobacter* clade, especially the *Roseobacter* clade-affiliated (RCA) cluster, are of particular significance and constitute this bacterial class to a great extent (Selje *et al.*, 2004; Giebel *et al.*, 2011; Sperling *et al.*, 2012). However, the vast majority of investigations focused on assessing total bacterioplankton community structures by DNA-based 16S rRNA gene analysis and did not consider the active members of the microbial communities (but see Teeling *et al.*, 2012; Wemheuer *et al.*, 2012), which can be examined, that is, by the analysis of 16S rRNA gene transcripts (e.g. West *et al.*, 2008; Wemheuer *et al.*, 2012).

Another approach to assess active bacterioplankton community structures and their functional response to ambient growth and environmental conditions is the application of large-scale sequencing of environmental mRNA. This approach has been applied in recent studies providing first insights into how marine bacterioplankton communities respond to nutrient inputs and changing environmental conditions (e.g. Frias-Lopez *et al.*, 2008; Gifford *et al.*, 2011; Lesniewski *et al.*, 2012; Ottesen *et al.*, 2013).

The aim of the present study was to investigate the phylogenetic composition of active bacterioplankton communities in the southern North Sea, the German Bight, and validate the impact of a phytoplankton spring bloom on the structure of these communities. We hypothesized (1) that the structure of the active bacterioplankton community is affected by the bloom, as the availability and the composition of nutrients are different between bloom and nonbloom conditions. We further hypothesized (2) that the bacterial richness is also influenced by bloom presence, as a specialized ecological niche is formed during a bloom in which only certain bacterioplankton members can thrive.

To assess bacterial community structures, we applied large-scale pyrosequencing of the 16S rRNA gene V2–V3 region. Moreover, we also employed direct sequencing of cDNA derived from rRNA-depleted environmental RNA to assess bacterial community structures via subsequent phylogenetic assignment of the recovered sequences. This is the first study using RNA-based next-generation sequencing approaches to monitor the impact of a phytoplankton bloom on active bacterioplankton communities. It provides evidence for changes in structure and activity of these communities during the bloom event.

Materials and methods

Sampling and sample preparation

Fourteen water samples were collected in the southern North Sea at 13 stations in May 2010 on board RV Heincke (cruise HE327) to investigate the impact of a phytoplankton bloom on the ambient bacterial community structure (Supporting Information, Table S1). The phytoplankton bloom examined in this study was localized by satellite data. Its presence at the sampling stations was then evaluated based on chlorophyll *a* content. Samples were taken using 5-L Niskin bottles mounted on a CTD rosette (Fig. 1, Table S1). To allow comparison with other studies using samples from the HE327 cruise, station and corresponding sample numbers were given according to the designation provided by the ICBM (Oldenburg, Germany). The water of at least eight Niskin bottles (*c.* 40 L) was pooled in an ethanol-rinsed PE barrel. It was prefiltered through a 10-μm nylon net and a filter sandwich consisting of a precombusted (4 h at 450 °C) glass fiber filter (47 mm diameter, Whatman GF/D; Whatman, Maidstone, UK) and 3.0-μm polycarbonate filter (47 mm diameter, Nuclepore; Whatman). Bacterioplankton was harvested from a prefiltered 1-L sample on a filter sandwich consisting of a glass fiber filter (47 mm diameter, Whatman GF/F; Whatman) and 0.2-μm polycarbonate filter (47 mm diameter, Nuclepore; Whatman). All filters were frozen using liquid nitrogen and stored at −80 °C until further use. Additionally, phytoplankton samples were collected for screening of the major phytoplankton taxa employing a plankton net (pore size, 55 μm).

For the determination of Chl *a* and phaeopigments, water samples were filtered onto glass fiber filters (47 mm diameter, Whatman GF/F; Whatman), immediately wrapped into aluminum foil, and kept frozen at −20 °C until further analysis within 2 weeks according to Giebel *et al.* (2011) and Nusch (1999). Water samples for the analysis of suspended particulate matter (SPM),

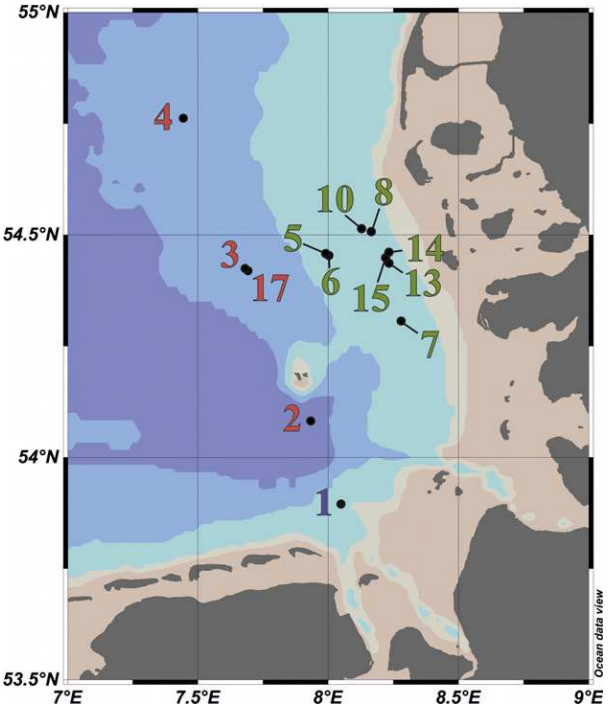


Fig. 1. Map of the German Bight showing the location of the 13 sampling stations visited in May 2010. Stations inside the examined algal bloom are depicted in green; those outside the bloom are in red. Station 1 is depicted in blue as it was located in a bloom outside of the area of the examined bloom. Bloom samples were defined as samples with Chl *a* concentrations of > 4.4 μg Chl *a* L^{-1} The map was generated using the Ocean Data View software package (Schlitzer, 2013).

particulate organic carbon (POC), and particulate organic nitrogen (PON), respectively, were filtered onto precombusted (2 h, 500 °C) and preweighed glass fiber filter (47 mm diameter, Whatman GF/F; Whatman). Filters for SPM determination were subsequently rinsed with distilled water to remove salt. All filters were kept frozen at −20 °C until the analysis as described in detail by Lunau *et al.* (2006). Samples for dissolved inorganic nutrients (nitrate, nitrite, and phosphate) were filtered through glass fiber filter (47 mm diameter, Whatman GF/F; Whatman) and measured with an autoanalyzer (Autoanalyzer Evolution 3; Alliance Instruments, Frepillon, France) following the protocol by Koroleff & Grasshoff (1983). Detection limits were 0.3 μM for N and 0.1 μM for P. Bacterioplankton cell numbers were determined by flow cytometry (BD FACSAria™ III; BD Biosciences, San Jose, CA) using SYBR Green I staining and internal bead calibration as described previously (Giebel *et al.*, 2009). For this purpose, water samples were taken, preserved with glutaraldehyde (final concentration 1%), and stored at −20 °C until further analysis. All measured environmental parameters are listed in Table S2.

Extraction and purification of environmental RNA

Total RNA was extracted as described by Weinbauer *et al.* (2002). One 47-mm-diameter filter (pore size, 0.2 μm) was used per sample. Subsequently, RNA was purified employing the RNeasy Mini Kit as recommended by the manufacturer (Qiagen, Hilden, Germany). Residual DNA was removed from RNA samples, and the absence of DNA was confirmed according to Wemheuer *et al.* (2012). Either DNA-free RNA was converted directly to cDNA for 16S rRNA gene analysis or mRNA was enriched by rRNA depletion. The RiboMinus™ transcriptome isolation kit for Bacteria (Invitrogen™, Carlsbad, CA) was employed for mRNA enrichment. The removal of rRNA was performed according to the manufacturer's instructions with one modification. For initial denaturation, the RNA was incubated at 70 °C for 10 min.

DNA was synthesized from total RNA and enriched mRNA by employing the SuperScript™ double-stranded cDNA synthesis kit (Invitrogen™) with slight modifications according to Wemheuer *et al.* (2012). Total RNA-derived cDNA was subjected to 16S rRNA gene PCR. The Göttingen Genomics Laboratory determined the sequences of the enriched mRNA-derived cDNA using a Roche GS-FLX 454™ pyrosequencer and titanium chemistry (Roche, Mannheim, Germany).

Amplification and sequencing of 16S rRNA gene

To analyze the bacterial diversity, the V2–V3 region of the bacterial 16S rRNA gene was amplified by PCR. The PCR (25 μL) contained 5 μL of fivefold Phusion GC buffer (Finnzymes, Vantaa, Finland), 200 μM of each of the four deoxynucleoside triphosphates, 1.5 mM $MgCl_2$, 4 μM of each primer, 2.5% DMSO, 1 U of Phusion Hot Start DNA polymerase (Finnzymes), and *c.* 50 ng of cDNA. The following thermal cycling scheme was used: initial denaturation at 98 °C for 5 min, 25 cycles of denaturation at 98 °C for 45 s, annealing at 68 °C for 45 s, followed by extension at 72 °C for 30 s. The final extension was carried out at 72 °C for 5 min. Negative controls were performed using the reaction mixture without template. The V2–V3 region was amplified with the following set of primers according to Nacke *et al.* (2011) containing the Roche 454 pyrosequencing adaptors, keys, and one unique MID per sample (underlined): V2for 5′-GCCTCCCTCGCGCCATCAG-$(dN)_{10}$-AGTGGCGGACG GGTGAGTAA-3′ and V3rev 5′-GCCTTGCCAGCCCGCT CAGACAGTATATA-CGTATTACCGCGGCTGCTG-3′.).

The resulting PCR products were checked for appropriate size and purified by employing the peqGOLD gel

extraction kit (Peqlab, Erlangen, Germany) as recommended by the manufacturer. Three independent PCRs were performed per sample, purified by gel extraction, and pooled in equal amounts. Quantification of the PCR products was performed using the Quant-iT dsDNA HS assay kit and a Qubit fluorometer (Invitrogen) as recommended by the manufacturer. The Göttingen Genomics Laboratory determined the 16S rRNA gene sequences using a Roche GS-FLX 454™ pyrosequencer and titanium chemistry (Roche).

Processing and analysis of pyrosequencing derived data sets

Generated 16S rRNA gene data sets were processed and analyzed employing the QIIME software package (version 1.6; Caporaso *et al.*, 2010) and other tools. A detailed workflow is depicted and described in Fig. S1. Sequence statistics are listed in Table S3. Generated mRNA data sets were analyzed by employing the MG-RAST platform (Meyer *et al.*, 2008). Sequence statistics are listed in Table S4. Community compositions was determined by best hit classification using the nonredundant multisource annotation database (M5NR) as annotation source and a minimum *e*-value of 1e-5, a minimal identity of 60%, and a minimum alignment length of 50 bp. The abundance of different phylogenetic groups was defined by raw sequence numbers of classified sequences. Functional analysis was performed by hierarchical classification using MG-RAST subsystems as annotation source and a minimum e-value of 1e-5, a minimal identity of 60%, and a minimum alignment length of 50 amino acids.

Statistical analysis

To validate the impact of phytoplankton bloom presence on the measured environmental parameters as well as on calculated diversity indices, we tested for significance of the parameters/index with respect to bloom presence by employing the Student's two-sample *t*-test (homogenous variances) or with the Welsh's two-sample *t*-test (heterogeneous variances) for normally distributed samples and with the Wilcoxon–Mann–Whitney test for non-normally distributed samples. The effect of the algal bloom onto relative abundances of prominent bacterial groups was tested by Dirichlet regression in *R* using the Dirichlet Reg package (Meier, 2012). The relative abundances were thereby treated as compositional data and tested against the presence/absence of the investigated bloom. Effects were regarded as statistically significant by $P \leq 0.05$. All statistical analyses were conducted employing R (version 2.15.0; R Development Core Team, 2009).

Deposition of sequence data

Sequence data were deposited in the sequence read archive of the National Center for Biotechnology Information under the accession number SRA061816.

Results and discussion

Study area and characteristics of sampling sites

In this survey, we examined the impact of a phytoplankton bloom on bacterial community structures in the southern North Sea of samples taken in and outside of a phytoplankton spring bloom by assessing the metatranscriptomes using direct sequencing of rRNA-depleted RNA and especially pyrosequencing-based analysis of 16S rRNA gene transcripts. During the survey in southern North Sea, *in situ* temperatures ranged from 8.18 to 11.83 °C with slightly colder temperatures in the western compared with the eastern region (Table S1). Salinity ranged from 30.64 to 32.71. Samples were taken at different locations and different depths within the German Bight (Fig. 1, Table S1). In the western region, four samples were taken at stations 2–4 and 17 with Chl *a* concentrations below 3.4 μg Chl *a* L^{-1}. Nine samples were taken in phytoplankton bloom areas with Chl *a* concentrations of > 4.4 μg Chl *a* L^{-1} (Fig. 1, Table S2), one at the southernmost station (station 1) and eight in the northeastern region (stations 5–8, 10, and 13–15). Nine samples were collected from 2 m depth, and the other 5, from 9 to 20 m (Table S1). At station 3, samples were collected from 2 and 12 m. Due to the shallow water depth of < 22 m at 12 of the 13 stations visited (Table S1) and the weak stratification of the water column, no systematic differences between values of any parameter from 2 m compared with 9–20 m were observed.

The phytoplankton bloom examined in this study was dominated by *Phaeocystis globosa* and various diatoms typical for spring in the southern North Sea such as *Thalassiosira* spp., *Chaetoceros* spp., and *Rhizosolenia* spp. (Wiltshire *et al.*, 2010). Mean concentrations of POC, PON, Chl *a*, and phaeopigments of stations in the bloom area were significantly higher than that outside the bloom area (Table S2). The opposite was recorded for bacterial cell numbers, which were significantly lower in bloom samples. Concentrations of SPM ranged between 2.35 and 11.3 mg L^{-1} without significant differences between the bloom and nonbloom areas. The same was recorded for the nitrite concentrations, whereas nitrate and phosphate concentrations were significantly correlated with bloom presence (Table S2).

16S rRNA gene-based assessment of active bacterioplankton community structures

Active bacterioplankton community compositions were assessed by pyrosequencing-based analysis of the V2–V3 region of the 16S rRNA gene amplified from environmental RNA by two-step RT-PCR. A total of 211 769 high-quality bacterial 16S rRNA gene sequences were recovered across all 14 analyzed samples (Table S3). Calculated rarefaction curves as well as diversity indices revealed that the majority of the bacterial community was recovered by the surveying effort (Fig. S2, Table S5). Approximately 93% of all estimated OTUs were recovered at 1% and 3% genetic divergence. At 20% genetic divergence, the OTU coverage was *c.* 85%, indicating that the majority but not all bacterial phyla were recovered by the surveying effort. Calculated Shannon indices ranged from 4.6 to 5.5, 4.4 to 5.4, and 1.4 to 2.2 at 1%, 3%, and 20% genetic distance, respectively.

Classification of the 16S rRNA gene sequences revealed that *Proteobacteria* was the most abundant bacterial phylum across all samples (*c.* 98.6% of all sequences). Other identified phyla present in all samples were *Bacteroidetes* (0.9%) and *Actinobacteria* (0.3%). The majority of all 16S rRNA gene sequences (*c.* 95%) were further affiliated to 13 bacterial groups, clades, and genera (Fig. 2). The low abundance of *Bacteroidetes* was surprising, as the prominence of this phylum in bacterioplankton communities has been shown in other studies (e.g. Jamieson *et al.*, 2012).

Most of the sequences were assigned to *Gammaproteobacteria.* The SAR92 clade (21.8%), the SAR86 clade (5.9%), the OM60 clade (10.3%), the OM182 clade (13.4%), and the genus *Pseudospirillum* (3.5%) were identified as the most abundant phylogenetic groups in this class (Fig. 2). *Alphaproteobacteria* were the second most abundant proteobacterial class in all samples. The majority of 16S rRNA gene sequences (19.5%) were assigned to the *Roseobacter* RCA cluster (Newton *et al.*, 2010; Giebel *et al.*, 2013; Hahnke *et al.*, 2013), the largest cluster of the *Roseobacter* clade (Buchan *et al.*, 2005). It has been consistently detected at various locations and over several years in the German Bight and in the entire North Sea (Selje *et al.*, 2004; Giebel *et al.*, 2011; Sperling *et al.*, 2012).

Another lineage found in all samples was the SAR11 clade (5%) (Fig. 2), which is abundant globally in the ocean and also in the North Sea (Morris *et al.*, 2002; Giebel *et al.*, 2011; Teeling *et al.*, 2012). The high abundance of the *Roseobacter* clade was also supported by our results, but the SAR11 clade was substantially less abundant compared with previous DNA-based studies (Giebel *et al.*, 2011; Sperling *et al.*, 2012; Teeling *et al.*, 2012). The overall low abundance of SAR11 on RNA level might be explained by a low metabolic activity. Observed

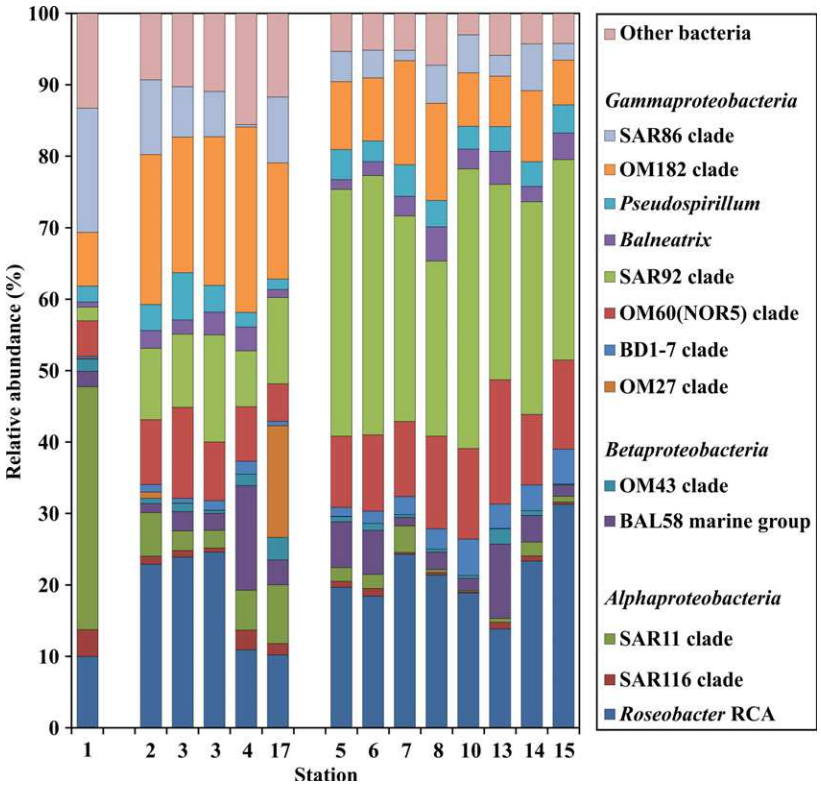


Fig. 2. Relative distribution of abundant bacterial lineages at stations outside (1–4 and 17) and inside (5–8, 10, and 13–15) the examined phytoplankton bloom. Only groups with an average abundance of more than 1% are shown. Station 1 is separated, as it was located in a bloom outside of the examined bloom.

changes in relative abundance in RNA-based studies are not necessarily solely caused by changes in the total number of corresponding bacterial cells. These changes can also partly result from increasing or decreasing metabolic activity, which directly affects the number of ribosomes present in the cell. A low activity of SAR11 has been reported in other studies. For example, Alonso & Pernthaler (2006) showed that SAR11 is highly abundant, but not very active in costal North Sea waters. In addition, West *et al.* (2008) demonstrated that SAR11 was more abundant in the Southern Ocean on DNA level than on RNA level.

The third most abundant proteobacterial class was *Betaproteobacteria* (5.7%). Sequences were mainly affiliated to BAL58 marine group (4.3%). The name of this group originated from strain BAL58, which is an obligate oligotrophic marine bacterium (Simu & Hagström, 2004). This strain was also detected by DGGE-based analysis of the bacterioplankton community in the central Baltic Sea (Riemann *et al.*, 2008).

Assessment of bacterioplankton composition based on direct sequencing of mRNA

To gain initial insights into gene expression patterns, three samples from nonbloom and bloom locations were selected for direct metatranscriptome sequencing and assessment of bacterioplankton community structures by phylogenetic classification of the mRNA reads: one sample at the southernmost station with high concentrations of SPM (station 1), one sample from the northeastern bloom with the highest chlorophyll *a* concentration (station 5), and one sample from outside the bloom (station 17). For this purpose, enriched mRNA was converted to cDNA and directly sequenced. Different from other studies (e.g. McCarren *et al.*, 2010; Gifford *et al.*, 2013), we omitted amplification steps prior to sequencing to avoid bias formation. Generated mRNA sets contained a total of 988 202 sequences (Table S4). The amount of residual rRNA ranged from 5.8% to 76.8%. Between 46 238 and 360 032 hits in the M5 nonredundant protein database were used for taxonomic classification.

Based on the phylogenetic affiliation of the deduced proteins, *Bacteria* encompassed the major fraction (Fig. S3). *Alphaproteobacteria* and *Gammaproteobacteria* were the dominant active classes (Fig. 3a). *Betaproteobacteria*, *Actinobacteria*, and *Cyanobacteria* were of minor abundance. *Bacteroidetes* encompassed 34% of the transcripts at station 5, but < 8% at the other two stations. *Alphaproteobacteria* were predominant in the mRNA-based analysis of the latter samples. Different from the 16S rRNA gene analysis, *Gammaproteobacteria* (26%) were less abundant in the mRNA analysis than *Alphaproteobacteria*

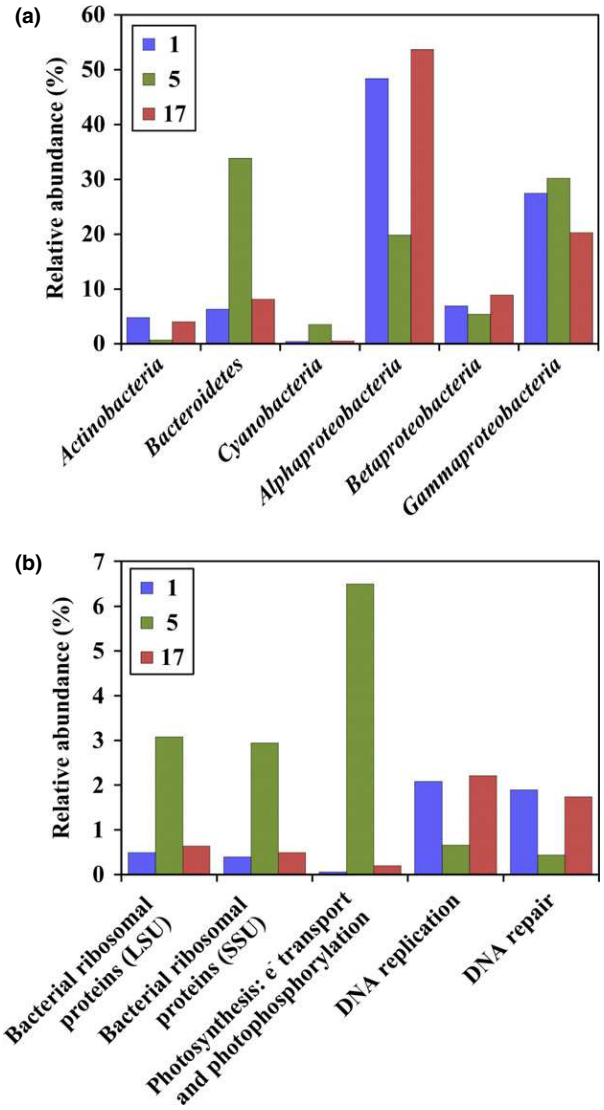


Fig. 3. MG-RAST-based phylogenetic (a) and gene expression analysis (b) of mRNA data sets derived from stations 1, 5, and 17. The sample from station 5 was taken in the center of a phytoplankton bloom with the highest chlorophyll *a* concentration measured during the survey (11.45 μg L^{-1}). Stations 1 and 17 were taken outside the examined bloom and exhibited chlorophyll *a* concentration of < 4.4 μg L^{-1}.

(Fig. 3a). Almost 41% of all identified proteobacterial mRNA sequences were affiliated to the latter class.

The differences in the relative abundance of certain bacterial groups between both approaches highlight common problems of both analysis types. PCR-based methods usually result in a biased picture of bacterial communities, for example, by primer pair mismatches (Klindworth *et al.*, 2013). Direct sequencing approaches do not introduce bias like PCR-based methods, but one disadvantage is the reliability on the corresponding and well-annotated

entries in sequence databases. For example, the SAR11 clade was abundant in the 16S rRNA gene analysis of sample 1, but was not detected during mRNA analysis of the corresponding sample. This could be explained by the low number of entries for SAR11 in public databases. The M5NR database used by MG-RAST (Meyer *et al.*, 2008) is an integration of many sequence databases including NCBI, RDP, and KEGG databases into one single, searchable database. Despite its size, this database harbors only 16 477 entries for SAR11, of which many derive from *Pelagibacter ubique*. Another example for the reliability on the corresponding and well-annotated entries in sequence databases is the study by Ottesen *et al.* (2013). Here, the authors focused on dynamics among five abundant microbial lineages and genera in the investigated marine community, that is, *Ostreococcus*, *Synechococcus*, *Pelagibacter*, SAR86 cluster (*Gammaproteobacteria*), and the marine group II (*Euryarchaeota*). Interestingly, the majority of the bacterial community was not studied as < 31% of the reads with matches in the NCBI nonredundant peptide database were affiliated to the five groups. We avoided performing a phylogenetic analysis of the transcripts at high taxonomic resolution because most phylogenetic lineages detected by the 16S rRNA gene amplicon sequencing approach are not represented in the M5NR protein database.

Other bacterial groups, for example, the *Roseobacter* clade, are relatively well represented in public databases due to the high number of physiologically diverse isolates and corresponding genome sequences (González *et al.*, 2000; Brinkhoff *et al.*, 2008). However, recent studies showed that also abundant *Roseobacter* clade members thriving in bacterioplankton communities are still largely missing in these databases, as corresponding isolates are not yet available (Luo *et al.*, 2012). Further analysis of the mRNA data sets, however, revealed that the majority of the annotated protein hits derived from the three samples were affiliated to the *Rhodobacterales* and, in particular, to different members of the *Roseobacter* clade. The 16S rRNA gene analysis of the corresponding samples revealed that *c.* 91% of the 16S rRNA gene sequences assigned to the *Rhodobacterales* were affiliated to the RCA cluster. As this cluster is not represented in the genome databases so far, protein hits assigned to the *Rhodobacteraceae* putatively belong mainly to the RCA cluster, the predominant *Roseobacter* cluster in the German Bight.

Impact of environmental conditions on active bacterial community structures

We further investigated the impact of environmental conditions on active bacterioplankton community structures. For this purpose, the community structure response to different environmental parameters was examined by principal coordinate analysis (PCoA; Fig. 4).

As environmental conditions might be linked to bloom presence, we initially tested for correlations between the recorded parameters and the bloom. Only the sampling depth, the suspended particulate matter content (SPM), and the nitrite concentration exhibited no direct correlation with bloom presence and were further examined by PCoA. They exhibited no effect on diversity and community structure (data not shown). Nevertheless, PCoA with respect to bloom presence revealed that all samples taken in the presence of the phytoplankton bloom tend to cluster and, thus, indicate similarity in bacterial community composition (Fig. 4). Samples taken outside the algal bloom did not show cluster formation. Thus, these results support our first hypothesis that bacterioplankton community composition is affected by the bloom and the environmental conditions found during the bloom. Sample 1 taken in a bloom separated from the examined bloom area exhibits a different community structure than the other bloom samples.

Impact of the phytoplankton bloom onto bacterioplankton composition and richness

As bacterioplankton community structures and phytoplankton bloom presence exhibit a strong correlation in the PCoA, we investigated the effect of the bloom on bacterial richness as well as the abundances and activity of certain bacterial groups. We recorded differences in the number of OTUs by direct comparison of rarefaction curves calculated with regard to bloom presence (Fig. 5). Nonbloom samples exhibited a higher diversity at 1% and 3% genetic distance. In addition, Shannon indices calculated at all three genetic distance levels were significantly reduced in the bloom samples (Table S5). This indicates a lower bacterial diversity in the bloom than outside the bloom and confirms our second hypothesis that the bacterial richness is affected by the presence of the phytoplankton bloom. The overall reduction in diversity during the bloom indicates that certain members of the bacterioplankton community benefitted from the substrates provided by primary production of the phytoplankton species present. A specialized ecological niche is formed in which only certain members of the bacterioplankton can grow. This is in accordance with other studies (West *et al.*, 2008; Teeling *et al.*, 2012).

We further examined which bacterial groups were significantly affected by bloom presence using statistical analysis. In accordance with our first hypothesis, eight of the 13 identified abundant bacterial groups and genera were significantly influenced by the bloom (Table 1). The strongest response to phytoplankton bloom presence was

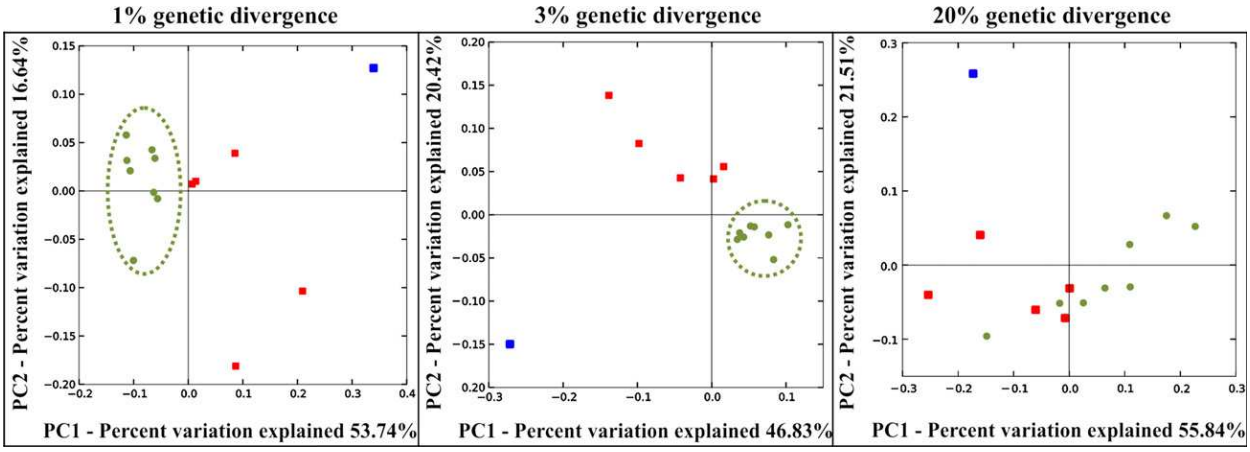


Fig. 4. Weighted UniFrac 2D PCoA plots calculated at 1% and 3% genetic distance employing QIIME (Caporaso *et al.*, 2010). To compare community structures, 10 353 randomly selected sequences from each sample were used for the calculation. Stations located inside of the phytoplankton bloom area are depicted as green dots, those outside the bloom, as red squares. Station 1 is shown as it was taken in a bloom separated from the examined bloom. Circles are drawn to better visualize putative structural similarities of the studied communities.

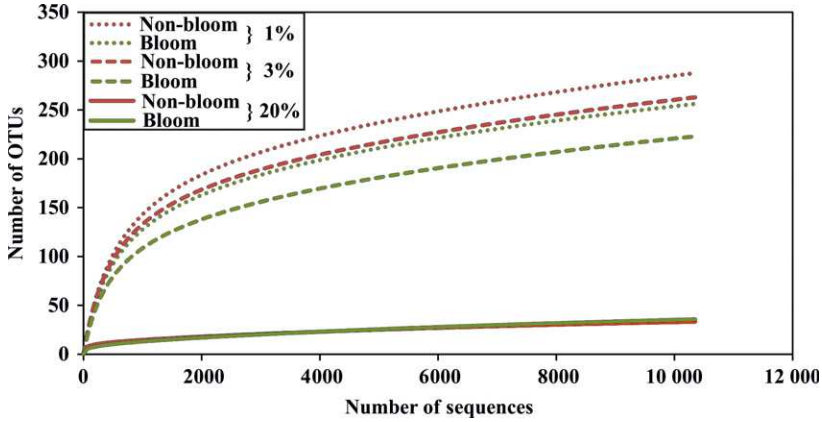


Fig. 5. Average rarefaction curves calculated for bloom and nonbloom samples at 1%, 3%, and 20% genetic distance.

recorded for the SAR92 clade. On average, the relative abundance of this clade increased threefold in bloom presence. It was shown that members of the SAR92 clade benefit from nutrient-rich conditions (Stingl *et al.*, 2007) as found during a phytoplankton bloom. Interestingly, Teeling *et al.* (2012) recorded no increase in abundance for the SAR92 clade during a bloom in the southern North Sea applying catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH). This might indicate that not the number of cells, but the activity of the SAR92 clade is increased under bloom conditions.

We also recorded significantly higher abundances of the OM60 (NOR5) clade and the BD1-7 clade in the bloom area. These two clades and the SAR92 clade do belong to *Oligotrophic Marine Gammaproteobacteria* (OMG) group (Cho & Giovannoni, 2004). Although genomic information on diverse OMG isolates is available (e.g. Stingl *et al.*, 2007; Huggett & Rappé, 2012), little is known about their ecological role.

Another bacterial group positively correlated with bloom presence was the RCA cluster (Table 1). Giebel *et al.* (2011) and Sperling *et al.* (2012) showed that the abundance of the RCA cluster was positively correlated with concentrations of Chl *a* and other biogeochemical properties characterizing phytoplankton blooms in the North Sea. Also West *et al.* (2008) reported that the RCA cluster was abundant in and outside of a phytoplankton bloom in the Southern Ocean. Moreover, two other *Roseobacter* clusters, NAC11-6 and NAC11-7, had higher relative abundances and higher activities in the bloom. González *et al.* (2000) showed that members of the *Roseobacter* clade, the SAR11 clade, and the SAR86 clade accounted for more than 50% of the bacterial 16S rRNA genes in a phytoplankton bloom in the North Atlantic. Our data suggest that members of RCA cluster were slightly more abundant or active during the bloom. According to Teeling *et al.* (2012), the abundance of phosphorus-related ABC-type transporters affiliated to the

Table 1. Relative abundances of dominant bacterial groups and correlations with phytoplankton bloom presence. Significance was tested by Dirichlet regression in *R* (significance level $P \leq 0.05$)

	Mean relative abundance (%)			
Bacterial group	Nonbloom	Bloom	Correlation	*P*
Alphaproteobacteria				
Roseobacter RCA	18.5 ± 7.3	21.4 ± 5.1	Positive	< 0.001
SAR11 clade	1.4 ± 0.86	0.58 ± 0.35	–	–
SAR116 clade	5.0 ± 2.4	1.4 ± 1.2	–	–
Betaproteobacteria				
BAL58 marine group	4.9 ± 5.5	4.2 ± 3.2	–	< 0.05
OM43 clade	1.4 ± 1.1	0.73 ± 0.64	–	–
Gammaproteobacteria				
Balneatrix	2.4 ± 0.92	3.0 ± 1.3	Positive	< 0.01
BD1-7 clade	0.5 ± 3.2	1.4 ± 0	Positive	< 0.001
OM182 clade	20.6 ± 3.6	9.7 ± 3.0	–	–
OM27 clade	3.3 ± 6.9	> 0.01	–	–
OM60 (NOR5) clade	8.65 ± 2.7	12.1 ± 2.5	Positive	< 0.001
Pseudospirillum	3.51 ± 2	3.6 ± 0.51	Positive	< 0.01
SAR86 clade	6.7 ± 3.9	4.0 ± 1.7	Negative	< 0.05
SAR92 clade	11.0 ± 2.7	31.1 ± 5.0	Positive	< 0.001

Rhodobacterales is enhanced in the late bloom stage, whereas the abundance of the *Roseobacter* clade, comprising mainly this order, was not affected. As for SAR92, not the abundance of members of the RCA cluster, but the activity might be affected by bloom presence in the German Bight. Interestingly, Teeling *et al.* (2012) recorded a slight decrease in SAR11 abundance during the bloom by CARD-FISH analysis. We also recorded 2.5-fold decrease in our analysis (Table 1). However, a significant correlation between SAR11 abundance and bloom presence was not recorded.

Impact of a phytoplankton bloom onto bacterial community structure and activity as assessed by direct sequencing of mRNA

Generated mRNA data sets were further analyzed to evaluate the impact of the phytoplankton bloom on structure as well as metabolic activity of the bacterioplankton. Phylogenetic analysis of the generated mRNA data sets revealed an increase in abundance of the *Bacteroidetes* in the sample taken in the area of the investigated bloom (sample 5) (Fig. 3a). *Bacteroidetes* are widespread in marine systems and play an important role in organic matter degradation (Gomez-Pereira *et al.*, 2010). The higher abundance of this phylum corresponds to the results of Teeling *et al.* (2012). They recorded a higher cell abundance of different flavobacterial genera during and after the bloom event. Therefore, *Bacteroidetes* seem to be one bacterial group benefitting from the conditions provided by the investigated phytoplankton bloom.

Functional analysis of the generated mRNA data sets also revealed differences between samples collected at stations with low and high Chl *a* concentrations. As expected, expression levels of genes affiliated to photosynthesis were enhanced at station 5, inside the examined bloom (Figs 3b and S4). More than 80% of these were further affiliated to different genes encoding proteins of the photosystem II such as PsbA, PsbB, PsbC, PsbD, and PspE (data not shown).

The number of transcripts assigned to protein metabolism was higher at this station compared with the other two stations (Fig. S4). A deeper analysis revealed that especially the number of transcripts affiliated to bacterial ribosomal protein-encoding genes was higher (Fig. 3b). An increased number of ribosomal proteins indicates an increased ribosome formation, which in turn might correspond to a higher bacterial activity during the bloom event. Interestingly, the number of transcripts affiliated to DNA metabolism, in particular DNA replication and repair (Fig. 3b), as well as the number of transcripts assigned to the cell cycle (Fig. S4) was lower at station 5. This might indicate a reduced bacterioplankton growth rate within the bloom and complies with the lower bacterial numbers inside the bloom area relative to outside the bloom. This corresponds to our second hypothesis and also to the hypothesis proposed by Teeling *et al.* (2012) that specialized ecological niches are formed during a phytoplankton bloom in which only certain members of the bacterioplankton can grow, whereas others are inhibited.

In conclusion, the investigated active bacterial communities in the southern North Sea were dominated by only a few marine groups such as the SAR92 clade and *Roseobacter* RCA cluster. The metatranscriptomic approaches revealed that most of the environmentally occurring abundant and active marine groups found in the bacterioplankton are underrepresented by isolates or type strains and correspondingly by reference genome sequences. In addition, the number of field studies targeting the active bacterial community either by metatranscriptomic or by metaproteomic approaches is limited. This study provides first insights into structural and functional changes in the active bacterioplankton community as response to a phytoplankton bloom. More studies targeting active bacterioplankton communities in combination with isolation and characterization of environmentally relevant strains are required to unravel the ecological role and ecosystem function of bacterioplankton community members in different marine ecosystems.

Acknowledgements

We thank the crew of RV Heincke for their valuable support; B. Schmidt from the Experimental and Applied Mineralogy Department of the Geoscience Center Göttingen for providing his muffle furnace; B. Kuerzel, A. Schlingloff, and R. Weinert for analysis of chlorophyll, SPM, POC, and PON, as well as M. Wurst, N. Händel, G. Wienhausen, and M. Wolterink for analysis of inorganic nutrients. This work was funded by Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center TRR 51.

References

Alderkamp A-C, Sintes E & Herndl GJ (2006) Abundance and activity of major groups of prokaryotic plankton in the coastal North Sea during spring and summer. *Aquat Microb Ecol* **45**: 237–246.

Alonso C & Pernthaler J (2006) Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environ Microbiol* **8**: 2022–2030.

Brinkhoff T, Giebel HA & Simon M (2008) Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch Microbiol* **189**: 531–539.

Buchan A, Gonzalez JM & Moran MA (2005) Overview of the marine Roseobacter lineage. *Appl Environ Microbiol* **71**: 5665–5677.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Cho J-C & Giovannoni SJ (2004) Cultivation and growth characteristics of a diverse group of oligotrophic marine gammaproteobacteria. *Appl Environ Microbiol* **70**: 432–440.

Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW & DeLong EF (2008) Microbial community gene expression in ocean surface waters. *P Natl Acad Sci USA* **105**: 3805–3810.

Giebel HA, Brinkhoff T, Zwisler W, Selje N & Simon M (2009) Distribution of *Roseobacter* RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environ Microbiol* **11**: 2164–2178.

Giebel HA, Kalhoefer D, Lemke A, Thole S, Gahl-Janssen R, Simon M & Brinkhoff T (2011) Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J* **5**: 8–19.

Giebel HA, Kalhoefer D, Gahl-Janssen R *et al.* (2013) *Planktomarina temperata* gen. nov., sp. nov., belonging to the globally distributed RCA cluster of the marine *Roseobacter* clade, isolated from the German Wadden Sea. *Int J Syst Evol Microbiol*, DOI: 10.1099/ijs.0.053249-0 [Epub ahead of print].

Gifford SM, Sharma S, Rinta-Kanto JM & Moran MA (2011) Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *ISME J* **5**: 461–472.

Gifford SM, Sharma S, Booth M & Moran MA (2013) Expression patterns reveal niche diversification in a marine microbial assemblage. *ISME J* **7**: 281–298.

Giovannoni SJ & Stingl U (2005) Molecular diversity and ecology of microbial plankton. *Nature* **437**: 343–348.

Giovannoni SJ, Britschgi TB, Moyer CL & Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60–63.

Gomez-Pereira PR, Fuchs BM, Alonso C, Oliver M, van Beusekom J & Amann R (2010) Distribution patterns and diversity of planktonic Flavobacterial clades in contrasting water masses of the North Atlantic Ocean. *ISME J* **4**: 472–487.

González JM, Simó R, Massana R, Covert JS, Casamayor EO, Pedrós-Alió C & Moran MA (2000) Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl Environ Microbiol* **66**: 4237–4246.

Hahnke S, Brock NL, Zell C, Simon M, Dickschat JS & Brinkhoff T (2013) Physiological diversity of *Roseobacter* clade bacteria co-occurring during a phytoplankton bloom in the North Sea. *Syst Appl Microbiol* **36**: 39–48.

Huggett MJ & Rappé MS (2012) Genome sequence of strain HIMB55, a novel marine gammaproteobacterium of the OM60/NOR5 clade. *J Bacteriol* **194**: 2393–2394.

Jamieson RE, Rogers AD, Billett DS, Smale DA & Pearce DA (2012) Patterns of marine bacterioplankton biodiversity in the surface waters of the Scotia Arc, Southern Ocean. *FEMS Microbiol Ecol* **80**: 452–468.

Kirchman DL, Cottrell MT & Lovejoy C (2010) The structure of bacterial communities in the western Arctic Ocean as revealed by pyrosequencing of 16S rRNA genes. *Environ Microbiol* **12**: 1132–1143.

Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M & Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**: e1.

Koroleff F & Grasshoff K (1983) Determination of nutrients. *Methods of Seawater Analysis*, Vol. 2 (Grasshoff K, Ehrhardt M & Kremling K, eds), pp. 125–188. Verlag Chemie, Weinheim, Germany.

Lesniewski RA, Jain S, Anantharaman K, Schloss PD & Dick GJ (2012) The metatranscriptome of a deep-sea hydrothermal plume is dominated by water column methanotrophs and lithotrophs. *ISME J* **6**: 2257–2268.

Lunau M, Lemke A, Dellwig O & Simon M (2006) Physical and biogeochemical controls of microaggregate dynamics in a tidally affected coastal ecosystem. *Limnol Oceanogr* **51**: 847–859.

Luo H, Löytynoja A & Moran MA (2012) Genome content of uncultivated marine *Roseobacters* in the surface ocean. *Environ Microbiol* **14**: 41–51.

McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW & DeLong EF (2010) Microbial community transcriptomes reveal microbes and

metabolic pathways associated with dissolved organic matter turnover in the sea. *P Natl Acad Sci USA* **107**: 16420–16427.
McQuatters-Gollop A, Raitsos DE, Edwards M, Pradhan Y, Mee LD, Lavender SJ & Attrill MJ (2007) A long-term chlorophyll data set reveals regime shift in North Sea phytoplankton biomass unconnected to nutrient trends. *Limnol Oceanogr* **52**: 635–648.
Meier MJ (2012) DirichletReg. Available at: http://cran.r-project.org/web/packages/DirichletReg.
Meyer F, Paarmann D, D'Souza M *et al.* (2008) The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.
Morris RM, Rappe MS, Connon SA, Vergin KL, Siebold WA & Carlson CA (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806–810.
Muyzer G, de Waal EC & Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695–700.
Nacke H, Thurmer A, Wollherr A, Will C, Hodac L, Herold N, Schoning I, Schrumpf M & Daniel R (2011) Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils. *PLoS ONE* **6**: e17000.
Newton RJ, Griffin LE, Bowles KM *et al.* (2010) Genome characteristics of a generalist marine bacterial lineage. *ISME J* **4**: 784–798.
Nusch EA (1999) Chlorophyllbestimmung. *Biologische Gewässeruntersuchung* (von Tuempling W & Friedrich G, eds), pp. 368–375. G Fischer, Stuttgart, Germany.
Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA & DeLong EF (2013) Pattern and synchrony of gene expression among sympatric marine microbial populations. *P Natl Acad Sci USA* **110**: 488–497.
R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
Riemann L, Leitet C, Pommier T, Simu K, Holmfeldt K, Larsson U & Hagström Å (2008) The native bacterioplankton community in the Central Baltic Sea is influenced by freshwater bacterial species. *Appl Environ Microbiol* **74**: 503–515.
Rink B, Seeberger S, Martens T, Duerselen C-D, Simon M & Brinkhoff T (2011) Regional patterns of bacterial community composition and biogeochemical properties in the southern North Sea. *Aquat Microb Ecol* **63**: 207–222.
Schlitzer R (2013) Ocean data view. Available at: http://odv.awi.de.
Schmidt TM, DeLong EF & Pace NR (1991) Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J Bacteriol* **173**: 4371–4378.
Selje N, Simon M & Brinkhoff T (2004) A newly discovered *Roseobacter* cluster in temperate and polar oceans. *Nature* **427**: 445–448.
Simu K & Hagström Å (2004) Oligotrophic bacterioplankton with a novel single-cell life strategy. *Appl Environ Microbiol* **70**: 2445–2451.
Sintes E, Witte H, Stodderegger K, Steiner P & Herndl GJ (2013) Temporal dynamics in the free-living bacterial community composition in the coastal North Sea. *FEMS Microbiol Ecol* **83**: 413–424.
Sperling M, Giebel HA, Rink B, Grayek S, Staneva J, Stanev E & Simon M (2012) Differential effects of hydrographic and biogeochemical properties on the SAR11 clade and *Roseobacter* RCA cluster in the North Sea. *Aquat Microb Ecol* **67**: 25–34.
Stingl U, Desiderio RA, Cho JC, Vergin KL & Giovannoni SJ (2007) The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Appl Environ Microbiol* **73**: 2290–2296.
Teeling H, Fuchs BM, Becher D *et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608–611.
Venter JC, Remington K, Heidelberg JF *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.
Vila-Costa M, Gasol JM, Sharma S & Moran MA (2012) Community analysis of high- and low-nucleic acid-containing bacteria in NW Mediterranean coastal waters using 16S rDNA pyrosequencing. *Environ Microbiol* **14**: 1390–1402.
Weinbauer MG, Fritz I, Wenderoth DF & Höfle MG (2002) Simultaneous extraction from bacterioplankton of total RNA and DNA suitable for quantitative structure and function analyses. *Appl Environ Microbiol* **68**: 1082–1087.
Wemheuer B, Wemheuer F & Daniel R (2012) RNA-based assessment of diversity and composition of active archaeal communities in the German Bight. *Archaea* **2012**: 695826.
West NJ, Obernosterer I, Zemb O & Lebaron P (2008) Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. *Environ Microbiol* **10**: 738–756.
Will C, Thurmer A, Wollherr A, Nacke H, Herold N, Schrumpf M, Gutknecht J, Wubet T, Buscot F & Daniel R (2010) Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. *Appl Environ Microbiol* **76**: 6751–6759.
Wiltshire K, Kraberg A, Bartsch I, Boersma M, Franke H-D, Freund J, Gebühr C, Gerdts G, Stockmann K & Wichels A (2010) Helgoland roads, North Sea: 45 years of change. *Estuaries Coasts* **33**: 295–310.

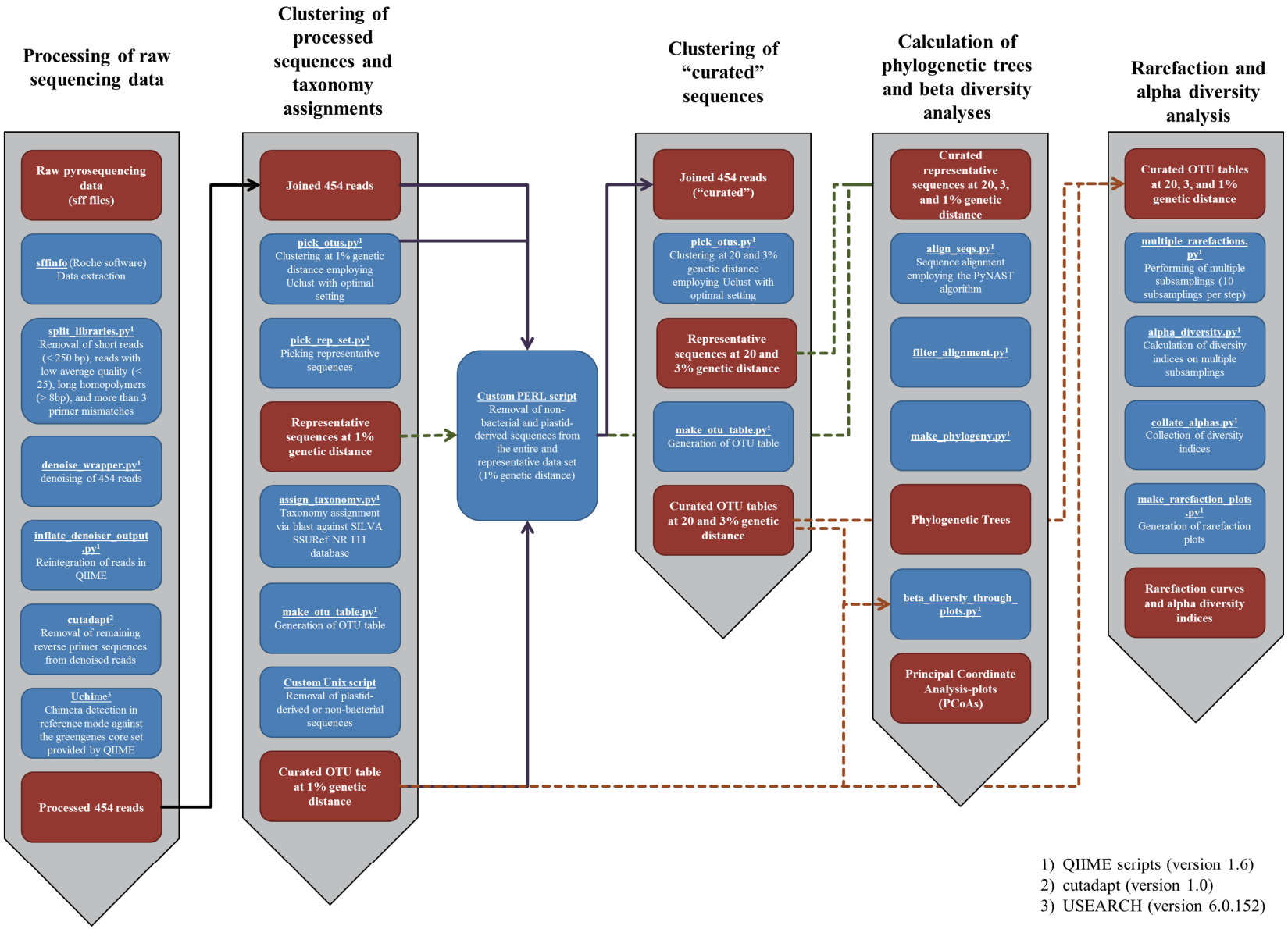
Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Processing of amplicon-based 454 pyrosequencing reads with QIIME and other tools.
Fig. S2. Rarefaction curves for all 14 samples at 1% (a), 3% (b), and 20% (c) genetic divergence.
Fig. S3. MG-RAST-based taxonomic assignment of mRNA reads derived from stations 1, 5 and 17.
Fig. S4. MG-RAST-based hierarchical protein annotation using subsystems as annotation source.
Table S1. Sampling time and position as well as CTD measured parameters of the sampling sites.
Table S2. Environmental parameters determined for the samples.
Table S3. Sequence statistics of the 16S rRNA gene data sets.
Table S4. Statistics of the mRNA analysis according to MG-RAST (Meyer *et al.*, 2008).
Table S5. Bacterial diversity and richness at 1%, 3% and 20% genetic distance.

SUPPORTING INFORMATION FOR STUDY 2

CONTENTS:



1) QIIME scripts (version 1.6)
2) cutadapt (version 1.0)
3) USEARCH (version 6.0.152)

Fig. S1. Processing of amplicon-based 454 pyrosequencing reads with QIIME and other tools. After raw data extraction, reads shorter than 250 bp, with an average quality value below 25, possessing long homopolymer stretches (>8 bp), or primer mismatches (>3) were removed. Subsequently, sequences were denoised employing QIIME (Caporaso *et al.*. 2010). Remaining primer sequences were truncated employing cutadapt (Martin. 2011). Chimeric sequences were removed using UCHIME and the most recent Greengenes core set as reference dataset (DeSantis *et al.*. 2006. Edgar *et al.*. 2011). Processed sequences of all samples were joined, sorted by decreasing length, and clustered employing the UCLUST algorithm (Edgar 2010). Sequences were clustered in operational taxonomic units (OTUs) at 1%, 3% and 20% genetic dissimilarity according to Simon *et al.* (2009). OTUs at 3 and 20% sequence divergence represent species and phylum level, respectively (Schloss & Handelsman. 2005). Phylogenetic composition was determined using the QIIME assign_taxonmy.py script. A BLAST alignment against the Silva SSURef 111 NR database (Pruesse *et al.*. 2007) was thereby performed. Sequences were classified with respect to the silva taxonomy of their best hit. Rarefaction curves, Shannon indices (Shannon, 2001) and Chao1 indices (Chao & Bunge. 2002) were calculated. In addition, the maximal number of OTUs (n_{max}) was estimated for each sample using the Michaelis-Menten-fit alpha diversity metrics included in the QIIME software package. To compare bacterial community structures across all samples based on phylogenetic or count-based distance metrics, Principal Coordinate Analysis (PCoA) plots were generated. A phylogenetic tree was calculated prior to PCoA generation. For this purpose, sequences were aligned using the PyNAST algorithm. The phylogenetic tree and the corresponding OTU table were subsequently used to generate PCoA plots.

References

Caporaso JG, Kuczynski J, Stombaugh J, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat methods* **7**: 335-336.

Chao A & Bunge J (2002) Estimating the number of species in a stochastic abundance model. *Biometrics* **58**: 531-539.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P & Andersen GL (2006) Greengenes. a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069-5072.

Edgar RC, Haas BJ, Clemente JC, Quince C & Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194-2200.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17: 10-12.**

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J & Glockner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188-7196.

Schloss PD & Handelsman J (2005) Introducing DOTUR. a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71**: 1501-1506.

Shannon CE (2001) A mathematical theory of communication. *SIGMOBILE Mob Comput Commun Rev* **5**: 3-55.

Simon C, Wiezer A, Strittmatter AW & Daniel R (2009) Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. *Appl Environ Microbiol* **75**: 7519-7526.

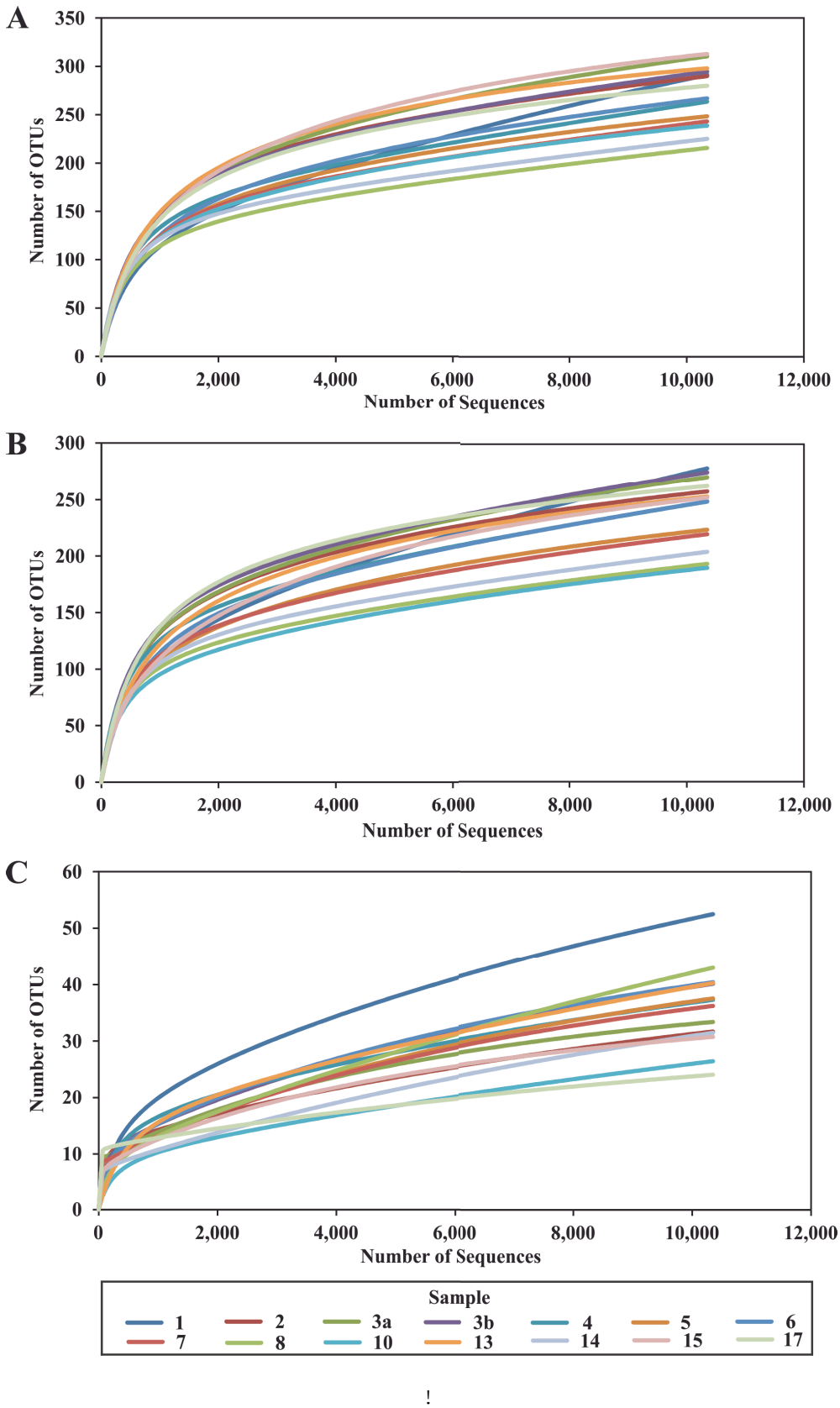


Fig. S2. Rarefaction curves for all 14 samples at 1 (A), 3 (B), and 20% (C) genetic divergence. Rarefaction analysis was performed with QIIME (Caporaso *et al.*. 2010).

References

Caporaso JG, Kuczynski J, Stombaugh J, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335-336.

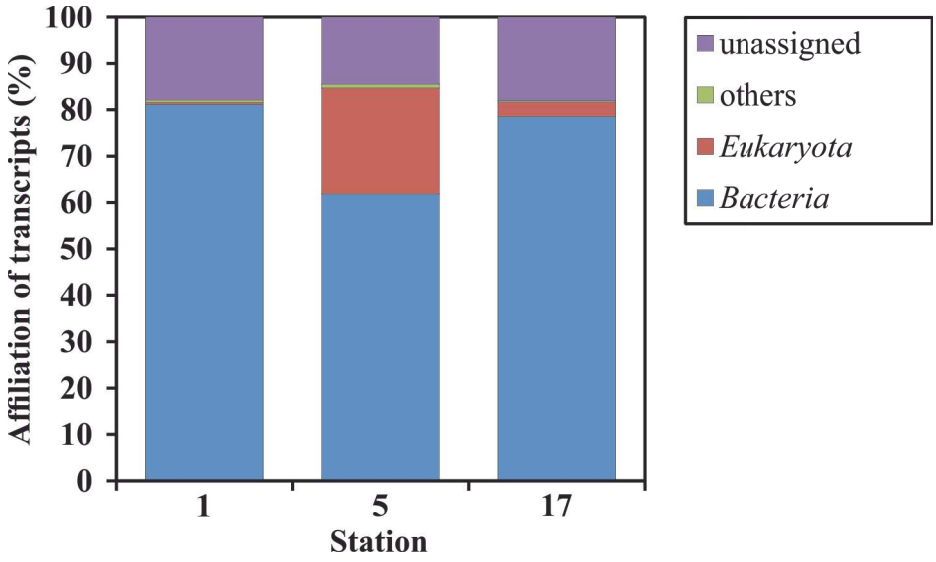


Fig. S3. MG-RAST-based taxonomic assignment of mRNA reads derived from stations 1, 5 and 17. The bioinformatic tools used were all part of MG-RAST (Meyer *et al.*, 2008). Sample from station 5 was taken in the center of a phytoplankton bloom with the highest chlorophyll *a* concentration measured during the survey (11.45 µg L^{-1}). Stations 1 and 17 were taken outside the examined bloom and exhibited chlorophyll *a* concentrations <4.4 µg L^{-1}.

Reference

Meyer F, Paarmann D, D'Souza M, *et al.* (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.

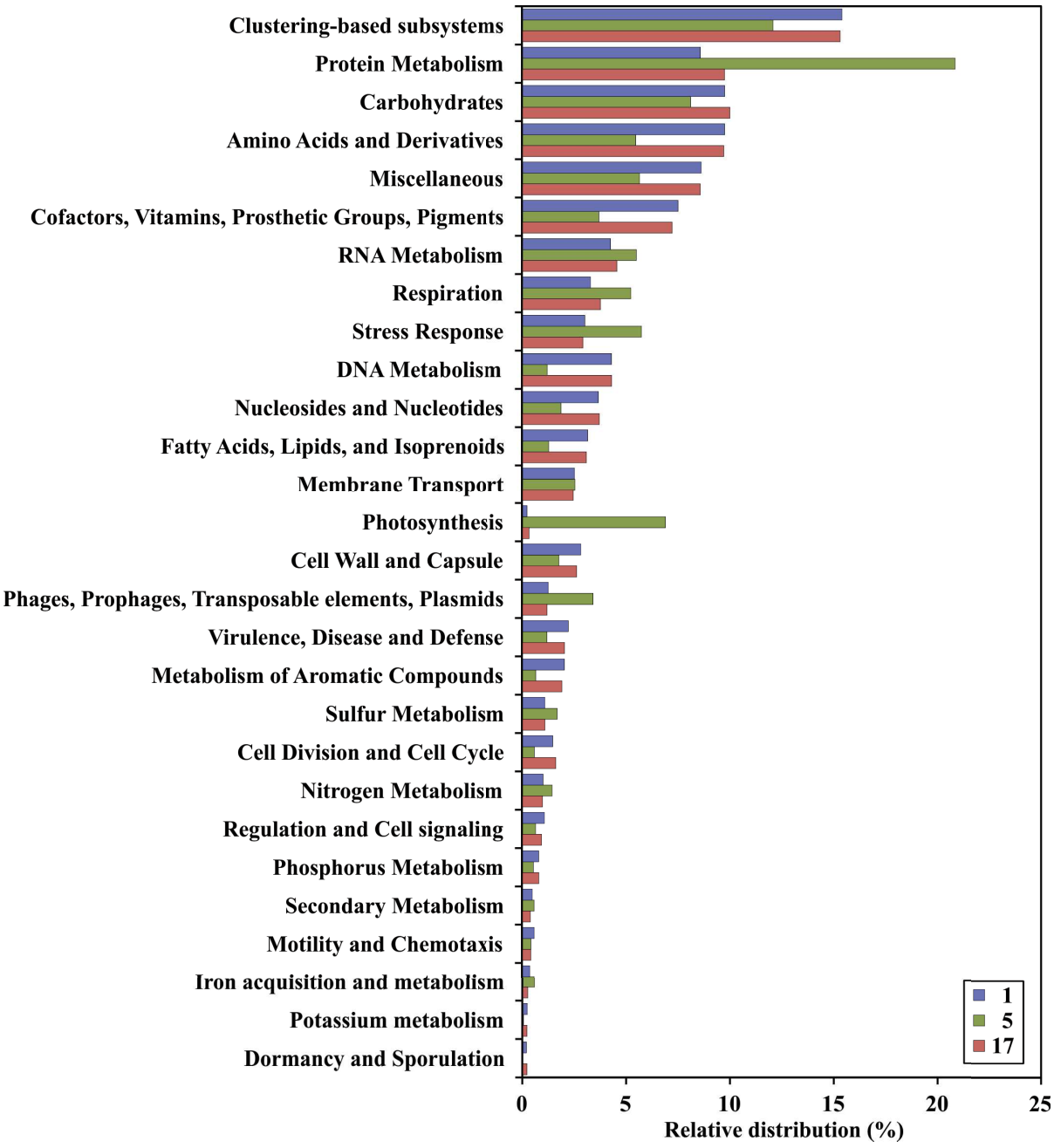


Fig. S4. MG-RAST-based hierarchical protein annotation using subsystems as annotation source. The bioinformatic tools used were all part of MG-RAST (Meyer *et al.*, 2008). The sample from station 5 (green) was taken in the center of a phytoplankton bloom with the maximum chlorophyll *a* concentration of all samples analyzed in this survey (11.45 µg L^{-1}). The samples from stations 1 and 17 exhibited chlorophyll *a* concentrations <4.4 µg L^{-1} and were taken outside the examined bloom.

Reference

Meyer F, Paarmann D, D'Souza M, *et al.* (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.

Table S1. Sampling time and position as well as CTD measured parameters of the sampling sites. Significant differences between bloom and non-bloom samples are shown in bold type. The significance was tested either with the Student's two-sample t-test (homogenous variances) or with the Welsh Two Sample t-test (heterogeneous variances) for normally distributed samples and with the Wilcoxon–Mann–Whitney test for not-normally distributed samples. Sample 1 was excluded from statistical testing as it was taken in a bloom different to the examined phytoplankton bloom.

Sample	Ship station	Date	Latitude (°N)	Longitude (°E)	Depth (m)	Bottom Depth (m)	Temperature (°C)	Salinity (psu)	Fluorescence (FU)	Transmission (%)	Density (g/L)
1	655	05/25/2010	53.8955	8.0496	2	15.5	11.09	30.24	1.21	57.20	1023.10
Non-bloom											
2	656	05/25/2010	54.0813	7.9338	12	34.5	9.96	30.81	0.27	84.74	1023.77
3a	657	05/26/2010	54.4223	7.6833	2	22.1	9.43	31.42	0.20	87.98	1024.27
3b	657	05/26/2010	54.4223	7.6833	12	22.1	8.18	32.01	0.77	84.83	1024.93
4	658	05/26/2010	54.7626	7.4463	2	20	9.73	32.71	0.49	81.23	1025.20
17	673	05/30/2010	54.4202	7.6922	2	22	10.75	31.6	NA(*)	87.74	1024.20
Mean	-	-	-	-	-	-	**9.61**	**31.71**	**0.43**	**85.30**	**1024.47**
SD	-	-	-	-	-	-	**0.94**	**0.71**	**0.26**	**2.75**	**0.58**
CV	-	-	-	-	-	-	**0.10**	**0.02**	**0.59**	**0.03**	**0.00**
Bloom											
5	659	05/26/2010	54.4575	7.9893	9	12.5	10.80	30.64	2.76	60.14	1023.50
6	660	05/27/2010	54.4542	8.0018	2	12.5	10.83	30.65	1.89	72.78	1023.40
7	661	05/27/2010	54.3054	8.2807	10	13.5	11.37	30.68	2.19	66.10	1023.40
8	663	05/27/2010	54.5065	8.1645	2	12	11.31	31.05	2.39	73.92	1023.67
10	665	05/28/2010	54.5135	8.128	2	11	11.40	31.11	2.27	74.79	1023.70
13	668	05/29/2010	54.4365	8.2328	10	12	11.83	31.18	2.80	67.83	1023.70
14	670	05/29/2010	54.4595	8.2328	2	11	11.43	30.83	NA(*)	75.72	1023.50
15	671	05/30/2010	54.449	8.22	2	12	11.70	31.04	NA(*)	76.59	1023.60
Mean	-	-	-	-	-	-	**11.33**	**30.90**	**2.38**	**70.98**	**1023.56**
SD	-	-	-	-	-	-	**0.37**	**0.22**	**0.35**	**5.75**	**0.13**
CV	-	-	-	-	-	-	**0.03**	**0.01**	**0.15**	**0.08**	**0.00**

(*) not measured.

Table S2. Environmental parameters determined for the samples. Significant differences between bloom and non-bloom samples are shown in bold type. The significance was tested either with the Student's two-sample t-test (homogenous variances) or with the Welsh Two Sample t-test (heterogeneous variances) for normally distributed samples and with the Wilcoxon–Mann–Whitney test for not-normally distributed samples. Sample 1 was excluded from statistical testing as it was taken in a bloom different to the examined phytoplankton bloom.

Station	Chlorophyll a ($\mu g\ L^{-1}$)	Phaeo-pigments ($\mu g\ L^{-1}$)	Phaeo-pigments (% total Chl)	Bacterial abundance ($10^6\ mL^{-1}$)	Suspended particulate matter ($mg\ L^{-1}$)	Particulate organic carbon ($\mu g\ L^{-1}$)	Particulate organic nitrogen ($\mu g\ L^{-1}$)	Nitrate (μM)	Nitrite (μM)	Mono-nitrogen oxides (μM)	Phosphate (μM)
1	4.38	2.11	32.5	4.00	9.80	997.4	152.1	8.5	0.19	8.65	
Non-bloom											
2	1.01	0.37	26.8	1.96	8.73	282.4	42.0	9.8	0.28	10.07	0.050
3a	1.12	0.25	18.3	2.57	4.60	291.2	43.0	7.4	0.25	7.65	0.020
3b	3.37	1.08	24.3	2.54	7.15	496.4	82.2	5.9	0.27	6.17	0.040
4	2.55	0.37	12.7	0.83 (*)	6.15	290.4	46.9	6.2	0.24	6.41	0.030
17	1.07	0.18	14.4	2.22	2.35	332.2	49.4	7.6	0.19	7.76	0.040
Mean	**1.82**	**0.45**	**19.30**	**2.02**	5.80	**338.52**	**52.70**	**7.36**	0.25	**7.61**	**0.036**
SD	**1.08**	**0.36**	**6.12**	**0.71**	2.44	**90.37**	**16.76**	**1.54**	0.04	**1.55**	**0.011**
CV	**0.59**	**0.80**	**0.32**	**0.35**	0.42	**0.27**	**0.32**	**0.21**	0.14	**0.20**	**0.317**
Bloom											
5	11.45	7.03	38.0	1.84	3.27	1673.0	213.6	5.0	0.24	5.27	0.100
6	7.34	2.77	27.4	2.00	11.30	728.2	106.2	9.1	0.42	9.53	0.080
7	7.28	2.91	28.6	1.82	7.50	1079.9	133.8	2.7	0.30	3	0.060
8	6.81	2.18	24.3	1.12	10.40	958.0	174.1	3.6	0.29	3.86	0.060
10	6.93	2.10	23.3	1.21	7.50	737.5	95.3	3.7	0.29	3.98	0.070
13	5.53	3.16	36.4	1.66	9.91	936.5	123.5	2.1	0.21	2.29	0.100
14	4.44	1.43	24.4	2.00	9.05	638.4	80.1	3.2	0.21	3.37	0.070
15	5.33	1.58	22.9	1.37	6.25	624.4	83.9	3.0	0.30	3.27	0.080
Mean	**6.89**	**2.90**	**28.16**	**1.66**	8.15	**921.99**	**126.31**	**4.04**	0.28	**4.32**	**0.078**
SD	**2.12**	**1.78**	**5.93**	**0.36**	2.60	**344.56**	**46.70**	**2.22**	0.07	**2.28**	**0.016**
CV	**0.31**	**0.62**	**0.21**	**0.22**	0.32	**0.37**	**0.37**	**0.55**	0.24	**0.53**	**0.204**

(*) excluded as outlier.

Table S3. Sequence statistics of the 16S rRNA datasets.

Sample	Before preprocessing		After preprocessing		After denoising and removal of non-bacterial or chimeric sequences	
	No. of sequences	Average length	No. of sequences	Average length	No. of sequences	Average length
1	23,400	355.2 bp	15,958	424.2 bp	12,202	398.8 bp
2	34,759	342.6 bp	23,176	433.9 bp	17,866	404.3 bp
3a	31,754	349.3 bp	21,727	434.4 bp	15,882	406.1 bp
3b	24,831	364.2 bp	18,074	434.3 bp	13,189	406.8 bp
4	25,020	358.2 bp	16,908	431.8 bp	11,529	408.7 bp
5	23,768	359.5 bp	16,330	432.0 bp	12,963	407.8 bp
6	21,200	360.3 bp	14,584	432.3 bp	10,710	409.1 bp
7	20,507	358.9 bp	14,495	436.2 bp	11,827	408.4 bp
8	17,318	361.3 bp	12,368	437.5 bp	10,354	408.2 bp
10	26,104	356.7 bp	15,865	439.4 bp	13,075	408.9 bp
13	43,113	355.5 bp	28,777	439.5 bp	21,798	413.4 bp
14	21,040	368.0 bp	14,618	433.6 bp	10,633	406.9 bp
15	80,326	349.8 bp	54,407	437.0 bp	32,574	406.3 bp
17	34,917	373.3 bp	25,973	435.0 bp	17,167	404.7 bp
Sum	428,057	357 bp	293,260	435 bp	211,769	407 bp

Table S4. Statistics of the mRNA analysis according to MG-RAST (Meyer *et al.*. 2008).

Sample	Number of sequences	Total number of basepairs	Average read length (bp)	Number of ribosomal RNAs	Number of bacterial hits in the M5NR database used for taxonomic classification	Number of bacterial hits annotated to MG-RAST subsystems
1	478,024	173,971,335	363	27,926 (5.8%)	359,387	235,139
5	389,593	139,023,923	357	299,333 (76.8%)	46,238	37,620
17	120,585	39,448,347	327	8,884 (7.4%)	62,389	37,489

Reference

Meyer F, Paarmann D, D'Souza M*, et al.* (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.

STUDY 3:

FROM GERMANY TO NORWAY: SIMULTANEOUS ASSESSMENT OF TOTAL AND ACTIVE BACTERIAL COMMUNITY STRUCTURES IN THE NORTH SEA ALONG A LATITUDINAL GRADIENT

WEMHEUER B[1], MEIER D[1,4], BILLERBECK S[2], GIEBEL HA[2], SCHERBER C[3], SIMON M[2], AND DANIEL R[1]

(IN PREPARATION)

[1]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY; [2]INSTITUTE FOR CHEMISTRY AND BIOLOGY OF THE MARINE ENVIRONMENT (ICBM), CARL-VON-OSSIETZKY-UNIVERSITY OF OLDENBURG, CARL-VON-OSSIETZKY-STR. 9-11, D-26111 OLDENBURG, GERMANY; [3]DEPARTMENT FOR CROP SCIENCES, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 6, D-37077 GÖTTINGEN, GERMANY; [4]PRESENT ADDRESS: MAX PLANCK INSTITUTE FOR MARINE MICROBIOLOGY, CELSIUSSTRASSE 1, D-28359 BREMEN, GERMANY

Author contributions to the work:

Performed the experiments: BW, DM

Analyzed data: BW, DM, CS

Contributed data on water properties and analysis of these data: SB, HAG, MS

Wrote the publication: BW, RD

Conceived and designed the experiments: BW, RD

ABSTRACT

Marine bacterioplankton communities play major roles in many ecological important pathways and nutrient cycles. However, the contribution of many marine groups to ecosystem functioning is largely unexplored. Little is known about the active community members, as most studies were mainly focused on structure of total bacterial communities. In this survey, we assessed total and active bacterial community structures in North Sea water samples along a latitudinal gradient from Germany to Norway by pyrosequencing-based analysis of 16S rRNA amplicons generated from environmental DNA and RNA, respectively. The generated 16S rRNA gene datasets comprised 382,507 sequences across all 13 samples with an average read length of 519 bp. *Proteobacteria* and *Bacteroidetes* were identified as the most abundant phylogenetic groups across all samples. *Cyanobacteria* and *Verrucomicrobia* were found in minor abundances. Sequences were mainly affiliated to different flavobacterial groups and genera such as the marine groups NS5 and NS7, different clusters of the *Roseobacter* clade such as CHAB-I-5 and NAC11-7, the SAR116 clade, and the genus *Synechococcus*. Although the bacterial richness remained almost constant along the gradient, the abundance of several of these bacterial groups was significantly affected by latitudinal change. A comparison of both DNA-based and RNA-based approach revealed significant differences between total and active bacterial community structures. The number of phylotypes at 1, 3, and 20% genetic distance was significantly higher on RNA level than on DNA level. Thus, the diversity of active members in marine bacterioplankton communities appears to be underestimated by previous DNA-based studies.

IMPORTANCE

The ecological importance of marine microbial communities is generally accepted, as they are major components of global nutrient cycles. Understanding the factors controlling their distribution and versatile ways of nutrient transformations is a major challenge in marine microbiology. Moreover, the ecological role of single marine groups or lineages is still largely unexplored, as the majority of field studies targeted the total community (active and inactive) but not in particular the active community members. This study provides a yet unique insight into compositional variations of marine bacterial communities along a latitudinal gradient as total and active bacterial community structures were simultaneously assessed. The bacterial diversity of the active community was higher compared to the total community. Consequently, the number of active members contributing to ecosystem functioning might have been underestimated by previous DNA-based studies.

INTRODUCTION

Oceans cover more than 70% of the world's surface. According to the National Oceanic and Atmospheric Administration (http://www.noaa.gov/ocean.html), they harbor an enormous biodiversity and support the life of nearly 50 percent of all species on Earth. Recent estimates revealed that up to 10^6 different microorganisms inhabit 1 ml of oceanic seawater (1). These archaea, bacteria, protists, and unicellular fungi contribute 98% to the primary biomass production. Moreover, marine plankton is a major component of global nutrient cycles (2, 3). Culture-independent approaches have greatly advanced our understanding of the diversity of marine microbial communities (4). Early studies used traditional techniques such as Sanger sequencing-based analysis of 16S rRNA gene libraries to assess bacterial community structures (5, 6). In this way, due to the richness and size of marine microbial communities, only a small fraction of the diversity was unraveled. The application of next-generation sequencing revolutionized this field of research. Novel sequencing techniques have been applied for in-depth investigation of bacterial communities in diverse ecosystems such as soil (7, 8) or seawater (9, 10). However, metagenomic approaches fail to distinguish between total and active bacterial communities as environmental DNA is used as starting material. Thus, the vast majority of investigations targeting marine microbial communities did not focus on the active members (but see 11, 12), which can be investigated, i.e., by analysis of 16S rRNA transcripts (e.g., 11, 13).

A recent study investigated community patterns of marine planktonic bacteria worldwide by amplified ribosomal intergenic spacer analysis (ARISA) (14). Geographic patterns similar to those of higher organisms were recorded. The authors concluded that the kinetics of metabolisms strongly influenced bacterial diversity and global distribution (14). Other studies focused on assessing bacterial communities structures along latitudinal gradients. Two recent studies analyzing bacterial community

structures along such gradients found high numbers of *Proteobacteria* in coastal bacterioplankton communities in Latin America (15) and the SAR11 clade and the genus *Prochlorococcus* as dominant phylogenetic groups in oceanic surface waters from South Africa to the English Channel (16).

Coastal shelf seas of the temperate zone are highly productive because of the continuous nutrient supply by rivers. The North Sea, a typical costal shelf sea, is connected to the Atlantic Ocean via the English Channel in the South and the Norwegian Sea in the North. The water depth of the southern part is less than 50 m and is subjected to strong tidal currents. Nutrient suspension from the sediment and loss of water stratification are results of tides. This region underwent high nutrient loading and warming during the last 40 years (17, 18). The northern part of the North Sea is deeper (up to 725 m) and strong tidal currents are not occurring. Studies concerning the bacterioplankton community in the North Sea showed that distinct lineages of *Alpha*proteobacteria and *Gammaproteobacteria* as well as *Flavobacteria* constitute the major parts of these communities (12, 19-21).

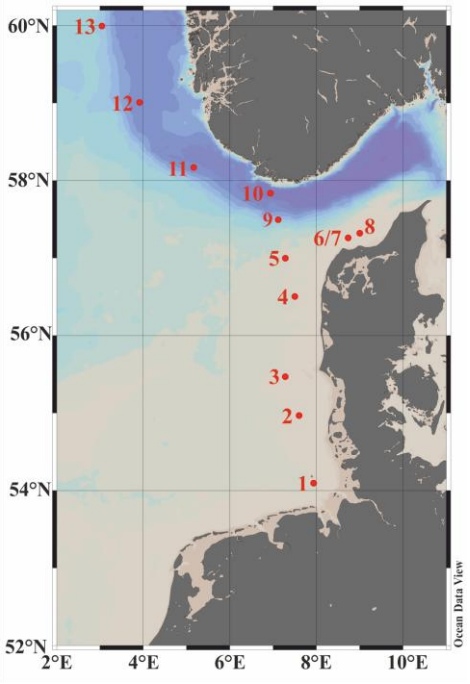


FIG 1 Map of the North Sea showing the positions of all 13 samplings sites. The map was generated using the Ocean Data View software package (version 4.5.2, Alfred Wegener Institute, Bremerhaven, Germany, [http://odv.awi.de/]).

In this study, we simultaneously investigated entire and active bacterioplankton community structures in the North Sea along a latitudinal gradient from Germany to Norway. Community structures were assessed by large-scale pyrosequencing-based analysis of the V3-V5 region of the 16S rRNA amplified by gene-specific PCR or RT-PCR using environmental DNA and RNA as template, respectively. We hypothesized that total and active bacterial communities undergo significant structural changes along the studied gradient. We further hypothesized that community structures of the total bacterial community assessed by DNA-based analysis differ from those of the active bacterial community assessed by RNA-based analysis. To our knowledge, this is the first investigation using both 16S rRNA and 16S rRNA gene analysis combined with large-scale next-generation sequencing to study total and active bacterioplankton community structures simultaneously.

RESULTS AND DISCUSSION

Sampling and sampling site characteristics

In this survey, we examined the bacterial community composition of thirteen marine water samples derived from the North Sea. All analyzed samples were taken at 3 m or 4 m depth within 6 days (see Table S1 in the supplemental material). Samples were taken on a latitudinal gradient from Germany to Norway. Five samples (1, 2, 3, 4, and 5) were taken near the German and Danish coast, three samples (6, 7, and 8) in the Skagerrak, and 5 samples (9, 10, 11, 12, and 13) in the Norwegian trench (Fig. 1). Temperature and salinity ranged from 14.5°C to 18.3°C and from 29.8 psu to 33.4 psu, respectively. Fluorescence varied from 0.36 to 1.54 mg/m^3 and oxygen content from 5.3 to 6.1 mL/L (see Table S1 in the supplemental material). As samples were taken along a North-South gradient, we tested for correlations between latitude and other measured environmental parameters. Temperature, bacterial cell counts, biomass production (BP), as well as turnover of glucose and dissolved free amino acids were significantly correlated to the latitudinal change and decreased with rising latitude (see Table S2 in the supplemental material). In addition, dissolved organic carbon (DOC) and total nitrogen (TN) content decreased in northern samples (Osterholz H and Dittmar T, {personal communication}).

Bacterial community composition

To assess total and active bacterioplankton community structures, DNA-based and RNA-based analysis, respectively, of 16S rRNA sequences were performed. For this purpose, DNA and RNA were extracted from the samples. The isolated DNA and the cDNA generated from the RNA served as templates for the amplification of the V3 to V5 region of 16S rRNA gene region. Amplification of the V3-V5 region is commonly used for DGGE analysis of marine bacterial communities and allows a high coverage and taxonomic resolution of bacterial phylogenetic groups (22, 23). The high coverage was supported by analysis with TestPrime (24). After quality filtering, denoising, and removal of potential chimeras and non-bacterial sequences, 382,507 high-quality sequences with an average read length of 519 bp (DNA 155,919 sequences, RNA 226,588 sequences) were used to study total and active bacterial community compositions (see Table S3 in the supplemental material). The number of sequences per sample ranged from 4,762 to 20,981 (DNA-based) and 1,798 to 37,444 (RNA-based). We were able classify all sequences below phylum level.

Calculated rarefaction curves as well as diversity indices revealed that the majority of the bacterial community was covered by the surveying effort (see Table S4 and Fig. S1 in the supplemental material). On DNA level, more than 80% of all estimated OTUs were regained at 1, 3, and 20% genetic distance. On RNA level, almost 88% of all OTUs at 20% genetic distance and 67% of all OTUs at 1 and 3% genetic distance were recovered. Classification of the sequences revealed the presence of 18 and 22 bacterial phyla and candidate divisions on DNA and RNA level, respectively. *Proteobacteria* were predominant across all samples (DNA 50.03%, RNA 54.83%). *Bacteroidetes* (DNA 38.36%, RNA 36.31%) and *Cyanobacteria* (DNA 7.66%, RNA 5.96%) were the second and third most abundant bacterial phyla. Other abundant bacterial phyla, present in all samples, were *Actinobacteria* (DNA 1.60%, RNA 0.19%), *Deferribacteres* (DNA 0.49%, RNA 0.10%), and *Verrucomicrobia* (DNA 1.45%, RNA 2.54%).

Alphaproteobacteria were the most abundant bacterial class in all samples (DNA 44.02%, RNA 50.10%). Sequences were mainly assigned to the *Rhodobacterales*, in particular, to different clusters of the *Roseobacter* clade (Fig. 2), i.e., the CHAB.I-5, the RCA (25), the NAC11-6, and the NAC11-7 clusters. These four clusters comprised almost 30% of the bacterial community (DNA 25.59%, RNA 27.54%). The marine *Roseobacter* clade is the only marine lineage of which cultured representatives are closely related to environmental clones (26). Members of the *Roseobacter* clade play important roles in the global carbon and sulfur cycles, as they are capable of diverse metabolic traits such as aerobic anoxygenic photosynthesis or carbon monoxide oxidation. The high abundance of this clade in the North Sea is in accordance previous studies (27).

In a previous study, Giebel et al. investigated the abundances of the *Roseobacter* RCA cluster and the SAR11 clade by qPCR along a similar gradient in the North Sea employing specific primers for RCA and SAR11 (28). The RCA cluster accounted for approximately 4.4% of the total bacterial community, which corresponds to our results (DNA 3.69%, RNA 5.58%). The abundance of the SAR11 clade was also analyzed and constituted approximately 20% of the total bacterial community. In another study from 2009 SAR11 abundances of up to 40% in the North Atlantic have been detected by CARD-FISH analysis (29). Surprisingly, we only found members of the SAR11 clade in low abundances (< 1%). This difference could be due to the different detection method employed or seasonal effects (30). Giebel et al. analyzed samples taken in late 2005 and early 2006, whereas our samples were taken in summer 2011. We also studied samples taken in the German Bight in May 2010 in which SAR11 was highly abundant in only one of 14 analyzed samples (data not shown). Thus, seasonal changes are most likely not the reason for this discrepancy. Primer evaluation of the V3-V5 primer set used in this study revealed that more than 93% of the SAR11 clade should be covered (SSURef 114, 95.8%; SSURef114 NR, 93.8%). This is even higher as the coverage calculated for the entire *Bacteroidetes* and the *Proteobacteria* (data not shown). Interestingly, evaluation of the SAR11 clade-specific primer used by Giebel et al. (28) revealed that only 31.1% (SSURef 114NR) and 19% (SSURef 114) of the SAR11 clade were theoretically covered.

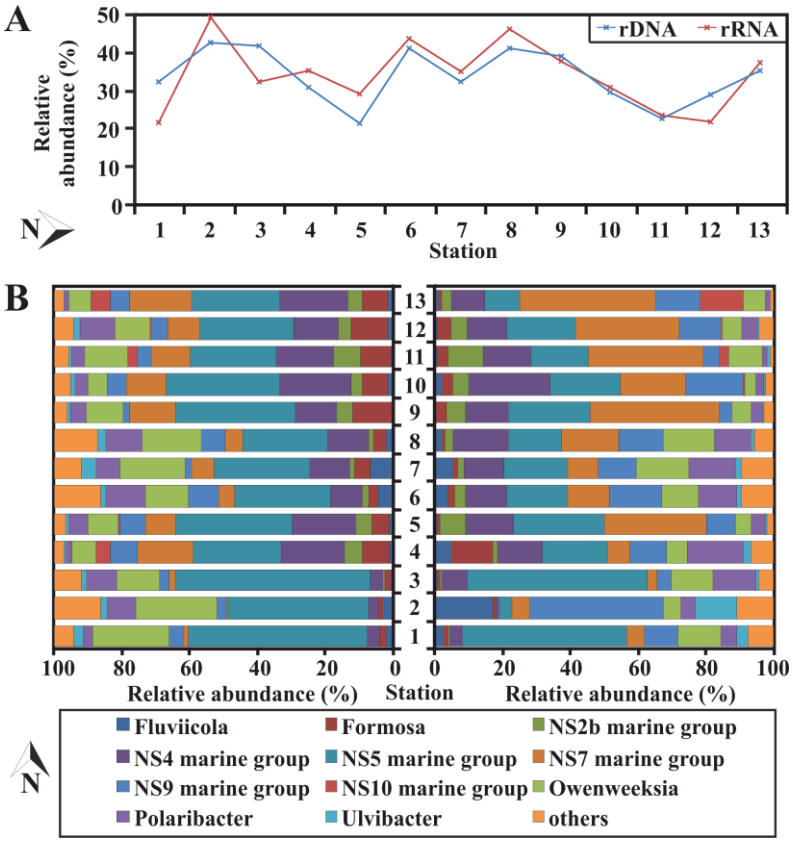


FIG 2 Relative abundance (A) and composition (B) of the *Rhodobacterales* along the latitudinal gradient. The *Roseobacter* NAC11-6 and NAC11-7 lineages were significantly correlated to latitude. The abundance of the *Roseobacter* clade affiliated (RCA) and CHAB-I-5 clusters were correlated to nucleic acid type (DNA/RNA). Significance was tested by Dirichlet regression in R.

Almost 10% of all sequences were affiliated to members of the *Rickettsiales* with the SAR116 clade as the most abundant marine group (DNA 7.02%, RNA 11.81%). This clade has been described as ubiquitous in marine environments (31). Members of SAR116 were found as a major fraction of the bacterioplankton in coastal regions as well as in the open ocean (32). Interestingly, the abundances of the clade were reported to be relatively low (< 5%) compared to those of SAR11 or *Bacteroidetes* in the Sargasso Sea (4). Only a few isolates of SAR116 exist with first genomes published in 2010 and 2011 (33, 34). Members of the SAR116 clade are proposed to be metabolic generalists in ocean nutrient cycling, as the genome sequences of the SAR116 member Candidatus *Puniceispirillum marinum* contains genes for proteorhodopsin, aerobic-type carbon monoxide dehydrogenase, dimethylsulfoniopropionate demethylase, and metabolism of C1 compounds (33). Another abundant bacterial group was the genus *Synechococcus* comprising the major fraction of the *Cyanobacteria*.

Sequences affiliated to the *Bacteroidetes* were mainly affiliated to different genera and marine groups within the *Flavobacteriales* (Fig. 3). *Bacteroidetes* are well known as one of the main groups of marine bacteria. Members are present in various habitats from the Southern Ocean (35) to the North Atlantic (36). *Bacteroidetes* are heterotrophs and important for the degradation of organic matter in the oceans. Their activity increases after algal blooms, as these organisms are able to effectively degrade algal remains (37). As genome sequences of isolates contain high numbers of genes involved in CO_2 fixation (38), *Bacteroidetes* might play an important role in the global carbon cycle. However, a lack between diversity of *Bacteroidetes* assessed by culture-dependent and independent approaches has been previously reported. All available isolates have not yet been found in studies using culture-independent approaches (39). In our study, less than 1% of all sequences were affiliated to known cultured members of the *Bacteroidetes*.

Impact of the latitudinal gradient on bacterial community composition

As samples for community analysis were taken along a latitudinal gradient from Germany to Norway, we further studied the composition of the bacterial communities along the gradient. We initially tested for correlations between calculated diversity indices and latitude. None of the calculated diversity indices was affected by latitudinal change.

As 16S rRNA sequences were mainly affiliated to the two bacterial orders

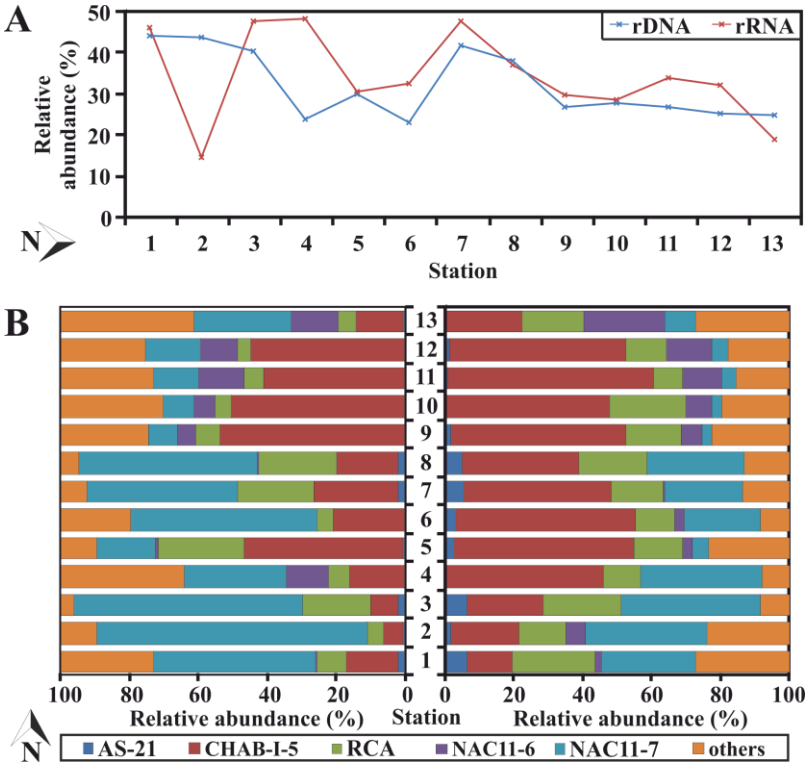


FIG 3 Relative abundance (A) and composition (B) of the *Flavobacterales* along the latitudinal gradient. Only the genera *Fluviicola* and *Ulvibacter* showed no correlation to latitude or nucleic acid type (DNA/RNA). The NS2b, NS7, NS9 and NS10 marine groups were correlated to latidutinal change, whereas the NS5 marine group and the genera *Owenweeksia* and *Polaribacter* were significantly affected by nucleic acid type. The genus *Formosa* and the NS4 marine group were significantly correlated to both parameters. Significance was tested by Dirichlet regression in R.

Flavobacteriales and *Rhodobacterales*, we further examined compositional changes of these groups along the latitudinal gradient (Figs. 2 and 3). Analysis of the composition of the *Rhodobacterales* revealed that the abundance of two of the five abundant groups was significantly correlated to latitude. The *Roseobacter* NAC11-7 cluster was more abundant in the southern samples whereas the abundance of the NAC11-6 cluster increased in northern samples. A similar result was recorded for the *Flavobacteriales* (Fig. 3). Six marine groups and genera, namely the genus *Formosa*, and the NS2b, NS4, NS7, NS9, and NS10 marine groups, were significantly correlated to latitudinal change.

In other studies, variations in community structures were correlated to alterations of environmental parameters. Schattenhofer et al. recorded changes in community structures with the latitude by studying a large South-North transect in the Atlantic (16). However, the surface water temperatures as well as other environmental factors drastically altered along the latitudinal gradient. The in our study recorded temperature decrease of approximately 1 °C from North to South might not explain the changes in bacterial community structures across the latitudinal gradient. As samples 6 to 13 were taken in the Skagerrak and Norwegian trench, the general circulation in the North Sea might explain latitudinal changes. Due to the topography of this area, a large inflow of Atlantic seawater occurs (40). The anti-clockwise circulation of the North Sea causes most of the water in the North Sea to pass through this area. Moreover, all the water from the Baltic Sea passes this area as well. Therefore, recorded changes in bacterial community structures might be explained by regional differences as water masses differ between the southern and northern sampling sites. These different water masses have different nutrient availabilities, as indicated, for example, in the recorded decrease of DOC and TN with arising latitude (Osterholz H and Dittmar T, {personal communication}). Structural changes in bacterioplankton communities caused by changes of DOC and nutrient availability have been reported previously (e.g., 41). In a study on a phytoplankton bloom in the Southern Ocean, the *Roseobacter* NAC11-6 and NAC11-7 clusters were found as an active part of the bloom community but dwell in different stages of the bloom (42). The authors suggested that both clusters are important contributors to carbon cycling. In this study, NAC11-7 was more abundant in the southern samples compared to the northern samples. The opposite was recorded for NAC11-6. The presence in different bloom stages reported by West et al. (13) might indicate that both groups are adapted to the different nutrient availabilities, which are also found along latitudinal gradients.

Comparision between RNA-based and DNA-based approch

We assessed bacterial community structures and relative abundances of phylogenetic groups by simultaneous 16S rRNA and 16S rRNA gene analysis. The advantage of the RNA-based approach is the ability to gain insight into active part of a microbial community whereas DNA-based approach allows assessment of structure of the total microbial community present in an environment (43). We recorded a lower diversity coverage on RNA level than on DNA level (see Table S4 in the supplemental material). To further evaluate this discrepancy, we caluculated rarefaction curves with respect to RNA and DNA (see Fig. S2 in the supplemental material). Curves calculated for RNA samples at all three genetic distance levels were less saturated than the corresponding curves generated for DNA samples. This indicates a higher bacterial diversity on RNA level. Further analysis revealed that the calculated genetic richness (number of operational taxonomic units (OUTs)) was significantly higher at all three genetic distances on RNA level (see Table S4 in the supplemental material). This is in accordance with the result of the rarefaction analysis and indicates a higher diversity of the active bacterial community.

The number of large comparative studies analyzing both active and total microbial communities is limited. Baldrian et al. used a similar approach to study microbial communities in forest soils (44). Similarities between total and active community structure were identified. Moreover, a higher fungal diversity on RNA level was recorded. Around 18% of the fungal OTUs identified in the forest soil were exclusively found in the active community. The higher genetic richness is in accordance with our findings. The active bacterial communities showed a higher richness and diversity compared to the total community. Although similar major groups were identified in the active and total community, many rare groups were exclusively identified on RNA level as indicated by the increased number of singleton OTUs with only one affiliated sequence (see Table S3 in the supplemental material).

We further examined compositional changes of the two abundant orders with respect to DNA-based and RNA-based analysis. This analysis revealed that the abundances of many predominant bacterial groups significantly differed between total (DNA-based) and active (RNA-based) bacterial community (Figs. 2 and 3). The *Roseobacter* CHAB-I-5 and RCA clusters as well as the genera *Formosa*, *Owenweeksia*, and *Polaribacter* and the NS4 marine group were significantly affected by nucleic acid type. For example, the RCA cluster was more abundant on RNA level than on DNA leve, indicating a higher activity.

Whereas abundance on DNA level is linked to cell abundance and to the number of rRNA operons in the genome, the abundance on RNA level is linked to cell abundance and to the number of rRNA transcripts. In *Escherichia coli*, rRNA promoter activity is regulated by a negative-feedback control loop responding to the translation activity (45). Consequently, low cell activity leads to a lower rRNA transcription rate in the cell and, thus, to a decreased number of rRNA transcripts. Thus, even rare species can be detected on RNA level as long as they are highly active. Moreover, the number of 16S rRNA copies in a PCR reaction is higher when RNA is used as starting material as 16S rRNA transcripts constitute a major fraction of the total RNA. As a consequence, the detection limit for single species in a PCR reaction is lower in RNA-based studies than in DNA-based metagenomic approaches.

Metagenomic studies greatly advanced our knowledge on the composition of bacterial communities in diverse habitats. However, the results of this study suggest that the diversity of marine bacterial communities and especially of the active members contributing to ecosystem functioning might be higher as previously reported by DNA-based studies. Consequently, the richness and diversity of marine communities, especially of the rare biosphere, might still be largely unexplored.

MATERIAL AND METHODS

Sampling and sample preparation

Water samples for community analyses were collected in the North Sea along a latitudinal gradient from Germany to Norway at 13 stations in July 2011 on board RV Heincke by 5 l-Niskin bottles mounted on a CTD (conductivity, temperature, and depth) rosette (Fig. 1, see Table S1 in the supplemental material). The water of at least 8 Niskin bottles (approximately 40 liters) was pooled in an ethanol-rinsed PE barrel. It was prefiltered through a 10-µm nylon net and a precombusted (4 h at 450°C) glass fiber filter (142 mm diameter, Whatman GF/D, Whatman, Maidstone, UK). Bacterioplankton was harvested from a prefiltered 10 liter sample using a filter sandwich consisting of a glass fiber filter (142 mm diameter, Whatman GF/F, Whatman) and 0.2-µm

polycarbonate filter (142 mm diameter, Nuclepore, Whatman).

For determination of chlorophyll *a* (Chl *a*) and phaeopigments, water samples were filtered onto glass fiber filters (47 mm diameter, Whatman GF/F, Whatman), immediately wrapped into aluminum foil and kept frozen at -20°C until further analysis within two weeks according to (28) and (46). Biomass production (BP) of heterotrophic prokaryotes was determined employing the ^{14}C-leucine incorporation method as described in (28) and (47). Turnover rates of free amino acids and glucose were determined by the radiotracer technique according to (28) and (48). Subsamples for DFAA analysis were filtered on board through 0.2 µm low protein binding filters (Tuffrin Acrodisc, Whatman) and kept frozen at -20°C until analysis. Bacterioplankton cell numbers were determined by flow cytometry (BD Accuri™ C6, BD Biosciences, Heidelberg, Germany) using SybrGreen I staining and internal bead calibration as described previously (28). For this purpose, water samples were taken, preserved with glutardialdehyde (final concentration 1%), and stored at -20°C until further analysis.

Extraction and purification of environmental DNA and RNA

Total DNA and RNA were extracted using acidic phenol as described by Weinbauer et al. (49). One sixth of a filter (pore size 0.2 μm) was used per sample in the extraction. Subsequently, RNA was purified employing the RNeasy Mini kit as recommended by the manufacturer (Qiagen, Hilden, Germany). Residual DNA was removed from RNA samples by treatment with Turbo™ DNase treatment (Ambion) and the absence of DNA was confirmed by 16S rRNA gene PCR according to Wemheuer et al. (11). Purified RNA was converted to cDNA by employing the SuperScript™ double-stranded cDNA synthesis kit (Invitrogen™) as modified by Wemheuer et al. (11). Total DNA was purified employing the peqGOLD Gel extraction kit (Peqlab, Erlangen Germany).

Amplification and sequencing of 16S rRNA

To assess bacterial community structures, the V3-V5 region of the bacterial 16S rRNA was amplified by PCR. The PCR reaction (50 µl) contained 10 µl of 5-fold Phusion HF buffer (Finnzymes, Vantaa, Finland), 200 µM of each of the four desoxynucleoside triphosphates, 1.5 mM $MgCl_2$, 4 µM of each primer, 2.5% DMSO, 2 U of Phusion high fidelity hot start DNA polymerase (Finnzymes), and approximately 50 ng of DNA or 25 ng of cDNA as template. The following thermal cycling scheme was used: initial denaturation at 98°C for 5 min, 25 cycles of denaturation at 98°C for 45 s, annealing at 63°C for 45 s, followed by extension at 72°C for 30 s. The final extension was carried out at 72°C for 5 min. Negative controls were performed by using the reaction mixture without template. The V3-V5 region was amplified with the following set of primers according to Muyzer et al. (50) containing the Roche 454 pyrosequencing adaptors, keys and one unique MID per sample (underlined): 341f 5′-CCATCTCATCCCTGCGTGTCTCCGAC-TCAG-$(dN)_{10}$-CCTACGGRAGGCAGCAG-3′ and 907r 5′-CCTATCCCCTGTGTGCCTTGGCAGTC-TCAG-CCGTCAATTCMTTTGAGT-3′. Primer coverage was evaluated using TestPrime (24) with a maximum number of 0 mismatches and the most recent Silva database (SSURef 114 and SSURef NR) as reference (51). Obtained PCR products were checked for appropriate size and then purified by using the peqGOLD gel extraction kit (Peqlab) as recommended by the manufacturer. Three independent PCR reactions were performed per sample, purified by gel extraction, and pooled in equal amounts. Quantification of the PCR products was performed using the Quant-iT dsDNA HS assay kit and a Qubit fluorometer (Invitrogen) as recommended by the manufacturer. The Göttingen Genomics Laboratory determined the sequences by using a Roche GS-FLX+ 454 pyrosequencer with Titanium chemistry (Roche, Mannheim, Germany).

Processing and analysis of datasets

Sequence data were deposited in the sequence read archive of the National Center for Biotechnology Information under the accession number SRA082674. Generated 16S rRNA gene and rRNA datasets were processed and further analyzed employing the QIIME 1.6.0 software package (52) and other tools. A detailed workflow is depicted in Fig. S3 in the supplemental material.

After raw data extraction, reads shorter than 300 bp, with an average quality value below 25, possessing long homopolymer stretches (> 8 bp), or primer mismatches (> 3) were removed. Subsequently, sequences were denoised employing QIIME. Remaining primer sequences were truncated employing cutadapt (53). Chimeric sequences were removed using UCHIME and the most recent Greengenes core set as reference dataset (54, 55).

Processed sequences of all samples were joined, sorted by decreasing length, and clustered employing the UCLUST algorithm (56). Sequences were clustered in operational taxonomic units (OTUs) at 1%, 3% and 20% genetic dissimilarity according to (57). OTUs at 3 and 20% sequence divergence represent species and phylum level, respectively (58). Phylogenetic composition was determined using the QIIME assign_taxonmy.py script. A BLAST alignment against a modified version of the Silva SSURef 111 NR database (51) was thereby performed. Modification included the integration and/or reclassification of sequences affiliated to the *Roseobacte*r OCT lineage, which consists of the *Roseobacter* NAC1-2, NAC11-6, and CHAB-I-5 clusters (see Table S5 in the supplemental material). Sequences were classified with respect to the modified Silva taxonomy of their best hit. Rarefaction curves, Shannon indices (59) and Chao1 indices (60) were calculated. In addition, the maximal number of OTUs (n_{max}) was estimated for each sample using the Michaelis-Menten-fit alpha diversity metrics included in the QIIME software package.

Statistical analysis

All statistical analyses were conducted employing R (version 2.15.0; R Development Core Team 2011 [http://cran.r-project.org/]). Comparisons of community structures were performed at the same level of surveying effort (4,762 randomly selected sequences). The RNA sample derived from station 6 (sample 4) was excluded from the analysis and any statistical testing, as it contains less than 1800 sequences. Possible correlations between latitude and other environmental parameters were determined by linear modeling in R (lm function). Possible impacts of latitude or nucleic acid type (DNA/RNA) on bacterial richness and diversity were tested by Analysis of Covariance (ANCOVA). Changes in the composition of the two abundant orders with respect to nucleic acid type and latitude were further examined by Dirichlet regression in R using the Dirichlet Reg package. Effects were assumed as statistically significant when $P \leq 0.05$.

ACKNOWLEDGEMENTS

We thank the crew of RV Heincke for their valuable support and B. Schmidt from the Experimental and Applied Mineralogy Department of the Geoscience Center Göttingen for providing his muffle furnace as well as Helena Osterholz and Thorsten Dittmar from the Marine Geochemistry Group of the Max Planck Institute for Marine Microbiology in Bremen (Germany) for providing information about DOC and TN. This work was funded by Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center TRR 51.

REFERENCES

1. **Curtis TP, Sloan WT, Scannell JW.** 2002. Estimating prokaryotic diversity and its limits. P. Natl. Acad. Sci. USA **99:**10494-10499.
2. **Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, Arrieta JM, Herndl GJ.** 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". P. Natl. Acad. Sci. USA **103:**12115-12120.
3. **Arrigo KR.** 2005. Marine microorganisms and global nutrient cycles. Nature **437:**349-355.
4. **Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:**66-74.
5. **Giovannoni SJ, Britschgi TB, Moyer CL, Field KG.** 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature **345:**60-63.
6. **Schmidt TM, DeLong EF, Pace NR.** 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J. Bacteriol. **173:**4371-4378.
7. **Will C, Thurmer A, Wollherr A, Nacke H, Herold N, Schrumpf M, Gutknecht J, Wubet T, Buscot F, Daniel R.** 2010. Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. Appl. Environ. Microb. **76:**6751-6759.
8. **Nacke H, Thurmer A, Wollherr A, Will C, Hodac L, Herold N, Schoning I, Schrumpf M, Daniel R.** 2011. Pyrosequencing-based assessment of

bacterial community structure along different management types in German forest and grassland soils. PloS one **6:**e17000.

9. **Vila-Costa M, Gasol JM, Sharma S, Moran MA.** 2012. Community analysis of high- and low-nucleic acid-containing bacteria in NW Mediterranean coastal waters using 16S rDNA pyrosequencing. Environ. Microbiol. **14:**1390-1402.
10. **Kirchman DL, Cottrell MT, Lovejoy C.** 2010. The structure of bacterial communities in the western Arctic Ocean as revealed by pyrosequencing of 16S rRNA genes. Environ. Microbiol. **12:**1132-1143.
11. **Wemheuer B, Wemheuer F, Daniel R.** 2012. RNA-Based Assessment of Diversity and Composition of Active Archaeal Communities in the German Bight. Archaea **2012:**8.
12. **Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM.** 2012. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science **336:**608-611.
13. **West NJ, Obernosterer I, Zemb O, Lebaron P.** 2008. Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. Environ. Microbiol. **10:**738-756.
14. **Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL, Brown JH.** 2008. A latitudinal diversity gradient in planktonic marine bacteria. P. Natl. Acad. Sci. USA **105:**7774-7778.
15. **Thompson F, Bruce T, Gonzalez A, Cardoso A, Clementino M, Costagliola M, Hozbor C, Otero E, Piccini C, Peressutti S, Schmieder R, Edwards R, Smith M, Takiyama L, Vieira R, Paranhos R, Artigas L.** 2011. Coastal bacterioplankton community diversity along a latitudinal gradient in Latin America by means of V6 tag pyrosequencing. Arch. Microbiol. **193:**105-114.
16. **Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA, Pernthaler J.** 2009. Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. Environ. Microbiol. **11:**2078-2093.
17. **Wiltshire K, Kraberg A, Bartsch I, Boersma M, Franke H-D, Freund J, Gebühr C, Gerdts G, Stockmann K, Wichels A**. 2010. Helgoland Roads, North Sea: 45 Years of Change. Estuar. Coasts. **33**:295-310.
18. **McQuatters-Gollop A, Raitsos DE, Edwards M, Pradhan Y, Mee LD, Lavender SJ, Attrill MJ.** 2007. A long-term chlorophyll data set reveals regime shift in North Sea phytoplankton biomass unconnected to nutrient trends. Limnol. Oceanogr. **52**:635–648.
19. **Alderkamp AC, Sintes E, Herndl GJ.** 2006. Abundance and activity of major groups of prokaryotic plankton in the coastal North Sea during spring and summer. Aquat. Microb. Ecol. **45:**237-246.
20. **Rink B, Seeberger S, Martens T, Duerselen C-D, Simon M, Brinkhoff T.** 2011. Regional patterns of bacterial community composition and biogeochemical properties in the southern North Sea. Aquat. Microb. Ecol. **63:**207-222.
21. **Sintes E, Witte H, Stodderegger K, Steiner P, Herndl GJ.** 2013. Temporal dynamics in the free-living bacterial community composition in the coastal North Sea. FEMS Microbiol. Ecol. **83:**413-424.
22. **Muyzer G, Brinkhoff T, Nübel U, Santegoeds C, Schäfer H, Wawer C.** 2004. Denaturing gradient gel electrophoresis (DGGE) in microbial ecology, p. 743-769. Kluwer Academic Publishers, Dordrecht.
23. **Sánchez O, Gasol JM, Massana R, Mas J, Pedrós-Alió C.** 2007. Comparison of Different Denaturing Gradient Gel Electrophoresis Primer Sets for the Study of Marine Bacterioplankton Communities. Appl. Environ. Microb. **73:**5962-5967.
24. **Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO.** 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. **41**:e1-e1.

25. **Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE, Howard EC, King E, Oakley CA, Reisch CR, Rinta-Kanto JM, Sharma S, Sun S, Varaljay V, Vila-Costa M, Westrich JR, Moran MA.** 2010. Genome characteristics of a generalist marine bacterial lineage. ISME J **4:**784-798.

26. **Wagner-Dobler I, Biebl H.** 2006. Environmental biology of the marine Roseobacter lineage. Annu. Rev. Microbiol. **60:**255-280.

27. **González JM, Simó R, Massana R, Covert JS, Casamayor EO, Pedrós-Alió C, Moran MA.** 2000. Bacterial Community Structure Associated with a Dimethylsulfoniopropionate-Producing North Atlantic Algal Bloom. Appl. Environ. Microb. **66:**4237-4246.

28. **Giebel HA, Kalhoefer D, Lemke A, Thole S, Gahl-Janssen R, Simon M, Brinkhoff T.** 2011. Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. ISME journal **5:**8-19.

29. **Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA, Pernthaler J.** 2009. Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. Environ Microbiol. **11:**2078-2093.

30. **Giovannoni SJ, Vergin KL.** 2012. Seasonality in ocean microbial communities. Science **335:**671-676.

31. **Oh HM, Kwon KK, Kang I, Kang SG, Lee JH, Kim SJ, Cho JC.** 2010. Complete genome sequence of "Candidatus Puniceispirillum marinum" IMCC1322, a representative of the SAR116 clade in the Alphaproteobacteria. J. Bacteriol. **192:**3240-3241.

32. **Morris RM, Frazar CD, Carlson CA.** 2012. Basin-scale patterns in the abundance of SAR11 subclades, marine Actinobacteria (OM1), members of the Roseobacter clade and OCS116 in the South Atlantic. Environ. Microbiol. **14:**1133-1144.

33. **Oh HM, Kwon KK, Kang I, Kang SG, Lee JH, Kim SJ, Cho JC.** 2010. Complete genome sequence of "Candidatus Puniceispirillum marinum" IMCC1322, a representative of the SAR116 clade in the Alphaproteobacteria. J. Bacteriol. **192:**3240-3241.

34. **Grote J, Bayindirli C, Bergauer K, Carpintero de Moraes P, Chen H, D'Ambrosio L, Edwards B, Fernández-Gómez B, Hamisi M, Logares R, Nguyen D, Rii YM, Saeck E, Schutte C, Widner B, Church MJ, Steward GF, Karl DM, Delong EF, Eppley JM, Schuster SC, Kyrpides NC, Rappé MS.** 2011. Draft genome sequence of strain HIMB100, a cultured representative of the SAR116 clade of marine Alphaproteobacteria. Stand. Genomic Sci. **5:**269-278.

35. **Jamieson RE, Rogers AD, Billett DS, Smale DA, Pearce DA.** 2012. Patterns of marine bacterioplankton biodiversity in the surface waters of the Scotia Arc, Southern Ocean. FEMS Microbiol. Ecol. **80:**452-468.

36. **Gómez-Pereira PR, Schüler M, Fuchs BM, Bennke C, Teeling H, Waldmann J, Richter M, Barbe V, Bataille E, Glöckner FO, Amann R.** 2012. Genomic content of uncultured Bacteroidetes from contrasting oceanic provinces in the North Atlantic Ocean. Environ. Microbiol. **14:**52-66.

37. **Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang S, Mann AJ, Waldmann J, Weber M, Klindworth A, Otto A, Lange J, Bernhardt J, Reinsch C, Hecker M, Peplies J, Bockelmann FD, Callies U, Gerdts G, Wichels A, Wiltshire KH, Glöckner FO, Schweder T, Amann R.** 2012. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science **336:**608-611.

38. **Fernandez-Gomez B, Richter M, Schuler M, Pinhassi J, Acinas SG, Gonzalez JM, Pedros-Alio C.** 2013. Ecology of marine Bacteroidetes: a comparative genomics approach. ISME J, in press.

39. **Alonso C, Warnecke F, Amann R, Pernthaler J.** 2007. High local and global diversity of Flavobacteria in marine plankton. Environ. Microbiol. **9:**1253-1266.

40. **OSPAR commission.** 2000. Quality Status Report 2000; Region II – Greater North Sea. OSPAR Commission, London.
41. **Morris RM, Frazar CD, Carlson CA.** 2012. Basin-scale patterns in the abundance of SAR11 subclades, marine Actinobacteria (OM1), members of the Roseobacter clade and OCS116 in the South Atlantic. Environ. Microbiol. **14:**1133-1144.
42. **West NJ, Obernosterer I, Zemb O, Lebaron P.** 2008. Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. Environ. Microbiol. **10:**738-756.
43. **Simon C, Daniel R.** 2011. Metagenomic analyses: past and future trends. Appl. Environ. Microbiol. **77:**1153-1161.
44. **Baldrian P, Kolarik M, Stursova M, Kopecky J, Valaskova V, Vetrovsky T, Zifcakova L, Snajdr J, Ridl J, Vlcek C, Voriskova J.** 2012. Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. ISME J **6:**248-258.
45. **Schneider DA, Gourse RL.** 2003. Changes in Escherichia coli rRNA Promoter Activity Correlate with Changes in Initiating Nucleoside Triphosphate and Guanosine 5′ Diphosphate 3′-Diphosphate Concentrations after Induction of Feedback Control of Ribosome Synthesis. J. Bacteriol. **185:**6185-6191.
46. **Nusch EA.** 1999. Chlorophyllbestimmung. In: von Tuempling W, Friedrich G. (eds). *Biologische Gewässeruntersuchung*. G Fischer, Stuttgart; Germany.
47. **Simon M, Rosenstock B, Zwisler, W.** 2004. Coupling of epipelagic and mesopelagic heterotrophic picoplankton production to phytoplankton biomass in the Antarctic Polar Frontal Region. Limnol. Oceanogr. **49:** 1035–1043.
48. **Simon, M, Rosenstock, B.** 2007. Different coupling of dissolved amino acid, protein, and carbohydrate turnover to heterotrophic picoplankton production in the Southern Ocean in austral summer and fall. Limnol. Oceanogr. **52:**85–95.
49. **Weinbauer MG, Fritz I, Wenderoth DF, Höfle MG.** 2002. Simultaneous extraction from bacterioplankton of total RNA and DNA suitable for quantitative structure and function analyses. Appl. Environ. Microbiol. **68:**1082-1087.
50. **Muyzer G, Teske A, Wirsen CO, Jannasch HW.** 1995. Phylogenetic relationships of Thiomicrospira species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. Arch. Microbiol. **164:**165-172.
51. **Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J.** 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. **35:**7188-7196.
52. **Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R.** 2010. QIIME allows analysis of high-throughput community sequencing data. Nature methods **7:**335-336.
53. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal **17**:10-12.
54. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R.** 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics **27:**2194-2200.
55. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL.** 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. **72:**5069-5072.
56. **Edgar RC.** 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26**:2460-2461.
57. **Simon C, Wiezer A, Strittmatter AW, Daniel R.** 2009. Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. Appl. Environ. Microbiol. **75**:7519-7526.
58. **Schloss PD, Handelsman J.** 2005. Introducing DOTUR, a computer program for defining operational taxonomic units

and estimating species richness. Appl. Environ. Microbiol. **71:**1501-1506.

59. **Shannon CE.** 2001. A mathematical theory of communication. SIGMOBILE Mob. Comput. Commun. Rev. **5:**3-55.

60. **Chao A, Bunge J.** 2002. Estimating the number of species in a stochastic abundance model. Biometrics **58:**531-539.

SUPPORTING INFORMATION FOR STUDY 3

CONTENTS:

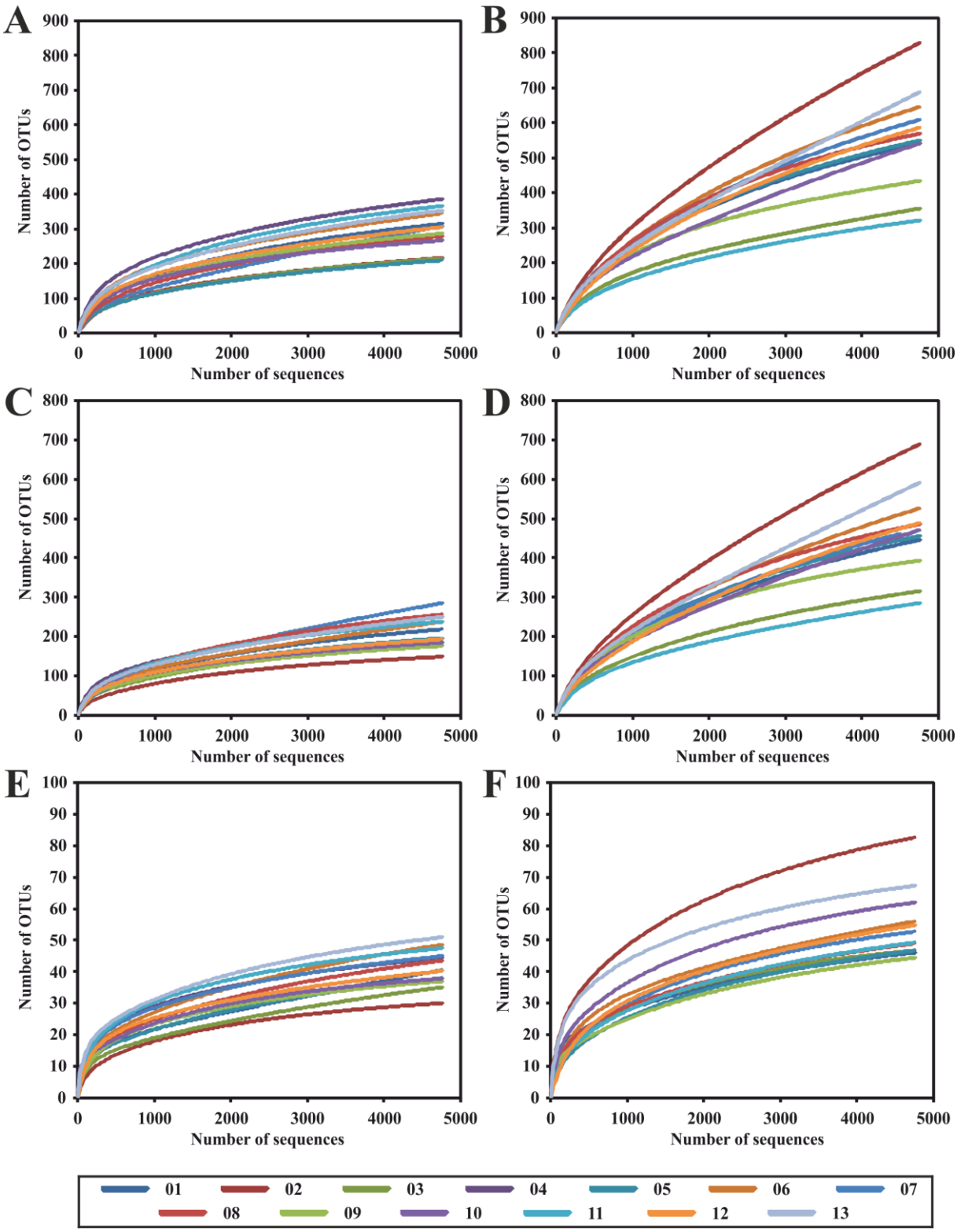


FIG S1 Rarefaction curves at 1 (A,B), 3 (C,D), and 20% (E,F) genetic divergence calculated for DNA-based (A,C,E) and RNA-based (B, D, F) approaches. Rarefaction analysis was performed with QIIME (1).

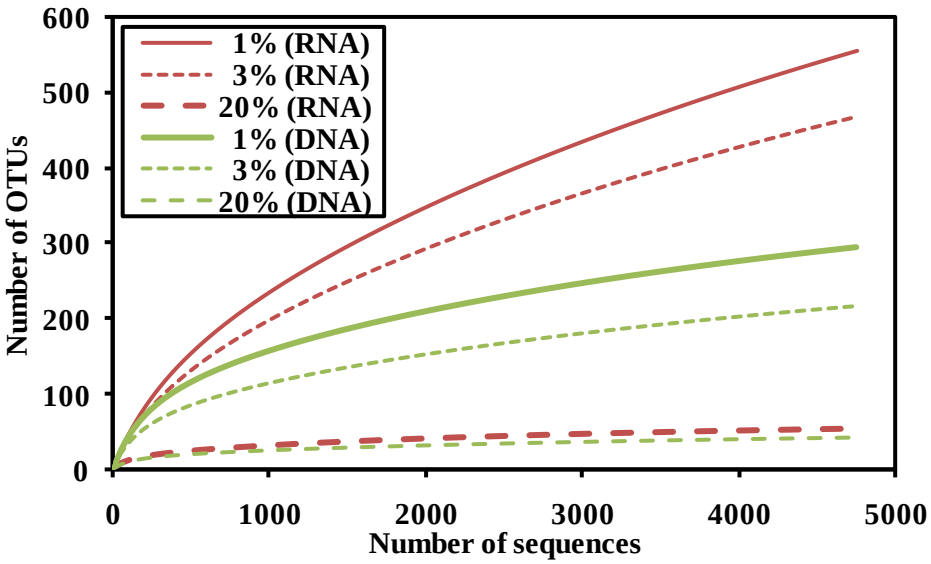
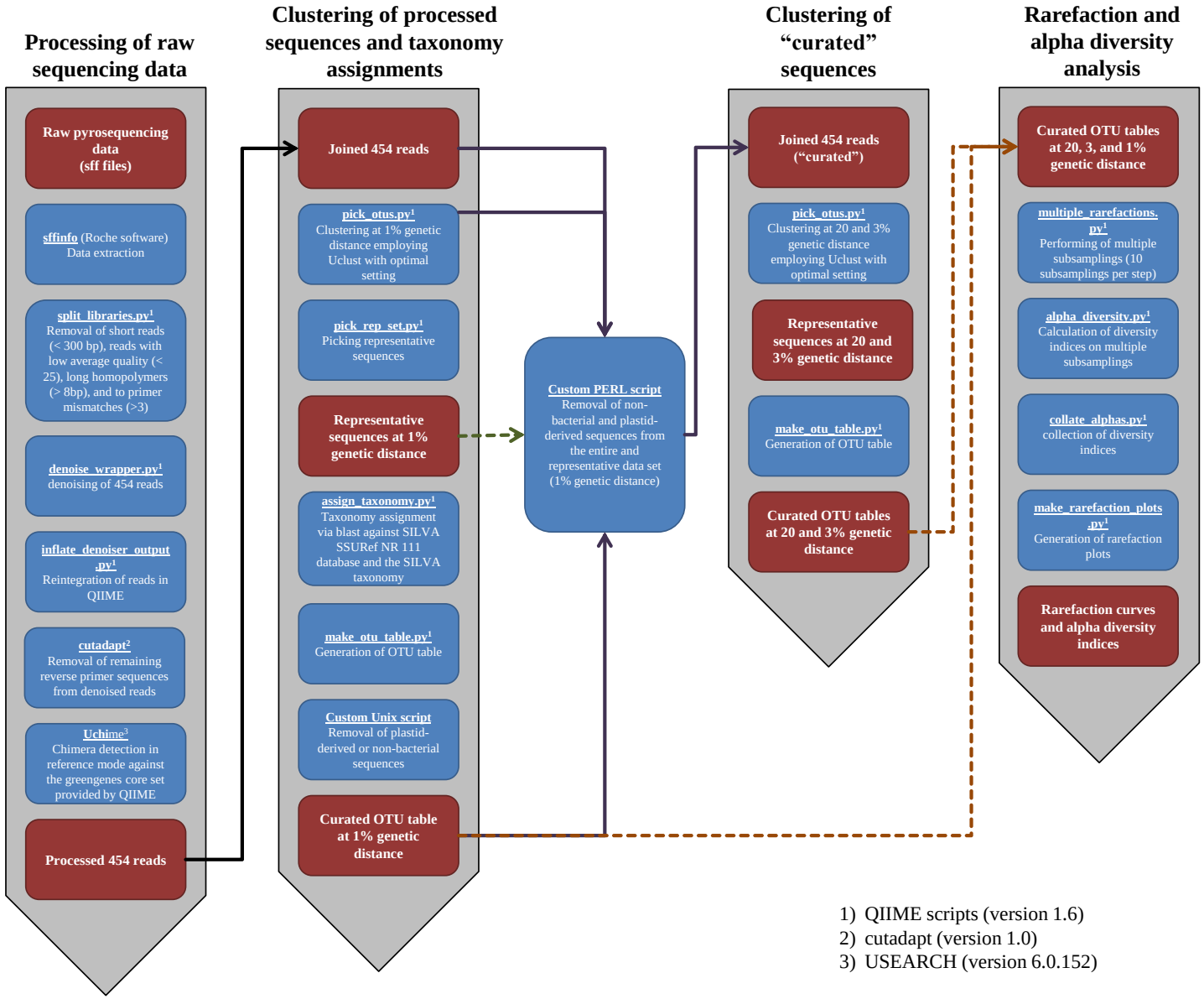


FIG S2 Rarefaction analysis with respect to nucleic acid type used for 16S rRNA analysis. Curves were calculated at 1, 3 and 20% genetic divergence. Rarefaction analysis was performed with QIIME (1).



1) QIIME scripts (version 1.6)
2) cutadapt (version 1.0)
3) USEARCH (version 6.0.152)

FIG S3 Processing of 454 pyrosequencing reads with QIIME (1) and other tools. After raw data extraction, reads shorter than 300 bp, with an average quality value below 25, possessing long homopolymer stretches (> 8 bp), or primer mismatches (> 3) were removed. Subsequently, sequences were denoised employing QIIME. Remaining primer sequences were truncated employing cutadapt (2). Chimeric sequences removed using UCHIME and the most recent Greengenes core set as reference dataset (3, 4). Processed sequences of all samples were joined, sorted by decreasing length, and clustered employing the UCLUST algorithm (5). Sequences were clustered in operational taxonomic units (OTUs) at 1%, 3% and 20% genetic dissimilarity according to Simon et al (6). Phylogenetic composition was determined using the QIIME assign_taxonmy.py script. A BLAST alignment against a modified version of the Silva SSURef 111 NR database (7) was thereby performed. Modification included the integration and/or reclassification of sequences affiliated to the Roseobacter OCT lineage, which consist of the Roseobacter NAC1-2, NAC11-6, and CHAB-I-5 clusters (see Table S4 in the supplemental material). Sequences were classified with respect to the modified Silva taxonomy of their best hit. Rarefaction curves, Shannon indices (8) and Chao1 indices (9) were calculated. In addition, the maximal number of OTUs (nmax) was estimated for each sample using the Michaelis-Menten-fit alpha diversity metrics included in the QIIME software package.

TABLE S1 CTD measured parameters for the sampling sites.

Sample	*Ship station*	*Date (dd.mm.yyyy)*	*Latitude (°N)*	*Longitude (°E)*	*Depth (m)*	*Bottom depth (m)*	*Temperature (°C)*	*Salinity (psu)*	*Transmission (%)*	*Fluorescence (mg/m^3)*	*Oxygen (mL/L)*
1	01	15.07.2011	54.0893	7.9327	3	32	18.308	29.826	78.563	0.757	6.118
2	02	15.07.2011	54.9667	7.6058	3	21	15.804	32.248	92.506	0.621	5.314
3	03	16.07.2011	55.4662	7.2852	3	25	14.775	33.420	93.706	0.377	5.277
4	06	16.07.2011	56.5018	7.5015	3	29	15.700	33.225	90.844	0.388	5.718
5	07	16.07.2011	56.994	7.2925	3	32	15.192	33.110	92.886	0.391	5.730
6	08	17.07.2011	57.261	8.732	3	23	14.782	33.408	83.597	1.490	5.943
7	10	17.07.2011	57.3173	9.0008	3	23	15.335	33.169	77.244	1.540	5.874
8	11	18.07.2011	57.2575	8.7438	3	24	15.252	33.259	82.817	1.366	5.696
9	12	18.07.2011	57.4887	7.1	3	181	15.453	31.873	92.874	0.361	5.696
10	13	19.07.2011	57.8305	6.9185	3	384	16.506	30.062	93.006	0.504	5.586
11	14	19.07.2011	58.1657	5.1693	4	290	14.652	32.566	90.982	0.572	5.751
12	15	20.07.2011	58.9985	3.9232	3	275	14.559	31.857	92.501	0.405	5.752
13	16	20.07.2011	59.9848	3.036	4	119	14.507	32.167	91.153	0.466	5.861

TABLE S2 Environmental parameters measured for the sampling stations.

Station	*Chlorophyll a (µg/L)*	*Phaeopigments (µg/l)*	*No. bacterial cells (10^9 cells/L)*	*Bacterial biomass production (µg C/L*h)*	*Glucose turnover (day^{-1})*	*Dissolved free amino acids turnover (day^{-1})*
1	2.013	0.788	1.66	712.7	2.68	0.17
2	1.302	0.492	1.60	321.6	3.08	0.29
3	0.533	0.382	0.76	155.2	3.49	0.29
4	0.789	0.251	0.97	420.0	3.91	0.27
5	0.281	0.187	0.86	149.4	0.92	0.06
6	1.598	0.606	0.72	252.6	0.67	0.16
7	3.157	0.462	0.67	316.4	1.95	0.20
8	2.625	0.121	0.86	-	-	-
9	0.222	0.132	0.58	161.8	0.92	0.10
10	0.474	0.026	0.82	147.0	1.01	0.10
11	0.612	0.065	1.10	108.3	0.19	0.05
12	0.304	0.100	0.91	183.6	1.37	0.11
13	0.459	0.040	1.14	196.8	1.08	0.09

-, not measured

TABLE S3 Statistics of the 16S rRNA sequence datasets.

Sample	***No. of sequences***	***Average length (bp)***	***No. of centroids***	***No. of centroids /10,000 sequences***	***No. of singeltons***	***No. of singeltons /10,000 sequences***
			DNA-based			
1	14,360	515	270	188.0	208	144.8
2	11,566	522	189	163.4	105	90.8
3	20,981	508	189	90.1	230	109.6
4	11,536	505	325	281.7	213	184.6
5	9,509	513	159	167.2	116	122.0
6	13,133	509	276	210.2	253	192.6
7	4,762	452	108	226.8	200	420.0
8	13,260	505	212	159.9	237	178.7
9	12,094	516	253	209.2	145	119.9
10	12,930	504	234	181.0	141	109.0
11	9,657	502	301	311.7	169	175.0
12	10,976	508	255	232.3	154	140.3
13	11,155	486	250	224.1	282	252.8
Sum/Avg	**155,919**	**506.2**	**188.7**	**203.5**	**232.4**	**172.3**
			RNA-based			
1	22,110	535	558	252.4	473	213.9
2	7,816	518	339	433.7	829	1060.6
3	15,171	538	244	160.8	422	278.2
4	1,798	513	101	561.7	86	478.3
5	16,779	534	471	280.7	499	297.4
6	33,857	516	693	204.7	1224	361.5
7	37,444	518	558	149.0	1469	392.3
8	23,986	542	574	239.3	574	239.3
9	16,363	544	397	242.6	265	162.0
10	11,087	505	222	200.2	778	701.7
11	14,748	532	220	149.2	387	262.4
12	17,472	533	457	261.6	820	469.3
13	7,957	507	218	274.0	801	1006.7
Sum/Avg	**226,588**	**527.2**	**388.6**	**262.3**	**663.6**	**455.7**

TABLE S4 Bacterial diversity and richness at 1%, 3%, and 20% genetic distance. Coverage was determined based on observed clusters and n_{max}. To compare community structures 4,762 randomly selected sequences from each sample were used for calculation. RNA sample 4 was removed from analysis, as it comprised only 1800 sequences. Significant differences between RNA and DNA samples are indicated by bold type. Significant correlations between any calculated diversity index and latitude were not observed. Significance was tested by Analysis of Covariance (ANCOVA). Diversity inices were calculated with QIIME (1).

Sample	***Observed clusters***			***Max. clusters (n_{max})***			***Coverage (%)***			***Shannon index (H')***			***Chao1***		
	1%	***3%***	***20%***	***1%***	***3%***	***20%***	***1%***	***3%***	***20%***	***1%***	***3%***	***20%***	***1%***	***3%***	***20%***
DNA															
1	311.4	216.3	38.5	388.1	267.3	44.3	80.23	80.93	86.83	4.10	3.33	1.42	465.3	351.2	62.7
2	216.0	150.5	30.5	262.0	183.0	35.3	82.43	82.23	86.42	3.33	2.79	1.19	304.7	229.0	45.8
3	213.0	182.8	34.7	257.9	219.9	40.2	82.59	83.14	86.31	3.41	3.23	1.34	336.4	296.1	66.6
4	386.5	248.0	45.8	452.9	288.9	48.3	85.34	85.85	94.79	4.94	4.22	2.18	605.6	445.9	67.3
5	207.6	197.7	37.6	250.8	238.6	42.2	82.78	82.87	89.05	3.24	3.19	1.74	324.7	303.9	47.7
6	340.1	232.5	48.4	395.8	286.1	56.1	85.93	81.26	86.25	4.48	3.74	1.87	624.7	463.3	81.1
7	308.0	286.0	45.0	447.8	396.1	48.7	68.78	72.21	92.38	3.78	3.72	1.83	1634.7	870.5	54.4
8	281.3	262.3	44.2	350.0	326.0	50.2	80.37	80.46	88.08	3.87	3.78	1.68	511.7	454.3	89.0
9	288.8	175.1	36.2	343.7	201.5	39.8	84.03	86.90	91.06	4.22	3.63	1.66	434.5	251.8	42.4
10	266.8	185.6	39.0	308.1	210.9	42.6	86.59	87.99	91.59	4.14	3.67	1.82	405.6	276.6	53.1
11	367.0	238.9	46.8	449.8	283.5	50.5	81.59	84.28	92.75	4.53	3.81	1.93	530.7	333.5	64.9
12	302.2	191.4	39.5	356.5	221.9	41.9	84.76	86.26	94.17	4.30	3.59	1.84	465.8	296.0	51.4
13	356.0	249.3	51.5	427.0	301.3	56.2	83.37	82.75	91.58	4.71	4.09	2.21	753.8	523.1	69.3
Mean	**295.7**	**216.6**	**41.4**	**360.8**	**263.5**	**45.9**	**82.21**	**82.86**	**90.10**	4.21	**3.66**	**1.79**	**569.1**	**391.9**	**61.2**
SD	**58.8**	**39.7**	**6.1**	**74.1**	**59.0**	**6.4**	**4.49**	**3.99**	**3.09**	0.53	**0.39**	**0.30**	**345.1**	**170.5**	**13.8**
CV	**0.20**	**0.18**	**0.15**	**0.21**	**0.22**	**0.14**	**0.05**	**0.05**	**0.03**	0.13	**0.11**	**0.17**	**0.61**	**0.43**	**0.23**

TABLE S4 continued

Sample	***Observed clusters***			***Max. clusters (n_{max})***			***Coverage (%)***			***Shannon index (H')***			***Chao1***		
	1%	***3%***	***20%***	***1%***	***3%***	***20%***	***1%***	***3%***	***20%***	***1%***	***3%***	***20%***	***1%***	***3%***	***20%***
RNA															
1	537.1	440.5	45.1	755.5	618.0	50.0	71.10	71.28	90.24	4.60	4.07	1.52	849.0	710.0	65.3
2	830.2	684.9	82.8	1541.4	1226.4	94.0	53.86	55.85	88.13	5.05	4.71	2.48	3465.6	2115.1	105.1
3	353.8	317.3	46.4	477.8	424.4	52.2	74.05	74.76	88.89	3.86	3.67	1.45	772.1	658.7	62.2
4	548.9	456.9	47.7	765.1	642.1	57.7	71.74	71.16	82.64	4.47	4.15	1.68	937.1	834.4	64.9
5	650.4	512.9	55.3	1041.3	798.9	62.3	62.46	64.20	88.77	4.70	4.31	1.91	1242.3	1042.7	77.5
6	603.5	476.2	54.8	896.3	713.6	64.3	67.33	66.73	85.29	4.63	4.08	1.59	1165.8	1096.8	74.4
7	578.5	487.9	49.1	820.8	678.0	56.6	70.48	71.96	86.70	4.58	4.33	1.60	903.7	820.0	62.7
8	431.4	387.2	44.6	558.6	499.5	51.2	77.23	77.52	87.12	4.28	4.12	1.64	625.2	524.3	66.6
9	535.6	474.1	61.8	843.9	761.5	69.6	63.47	62.26	88.85	4.46	4.18	1.99	3465.0	2121.1	92.3
10	318.6	283.6	50.8	423.9	384.0	61.0	75.16	73.85	83.32	3.66	3.53	1.50	735.6	606.1	65.6
11	587.0	494.9	54.3	970.7	854.8	64.5	60.47	57.90	84.14	4.14	3.91	1.77	1324.4	1137.9	65.6
12	683.5	585.4	65.9	1310.3	1088.9	70.7	52.16	53.76	93.24	4.73	4.44	2.37	5208.8	2872.1	75.4
Mean	**554.9**	**466.8**	**54.9**	**867.1**	**724.2**	**62.8**	**66.63**	**66.77**	**87.28**	4.49	**4.17**	1.85	**1724.6**	**1211.6**	**73.1**
SD	**140.7**	**107.9**	**11.0**	**325.0**	**249.6**	**11.9**	**8.20**	**7.92**	**3.07**	0.39	**0.32**	0.34	**1478.9**	**747.1**	**13.2**
CV	**0.25**	**0.23**	**0.20**	**0.37**	**0.34**	**0.19**	**0.12**	**0.12**	**0.04**	0.09	**0.08**	0.18	**0.86**	**0.62**	**0.18**

TABLE S5 16S rRNA gene sequences used to modify the Silva SSURef 111 NR database. Sequences were integrated and/or reassigned to the respective *Roseobacter* subclusters CHAB-I-5, NAC1-2, and NAC11-6, to resolve the *Roseobacter* OCT lineage, which does not represents a single lineage but mainly comprises the three subclusters mentioned.

Taxonomic affiliation (Roseobacter subcluster)	***Name***	***Accession***
CHAB-I-5	Uncultured alpha proteobacterium KTc0993	AF235129
CHAB-I-5	Uncultured alpha proteobacterium CHAB-I-5	AJ240910
CHAB-I-5	Uncultured Roseobacter sp.	AY627366
CHAB-I-5	Uncultured Roseobacter sp.	AY627371
CHAB-I-5	Uncultured marine bacterium	DQ009284
CHAB-I-5	Uncultured marine bacterium	DQ071080
CHAB-I-5	Uncultured alpha proteobacterium	DQ810318
CHAB-I-5	Uncultured alpha proteobacterium	EF016462
CHAB-I-5	Uncultured bacterium	EF645958
CHAB-I-5	Uncultured alpha proteobacterium	EU268090
CHAB-I-5	Uncultured alpha proteobacterium	EU268095
CHAB-I-5	Uncultured bacterium	EU799191
CHAB-I-5	Uncultured bacterium	EU799372
CHAB-I-5	Uncultured bacterium	EU799470
CHAB-I-5	Uncultured bacterium	EU799319
CHAB-I-5	Uncultured bacterium	EU799544
CHAB-I-5	Uncultured bacterium	EU799596
CHAB-I-5	Uncultured bacterium	EU799604
CHAB-I-5	Uncultured bacterium	EU799640
CHAB-I-5	Uncultured bacterium	EU799809
CHAB-I-5	Uncultured bacterium	EU799837
CHAB-I-5	Uncultured bacterium	EU799994
CHAB-I-5	Uncultured bacterium	EU800023
CHAB-I-5	Uncultured bacterium	FJ545513
CHAB-I-5	Uncultured bacterium	FJ545522
CHAB-I-5	Uncultured bacterium	FJ545623
CHAB-I-5	Uncultured Sulfitobacter sp.	FJ744918
CHAB-I-5	Uncultured Sulfitobacter sp.	FJ745272
CHAB-I-5	Uncultured marine bacterium	FJ826232
CHAB-I-5	Uncultured marine bacterium	FJ826260
CHAB-I-5	Uncultured marine bacterium	FJ826501
CHAB-I-5	Uncultured alpha proteobacterium	GQ347827
CHAB-I-5	Uncultured alpha proteobacterium	GQ347863
CHAB-I-5	Uncultured alpha proteobacterium	GQ347787
CHAB-I-5	Uncultured alpha proteobacterium	GQ347819
CHAB-I-5	Uncultured alpha proteobacterium	GQ347872
CHAB-I-5	Uncultured alpha proteobacterium	GQ348856
CHAB-I-5	Uncultured alpha proteobacterium	GQ348786
CHAB-I-5	Uncultured alpha proteobacterium	GQ348929
CHAB-I-5	Uncultured marine bacterium	GU204725
CHAB-I-5	Uncultured alpha proteobacterium	HM057611
CHAB-I-5	Uncultured alpha proteobacterium	HM057744
CHAB-I-5	Uncultured bacterium	HQ166724
CHAB-I-5	Uncultured Rhodobacteraceae bacterium	HQ242009
CHAB-I-5	Uncultured Rhodobacteraceae bacterium	HQ241992
CHAB-I-5	Uncultured Rhodobacteraceae bacterium	HQ242334
CHAB-I-5	Uncultured Roseobacter sp.	JN233137
CHAB-I-5	Uncultured bacterium	JQ013156
NAC1-2	Uncultured Roseobacter NAC1-2	AF245615
NAC1-2	Marine bacterium SRF2	AJ002564

TABLE S5 continued

Taxonomic affiliation (Roseobacter subcluster)	***Name***	***Accession***
NAC1-2	Marine bacterium SRF1	AJ002563
NAC1-2	Uncultured organism	DQ395765
NAC1-2	Uncultured organism	DQ395728
NAC1-2	Uncultured organism	DQ395731
NAC1-2	Uncultured organism	DQ395864
NAC1-2	Uncultured organism	DQ395872
NAC1-2	Uncultured bacterium	EF645949
NAC1-2	Uncultured bacterium	EF646123
NAC1-2	Uncultured bacterium	EU799025
NAC1-2	Uncultured bacterium	EU799081
NAC1-2	Uncultured bacterium	EU799101
NAC1-2	Uncultured bacterium	EU799272
NAC1-2	Uncultured bacterium	EU799264
NAC1-2	Uncultured bacterium	EU800060
NAC1-2	Uncultured bacterium	FJ545544
NAC1-2	Uncultured bacterium	FJ545545
NAC1-2	Uncultured alpha proteobacterium	FJ615109
NAC1-2	Uncultured marine bacterium	FJ825934
NAC1-2	Uncultured marine bacterium	FJ826325
NAC1-2	Uncultured alpha proteobacterium	GQ347547
NAC1-2	Uncultured alpha proteobacterium	GQ347978
NAC1-2	Uncultured alpha proteobacterium	GQ348093
NAC1-2	Uncultured alpha proteobacterium	GQ348318
NAC1-2	Uncultured alpha proteobacterium	GQ348563
NAC1-2	Uncultured alpha proteobacterium	GQ348588
NAC1-2	Uncultured alpha proteobacterium	GQ349293
NAC1-2	Uncultured alpha proteobacterium	GQ350465
NAC1-2	Uncultured bacterium	GU061348
NAC1-2	Uncultured bacterium	GU061367
NAC1-2	Uncultured alpha proteobacterium	HQ163162
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242021
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242013
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ241978
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242010
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242028
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ241980
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242039
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242038
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242201
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242215
NAC1-2	Uncultured Rhodobacteraceae bacterium	HQ242314
NAC1-2	Uncultured bacterium	HQ671914
NAC1-2	Uncultured bacterium	HQ672014
NAC1-2	Uncultured bacterium	HQ672039
NAC1-2	Uncultured bacterium	HQ671987
NAC1-2	Uncultured bacterium	HQ672179
NAC11-6	Uncultured Roseobacter NAC11-6	AF245634
NAC11-6	Roseobacter sp. 3008	AM110967
NAC11-6	Uncultured alpha proteobacterium MB11C09	AY033324
NAC11-6	Uncultured organism	DQ395766
NAC11-6	Uncultured organism	DQ395672
NAC11-6	Uncultured organism	DQ395823
NAC11-6	Uncultured organism	DQ395796
NAC11-6	Uncultured organism	DQ395871
NAC11-6	Uncultured organism	DQ395831
NAC11-6	Uncultured alpha proteobacterium	EF016465
NAC11-6	Uncultured marine bacterium	EU005826

TABLE S4 continued

Taxonomic affiliation (Roseobacter subcluster)	***Name***	***Accession***
NAC11-6	Uncultured bacterium	EU035836
NAC11-6	Uncultured bacterium	EU265946
NAC11-6	Uncultured bacterium	EU570893
NAC11-6	Uncultured bacterium	EU799357
NAC11-6	Uncultured bacterium	EU799376
NAC11-6	Uncultured bacterium	EU799567
NAC11-6	Uncultured bacterium	FJ545521
NAC11-6	Uncultured bacterium	FJ545602
NAC11-6	Uncultured marine bacterium	FJ826399
NAC11-6	Uncultured Rhodobacteraceae bacterium	FM958453
NAC11-6	Uncultured Rhodobacteraceae bacterium	FN582321
NAC11-6	Uncultured Roseobacter sp.	FR681767
NAC11-6	Uncultured Roseobacter sp.	FR681768
NAC11-6	Uncultured Roseobacter sp.	FR681769
NAC11-6	Uncultured marine bacterium	FR683592
NAC11-6	Uncultured Roseobacter sp.	FR681770
NAC11-6	Uncultured alpha proteobacterium	GQ348520
NAC11-6	Uncultured alpha proteobacterium	GQ348598
NAC11-6	Uncultured alpha proteobacterium	GQ348534
NAC11-6	Uncultured alpha proteobacterium	GQ348722
NAC11-6	Uncultured alpha proteobacterium	GQ349969
NAC11-6	Uncultured alpha proteobacterium	GQ350050
NAC11-6	Uncultured alpha proteobacterium	GU061261
NAC11-6	Uncultured marine bacterium	GU235686
NAC11-6	Uncultured marine bacterium	GU235651
NAC11-6	Uncultured bacterium	HQ671749
NAC11-6	Uncultured bacterium	HQ671894
NAC11-6	Uncultured bacterium	JF451313
NAC11-6	Uncultured alpha proteobacterium	JN177665
NAC11-6	Uncultured bacterium	JX016532
NAC11-6	Uncultured bacterium	JX206779

References

1. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA,Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature methods **7:**335-336.
2. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. (2011) EMBnet.journal **17**:10-12.
3. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol **72:**5069-5072.
4. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics **27:**2194-2200.
5. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26**:2460-2461
6. Simon C, Wiezer A, Strittmatter AW, Daniel R. 2009. Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. Appl Environ Microbiol **75:**7519-7526.
7. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res **35:**7188-7196.
8. Shannon CE. 2001. A mathematical theory of communication. SIGMOBILE Mob. Comput. Commun. Rev. **5:**3-55.
9. Chao A, Bunge J. 2002. Estimating the Number of Species in a Stochastic Abundance Model. Biometrics **58:**531-539

STUDY 4:

COMMUNITY ANALYSIS OF FREE-LIVING AND ATTACHED *ROSEOBACTER*-AFFILIATED BACTERIA INDICATE AN OVERLAP IN DIVERSTY BETWEEN MATURED PARTICLES AND THE SEDIMENT SURFACE

KANUKOLLU S[1], **WEMHEUER B**[2], HERBER J[1], BILLERBECK S[1], LUCAS J[1], DANIEL R[2], SIMON M[1], CYPIONKA H[1], AND ENGELEN B[1]

(IN PREPARATION)

[1]INSTITUTE FOR CHEMISTRY AND BIOLOGY OF THE MARINE ENVIRONMENT (ICBM), CARL-VON-OSSIETZKY-UNIVERSITY OF OLDENBURG, CARL-VON-OSSIETZKY-STR. 9-11, D-26111 OLDENBURG, GERMANY; [2]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY; [3]PRESENT ADDRESS: BIOLOGICAL INSTITUTE HELGOLAND, KURPROMENADE 201, D-27498 HELGOLAND, GERMANY

Author contributions to the work:

Performed the experiments: SV, BW, HG, CB

Analyzed data: SV, MS, BW, HG, CB, JV, SD

Contributed data on water properties and analysis of these data: BW, TB, HG, CB, IB, SB

Wrote the publication: SV, MS, TB, RD

ABSTRACT

In this study, we compared the diversity of the *Roseobacter* clade within sediment and water samples from the eastern North Sea by cultivation-independent and cultivation-based methods. Cluster analysis of DGGE patterns revealed specific localizations of free-living and attached populations. *Roseobacter* communities from surface-near particles (3 m depth), clustered separately, while matured particles (10-40 m) and sediment surfaces (23-181 m) showed similar compositions. These results were confirmed by amplicon-based analysis of 16S rRNA gene sequences, indicating an increasing diversity of *Roseobacter*-populations from the sea surface to the seafloor. As roseobacters are known to contribute to sulfur transformations, MPN experiments were set up with media containing either dimethyl sulfide (DMS), dimethyl sulfonium propionate (DMSP) or dimethyl sulfoxide (DMSO). While anoxic, DMSO-containing enrichments showed highest MPN numbers for the sediment surface (up to 2.1 x 10^7 roseobacters/cm^3), growth of sediment-dwelling roseobacters was specifically stimulated in all DMSP and half of the DMS-containing dilution series. A total of 20 roseobacters (12 from sediments) was isolated from all oxic enrichments, representing 0.55% of all *Rhodobacteraceae* in the pyrosequence dataset. Thus, the isolates could be quantified in the environmental samples avoiding the laborious design of specific primer sets or probes for quantification.

INTRODUCTION

The *Roseobacter* clade is a major marine bacterial lineage (3) representing a significant part of pelagic and benthic microbial communities (4, 14, 38). The majority of *Roseobacter affiliated* bacteria (short: roseobacters) was detected in the water column as members of free-living and particle associated bacterial communities in various marine ecosystems and in different geographic regions. They can comprize more than 20% of all bacteria in coastal oceans (4). Additionally, about 3% of all roseobacter clones in 16S rRNA gene libraries originate from marine surface sediments (20, 26). The fact that marine sediments are a relevant but understudied environment for roseobacters is further documented by the observation that approximately 10% of all available pure cultures are derived from benthic origins (4). However, so far no systematic study was performed which examines differences among the roseobacters in the free-living, particle-associated and sediment-dwelling bacterial communities. Because of the great habitat diversity, roseobacters exhibit a broad metabolic versatility (47). At least 12 of the 41 major *Roseobacter* lineages are involved in sulfur transformation reactions (27, 52). For example, it has been shown that DMSP-degrading roseobacters live in symbiotic relationships with DMSP-producing dinoflagellates (25). While some of the DMSP is cleaved, most of it is demethylated, and a fraction of the volatile degradation product DMS is oxidized by pelagic bacteria (19). As high concentrations of DMSP are associated with decaying (and sinking) algal blooms (47), it is most likely, that this compound is a possible substrate for sediment-dwelling roseobacters. Interestingly, some of the benthic roseobacters were detected in permanently anoxic sediment layers, indicating an anaerobic metabolism (18, 28). An adaptation at least to periodic anaerobiosis was detected for *Dinoroseobacter shibae* $DFL12^T$ by the analysis of its genome showing the presence of a DMSO-reductase gene and pathways for denitrification and fermentation (49). However, it is still unknown whether roseobacters are adapted to and are able to dwell in permanently anoxic sediments.

The overall goal of our study was to evaluate the occurrence and diversity of free 72 living and particle-attached roseobacters in the water column and in both, oxic and anoxic sediments of a coastal sea. We hypothesized that distinct differences exist in the roseobacter communities from the near surface waters to deeper waters and in the sediment-dwelling microbial communities. *Roseobacter*-specific DGGE analyses gave a first overview on their distribution patterns. Pyrosequencing-based analysis of 16S rRNA gene amplicons allowed deeper insights into their diversity and relative abundance. Additionally, cultivation experiments were carried out with special emphasis on sulfur transformation reactions. Thus, serial dilution cultures were set up to quantify their most probable number (MPN) and isolate roseobacters that are involved in aerobic DMS and DMSP utilization or/and anaerobic DMSO reduction. The combination of microbiological and advanced molecular techniques allowed to determine the relevance of these isolates within the different compartments.

MATERIAL AND METHODS

Sample origin

The samples were collected in July 2011 during a cruise with RV 'Heincke' (expedition HE361) to the eastern North Sea (Fig. 1). Six sampling sites were chosen to be analyzed in detail by cultivation, DGGE and pyrosequencing. They represent two shallow sites (water depth: 35 and 26 meters below sea level; mbsl) at the German/Danish coast (station 36: 7°93.57'E, 54°09.02'N and station 5: 7°69.02'E, 56°01.18'N), a shallow (23 mbsl) and a deep site (114 mbsl) in the Skagerrak area (station 8: 8°73.20'E, 57°26.10'N and station 27: 8°35.41'E, 57°36.77'N), as well as two deep sites (181 and 119 mbsl) within the Norwegian trench (station 12: 7°10.00'E, 57°48.87'N and station 16: 3°03.60'E, 59°98.48'N). Intermediate sites (stations 1-3, 6, 7, 10, 11, 13-15) were analyzed by pyrosequencing of 3 m deep surface waters, only. Sediment samples and bottom seawater directly above the seafloor were taken by a multi-core sampler (MUC). Aliquots of the samples were transferred to cultivation media or stored at -20°C for further molecular analyses in the home lab. Other water samples were obtained from corresponding locations by 5 l-Niskin bottles mounted on a rosette and equipped with a conductivity-temperature-depth (CTD) probe. These samples were taken from the turbidity maxima at 3 mbsl and deep water layers between 23 and 181 mbsl, respectively. Up to several liters of water (depending on sample turbidity) were first filtered through 5-µm Nuclepore filters (to collect particle-associated bacteria) and then through 0.2104 µm Nuclepore filters (to collect free-living bacteria). All filters were stored at -80°C until further processing.

Cultivation

Immediately after sampling, water and sediments from all stations (Table 1) were inoculated in artificial seawater media (44). The media were slightly modified as Na2SO4 was substituted by NaCl. Four media variations were set up by adding i) DMS (100 µM) and lactate (5 mM), ii) DMS (500 µM), iii) DMSP (5 mM). and iv) DMSO (100 µM) and lactate (5 mM). For each station, water samples from the near surface and bottom seawater, as well as sediment samples from the seafloor (0-1 cmbsf) and the oxic/anoxic transition zone (4-5 cmbsf) were incubated in serial dilutions of the four media described above.

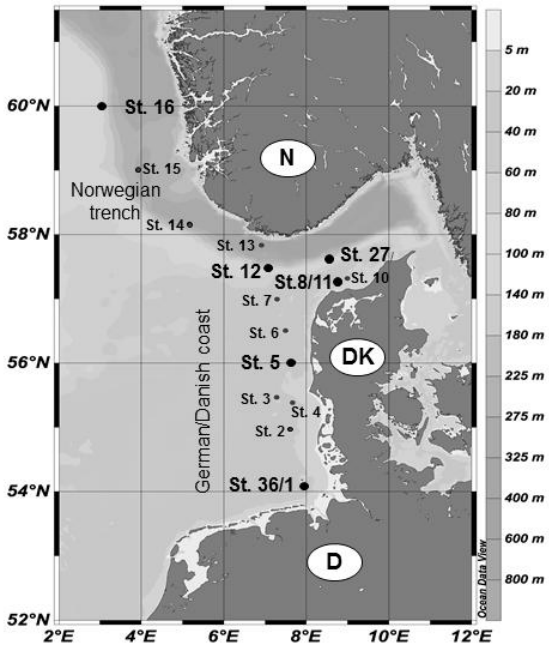


Fig. 1. Track of the Heincke cruise HE361 (June 2011). Samples were taken at six stations along the German/Danish coast and the Norwegian trench. The map was created with Ocean Data View *(Schlitzer, R., Ocean Data View, http://odv.awi-bremerhaven.de, 2004)*.

The dilutions were performed stepwise in three paralells from 10^{-2} to 10^{-7} in polypropylene 96-deepwell plates (Beckman, Fullerton, CA). Every well was filled with 900 µl medium, to which 100 µl inoculum was added. The plates containing media i-iii were incubated under oxic conditions and those containing medium iv were incubated anoxically for six months in the dark at 4°C as described in Süß et al (44).

Analysis of MPN series

Growth was tested according to Martens-Habbena and Sass (23). The cells within each well were stained with SYBR Green I (Molecular Probes, Eugene, OR) and growth was detected by using a microtiter plate reader (FLUOstar OPTIMA bMG, Offenburg/Germany). All MPN counts were calculated as described by de Man (10) and corrected for the values obtained from sample-free dilution series. Dilutions indicating growth were further subjected to DNA extraction and PCR to determine *Roseobacter*-specific MPN values. Additionally, grown dilution cultures served as isolation source for indigenous bacteria.

Isolation of bacteria

Aliquots (100 µl) from the MPN wells that were tested positive for growth of roseobacters were

spread on agar plates or transferred to deep-agar dilutions containing the respective media. The agar plates and tubes were incubated in the dark at 15°C under oxic and anoxic conditions for more than 2 and 6 months, respectively. Colonies from the agar plates were picked and further subcultivated at least 3 times to obtain pure cultures. Colonies from the anoxic agar tubes were picked and are under process to further subcultivations.

DNA extractions

For sediments, the DNA was extracted from 0.25 g of the samples, using the MoBio Power Soil DNA Isolation Kit (Carlsbad, California), following the manufacturer's instructions. PCR141 grade water was used to elute the purified DNA of the spin columns for downstream applications. The DNA of filtered water and particles was isolated after bead beating, phenol143 chloroform extraction, and isopropanol precipitation as described previously (22, 42). Treatment by sodium-dodecyl-sulfate (SDS, 1.75% of a 25% solution) was applied instead of lysozyme (43), precipitation was done with 30 µl NaAc and 750 µl isopropanol at –20°C overnight and 50 µl of PCR-grade water (Eppendorf, Hamburg, Germany) was used for resuspension at 4°C overnight. For cultivated samples, a freeze-thaw extraction protocol was used. From each well of the dilution series that was tested positively for growth, 100 µl were transferred into 1.5-ml reaction tubes (Eppendorf, Hamburg, Germany). The tubes were repeatedly heated to 60°C for 5 min, and then frozen in liquid nitrogen for 1min for 4 cycles. The cell lysate was stored at -20°C and used for further molecular analysis.

Amplification of 16S rRNA genes from environmental samples and enrichment cultures

PCR was used to amplify 16S rRNA gene fragments for DGGE analysis. For calculating *Roseobacter*-specific MPN from the serial dilution series, the microtiter wells that showed growth were tested for amplifiable DNA by using a general *Bacteria*-specific PCR according to Wilms et al. (49). To specifically target the *Roseobacter* clade in cultures and in environmental samples, the primer pair RoseoGC536f/GrB735r (15) was used. The PCR reaction mixture (50 µl) contained 0.2 mM dNTP´s, 1.5 mM MgCl2, 0.2 mM of each primer, 1 x Red *Taq* Buffer (Sigma, Munich, Germany), 0.2 ng μl^{-1} BSA, 2 U Red *Taq* DNA polymerase (Sigma, Munich, Germany) and 2-4 µl template DNA (2-6 ng μl^{-1}). The PCR was performed in a thermal cycler (Mastercycler, Eppendorf, Hamburg, Germany) for 10 cycles under the following conditions: 94°C for four minutes, 94°C for 30 seconds, 65°C for one minute, 72°C for one minute. After that, the samples were run for another 33 cycles with an annealing temperature of 63°C, followed by a final elongation step at 72°C for ten minutes. After amplification, the PCR products were visualized on a 1.5% (w/v) agarose gel.

DGGE analysis

DGGE was carried out as described by Süß et al. (44) using an INGENYphorU-2 system (Ingeny, Leiden, The Netherlands). The amplicons (ca. 200 bp) obtained by the *Roseobacter*- specific PCR were mixed with loading buffer (40 % [wt/vol] glycerol, 60 % [wt/vol] 1x Tris-acetate-EDTA [TAE], bromphenol blue) in a ratio of 1:2. PCR products were loaded onto polyacrylamide gels (6% wt/vol) stored in 1 x TAE (40 mmol l^{-1} Tris, 20 mmol l^{-1} acetate, 1 mmol l^{-1} EDTA), with a denaturing gradient from 50% to 70% (100% denaturant correspond to 7 mol l^{-1} urea and 40% formamide). DGGE-gels were run at a constant voltage of 100 V and at a temperature of 60 °C for 20 h. After electrophoresis, gels were stained for 2 h with 1×SYBRGold (Molecular Probes, Leiden, Netherlands) and destained in water for 20 min in distilled water prior to UV transillumination.

Cluster analysis of DGGE community patterns

The software package GelComparII, version 6.5 (Applied Maths, St-Martens-Latem, Belgium) was used for cluster analysis according to Wilms et al. (49). Since all lanes of a DGGE gel contain a characteristic degree of smear, a background subtraction was performed to make different lanes comparable. Therefore, a background scale of 20% was applied in the software package. The densitometric curves were compared using the Pearson coefficient (31). A position tolerance optimization was performed to fit the curves to the best possible match. Dendrograms were generated, using the UPGMA method (41).

Sequence analysis of DGGE bands and pure cultures

To identify the phylogenetic affiliation of pure cultures, genomic DNA was extracted from picked colonies. The cell pellet was resuspended in 100 µl of PCR-grade water (Eppendorf, Hamburg, Germany) and then treated by the freeze and thaw procedure as described above. Two µl of the final extract were added to 48 µl of PCR mixture. Nearly full-length bacterial 16S rRNA gene sequences

were amplified using the *Bacteria*-specific primers 8f/1492r according to Overmann and Tuschak (30). For sequence analysis of DGGE bands, those were excised, subjected to 50 µl of PCR-water and incubated over night at 4°C to elute the DNA. 2 µl of the DNA elute were taken for reamplification by using the *Roseobacter*-specific primer pair as described above without containing the GC-clamp. The reamplification comprised 26 PCR cycles (annealing temperature 55 °C). All PCR products were purified using the QIAquick PCR purification Kit (Qiagen GmbH, Hilden, Germany) and were commercially sequenced (GATC, Köln, Germany). The Göttingen Genomics Laboratory determined the sequences of the isolates. The partial 16S rRNA sequences were compared to those in GenBank using the BLAST function (1). All partial 16S rRNA gene sequences of DGGE bands and isolates have been deposited in the EMBL database under accession numbers HG423215 – HG423283.

Pyrosequencing of 16S rRNA genes

All 27 water, particle and sediment samples from this study were subjected to pyrosequencing. The V3-V5 region of the environmental DNA (50 ng/µl) was amplified with 15 different forward primers according to Schneider et al. (36) containing the Roche 454 pyrosequencing adaptors (underlined), the key (italic), one unique MID per sample and a template specific sequence: Forward primer V3for-MID-137-151: 5´-<u>CCA TCT CAT CCC TGC GTG TCT CCG AC</u>-*T CAG*-MID(137-151)-TAC GGR AGG CAG CAG-3′ and reverse primer V5rev: 5´-<u>CCT ATC CCC TGT GTG CCT TGG CAG TC</u>-*T CAG*-CCG TCA ATT CMT TTG AGT-3′. The PCR reactions were set up according to the manufacturer's instructions of the Phusion PCR Master Mix Kit (Thermo Scientific, Bremen, Germany). The reaction mixture (50 µl) contained 15.5 µl of PCR H2O, 25 µl of 2x Phusion Master Mix with HF Buffer, 2 µl of BSA, 1.5 µl of 100% DMSO, 2.5 µl of primers and 1 µl of template. The PCR was performed in a thermal cycler (Mastercycler, Eppendorf, Hamburg, Germany) under the following conditions: 98°C for 4 minutes, 26 cycles with 98°C for 30 seconds, 60°C for one minute, 72°C for one minute followed by a final elongation step at 72°C for 5 minutes. After amplification, the complete PCR mixture (50 µl) was loaded on 1% (w/v) agarose gels and excised by using peqGOLD Gel Extraction Kit (PEQLAB Biotechnologie GmbH, Erlangen, Germany) according to the manufacturer's instructions. The concentration and purity of the DNA samples were analyzed by using Nanodrop (Thermo Scientific, Bremen, Germany). For further sequencing, 5 ng/µl of DNA was required, if the desired concentration was not reached, the procedure was repeated for those samples. If necessary, 2-5 PCR reaction mixtures were pooled before purification. Finally, the samples were subjected to pyrosequencing at the Institute für Mikrobiologie und Genetik (Göttingen, Germany). Raw data of all 27 samples generated in this study were deposited in the NCBI Short Read Archive under accession *SRA*096062. Data from the 13 surface samples (3 m) were obtained from Wemheuer et al. (B. Wemheuer, D. Meier, P. Klembert, S. Billerbeck, H.-A. Giebel, C. Scherber, M. Simon, and R. Daniel, submitted for publication) and were deposited under accession number *SRA*082674.

Processing and analysis of pyrosequencing-derived datasets

After raw data extraction, reads shorter than 300 bp, with an average quality value below 25, possessing long homopolymer stretches (> 8 bp), or primer mismatches (> 5) were removed. Subsequently, sequences were denoised employing Acacia (2). Remaining primer sequences were truncated employing cutadapt (24). Chimeric sequences removed using UCHIME and the most recent greengenes core set as reference dataset (11, 13). Processed sequences of all samples were joined, sorted by decreasing length, and clustered employing the UCLUST algorithm (12). Sequences were clustered in operational taxonomic units (OTUs) at 1%, 3% and 20% genetic dissimilarity according to Simon et al. (40). OTUs at 3 and 20% sequence divergence represent species and phylum level, respectively (35). Phylogenetic composition was determined using the QIIME assign_taxonmy.py script (5). A BLAST alignment against the Silva SSURef 111 NR database (32) was thereby performed. Sequences were classified with respect to the silva taxonomy of their best hit. Rarefaction curves, Shannon indices (39) and Chao1 indices (7) were calculated. In addition, the maximal number of OTUs (nmax) was estimated for each sample using the Michaelis-Menten249 fit alpha diversity metrics included in the QIIME software package. To compare bacterial community structures across all samples based on phylogenetic or count-based distance metrics, Principal Coordinate Analysis (PCoA) plots were generated and a phylogenetic tree was calculated prior to PCoA generation. For this purpose, sequences were aligned using the PyNAST algorithm. The

phylogenetic tree and the corresponding OTU table were subsequently used to generate PCoA plots.

RESULTS

Free-living and particle-associated roseobacters show different diversity patterns

North Sea samples from the German/Danish coast and the Norwegian trench (Fig. 1) were analyzed by *Roseobacter*-specific DGGE to identify an overlap between free-living, particle260 associated and sediment-dwelling communities. Cluster analysis of DGGE patterns revealed specific differences between free-living and attached roseobacters (Fig. 2). Interestingly, particle-associated *Roseobacter*-communities from surface-near waters (3 m, young particles) clustered separately, while the community composition of matured particles from waters at the bottom of the mixed layer (10 to 30 m) showed high similarities to that of the sediment surface. Sequencing of representative DGGE bands indicated the presence of specific *Roseobacter* populations within the different pelagic and benthic compartments (Figs. S1, S2 and Tables S1, S2). Some bacteria such as relatives of *Sulfitobacter donghicola* were part of all free-living and particle-associated communities. Bacteria affiliated to *Nereida ignava* occurred in the entire water column down to the seafloor, whereas relatives of *Phaeobacter caeruleus* only occurred in shallow waters and sequences related to *Thalassobius mediterraneus* were exclusively detected in bottom waters. The material taken from the young particles (3 mbsl) was most notably different from the rest of the samples. Still some bacteria occurred on all particles, e.g. affiliates of *Sulfitobacter dubius*, whereas others were exclusively present in either the sediment such as relatives of *Roseovarius crassostreae* or the shallow water samples from the chlorophyll a maximum (bloom) like *Sulfitobacter pontiacus* affiliates. As the DGGE analysis gave a limited phylogenetic resolution (amplicon size: ~200 bp), pyrosequencing was performed on the same material to obtain a more detailed insight into the *Roseobacter* community structures. A total of app. 452,000 high quality 16S rRNA gene sequences were obtained by amplicon-based pyrosequencing of all sites and compartments.

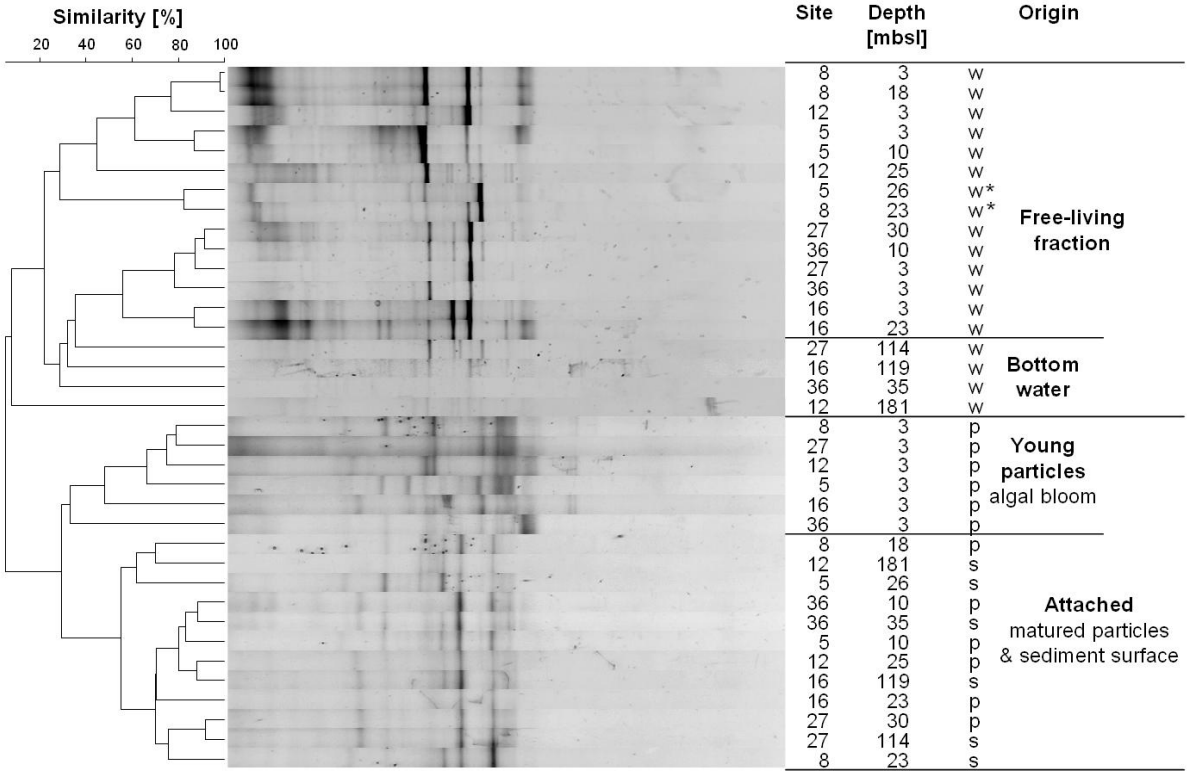


Fig. 2. Cluster analysis of *Roseobacter*-specific DGGE patterns. The densitometric curves of the DGGE community profiles were obtained by applying GelComparII and compared using the Pearson coefficient. The dendrogram was generated using the UPGMA method. Sampling sites, water depth in meters below sea level (mbsl) and origin (w = water, p = particles, s = sediments) are indicated. * = two bottom water samples (St. 5 & 8) clustering with other deep and surface waters.

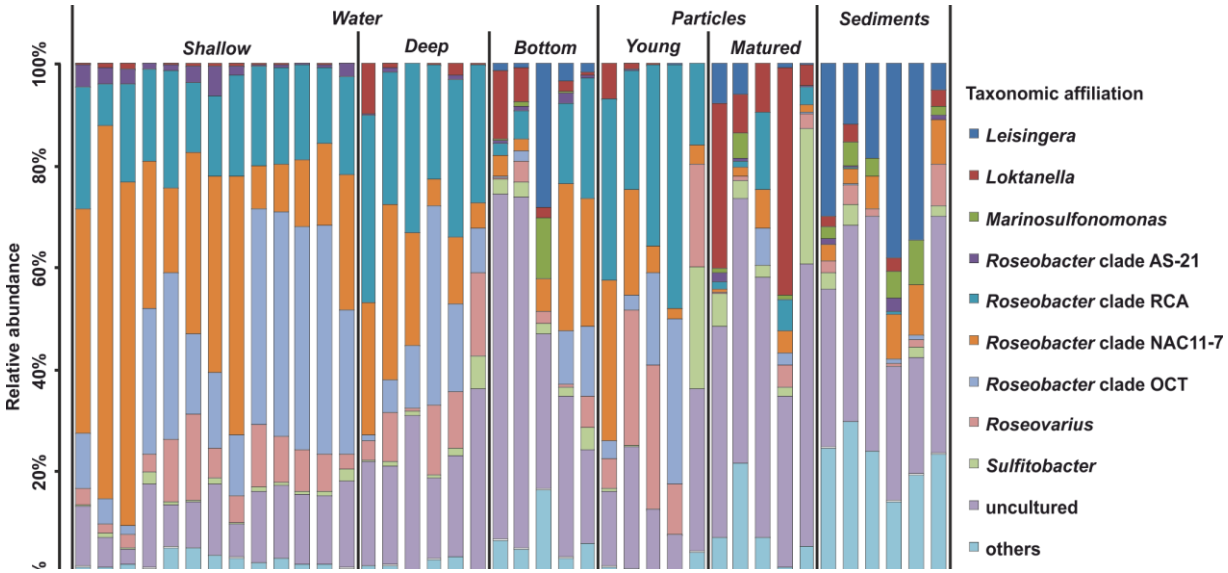


Fig. 3. Amplicon-based diversity patterns of *Rhodobacteraceae* within the different free-living and attached fractions. The sequences that were affiliated to the *Rhodobacteraceae* (93.000) were assigned to operational taxonomic units (OTUs) on a 99% sequence similarity level. The samples are named by sampling site (1-27), compartment (w = water, p = particles, s = sediment) and water depth (3-181 mbsl).

About 93,000 sequences were affiliated to the *Rhodobacteraceae* with an average of approximately 20% per sample. While approximately 25% of the free283 living fraction was affiliated to the *Rhodobacteraceae*, a proportion of about 19% and 2% were determined for particles and sediments, respectively. To display diversity patterns for this phylogenetic group in the different compartments, the sequences were assigned to operational taxonomic units (OTUs) on a 99% sequence similarity level. In general, the diversity of *Roseobacter* populations of sediments and matured particles was generally much higher than that of water samples and young particles (Fig. 3). Diversity patterns indicate increasing numbers of uncultured roseobacters from the sea surface to the seafloor. Water samples were dominated by members of the *Roseobacter* subcluster NAC 11-7, the *Roseobacter* Clade Affiliated (RCA) cluster, and the OCT lineage comprising the cluster NAC1-2, NAC11-6 (19) and CHAB-I-5. Sediments and matured particles harbored a variety of other lineages including *Sulfitobacter* and *Roseovarius* species. Matured particles were dominated by *Loktanella* species. In contrast, *Leisingera* species dominated sediments, only. One exception for the general separation of free-living and particle-associated roseobacters was the bottom water sample from station 16. The respective community composition was highly similar to that of the underlying sediment surface. This indicates a probable mixing during sample recovery which was not visible in the DGGE cluster analysis.

Next-generation sequencing confirms DGGE results

The composition of *Rhodobacteraceae*-related phylotypes from the pyrosequencing analysis that was identified for the different compartments was used to perform a principle coordinate analysis (PCoA). In this calculation, we have integrated an additional set of water samples from a depth of three meters that were recovered during the same sampling campaign (small black dots in Fig. 1). Two end-members of the *Rhodobacteraceae* community composition can be identified along the first principle coordinate (PC1) that explains 52% of the variation (Fig. 4). One comprises samples from shallow waters, the other is a combination of samples from sediments, matured particles and bottom waters. The samples from young particles and deep waters fall in between. While the shallow water samples are widespread along PC2 (11% of variation), indicating subtle differences among the free-living *Rhodobacteraceae* communities near the sea surface, the other end-member shows less variations. This was already indicated by the diversity plot shown in Figure 3. This general trend is in accordance to the clustering of DGGE community profiles of *Roseobacter* populations.

However, the clear separation between most bottom waters and the rest of the free-living fraction, as seen in the DGGE clusters is not as visible in the PCoA.

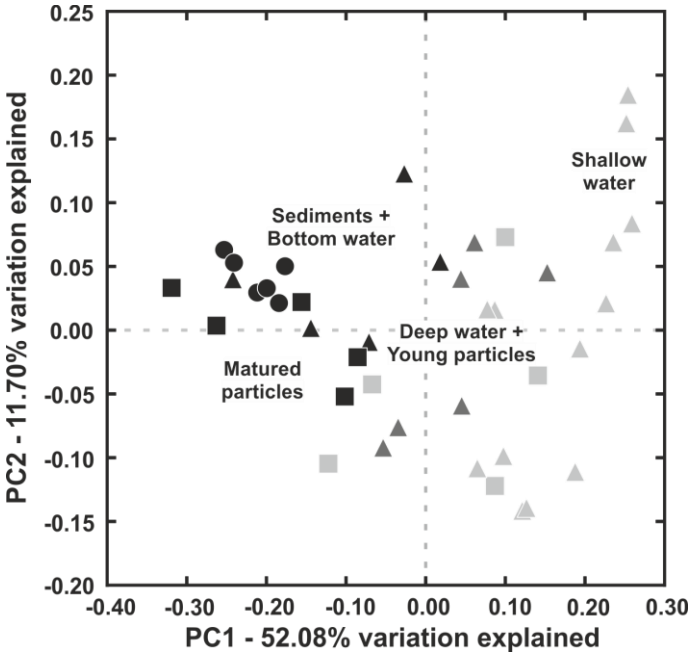


Fig. 4. Principle coordinate analysis of *Rhodobacteraceae* community compositions. Water samples are displayed as triangles (light gray = shallow, gray = deep, black = bottom), particles are displayed as squares (light gray = young, black = matured), sediments are shown as black circles.

High numbers of roseobacters were enriched in media amended with methylated sulfur compounds

Apart from the cultivation-independent methods, a subset of samples was inoculated in serial dilution cultures. These enrichments were performed with water samples of the chlorophyll a maximum (bloom) and bottom seawater, as well as with sediments from the upper centimeter of the seafloor and the oxic/anoxic transition zone between 4 and 10 centimeter below seafloor (cmbsf). The media were amended with DMS, DMSP or DMSO to quantify the amount of roseobacters utilizing organic sulfur compounds. In one medium, lactate was provided as a carbon source, while DMS served as electron donor. Two other media contained either DMS or DMSP as the only substrates. Here, the sulfur compounds served as both, electron donor and carbon source. In the fourth medium, DMSO was added as electron acceptor for anaerobic respiration. Due to technical problems during sampling, samples from station 36

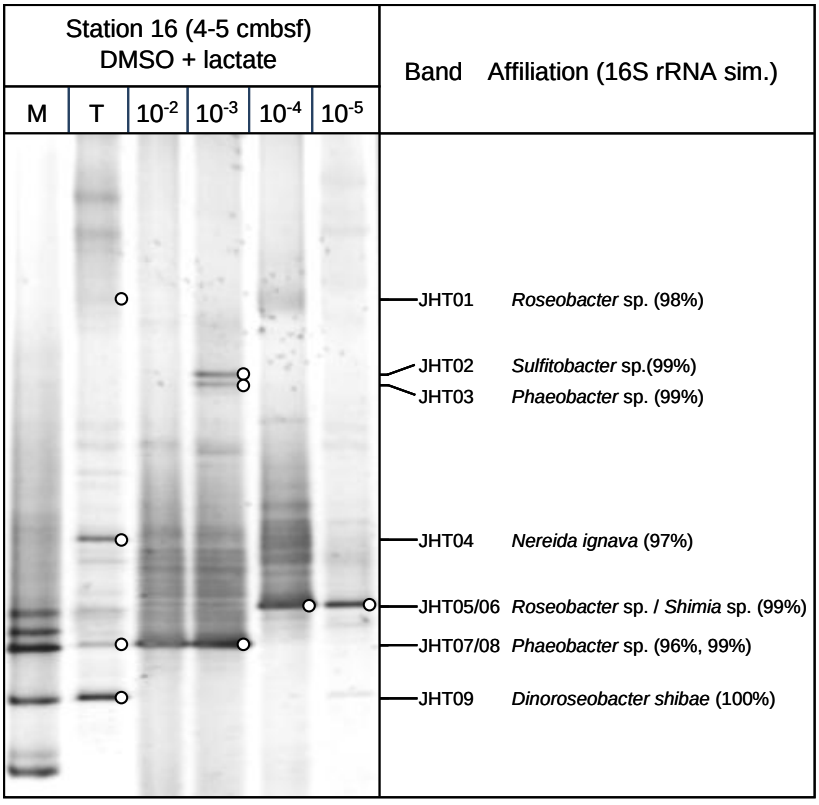


Fig. 5. DGGE analysis of one anoxic serial dilution series from the oxic/anoxic transition zone (4-5 cmbsf) of station 16. PCR amplicons (~200 bp) were generated by using *Roseobacter*-specific primers. Representative DGGE bands (white circles) were excised and sequenced. The next relatives in GenBank are indicated. M = marker, T = original sediment

Table 1: MPN analysis of pelagic and benthic North Sea samples. Samples were take from the chlorophyll a maximum (Bloom), bottom waters above the seafloor (Bottom), the sediment surface (Seafloor) and the oxic-anoxic transition zone (Transition) at six different stations. Media were amended with DMS and DMSP and DMSO as organic sulfur compounds. MPN numbers are given in 10^3 cells per milliliter which refers to 10^3 cells per cm^3 of sediment.

Station	***Origin***	***Depth***	**DMS+lactate [10^3 cells/ml]**		**DMS [10^3 cells/ml]**		**DMSP [10^3 cells/ml]**		**DMS0+lactate [10^3 cells/ml]**	
	Bloom	3 mbsl	14	9	9	9	0.4	0	-	-
36	Bottom	35 mbsl	300	0	0.4	0.4	0.15	0	-	-
	Seafloor	0-1 cmbsf	150	40	9	9	0.9	0.9	-	-
	Transition	4-5 cmbsf	9	4	0.9	0.4	0.023	0.023	-	-
	Bloom	3 mbsl	4	0.4	0	0	0	0	2000	1100
5	Bottom	26 mbsl	2.3	0.4	0.9	0	4	0	700	700
	Seafloor	0-1 cmbsf	1.5	1.1	4.3	4.3	0.4	0.4	7500	7500
	Transition	5-6 cmbsf	4	4	1.5	1.5	40	0	2000	2000
	Bloom	3 mbsl	0.4	0.4	300	0	0	0	2000	1100
8	Bottom	23 mbsl	0.4	0.4	0.4	0	0	0	15	7
	Seafloor	0-1 cmbsf	70	70	9	9	0.023	0.023	15000	2800
	Transition	5-6 cmbsf	4	4	0.9	0.9	0.023	0.023	230	230
	Bloom	30 mbsl	7	7	0.4	0	3	0	-	-
27	Bottom	114 mbsl	2.3	2.3	0.9	0.9	30	0	-	-
	Seafloor	0 cmbsf	200	200	70	70	0.023	0.023	-	-
	Transition	10-11 cmbsf	700	700	0.4	0.4	0.9	0.9	-	-
	Bloom	25 mbsl	0	0	0.9	0.9	0	0	300	300
12	Bottom	181 mbsl	0	0	0.023	0.023	0	0	15	15
	Seafloor	0-1 cmbsf	40	40	0.43	0.15	0.23	0.23	21000	21000
	Transition	5-6 cmbsf	40	40	2.4	0.43	0.023	0.023	1200	750
	Bloom	23 mbsl	40	0	0.4	0	0	0	0	0
16	Bottom	119 mbsl	0.4	0.4	400	0	0	0	4	4
	Seafloor	0-1 cmbsf	40	40	90	90	0.023	0.023	110000	300
	Transition	5-6 cmbsf	4	4	30	30	4	0	11000	30

Table 2: Origin and closest type strains of *Roseobacter*-affiliated isolates.

Strain affiliation	**Station**	**Water Depth [mbsl]**	**Sediment Depth [cmbsf]**	**Origin**	**Medium**	**Closest type strain**	**Max Ident. [%]**
SK003/38	36	3		Bloom	DMS+Lactate	*Sulfitobacter dubius*	*98-99*
SK025	36	3		Bloom	DMS+Lactate	*Sulfitobacter pontiacus*	*97*
SK024/29	36	35		Bottom	DMS	*Loktanella salsilacus*	*99*
SK009/10	36	35	0-1	Seafloor	DMS+Lactate	*Loktanella rosea*	*98-99*
SK031/32	36	35	4-5	Transition	DMS	*Celeribacter neptunius*	*96*
SK040	5	26	5-6	Transition	DMSP	*Phaeobacter arcticus*	*99*
SK042	5	26	5-6	Transition	DMSP	*Phaeobacter inhibens*	*97*
SK002/23	8	23		Bottom	DMS+Lactate	*Shimia isoporae*	*99*
SK012	8	23	0-1	Seafloor	DMS+Lactate	*Sulfitobacter mediterraneus*	*98*
SK013	27	114	0-1	Seafloor	DMS+Lactate	*Shimia isoporae*	*97*
SK015/21	12	181	0-1	Seafloor	DMS+Lactate	*Loktanella agnita*	*93*
SK033	12	181	0-1	Seafloor	DMS	*Roseobacter litoralis*	*94*
SK045	12	181	0-1	Seafloor	DMS	*Loktanella koreensis*	*94*
SK011	16	119		Bottom	DMS+Lactate	*Sulfitobacter delicatus*	*98*

and 27 could not be inoculated in the anoxic medium. All dilution series were screened for the diversity of roseobacters and used for their isolation. After detecting growth in the dilution series by using a SYBR Green I assay (23), all positive enrichments were analyzed by *Roseobacter*-specific PCR. The results were used to estimate most probable numbers of roseobacters for the different environmental samples (Table 1). Interestingly, the anoxic enrichments with DMSO as electron acceptor showed the highest MPN with 1.1 x 10^8 bacterial cells/cm^3 for the sediment surface of station 16 and up to 2.1 x 10^7 *Roseobacter*-affiliated cells/cm^3 (station 12, seafloor). All dilution series containing DMSP and half of the series containing DMS as sole carbon source stimulated growth of sediment-dwelling roseobacters, only.

Several roseobacters were enriched and isolated from serial dilution cultures

A total of 263 isolates was obtained from the oxic enrichments. From these strains, 45 showed a positive result by PCR-screening using *Roseobacter*-specific primers. These isolates were subcultivated on agar plates containing their respective media. After sequencing and BLAST analysis, 20 strains turned out to be affiliated to the *Roseobacter*
clade (Table 2). Some strains from the same samples but different dilution steps were affiliated to the same species. More than half of the isolates (12 strains) derived from sediment samples and most of them (6 strains) from media containing DMS + lactate. Seven strains belonged to four different species within the genus *Loktanella*. Strains affiliated to *L. rosea*, *L. agnita*, *L. koreensis* were isolated from the seafloor and *L. salsilacus* was obtained from bottom waters. Five strains were related to four different species of the genus *Sulfitobacter* (*S. dubius*, *S. pontiacus*, *S. medietrraneus*, *S. delicatus*). Those were obtained from both, water and sediment samples (bloom, bottom water and seafloor). Other isolates were affiliated to *Shimia isoporae*, *Roseobacter litoralis* and *Celeribacter neptunius*. The latter and two affiliates of the genus *Phaeobacter* were isolated from the oxic/anoxic transition zone. Interestingly, the two strains affiliated to *P. arcticus* and *P. inhibens*, isolated from the DMSP medium, were not found in sediments before. Genetic signatures of the isolates from the oxic enrichments were detected in low percentages in the pyrosequencing dataset. On average, their relative abundance in the environmental samples was 0.03% of the bacterial communities and 0.55% of the *Rhodobacteraceae*. Interestingly, their percentage for all detected *Rhodobacteraceae* within the sediment samples was much higher (2.45%). This was mainly due to strains SK033 and SK045 (affiliates of *R. litoralis* and *L. koreensis*, respectively) that were both isolated from the seafloor of site 12. In this sample, they comprised 2.9% of all detected *Rhodobacteraceae.* They seem to be typical members of the seafloor bacterial community as they were found in all sediment samples with up to 6.2% of all *Rhodobacteraceae* at site 27. Initial anoxic enrichments which showed a *Roseobacter*-specific PCR signal were subcultivated in deep-agar dilutions and additionally analyzed for their diversity by DGGE. Sequencing of DGGE bands revealed growth of various roseobacters in the different dilution steps. For instance, bands of the original sample from the oxic/anoxic transition zone of station 16 were affiliated to species of *Roseobacter*, *Shimia*, *Sulfitobacter*, *Phaeobacter*, *Nereida ignava*, and *Dinoroseobacter shibae* (Fig. 5). While a species affiliated to *Nereida ignava* did not grow in the dilution cultures, the *Phaeobacter* species was specifically enriched in dilutions of 10^{-2} and 10^{-3}. Additional bands that were affiliated to another *Phaeobacter* and a *Sulfitobacter* species showed up in the 10^{-3} dilution step. The last two dilution steps (10^{-4} and 10^{-5}) were dominated by relatives of *Roseobacter / Shimia* species and showed faint bands affiliated to *D. shibae*. This example displays the dilution to extinction of fast-growing, but less abundant species along serial dilution cultures. However, as roseobacters were not the majority within the entire microbial community, many colonies were present in the respective agar tubes. Unfortunately, after six months of incubation, no roseobacters were among the first 84 picked colonies of the anoxic enrichments as indicated by preliminary sequencing of the first subcultivation step.

DISCUSSION

As main outcome of our study, both, DGGE analysis and amplicon-based sequencing of 16S rRNA genes indicated a broad overlap of *Roseobacter* populations thriving on matured particles and the sediment surface which were clearly distinct from the free-living communities dwelling in the water column. The attached fraction

of roseobacters was much more diverse than the free-living fraction. A variety of roseobacters was enriched and isolated from all sampling sites and compartments with media containing different organic sulfur compounds.

Habitat specific characteristics trigger the separation of pelagic and benthic roseobacters.

Specific differences between free-living, particle-associated and sediment-associated bacterial fractions in the southern North Sea were previously identified by Stevens et al (43). Focusing on free-living, aggregate-associated and sediment-dwelling microbial communities from intertidal flat regions of the German Wadden Sea, distinctive patterns for the different compartments were obtained. In conclusion, particles represent an intermediate habitat that is characterized by bacterial transformation processes occurring in both, the water column and at the sediment surface. Also, Rink et al (33, 34) detected differences between free living and particle-associated bacterial communities in the southern North Sea by DGGE analysis and showed, by applying Alphaprotebacteria and *Roseobacter*-specific primer sets, a higher number of bands in the particle-associated relative to the free-living bacterial communties. So far, however, the present study is the first one which examines specifically differences in the composition of *Roseobacter* communities among the free-living, particle-and sediment-associated bacterial communities. The low diversity of roseobacters we have detected within surface water samples is in accordance with previous investigations on phytoplankton blooms (45, 48). Even though there is a significant difference between the free-living and the attached lifestyle, an overlap in microbial diversity between both compartments is visible. The release of algal exudates triggers growth of highly specialized microbial communities on the phytoplankton-born particles (17). As surface-near, young particles are composed of fresh algal material, the associated communities are mainly degrading their exudates (9). It was shown that roseobacters exhibit mutualistic interactions with planktonic algae and dinoflagellates (16, 47). In contrast, matured particles in deeper layers are constituted of decaying algal material and thus harbor different microbial communities (37, 51). Environmental conditions such as high nutrient concentrations and a complex organic-matter composition are comparable to those found at the sediment surface. Thus, both compartments exhibit similar microbial community patterns (21, 29). The recalcitrant organic matter found on matured particles and the sediment surface stimulates highly diverse, slow-growing microbial communities with various metabolic capacities. The variety of available ecological niches and interactions between different microbial community members is reflected in the high proportion of roseobacters that are assigned as "uncultured". This type of environment is hard to be mimicked under laboratory conditions, which results in a lack of respective isolates.

A variety of roseobacters are involved in the utilization of organic sulfur compounds.

DMSP is an osmolyte that is produced by marine macro- and microalgae (50). This compound is released from lysed, dead or grazed algae and subsequently degraded by various bacteria. High concentrations of DMSP are associated to algal blooms that in turn attract high numbers of roseobacters (16) which were among the first bacteria isolated from DMSP-containing media (47). Sedimenting algal material might lead to an enrichment of DMSP-consuming *Roseobacter* species at the seafloor. In our cultivation experiments, growth of roseobacters was mainly stimulated in sediment samples when DMSP and DMS were provided as a sole carbon source. Apart from several isolates that were obtained from DMS-cultures we have isolated two *Phaeobacter* species from DMSP-enrichments from this compartment. Unfortunately, several DMSP-enrichments were lost during subcultivation. When DMSP is demethylated, the degradation product DMS is oxidized to DMSO. The high numbers of *Roseobacter*-affiliated isolates we have obtained from DMS-amended media indicate their role in this degradation process. On matured particles and especially in sediments, oxic and anoxic microniches are present in close proximity. If aerobically produced DMSO diffuses into anoxic regions, this can be used as electron acceptor under anoxic conditions. In our enrichments, we found high numbers of roseobacters reducing DMSO in water and sediment samples. This indicates either the presence of facultative anaerobic roseobacters in both, oxic and anoxic compartments, or the existence of microniches. For the next relatives of some of our oxic isolates, e.g. *S. pontiacus*, it is known that they can perform aerobic and anaerobic respiration (e.g. nitrate reduction) and harbor the gene for DMSO reduction (6). The fact that *Sulfitobacter* species were isolated from aerobic cultures and detected to grow in anaerobic

enrichments points to their ability to switch between the two lifestyles. To prove this, we are currently subcultivating further colonies from our anoxic enrichments. Additionally, growth experiments with the aerobic isolates under anoxic conditions will reveal their potential for DMSO reduction.

Cultivation of Roseobacter-affiliated bacteria matches the results of the molecular investigations

The fact that sequences of all our isolates showed a match in the pyrosequence database was used to quantify their relative abundance in the different compartments. Due to their relatively low percentages, this would have not been possible by using DGGE as fingerprinting technique, only. In general, both molecular methods exhibited a similar grouping of *Roseobacter* populations from environmental samples with some exceptions. In both approaches, sediments and matured particles were separated from the water samples but clustering of free-living and attached roseobacters was much clearer in the DGGE analysis. However, the primers used in this study are specific for roseobacters but also target a limited number of *Rhodobacteraceae* and other *Alphaproteobacteria.* Additionally, they produce a PCR product with a length of app. 200 bp which only gives a low phylogenetic resolution. These limitations can be minimized by amplicon-based sequencing of 16S rRNA genes using universal primers to digitally extract information on certain groups from the whole dataset. In contrast to DGGE, pyrosequencing has a higher capacity to explore the entire bacterial richness including rare species (8). Even though pyrosequencing requires a high amount of data-processing and analytical expertise, the possibility to detect and determine the relative abundance of cultured organisms (46) is a big advantage over other techniques. Following this approach, the laborious design and testing of specific primers and probes for quantitative PCR or fluorescence in-situ hybridization (FISH) is circumvented. However, screening of enrichments by DGGE allows to specifically target strains that are growing in these cultures. A molecularly guided cultivation will lead to the isolation of strains even if their colonies are hidden under a broad diversity of other colonies within the subcultures.

ACKNOWLEDGEMENTS

We thank the crew and the scientific party of RV Heincke (expedition HE361) for their help during sampling. Michael Pilzen, Sonja Standfest, Thomas Schirdewahn and Nawras Ghanem are acknowledged for technical assistance during sample preparation and subcultivation. This work was financially supported by Deutsche Forschungsgemeinschaft (DFG) within the Transregional Collaborative Research Center TRR51.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389-3402.
2. **Bragg, L., G. Stone, M. Imelfort, P. Hugenholtz, and G. W. Tyson.** 2012. Fast,accurate error-correction of amplicon pyrosequences using Acacia. Nature Methods **9:**425-426.
3. **Brinkhoff, T., H.-A. Giebel, and M. Simon.** 2008. Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. Archives of Microbiology **189:**531-539.
4. **Buchan, A., J. M. Gonzalez, and M. A. Moran.** 2005. Overview of the marine*Roseobacter* lineage. Appl Environ Microbiol **71:**5665-77.
5. **Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K.Costello, N. Fierer, A. G. Pena, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T.Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters,J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight.** 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods **7:**335-6.
6. **Caspi, R., T. Altman, K. Dreher, C. A. Fulcher, P. Subhraveti, I. M. Keseler, A.Kothari, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, A.Pujar, A. G. Shearer, M. Travers, D. Weerasinghe, P. Zhang, and P. D. Karp**.2012. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Research **40:**D742-D753.
7. **Chao, A., and J. Bunge.** 2002. Estimating the number of species in a stochastic abundance model. Biometrics **58:**531-9.
8. **Chau, J. F., A. C. Bagtzoglou, and M. R. Willig.** 2011. The effect of soil texture onrichness and diversity of bacterial

communities. Environmental Forensics **12:**333-341.
9. **Cole, J. J.** 1982. Interactions between bacteria and algae in aquatic ecosystems.Annual Review of Ecology and Systematics **13:**291-314.
10. **de Man, J. C.** 1977. MPN tables for more than one test. European J. Appl. Microbiol.**4:**307-310.
11. **DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T**.**Huber, D. Dalevi, P. Hu, and G. L. Andersen.** 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol **72:**5069-72.
12. **Edgar, R. C.** 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26:**2460-2461.
13. **Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince, and R. Knight.** 2011.UCHIME improves sensitivity and speed of chimera detection. Bioinformatics **27:**2194-200.
14. **Eilers, H., J. Pernthaler, J. Peplies, F. O. Glockner, G. Gerdts, and R. Amann**.2001. Isolation of novel pelagic bacteria from the German bight and their seasonal contributions to surface picoplankton. Appl Environ Microbiol **67:**5134-42.
15. **Giebel, H.-A., T. Brinkhoff, W. Zwisler, N. Selje, and M. Simon.** 2009. Distributionof *Roseobacter* RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. Environmental Microbiology **11:**2164-2178.
16. **Gonzalez, J. M., R. Simo, R. Massana, J. S. Covert, E. O. Casamayor, C. Pedros-Alio, and M. A. Moran.** 2000. Bacterial community structure associated with adimethylsulfoniopropionate-producing North Atlantic algal bloom. Applied andEnvironmental Microbiology **66:**4237-4246.
17. **Grossart, H. P., F. Levold, M. Allgaier, M. Simon, and T. Brinkhoff.** 2005. Marinediatom species harbour distinct bacterial communities. Environmental Microbiology **7:**860-873.
18. **Inagaki, F., M. Suzuki, K. Takai, H. Oida, T. Sakamoto, K. Aoki, K. H. Nealson**,**and K. Horikoshi.** 2003. Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. Applied andEnvironmental Microbiology **69:**7224-7235.
19. **Kiene, R. P.** 1990. Dimethyl sulfide production from dimethylsulfoniopropionate incoastal seawater samples and bacterial cultures. Appl Environ Microbiol **56:**3292-7.
20. **Lenk, S., C. Moraru, S. Hahnke, J. Arnds, M. Richter, M. Kube, R. Reinhardt, T**.**Brinkhoff, J. Harder, R. Amann, and M. Mussmann.** 2012. *Roseobacter* clade bacteria are abundant in coastal sediments and encode a novel combination of sulfuroxidation genes. Isme J **6:**2178-87.
21. **Llobet-Brossa, E., R. Rossello-Mora, and R. Amann.** 1998. Microbial community composition of Wadden Sea sediments as revealed by fluorescence in situhybridization. Applied and Environmental Microbiology **64:**2691-2696.
22. **MacGregor, B. J., D. P. Moser, E. W. Alm, K. H. Nealson, and D. A. Stahl.** 1997.*Crenarchaeota* in Lake Michigan sediment. Appl Environ Microbiol **63:**1178-81.
23. **Martens-Habbena, W., and H. Sass.** 2006. Sensitive determination of microbial growth by nucleic acid staining in aqueous suspension. Appl Environ Microbiol **72:**87-95.
24. **Martin, M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. Bioinformatics in Action **17:**10-12.
25. **Miller, T. R., and R. Belas.** 2004. Dimethylsulfoniopropionate metabolism by Pfiesteria-associated Roseobacter spp. Applied and Environmental Microbiology**70:**3383-3391.
26. **Mills, H. J., C. Hodges, K. Wilson, I. R. MacDonald, and P. A. Sobecky.** 2003. Microbial diversity in sediments associated with surface-breaching gas hydratemounds in the Gulf of Mexico. Fems Microbiology Ecology **46:**39-52.
27. **Moran, M. A., R. Belas, M. A. Schell, J. M. Gonzalez, F. Sun, S. Sun, B. J. Binder**, **J. Edmonds, W. Ye, B. Orcutt, E. C. Howard, C. Meile, W. Palefsky, A**.**Goesmann, Q. Ren, I. Paulsen, L. E. Ulrich, L. S. Thompson, E. Saunders, and A**.**Buchan.** 2007. Ecological genomics of marine Roseobacters. Appl Environ Microbiol **73:**4559-69.
28. **Mouné, S., P. Caumette, R. Matheron, and J. C. Willison.** 2003. Molecularsequence analysis of prokaryotic diversity in the anoxic sediments underlyingcyanobacterial mats of two hypersaline ponds in Mediterranean salterns. FEMSMicrobiology Ecology **44:**117-130.
29. **Novitsky, J. A.** 1990. Evidence for sedimenting particles as the origin of the microbialcommunity in a coastal marine sediment. Marine Ecology Progress Series **60:**161-167.
30. **Overmann, J., and C. Tuschak.** 1997. Phylogeny and molecular fingerprinting ofgreen sulfur bacteria. Archives of Microbiology **167:**302-309.
31. **Pearson, K.** 1926. On the coefficient of racial likeness. Biometrika **18:**105-117.
32. **Pruesse, E., C. Quast, K. Knittel, B. M. Fuchs, W. Ludwig, J. Peplies, and F. O**. **Glockner.** 2007. SILVA: a comprehensive

online resource for quality checked andaligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res**35:**7188-96.

33. **Rink, B., N. Gruener, T. Brinkhoff, K. Ziegelmueller, and M. Simon.** 2011.Regional patterns of bacterial community composition and biogeochemical properties in the southern North Sea. Aquatic Microbial Ecology **63:**207-222.
34. **Rink, B., S. Seeberger, T. Martens, C.-D. Duerselen, M. Simon, and T. Brinkhoff**.2007. Effects of phytoplankton bloom in a coastal ecosystem on the composition ofbacterial communities. Aquatic Microbial Ecology **48:**47-60.
35. **Schloss, P. D., and J. Handelsman.** 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Applied and Environmental Microbiology **71:**1501-1506.
36. **Schneider, D., G. Arp, A. Reimer, J. Reitner, and R. Daniel.** 2013. Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of thekiritimati atoll, central pacific. PLoS One **8:**e66662.
37. **Schweitzer, B., I. Huber, R. Amann, W. Ludwig, and M. Simon.** 2001. Alpha- and beta-Proteobacteria control the consumption and release of amino acids on lake snow aggregates. Applied and Environmental Microbiology **67:**632-645.
38. **Selje, N., M. Simon, and T. Brinkhoff.** 2004. A newly discovered *Roseobacter* cluster in temperate and polar oceans. Nature **427:**445-8.
39. **Shannon, C. E.** 2001. A mathematical theory of communication. SIGMOBILE Mob. Comput. Commun. Rev. **5:**3-55.
40. **Simon, C., A. Wiezer, A. W. Strittmatter, and R. Daniel.** 2009. Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. Applied and Environmental Microbiology **75:**7519-7526.
41. **Sogin, M. L., and J. H. Gunderson.** 1987. Structural diversity of eukaryotic small subunit ribosomal RNAs. Annals of the New York Academy of Sciences **503:**125-139.
42. **Stahl, D. A., B. Flesher, H. R. Mansfield, and L. Montgomery.** 1988. Use of phylogenetically based hybridization probes for studies of ruminal microbial ecology. Applied & Environmental Microbiology **54:**1079-1084.
43. **Stevens, H., T. Brinkhoff, and M. Simon.** 2005. Composition of free-living, aggregate-associated and sediment surface-associated bacterial communities in the German Wadden Sea. Aquatic Microbial Ecology **38:**15-30.
44. **Süß, J., B. Engelen, H. Cypionka, and H. Sass.** 2004. Quantitative analysis of bacterial communities from Mediterranean sapropels based on cultivation-dependent methods. FEMS Microbiol Ecol **51:**109-21.
45. **Teeling, H., B. M. Fuchs, D. Becher, C. Klockow, A. Gardebrecht, C. M. Bennke, M. Kassabgy, S. Huang, A. J. Mann, J. Waldmann, M. Weber, A. Klindworth, A. Otto, J. Lange, J. Bernhardt, C. Reinsch, M. Hecker, J. Peplies, F. D. Bockelmann, U. Callies, G. Gerdts, A. Wichels, K. H. Wiltshire, F. O. Glöckner, T. Schweder, and R. Amann.** 2012. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. Science **336:**608619 611.
46. **Vaz-Moreira, I., C. Egas, O. C. Nunes, and C. M. Manaia.** 2011. Culture-dependent and culture-independent diversity surveys target different bacteria: a case study in a freshwater sample. Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology **100:**245-257.
47. **Wagner-Döbler, I., and H. Biebl.** 2006. Environmental biology of the marine *Roseobacter* lineage. Annu Rev Microbiol **60:**255-80.
48. **West, N. J., I. Obernosterer, O. Zemb, and P. Lebaron.** 2008. Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. Environmental Microbiology **10:**738629 756.
49. **Wilms, R., H. Sass, B. Kopke, H. Koster, H. Cypionka, and B. Engelen.** 2006. Specific bacterial, archaeal, and eukaryotic communities in tidal-flat sediments along a vertical profile of several meters. Applied and Environmental Microbiology **72:**2756633 2764.
50. **Yoch, D. C.** 2002. Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. Appl Environ Microbiol **68:**5804636 15.
51. **Zhu, M., G. Zhu, L. Zhao, X. Yao, Y. Zhang, G. Gao, and B. Qin.** 2013. Influence of algal bloom degradation on nutrient release at the sediment-water interface in Lake Taihu, China. Environ Sci Pollut Res Int **20:**1803-11.
52. **Zubkov, M. V., B. M. Fuchs, S. D. Archer, R. P. Kiene, R. Amann, and P. H. Burkill.** 2002. Rapid turnover of dissolved DMS and DMSP by defined bacterioplankton communities in the stratified euphotic zone of the North Sea. Deep643 Sea Research Part Ii-Topical Studies in Oceanography **49:**3017-3038.

SUPPORTING INFORMATION FOR STUDY 4

CONTENTS:

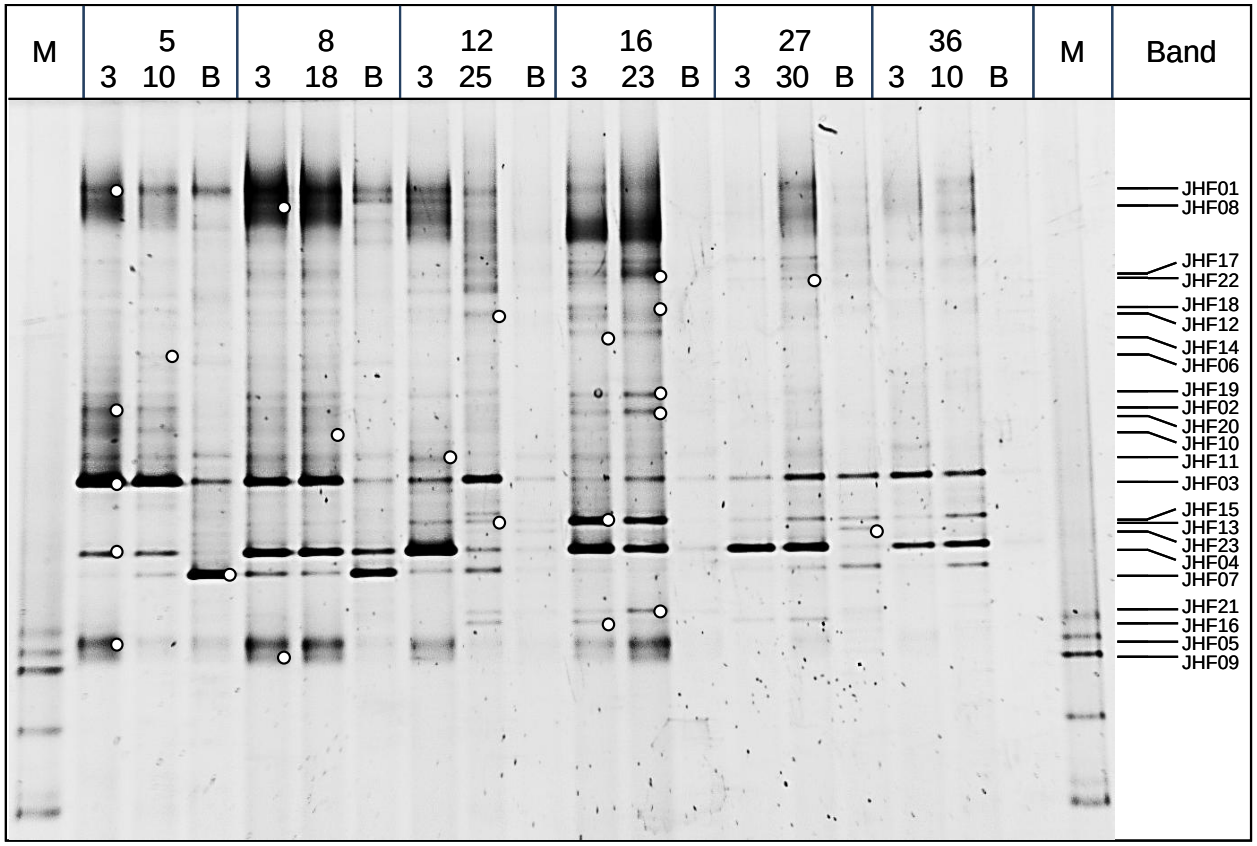


Fig. S1. DGGE community patterns of free-living roseobacters. Samples were obtained from surface waters (3 mbsl), deep waters (10-30 mbsl) and bottom waters (depth: see material and methods section). PCR amplicons (~200 bp) were generated by using *Roseobacter*-specific primers. Representative DGGE bands (white circles) were excised and sequenced. Phylogenetic affiliation and distribution, see table S1. M = marker, B = bottom water

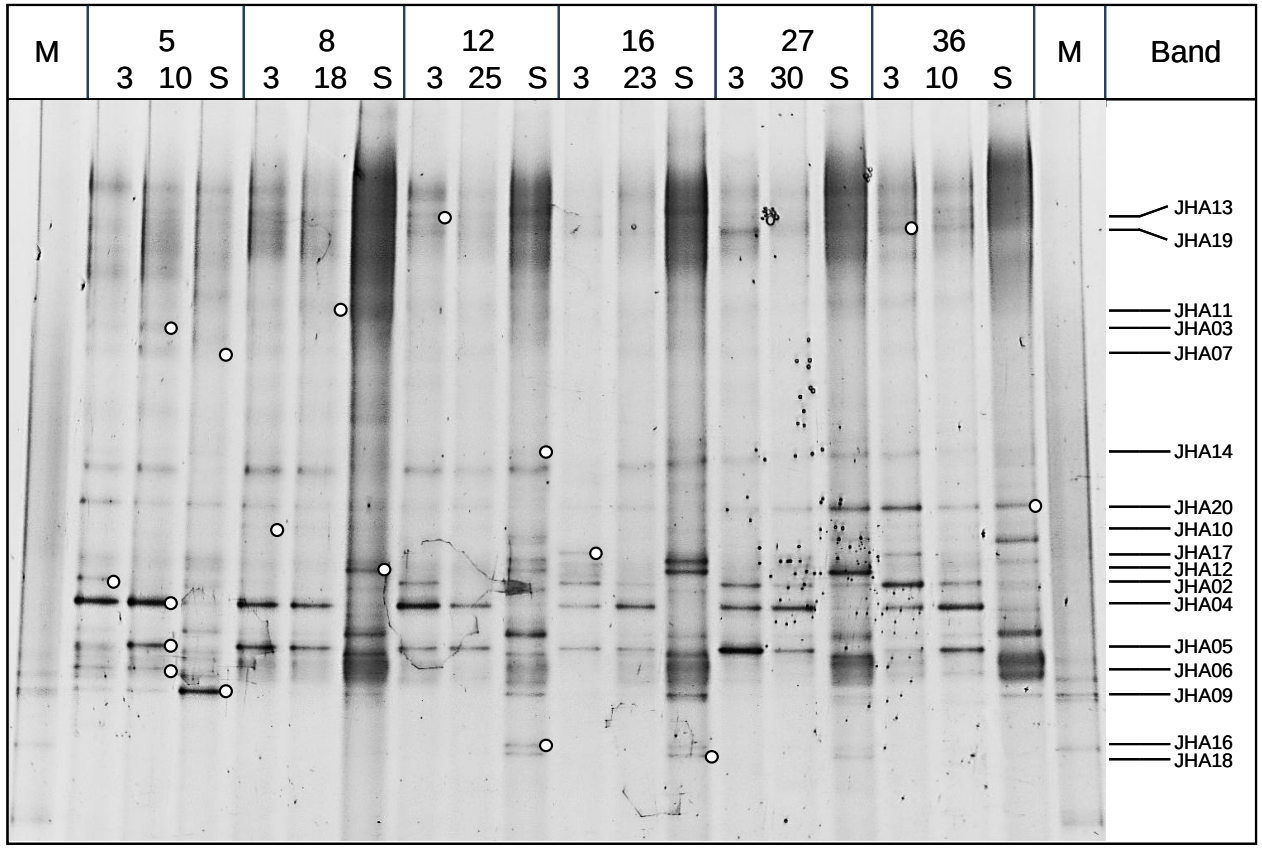


Fig. S2. DGGE community patterns of particle-associated and sediment-dwelling roseobacters. Samples were obtained from young particles (3 mbsl), matured particles (10-30 mbsl) and the sediment surface (depth: see material and methods section). PCR amplicons (~200 bp) were generated by using *Roseobacter*-specific primers. Representative DGGE bands (white circles) were excised and sequenced. Phylogenetic affiliation and distribution, see table S2. M = marker, S = seafloor

Table S1. Free-living fraction of *Roseobacter*-affiliated phylotypes detected by sequencing of DGGE bands (o) from the gel displayed in Figure S1. Site and depths related distribution is indicated by the presence of corresponding DGGE bands in other lanes (+).

Band No.	Affiliation	Sim [%]	Station and Depth [mbsl]																	
			3	5 10	B	3	8 18	B	3	12 25	B	3	16 23	B	3	27 30	B	3	36 10	B
JHF01	*Sulfitobacter donghicola*	92	o	+	+	+	+	+	+	+		+	+			+	+	+	+	
JHF02	*Nautella italica*	99	o	+	+	+	+	+				+	+							
JHF03	*Nereida ignava*	100	o	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+
JHF04	*Sulfitobacter pontiacus*	99	o	+		+	+	+	+	+	+	+	+		+	+		+	+	
JHF05	*Nereida ignava*	100	o	+	+	+	+		+			+	+							
JHF06	*Roseovarius crassostreae*	99		o																
JHF07	*Roseovarius aestuarii*	100		+	o	+	+	+	+	+						+	+		+	
JHF08	*Thalassobacter stenotrophicus*	100	+	+		o	+		+											
JHF09	*Thalassobacter stenotrophicus*	99	+	+	+	o	+		+			+	+							
JHF10	*Thalassobius mediterraneus*	100	+	+	+	+	o	+	+	+		+	+							
JHF11	*Sulfitobacter pontiacus*	96	+	+	+	+	+	+	o	+		+	+		+	+	+	+	+	+
JHF12	*Thalassobacter stenotrophicus*	97	+	+	+	+	+	+	+	o										
JHF13	*Thalassobacter stenotrophicus*	97							+	o		+	+		+	+	+	+	+	+
JHF14	*Sulfitobacter pontiacus*	95										o	+							
JHF15	*Nereida ignava*	98							+	+		o	+		+	+	+	+	+	+
JHF16	*Rhodobacteraceae bacterium*	95								+		o								
JHF17	*Nereida ignava*	98								+		+	o							
JHF18	*Sulfitobacter marinus*	95											o							
JHF19	*Phaeobacter caeruleus*	97	+	+		+	+		+	+		+	o							
JHF20	*Rhodobacteraceae bacterium*	94	+	+	+	+	+	+				+	o							
JHF21	*Sulfitobacter pontiacus*	98								+		+	o							
JHF22	*Thalassobacter stenotrophicus*	98								+		+	+			o				
JHF23	*Thalassobius mediterraneus*	98									+						o			

B: Bottom seawater (depth see material and methods section)

Table S2. Particle-asssociated fraction of *Roseobacter*-affiliated phylotypes detected by sequencing of DGGE bands (o) from the gel displayed in Figure S2. Site and depths related distribution is indicated by the presence of corresponding DGGE bands in other lanes (+).

Band No.	Affiliation	Sim [%]	Station and Depth [mbsl]																	
			5			8			12			16			27			36		
			3	10	S	3	18	S	3	10	S	3	18	S	3	10	S	3	18	S
JHA02	*Sulfitobacter pontiacus*	99	o	+	+				+			+			+			+	+	
JHA03	*Phaeobacter inhibens*	97		o																
JHA04	*Loktanella rosea*	97	+	o		+	+		+	+		+	+		+	+		+	+	
JHA05	*Pseudoruegeria aquimaris*	96	+	o		+	+		+	+		+	+		+	+		+	+	
JHA06	*Roseovarius crassostreae*	97	+	o		+	+	+	+		+			+			+			
JHA07	*Leisingera aquimarina*	97			o															
JHA09	*Tateyamaria omphalii*	99			o															
JHA10	*Jannaschia pohangensis*	96	+	+	+	o	+	+	+	+	+	+	+	+	+	+	+	+	+	+
JHA11	*Jannaschia pohangensis*	94					o	+												
JHA12	*Nereida ignava*	98						o			+						+			
JHA13	*Pelagicola litoralis*	94							o									+		
JHA14	*Nautella italica*	98									o			+			+	+	+	+
JHA16	*Phaeobacter inhibens*	100									o			+						
JHA17	*Thalassobacter stenotrophicus*	98	+									o						+		
JHA18	*Dinoroseobacter shibae*	100									+			o						
JHA19	*Sulfitobacter dubius*	95	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	o	+	+
JHA20	*Nereida ignava*	97	+	+	+	+	+	+	+	+	+		+		+	+	+	+	+	o

S: Seafloor (depth: see material and methods section)

STUDY 5:

ADAPTATION OF AN ABUNDANT *ROSEOBACTER* RCA ORGANISM TO PELAGIC SYSTEMS REVEALED BY GENOMIC AND TRANSCRIPTOMIC ANALYSES

VOGET S[1], **WEMHEUER B**[1], BRINKHOFF T[2], VOLLMERS J[1], DIETRICH S[1], GIEBEL HA [2], BEARDSLEY C[2], BAKENHUS I[2], BILLERBECK S[2], DANIEL R[1], AND SIMON M[2]

ISME J 2014 AUG 1.
(DOI: 10.1038/ISMEJ.2014.134)

[1]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY; [2]INSTITUTE FOR CHEMISTRY AND BIOLOGY OF THE MARINE ENVIRONMENT (ICBM), CARL-VON-OSSIETZKY-UNIVERSITY OF OLDENBURG, CARL-VON-OSSIETZKY-STR. 9-11, D-26111 OLDENBURG, GERMANY; [3]PRESENT ADDRESS: BIOZENTRUM KLEIN FLOTTBEK, MIKROBIOLOGIE, OHNHORSTSTR. 18, D-22609 HAMBURG, GERMANY

Author contributions to the work:

Performed the experiments: SV, BW, HG, CB

Analyzed data: SV, MS, BW, HG, CB, JV, SD

Contributed data on water properties and analysis of these data: BW, TB, HAG, CB, IB, SB

Wrote the publication: SV, MS, TB, RD

Conceived and designed the experiments:SV, MS, TB, RD

The ISME Journal (2014), 1–14

www.nature.com/ismej

ORIGINAL ARTICLE

Adaptation of an abundant *Roseobacter* RCA organism to pelagic systems revealed by genomic and transcriptomic analyses

Sonja Voget[1], Bernd Wemheuer[1], Thorsten Brinkhoff[2], John Vollmers[1], Sascha Dietrich[1], Helge-Ansgar Giebel[2], Christine Beardsley[2], Carla Sardemann[2], Insa Bakenhus[2], Sara Billerbeck[2], Rolf Daniel[1] and Meinhard Simon[2]

[1]*Institute of Microbiology and Genetics, Genomic and Applied Microbiology and Göttingen Genomics Laboratory, University of Göttingen, Göttingen, Germany and* [2]*Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany*

The RCA (*Roseobacter* clade affiliated) cluster, with an internal 16S rRNA gene sequence similarity of >98%, is the largest cluster of the marine *Roseobacter* clade and most abundant in temperate to (sub)polar oceans, constituting up to 35% of total bacterioplankton. The genome analysis of the first described species of the RCA cluster, *Planktomarina temperata* RCA23, revealed that this phylogenetic lineage is deeply branching within the *Roseobacter* clade. It shares not >65.7% of homologous genes with any other organism of this clade. The genome is the smallest of all closed genomes of the *Roseobacter* clade, exhibits various features of genome streamlining and encompasses genes for aerobic anoxygenic photosynthesis (AAP) and CO oxidation. In order to assess the biogeochemical significance of the RCA cluster we investigated a phytoplankton spring bloom in the North Sea. This cluster constituted 5.1% of the total, but 10–31% (mean 18.5%) of the active bacterioplankton. A metatranscriptomic analysis showed that the genome of *P. temperata* RCA23 was transcribed to 94% in the bloom with some variations during day and night. The genome of *P. temperata* RCA23 was also retrieved to 84% from metagenomic data sets from a Norwegian fjord and to 82% from stations of the Global Ocean Sampling expedition in the northwestern Atlantic. In this region, up to 6.5% of the total reads mapped on the genome of *P. temperata* RCA23. This abundant taxon appears to be a major player in ocean biogeochemistry.
The ISME Journal advance online publication, 1 August 2014; doi:10.1038/ismej.2014.134

Introduction

Our understanding of the role of abundant individual taxa in ocean biogeochemistry is hampered by the fact that so far very few of such taxa have been isolated and are available to genomic and postgenomic analyses (Yooseph *et al.*, 2010). It is extremely difficult to obtain representative isolates of the major players and only novel approaches made this effort more successful (Giovannoni and Stingl, 2007). However, still today detailed information on the role of abundant individual taxa in oceanic cycling of matter is available only for *Prochlorococcus* and for Cand. *Pelagibacter ubique* of the SAR11 clade. *Pelagibacter ubique* of the SAR11 clade (Giovannoni *et al.*, 2005; Tripp *et al.*, 2008; Sowell *et al.*, 2009; Thompson *et al.*, 2011). Isolates of these taxa are available, their genomes have been sequenced and thus form a basis for postgenomic studies and relating metagenomic and, more importantly, metatranscriptomic and metaproteomic information to individual taxa.

The *Roseobacter* and SAR11 clades are the most prominent subdivisions of Alphaproteobacteria in the ocean's near surface waters (Giovannoni and Stingl, 2005). The RCA (*Roseobacter* clade affiliated) cluster with an internal sequence similarity of the 16S rRNA gene of at least 98%, constitutes up to 35% of total bacterioplankton and is most abundant in temperate to (sub)polar oceans, but absent in tropical and subtropical regions (Selje *et al.*, 2004; Giebel *et al.*, 2009; Giebel *et al.*, 2011). It is divided into a subcluster with sequences from temperate regions and one with sequences of subpolar and polar origin (Giebel *et al.*, 2011). Because of the high abundance and distinct biogeography of the RCA cluster, there is great interest to elucidate its functional role and biogeochemical significance. However, no information on the genomic data and features of organisms of this important cluster are yet available because RCA organisms defy isolation and cultivation greatly.

Correspondence: M Simon, Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, C v Ossietzky Str 9-11, Oldenburg D-26111, Germany.
E-mail: m.simon@icbm.de
Received 24 April 2014; revised 17 June 2014; accepted 21 June 2014

An isolate of this cluster, retrieved from the southern North Sea, became recently available and was characterized as type species of the first described species of the genus *Planktomarina*, *Planktomarina temperata* RCA23 (Giebel *et al.*, 2013). In the North Sea, *P. temperata* RCA23 represents the most abundant ribotype of the RCA cluster that comprises persistently between 2% and 20% of total bacterioplankton (Selje *et al.*, 2004; Giebel *et al.*, 2011; Teeling *et al.*, 2012) and is a major representative of the *Roseobacter* clade in the active bacterioplankton (Wemheuer *et al.*, 2014). Based on these observations we hypothesized that the genome of this organism exhibits distinct features of its adaptation to life in the nutrient-poor pelagic environment with respect to genome organization and metabolic potential. Further, we hypothesized that the genome of *P. temperata* RCA23 is well represented in metagenomic and metatranscriptomic data in the North Sea because of its high abundance and activity. Therefore, we sequenced the genome of *P. temperata* RCA23 to elucidate its metabolic potential. In addition, we assessed its significance and functional role during a phytoplankton spring bloom in the southern North Sea by applying a metatranscriptomic approach. Finally, metagenomic data sets of pelagic marine systems were mined for the presence, abundance and genomic features of this organism.

Materials and methods

Origin and growth of P. temperata *RCA23*

P. temperata RCA23 was originally isolated from a water sample collected in the southern North Sea (Giebel *et al.*, 2011, 2013). It was grown in liquid culture of autoclaved sea water amended with marine broth (40% of peptone and yeast extract; Giebel *et al.*, 2013) to an optical density of 0.2. Biomass was harvested by 20 min centrifugation (Beckman JA10, Krefeld, Germany) at 7500 r.p.m. and 4 °C and DNA was extracted using the MasterPure DNA Purification Kit (Epicentre, Madison, Wisconsin, USA) according to the manufacturer's protocol.

Sequence determination, gene annotation and phylogenetic analyses

To sequence the genome of *P. temperata* RCA23 a pyrosequencing run was performed using a Roche GS-FLX 454 sequencer (Branford, CT, USA) with Titanium chemistry. All sequencing steps were performed according to the manufacturer's protocols and recommendations. In total 411 932 reads were generated and assembled to 78 contigs bigger than 500 bp with a 26-fold coverage. Furthermore, 576 fosmid Sanger-sequences were added to the data set to identify the contig order. Gap closure and polishing were carried out using the Staden software package (Staden, 1996) and PCR-based techniques on genomic DNA. Open reading frames were identified using YACOP (Tech and Merkl, 2003) and GLIMMER (Delcher *et al.*, 2007). The open reading frame finding was inspected manually with Artemis and open reading frames were corrected by checking the GC frame plot, ribosomal binding site and blast hits against the NCBI nr database.

Orthologous protein sequences were identified using reciprocal BLASTp-analysis combined with global alignments based on the Needleman–Wunsch algorithm. Only hits with *e*-values $<1e-20$ were considered and additionally filtered, based on sequence identity cutoffs of the respective global alignments. A cutoff value of 30% sequence identity was chosen to identify orthologs. For multilocus sequence analysis, protein sequences of 162 genes, for which one ortholog but no paralog was found in every comparison strain, were concatenated. The multilocus sequence analysis tree was constructed using ARB v5.1 (Ludwig *et al.*, 2004) and the evolutionary history was inferred using the neighbor-joining method and recovered reproducibly with the maximum-likelihood method (Figure 1).

Study area, sample collection and chlorophyll

The significance of the RCA cluster was investigated during a phytoplankton spring bloom in the southern North Sea. Samples were collected at 11 stations at 2 m depth between 25 and 31 May 2010 on board RV Heincke by 4 l Niskin bottles mounted on a CTD rosette (Sea-Bird, Bellevue, WA, USA). For pyrosequencing, metagenomic and metatranscriptomic analyses, 50 l of sea water were prefiltered through a 10-μm nylon net and a filter sandwich consisting of a precombusted glass fiber and 3-μm polycarbonate filter (47 mm diameter). Bacterioplankton was harvested from a prefiltered 1-liter sample on a filter sandwich consisting of a glass fiber and 0.2-μm polycarbonate filter (47 mm). One filter sandwich was used for RNA extraction. For gene expression analysis and metagenome sequencing, at least four filter sandwiches were subjected to RNA and DNA extraction. Chlorophyll *a* concentrations were determined as described (Giebel *et al.*, 2011).

Enumeration of bacteria, CARD-FISH and BrdU-FISH

Bacterial cell numbers were determined by flow cytometry after staining of subsamples with SybrGreen I (Sigma-Aldrich, Munich, Germany), preserved with glutardialdehyde (final concentration 1%) and stored at −80 °C until further analysis as described Giebel *et al.* (2011). The abundance of RCA cells was determined by catalyzed reporter deposition fluorescence *in situ* hybridization (CARD-FISH) and the proportion of DNA-synthesizing RCA cells by incorporation of BrdU (4 h incubation) according to Pernthaler *et al.* (2002) and applying an RCA-specific probe set (RCA996, C Beardsley and I Bakenhus, unpublished data). Cells were counted via epifluorescence microscopy using ×1000 magnification and suitable filter sets for DAPI-stained total cells, $Alexa_{488}$-stained RCA cells,

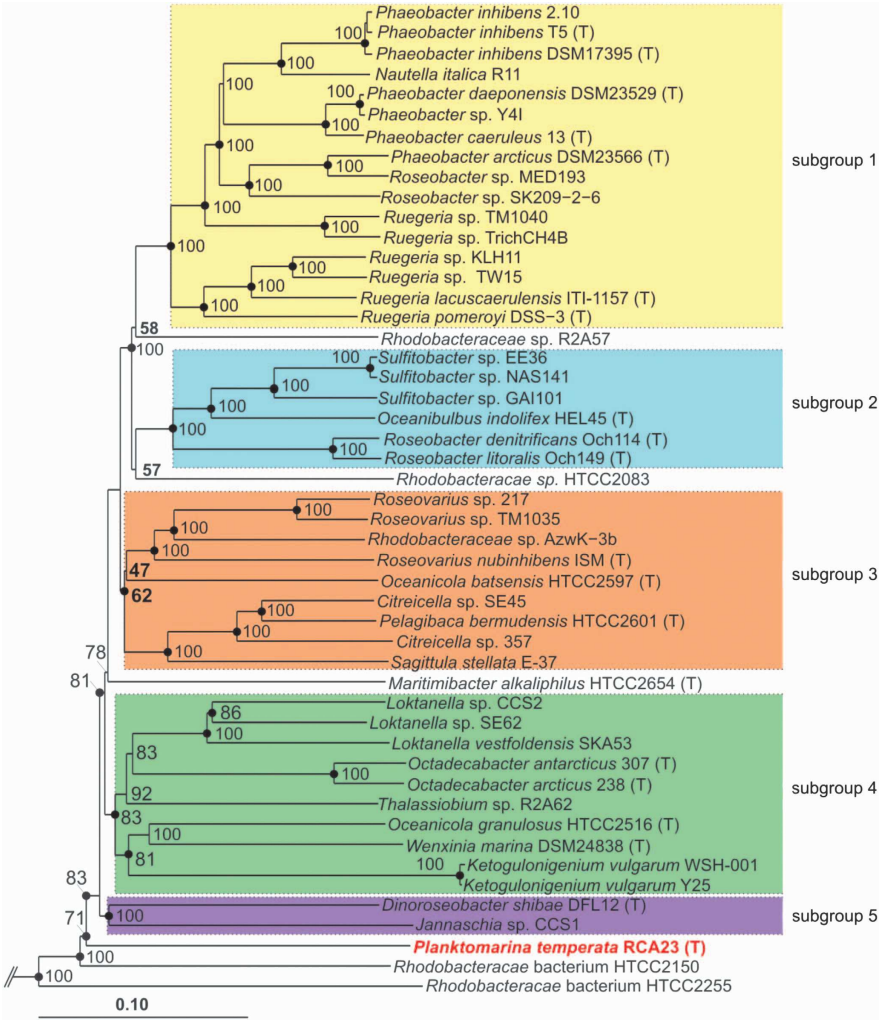


Figure 1 Neighbor-joining tree based on multilocus sequence analysis (MLSA) of genome-sequenced organisms of the *Roseobacter* clade. The tree was constructed using ARB v5.1 and is based on a similar tree of Newton *et al.* (2010) but includes additional genome sequences and an extended gene set. Filled circles indicate nodes also recovered reproducibly with maximum-likelihood calculation. Numbers at the nodes are bootstrap values (only >50% are shown) from 1000 replicates. Subclades of the *Roseobacter* clade are marked by different colors. (T) indicates type strains. *Escherichia coli* K12 MG1655 was used as outgroup.

and Cy3-stained (Perkin Elmer, Waltham, MA, USA) BrdU-positive cells, respectively.

16S rRNA gene amplicons, metagenomics and transcriptomics

Environmental DNA and RNA were coextracted and the composition of the active bacterial community was assessed by 16S rRNA PCRs as described before Wemheuer *et al.* (2014). For metagenomic and metatranscriptomic analyses DNA and cDNA were sequenced on an Illumina/Solexa GAIIx system (San Diego, CA, USA). In total, 54 334 282 paired-end sequences of 100 bp were generated for the metagenomic and 78 042 122 single-read sequences of 75–100 bp for the metatranscriptomic data sets, respectively (Supplementary Table S1). The sequences were quality trimmed and the Illumina adapter sequences were removed with Trimmomatic v0.30 (Bolger *et al.*, 2014) using the following parameters: adapter:2:40:15 leading:3 trailing:3 slidingwindow:4:15 minlen:50. Reads derived from ribosomal RNA gene fragments were filtered with SortMeRNA (Kopylova *et al.*, 2012). The remaining sequences were mapped with Bowtie 2 (Langmead and Salzberg, 2012) using the implemented end-to-end mode, which requires that the entire read align from one end to the other.

To compare the RNA-Seq results, the read counts were normalized to remove biases like the length of the transcript and the sequencing depth of a sample. We used the Nucleotide activity Per Kilobase of exon model per million Mapped reads (NPKM), a derivate of RPKM (reads per kilo base per million), as a normalized read count value (Wiegand *et al.*, 2013).

Data deposition

The genome sequence of *P. temperata* RCA23 has been deposited at the NCBI GenBank database with the accession number CP003984. The metagenomic

and metatranscriptomic data sets have been deposited at MG-Rast with the accession numbers: 4548721.3 (Station 3 gDNA), 4550305.3 (Station 3 cDNA), 4548722.3 (Station 9 gDNA), 4550303.3 (Station 9 cDNA), 4548723.3 (Station 13 gDNA) and 4550304.3 (Station 13 cDNA).

Results and discussion

Genomic features of P. temperata *RCA23*

P. temperata RCA23 is deeply branching within the *Roseobacter* clade, not affiliated to any of the known subclades, as shown by multilocus sequence analysis (Figure 1). The genome of *P. temperata* RCA23 encompasses 3.29 Mbp and carries 3101 genes of which 3054 encode proteins (Supplementary Table S2). It is the smallest of all closed and the third smallest of all genomes of the *Roseobacter* clade and carries the second lowest number of genes of all organisms of this clade (Figure 2a and Supplementary Table S2). The only organism of this clade with a significantly smaller genome is *Rhodobacterales* bacterium HTCC2255 (Luo *et al.*, 2013). Further, partial genomes of two other organisms of the *Roseobacter* clade with estimated genome sizes of <3 Mbp were recently reported from sequence analyses of single amplified genomes retrieved from coastal and oceanic surface waters (Swan *et al.*, 2013; Luo *et al.*, 2014). Members of the *Roseobacter* clade with a genome size of <3.8 Mbp are exclusively of pelagic origin (Figure 2a and Supplementary Table S2). The mean genome sizes of the pelagic members of this clade is 4.06 Mbp and 0.5 Mbp smaller than that of members associated to other organisms, surfaces or sediments (Figure 2b and Supplementary Table S3). The relatively small genome of *P. temperata* RCA23 and other pelagic members of the *Roseobacter* clade and the reduction in genomic traits is consistent with genomic streamlining and adaptation to nutrient-poor pelagic marine ecosystems (see below; Giovannoni *et al.*, 2005; Luo *et al.*, 2013; Swan *et al.*, 2013). The genomic difference between both groups of the *Roseobacter* clade obviously reflects the divergent evolutionary history of this clade with genomic

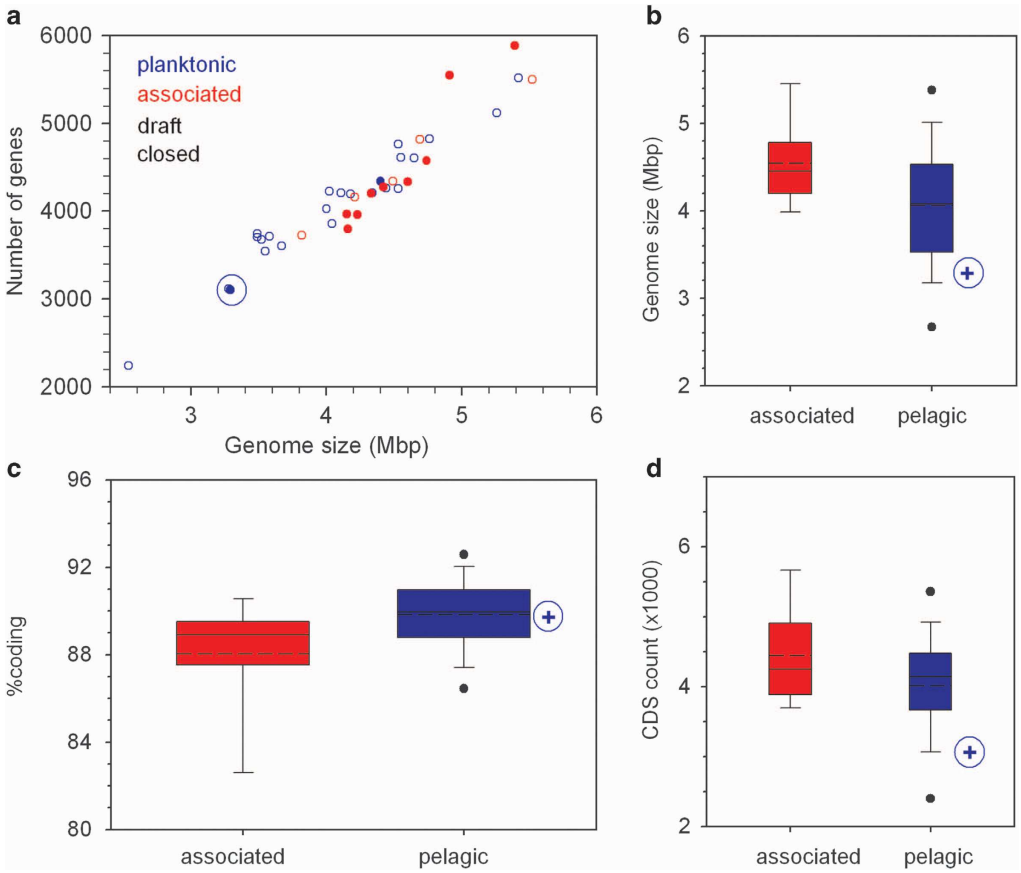


Figure 2 Genomic traits of *P. temperata* RCA23 in comparison to other organisms of the *Roseobacter* clade. Relation between genome size and number of genes (**a**). Box–Whisker plot of the GC content (**b**), percentage coding genes (**c**) and CDS (coding DNA sequences) count (**d**) of 14 members of the *Roseobacter* clade associated to other organisms, surfaces or sediments and of 24 members with a pelagic life style listed in Supplementary Table S2. Data include 12 closed and 26 draft genomes. Blue circle and +: *P. temperata* RCA23. The boxes show the median (solid line), mean (dashed line), the 25th and 75th percentile and the whiskers the 10th and 90th percentiles. •: Outlayers. For further details see Supplementary Tables S2 and S3.

streamlining of at least several pelagic members and gained genomic content by the associated members (Luo *et al*., 2013). The much larger genome size of *P. temperata* RCA23 and other pelagic members of the *Roseobacter* clade, however, harbour genomes much larger than highly streamlined marine pelagic bacteria such as Cand. *P. ubique* of the SAR11 clade. This feature points out that, besides streamlining, distinct aspects of their life history and adaptation to their niche in the marine pelagic realm were also important in shaping the evolution of these pelagic members of the *Roseobacter* clade (Giovannoni *et al*., 2014). These authors have identified a low number of σ-factors as a feature to distinguish between streamlined and non-streamlined genomes of bacteria with a more complex life style. The genome of *P. temperata* RCA23 harbors six σ-factors, a value much lower than that of other members of the *Roseobacter* clade and in the same range as that of many typical streamlined genomes of a size of <2 Mb (Giovannoni *et al*., 2014).

The GC content of 54% of *P. temperata* RCA23 is lower than that of all associated (mean $=60.07\pm$ 3.58%, $N=14$) and the fifth lowest of the 24 pelagic members of the *Roseobacter* clade included in this analysis (mean $=59.28\pm6.82$%; Supplementary Tables S2 and S3). However it is substantially higher than the genomic GC content of many other marine bacteria, ranging below 40%, such as *Rhodobacterales* bacterium HTCC2255 (Supplementary Table S2), Cand. *P. ubique* and pelagic bacteria subjected to single cell genome sequencing (Swan *et al*., 2013; Luo *et al*., 2014). It has been speculated that a reduced genomic GC content may be an adaptation to nitrogen limitation in marine systems (Swan *et al*., 2013), but clear-cut evidence is still missing, also because it is still unclear whether pelagic heterotrophic bacteria are carbon- or nitrogen-limited. The percentage of coding DNA of *P. temperata* RCA23 is very similar to the mean of all pelagic members of the *Roseobacter* clade which, however, is significantly higher than the mean of the organisms of this clade associated to other organisms, surfaces or sediment (Figure 2c and Supplementary Table S3).

The genome of *P. temperata* RCA23 shares not >65.7% of homologous genes (2009) with other members of the *Roseobacter* clade (Supplementary Tables S2 and S4). It harbors no prophage, which is consistent with the fact that prophage induction using mitomycin C and UV light was unsuccessful (H-A Giebel, unpublished data). Further, the genome carries no plasmid, no complete GTA (gene transfer agent) cluster (Supplementary Table S4) and no CRISPR (clustered regularly interspaced small palindromic repeats). CRISPR are uncommon in the genomes of most organisms of the *Roseobacter* clade and were so far detected in only two members of this clade, *Dinoroseobacter shibae* and *Maritimibacter alkaliphilus* HTCC2654. The lack of plasmids in the genome of *P. temperata* RCA23 and other pelagic members of this clade with a genome size of <3.5 Mb (Supplementary Table S2) appears to be an adaptation to the pelagic life style and exhibits another feature of streamlining not considered yet in other analyses. In contrast, plasmids are common genomic elements of all members of the *Roseobacter* clade associated to other organisms, surfaces and sediments and constitute between 2% and 33% of the genomic content (Supplementary Table S2; Pradella *et al*., 2010).

GTAs have been detected in Alphaproteobacteria and all genomes of the *Roseobacter* clade except in *Rhodobacterales* bacteria HTCC2083 and HTCC2255 (Zhao *et al*., 2009; Newton *et al*., 2010). GTA-related gene transfer was suggested as a potential adaptation mechanism of these bacteria to maintain the metabolic flexibility in the dynamic marine environment (Biers *et al*., 2008). The fact that the genome of *P. temperata* RCA23 encodes only three putative GTA-related genes (c18030, c18040 and c18050; Supplementary Table S4) implies that this mode of gene transfer was discarded during adaptation to the pelagic life style, presumably because of little benefits by this type of genetic exchange in the nutrient-poor pelagic environment. Hence, the mentioned missing genomic features, the absence of plasmids and GTA in the genome of *P. temperata* RCA23 indicate a less diversified genomic organization as compared to other organisms of the *Roseobacter* clade and suggests reduced genetic exchange with other members of this clade.

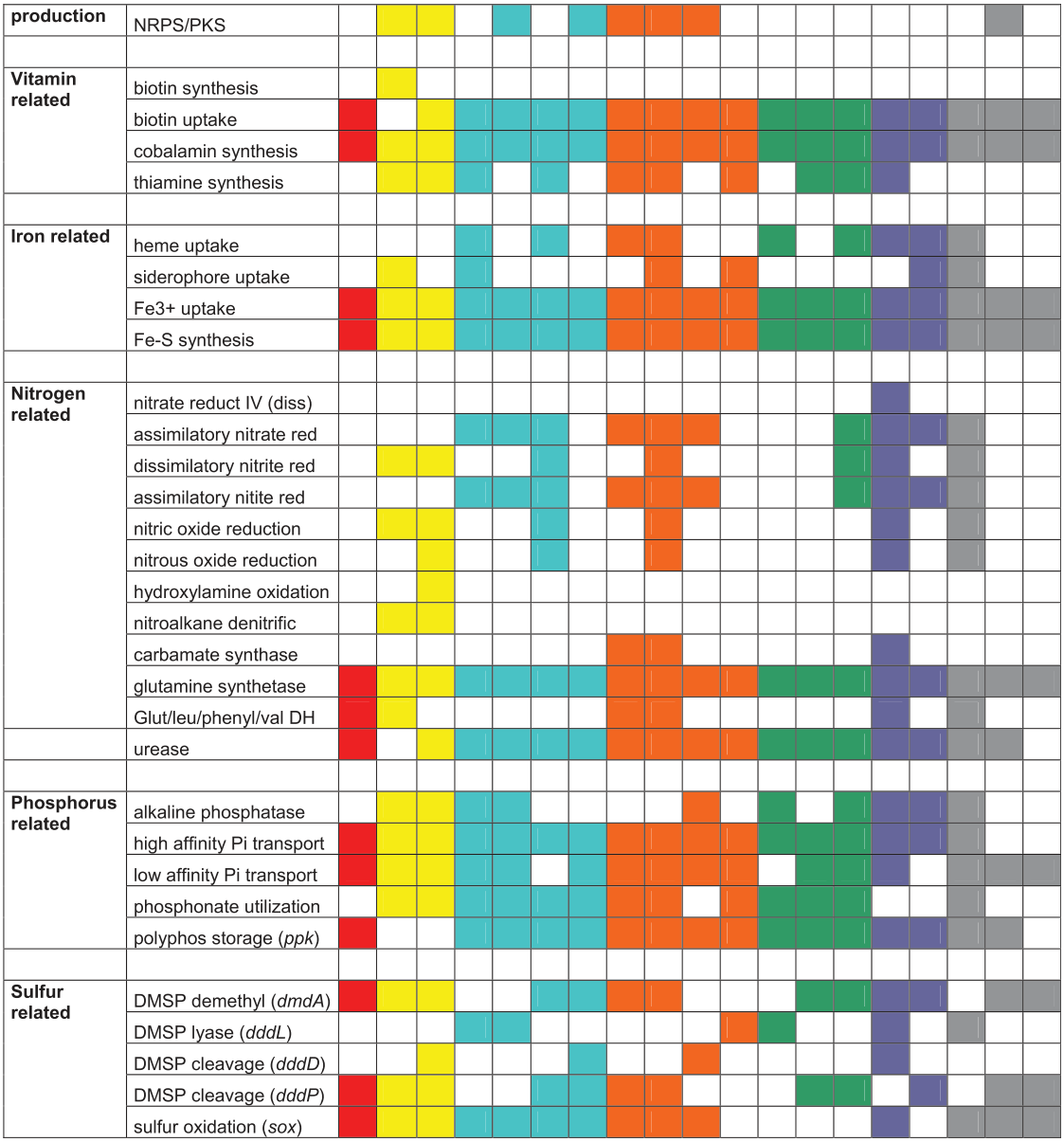
In comparison to typical groups of other *Roseobacter* clade organisms (Newton *et al*., 2010) 10 genomic islands (GIs), harboring 22.6% of all genes, are present in the genome of *P. temperata* RCA23 (Supplementary Figure S1). This bacterium encodes genes for chemotaxis, possesses a monotrichous flagellum and is motile (Table 1; Giebel *et al*., 2013). Phylogenetic analysis of the flagella synthesis cluster present in GI 3 shows that the flagellar gene sets of the *Roseobacter* clade are divided into two distinct groups (Supplementary Figure S2). The flagella genes of the majority of the *Roseobacter* clade organisms fall into group I, but those of *P. temperata* RCA23 into group II. This group contains relatively few sequences of *Roseobacter* clade members but additionally includes strains of the SAR116 clade and of *Rhodobacter sphaeroides*. The existence of these two groups of flagella gene clusters obviously reflects the evolutionary history of the *Roseobacter* clade with substantial lateral gene transfer (Luo *et al*., 2013). GI 5 harbours the genes for the group I CO dehydrogenase (*coxI*; Supplementary Table S4). The existence of the *coxI* gene cluster enables members of the *Roseobacter* clade to oxidize CO (Cunliffe, 2011).

The genome of *P. temperata* RCA23 encodes all basic metabolic functions. Carbohydrates are metabolized via the Entner–Doudoroff pathway that appears typical for the *Roseobacter* clade and also for the SAR11 clade (Giovannoni *et al*., 2005; Fuerch

Table 1 Comparison of selected biosynthetic/catabolic genes and pathways of representative members of the *Roseobacter* clade

Abbreviations: CO DH, carbon monoxide dehydrogenase; DMSP, dimethylsulphonium propionate; NRPS, non-ribosomal peptide synthase; PKS, polyketide synthetase; TDA, tropodithietic acid; TMA, trimethylamine. Colored boxes represent the presence of a gene or pathway within a genome. Data for *P. temperata* RCA23 were generated in this study; the data for representative members of the *Roseobacter* clade are taken from Newton *et al.* (2010) and colored according to the different subclades in Figure 1.

Table 1 (Continued)



et al., 2009). Dimethylsulphonium propionate can be degraded via the demethylation and the cleavage pathway (Table 1 and Supplementary Table S4) and enhances growth of *P. temperata* RCA23 when added additionally to the medium (Giebel *et al*., 2013). This organism is able to carry out AAP (Giebel *et al*., 2013). The organization of the photosynthetic operon, however, differs from that of other members of the *Roseobacter* clade (Supplementary Figure S3 and Supplementary Table S4). In accordance to many other *Roseobacter* clade organisms, nitrogen can only be taken up in reduced form as the genome harbors no genes encoding enzymes for reducing nitrate and nitrite, but gene clusters encoding enzymes for uptake of ammonium, amino acids, putrescine/spermidine and to utilize urea (Table 1 and Supplementary Table S4). These genomic features are in line with growth tests of *P. temperata* RCA23 (Giebel *et al*., 2013). Whether oxidized sulfur compounds can be reduced is unclear because a putative gene cluster for assimilatory sulfate reduction exists (c14410-c14360), but lacks the typical *cysH* gene known from other organisms of the *Roseobacter* clade. Instead a smaller variant harboring a phosphoadenosine phosphosulfate reductase domain was found (c14380).

The genome of *P. temperata* RCA23 encodes 287 transport proteins, a number lower than the mean of

the pelagic members of the *Roseobacter* clade (Supplementary Figure S4A and B; Supplementary Tables S2 and S3). A total of 208 transport proteins of *P. temperata* RCA23 belongs to the ATP-binding cassette (ABC) family (Supplementary Figure S4B and C; Supplementary Table S2). This number is lower than the mean of the pelagic members of the *Roseobacter* clade, whereas the proportion of ABC transporters of *P. temperata* RCA23 (72.5%) is very close to and the number of ABC transporters per Mb of *P. temperata* RCA23 (63.2) identical to the mean of all pelagic members (Supplementary Table S2). Interestingly, the other two deeply branching pelagic members of the *Roseobacter* clade, *Rhodobacterales* bacteria HTCC2255 and HTCC 2150, have the lowest numbers of ABC transporters per Mb of all roseobacters, 47.6 and 50.2, values close to that of Cand. *P. ubique*, 51.1 (Supplementary Table S2). The mean of the absolute numbers of ABC transporters of the pelagic members is significantly lower than that of the associated *Roseobacter* clade members (Student's *t*-test; $P = 0.014$; Supplementary Table S3), presumably a result of the different evolutionary history of both groups of this clade with respect to gene acquisition and loss (Luo *et al.*, 2013). A high proportion of ABC transporters, which exhibit high substrate affinities, was interpreted as an adaptation of pelagic bacteria to the nutrient-poor oceanic environment (Giovannoni *et al.*, 2005; Lauro *et al.*, 2009). As all members of the *Roseobacter* clade exhibit a rather high percentage of ABC transporters, this feature may not be an adaptation to the nutrient-poor oceanic environment. It may also reflect the generally highly dynamic nutrient supply to members of this clade, considered as opportunistic marine bacteria and being able to adapt to a variety of nutrient conditions (Brinkhoff *et al.*, 2008; Teeling *et al.*, 2012; Luo *et al.*, 2013). The number of tripartite ATP-independent periplasmic transporters of *P. temperata* RCA23 is lower than the mean of the pelagic as well as associated members of the *Roseobacter* clade (Figure 3d and Supplementary Table S3). The great majority of tripartite ATP-independent periplasmic transporters of *P. temperata* RCA23 mediates uptake of dicarboxylic acids (Supplementary Table S4). Physiological tests showed that *P. temperata* RCA23 is able to grow on a large variety of organic substrates including amino acids, monosaccharides and short-chain fatty acids (Giebel *et al.*, 2013). This high versatility appears to reflect the high number of transporter proteins.

Even though quite a few genomic features are similar to those of other members of the *Roseobacter* clade, the genome of *P. temperata* RCA23 is distinct in lacking genes encoding the Flp pilus (type IV) for attachment, the VirB system for discharge of genetic material and protein, and genes encoding enzymes for quorum sensing and synthesizing other secondary metabolites (Table 1). The only other organism of the *Roseobacter* clade also missing the genes encoding attachment properties is *Rhodobacterales* bacterium HTCC2255 (Table 1). In *Octadecabacter arcticus* these genes are present but fragmented into three partial clusters and not arranged homologously to the other organisms of this clade (Table 1; Vollmers *et al.*, 2013). The lacking feature for attachment in *P. temperata* RCA23 is consistent with the observation that the RCA cluster has not been detected in the fraction of particle-associated bacteria in the North Sea (Giebel *et al.*, 2011) and that growth on agar plates is weak and unreliable (Giebel *et al.*, 2013). This lacking attachment ability, however, is in contrast to the life style of another RCA cluster isolate, strain LE17, retrieved from Californian coastal waters that lives associated with a dinoflagellate (Mayali *et al.*, 2008).

In order to distinguish between genomic features of copiotrophic and oligotrophic bacteria Lauro *et al.* (2009) compared *Photobacterium angustum* S14 (copiotrophic; genome size 5.10 Mbp), and *Sphingopyxis alaskensis* RB 2256 (oligotrophic; genome size 3.37 Mbp). These authors found that categories of clusters of orthologous groups (COG) for motility (N), transcription (K), defense mechanisms (V) and signal transduction (T) constitute lower and COG categories for lipid transport and metabolism (I) and secondary metabolites, biosynthesis, transport and catabolism (Q) higher fractions in the oligotrophic as compared to the copiotrophic bacterium. The fractions of COG categories K, V and T in the genome of *P. temperata* RCA23 constitute even smaller fractions and COG categories I and Q higher fractions than in *S. alaskensis* (Supplementary Table S5), confirming the finding by Lauro *et al.* (2009) and emphasizing that *P. temperata* RCA23 is a truly oligotrophic bacterium in these respects. Interestingly, COG category I (lipid transport and metabolism) of the pelagic members of the *Roseobacter* clade constitutes a significantly higher fraction relative to the associated members (Supplementary Table S5) and *P. temperata* RCA23 has the third highest fraction of this COG category of all pelagic roseobacters (Supplementary Table S5). Genes affiliated to this COG category include short-chain fatty acid dehydrogenases, hydratases, dehydratases and acetyltransferases (Supplementary Table S4). The COG category of cell motility (N) of *P. temperata* RCA23 has a fraction not as low as that of *S. alaskensis*, suggesting that this feature is of adaptive significance for the life style of *P. temperata* RCA23 in the pelagic environment. The respective means of the associated and pelagic groups of the *Roseobacter* clade do not exhibit such a clear-cut difference between these COG categories and thus may reflect that members of this clade in general are more adapted to nutrient-poor environments than *P. angustum* (Supplementary Table S5).

Occurrence of the RCA cluster and P. temperata *RCA23 in the North Sea*

In order to assess the significance of the RCA cluster and *P. temperata* RCA23 among the active microbial

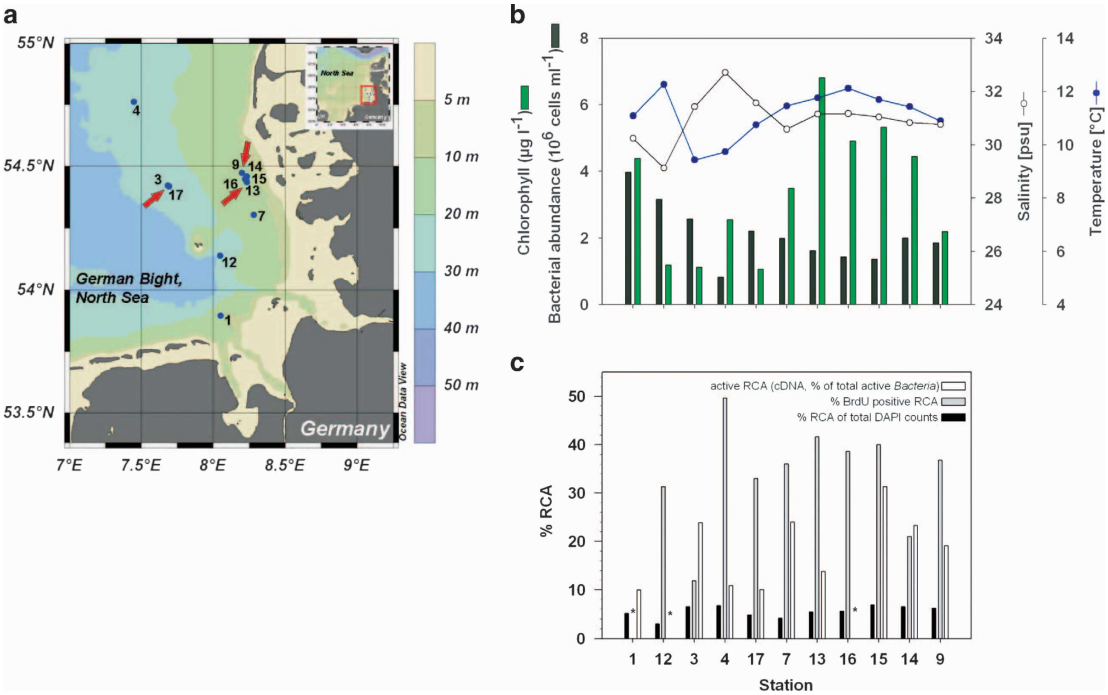


Figure 3 (**a**) Bathymetry of the southern North Sea (German Bight) and sampling stations in May 2010. Arrows indicate stations where the samples for the metagenomic and transcriptomic analyses were collected. (**b**) Chlorophyll *a*, bacterial abundance, salinity and temperature at 2 m depth. (**c**) Percentages of the RCA cluster as detected by CARD-FISH with an RCA-specific probe, of DNA-synthesizing RCA cells (BrdU positive) and of active RCA (cDNA). Samples in (**b**) and (**c**) are ordered as inside and outside the bloom except station 1, which was a separate bloom in the southernmost area. *: Missing data.

players in the southern North Sea, we investigated bacterioplankton abundance, composition and activity in and outside a diatom-dominated phytoplankton spring bloom in the German Bight (Figure 3). The RCA cluster dominated the *Roseobacter* clade by >90% (Wemheuer *et al.*, 2014), constituted 3.0–6.5% (mean = 5.1 ± 1.2%) of total bacterioplankton and the active bacterioplankton community to even higher proportions, to 10–31.3% (mean = 18.5 ± 7.7%), as determined by the cDNA derived from the 16S rRNA (Figure 3c; Wemheuer *et al.*, 2014). Between 11.9% and 49.6% (mean = 34.0 ± 10.7%) of the RCA cells were actively dividing as determined by incorporation of bromodeoxyuridine (BrdU; Figure 3c).

One sample collected outside (station 3) and two samples collected inside the bloom, during day (1100 hours; station 13) and at night (0330 hours; station 9), were subjected to a metagenomic and metatranscriptomic analysis. The genome of *P. temperata* RCA23 was retrieved from the metagenomes to >96% and accounted for 7.8–15.4% of total reads (Table 2; Figure 4). It was also retrieved from the metatranscriptomes, to 17.3% outside and to at least 93.4% inside the bloom (Table 2; Figure 4). Inside the bloom it accounted for 6.7% of total metatranscriptomic reads during night and for 8% during day (Table 2). These data, in line with the other above-mentioned data, show that *P. temperata* RCA23 was an abundant and in the bloom also highly active member of the bacterioplankton in the southern North Sea.

For comparison and to validate our genome mapping approach, we carried out similar genomic mappings in the metagenomes and transcriptomes for organisms of the SAR11 clade and SAR92 cluster, phylogenetic lineages also abundant at these stations (Wemheuer *et al.*, 2014). The genomes of Cand. *P. ubique* HTCC1062 of the SAR11 clade and of the Gammaproteobacterium HTCC2207 of the SAR92 cluster were retrieved to 92.7% and 44.6% in the metagenome outside the bloom, and to 50.2% and 95.6% inside the bloom, with some differences between day and night (Table 2). Both genomes accounted for not >1.7% of total metagenomic reads. The genome of Cand. *P. ubique* HTCC1062 had already been retrieved to 96% from metagenomic data of a North Sea phytoplankton spring bloom, but only from the pooled data of all combined samples (Teeling *et al.*, 2012). We retrieved the genomes of Cand. *P. ubique* HTCC1062 and of Gammaproteobacterium HTCC2207 also from metatranscriptomic reads, in the bloom sample collected during the day to 40.9% and 89.1%, respectively, and at night to 42.6% and 34.1%, respectively. These data show little variations between day and night for the transcriptome of Cand. *P. ubique* HTCC1062, but a much lower fraction of the genome of HTCC2207 transcribed at night. The transcriptomic reads of both organisms

Table 2 Days of sampling, mapped reads, percentage of genomic coverage and percentage of total reads of *P. temperata* RCA23, Cand. *Pelagibacter ubique* HTCC1062 of the SAR11 clade and of Gammaproteobacterium HTCC2207 of the SAR92 cluster of the metagenome (DNA) and metatranscriptome (RNA) retrieved at stations outside (non-bloom) and inside the phytoplankton spring bloom during day (1100 hours) and night (0330 hours)

Station	*Target*	*No. of reads*	*Organism (lineage)*	*Mapped reads*	*Percentage of genome coverage*	*Percentage of total reads*
3						
Non-bloom	DNA	23 381 812	*P. temperata* RCA23	1 819 024	97.1	7.8
26 May 2010			Cand. *P. ubique* HTCC1062 (SAR11)	407 611	92.7	1.7
			HTCC2207 (SAR92)	50 361	44.6	0.2
	RNA	2 223 804	*P. temperata* RCA23	19 435	17.3	0.9
			Cand. *P. ubique* HTCC1062 (SAR11)	16 947	17.9	0.8
			HTCC2207 (SAR92)	4021	2.6	0.2
9						
bloom, night	DNA	12 671 944	*P. temperata* RCA23	1 783 951	96.4	14.1
29 May 2010			Cand. *P. ubique* HTCC1062 (SAR11)	112 728	89.0	0.9
			HTCC2207 (SAR92)	45 926	50.2	0.4
	RNA	8 101 537	*P. temperata* RCA23	543 596	93.4	6.7
			Cand. *P. ubique* HTCC1062 (SAR11)	35 853	42.6	0.4
			HTCC2207 (SAR92)	35 853	34.1	0.3
13						
bloom, day	DNA	16 156 091	*P. temperata* RCA23	2 480 608	96.0	15.4
29 May 2010			Cand. *P. ubique* HTCC1062 (SAR11)	23 806	66.3	0.1
			HTCC2207 (SAR92)	238 054	95.5	1.5
	RNA	15 349 574	*P. temperata* RCA23	1 222 858	94.6	8.0
			Cand. *P. ubique* HTCC1062 (SAR11)	27 026	40.9	0.2
			HTCC2207 (SAR92)	179 429	89.1	1.2
GOS02-13	DNA	1 083 836	*P. temperata* RCA23	12 254	81.9	1.1
Norwegian fjord	DNA	863 687	*P. temperata* RCA23	54 895	83.9	6.4
	RNA	256 537		4176	23.7	1.6

For exact locations of the stations see Figure 3a.

accounted for not >1.2% of total metatranscriptomic reads, indicating that these organisms were far less abundant and active than *P. temperata* RCA23.

The fact that at least 93% of the genome of *P. temperata* RCA23 were transcribed in the bloom is remarkable because it reflects that the population of this abundant organism was highly active and used basically the entire potential of its metabolic properties. The largest fraction of the remaining unmapped regions refers to hypothetical proteins mostly located in GIs (Figure 4). To the best of our knowledge this is the first report of such a high fraction of the genome of an abundant marine bacterium transcribed and retrieved from a metatranscriptome. A similarly high genome coverage in a metatranscriptome analysis was recently found in a sediment-derived microbial community degrading hexadecane under anoxic conditions. The genome of *Smithella* spec., largely dominating this community, was transcribed to 94% (Embree *et al.*, 2014). Transcribed genomic fractions of 52–95% have been reported from pure culture experiments with various bacteria such as *Helicobacter pylori* (Sharma *et al.*, 2010), *Bacillus anthracis* (Passalacqua *et al.*, 2009), *Bacillus subtilis* (Nicolas *et al.*, 2012) and *Prochlorococcus* MED4 (Wang *et al.*, 2014). Hence, the high fraction of the genome of *P. temperata* RCA23 transcribed is in line with these reports but may indicate that the ambient population of this organism consisted of cells in different metabolic stages with respect to growth and substrate utilization.

A more detailed analysis of the transcriptomic patterns of *P. temperata* RCA23 shows that transcripts of COG category E (amino-acid transport and metabolism) exhibited the highest NPKM-normalized reads of all COG categories, followed by COG category R (general function) and J (translation, ribosomal structure and biogenesis) (Table 3, Supplementary Table S6). The high values of normalized transcripts of amino-acid and carbohydrate transport and metabolism (categories E and G) appear to reflect the significance of these substrates for growth of *P. temperata* RCA23, and is consistent with the broad substrate spectrum of this organism (Giebel *et al.*, 2013). Pronounced differences between the day and night transcriptome occurred in COG categories for translation, ribosomal structure and biogenesis (category J) and for cell motility, but not for chemotaxis (both category N), with higher fractions at night than at day (Supplementary Table S6). Further, the genes c29660, c29670 and c29760, encoding the light-harvesting protein B-870 beta chain PufB and alpha chain PufA and cytochrome c551, respectively, were highly overexpressed at night (Supplementary Figure S5 and Supplementary Table S6). This observation is in line with the well-known fact that bacteriochlorophyll *a* is only produced during night in AAP bacteria (Wagner-Döbler and Biebl, 2006).

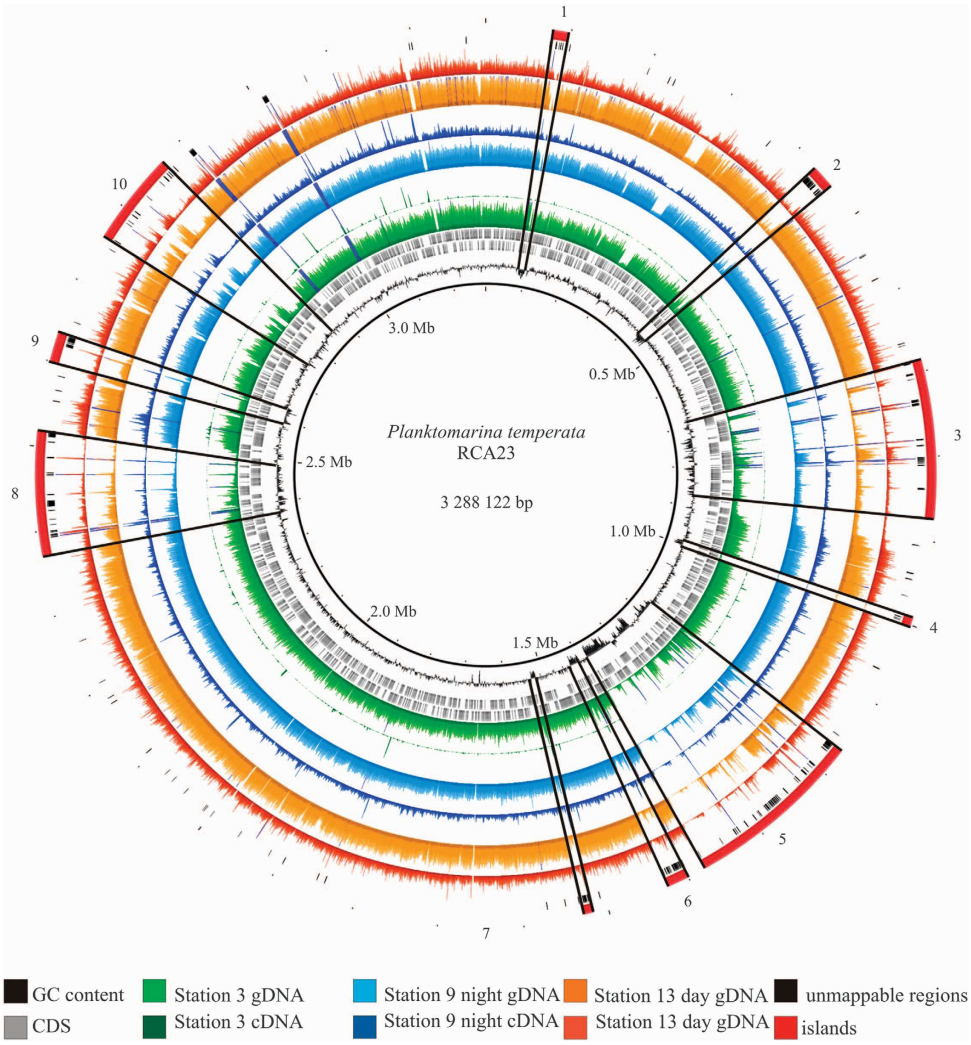


Figure 4 Circular plot of metagenome and transcriptome read mappings onto the genome of *P. temperata* RCA23. Read counts from the Bowtie 2 mappings are shown. From innermost to outer circle: GC content, CDS counts, stations 3, 9 and 13 (DNA and cDNA), unmappable regions and genomic islands.

Genes encoding small heat-shock protein IbpA (c1240, c20690), RNA polymerase σ-32 factors RpoH1 and RpoH2 (c9960, c21310), chaperonin GroEL (c20740), GroS (c20750) and chaperone protein DnaK (c30090) were also highly overexpressed at night. The small heat-shock protein DnaK and chaperonin GroS, GroEL are important in preventing stress-induced aggregation of various proteins in the cytosol of bacteria (Mogk *et al.*, 2003; Stolyar *et al.*, 2007; Ting *et al.*, 2010). Hence, *P. temperata* RCA23 obviously undergoes some stress during night. So far we have no clue about a more specific type of stress that upregulates the transcription of these genes during night.

As a conclusion from this analysis the population of *P. temperata* RCA23 undergoes intense metabolic reconstruction during night including protein synthesis and stress response, but also enhances the synthesis of flagella proteins. This intense metabolic reconstruction may be a result of the different modes of energy conservation of this AAP bacterium during day and night. It has been shown that *Roseobacter litoralis*, another AAP bacterium of the *Roseobacter* clade, changes its proteome drastically, mainly by downregulating protein synthesis, when shifted from dark to light conditions (Zong and Zhao, 2012). Hence the mode of energy conservation appears to have a pronounced effect on the global regulation of metabolic networks in AAP bacteria of the *Roseobacter* clade, and presumably beyond, with implications of their diurnal participation in turnover of organic matter in marine ecosystems.

Global and biogeochemical significance of the RCA cluster and P. temperata *RCA23*

Because of the known global distribution of the RCA cluster (Selje *et al.*, 2004) we examined publicly available metagenomic data sets for the presence of genomic features of *P. temperata* RCA23. The photosynthetic operon was present in an

Table 3 NPKM-normalized reads in the COG categories with type of metabolism of the genome of *P. temperata* RCA23 retrieved in the metatranscriptome at stations 9 and 13 in the German Bight of the North Sea

COG category	*Type of metabolism*	*Station 9 (night) NPKM*	*Station 13 (day) NPKM*
B	Chromatin structure and dynamics	661	954
C	Energy production and conversion	61 408	57 017
D	Cell cycle control, cell division and chromosome partitioning	7190	7533
E	Amino acid transport and metabolism	101 802	107 628
F	Nucleotide transport and metabolism	24 455	26 003
G	Carbohydrate transport and metabolism	59 872	59 448
H	Coenzyme transport and metabolism	36 820	42 486
I	Lipid transport and metabolism	45 455	49 575
J	Translation, ribosomal structure and biogenesis	71 558	59 835
K	Transcription	44 710	41 837
L	Replication, recombination and repair	36 761	39159
M	Cell wall/membrane/envelope biogenesis	37 299	39 404
N	Cell motility	8425	4975
O	Post-translational modification, protein turnover and chaperones	50 598	45 863
P	Inorganic ion transport and metabolism	32 016	32 078
Q	Secondary metabolites biosynthesis, transport and catabolism	16 285	16 089
R	General function prediction only	84 440	90 432
S	Function unknown	68 057	68 212
T	Signal transduction mechanisms	20 797	19 875
U	Intracellular trafficking, secretion and vesicular transport	7431	7254
V	Defense mechanisms	8258	8799

Abbreviations: COG, clusters of orthologous groups; NPKM, Nucleotide activity Per Kilobase of exon model per million Mapped reads.

environmental bacterial artificial chromosome clone derived from Californian coastal waters (Supplementary Figure S3; Béjà *et al*., 2002; Yutin and Béjà, 2005) and in metagenomic data sets from a Norwegian fjord, the western English Channel, the western coastal Atlantic visited during the Global Ocean Sampling expedition (GOS), and in the eastern (Monterey Bay) and southwestern Pacific in Australian waters (Rusch *et al*., 2007; Yutin *et al*., 2007; Gilbert *et al*., 2008, 2010; Thomas *et al*., 2010; Rich *et al*., 2011).

The genome of *P. temperata* RCA23 was retrieved from metagenomic and metatranscriptomic data of a Norwegian fjord to 83.9% and 23.7%, respectively (Gilbert *et al*., 2008; Table 2). At the GOS stations, the genome of *P. temperata* RCA23 was retrieved from the metagenomic data to 81.9% (Table 2) and accounted for 0.7% to 6.5% of the mapped reads (Supplementary Figure S6B). This finding, together with the observation that the structure of the photosynthetic operon of the AAP bacteria at these stations was similar to that of *P. temperata* RCA23 (see above), is consistent with the assumption that this organism constituted largely the AAP communities at these stations. In the GOS and the Norwegian fjord data sets, 80% and 81%, respectively, of the genes of *P. temperata* RCA23 showed a coverage of $\geqslant$80%. In contrast, 62.5% of the genes of the GIs were not mapped or showed a low coverage (Supplementary Table S7).

Conclusion

This study, on the basis of the sequenced genome of *P. temperata* RCA23 and the North Sea metagenome and metatranscriptome, sheds new light on the significance of the RCA cluster for biogeochemical processes in marine pelagic systems, considering the high abundances of this organism and the entire RCA cluster in temperate to polar oceans (Selje *et al*., 2004; West *et al*., 2008; Giebel *et al*., 2009, 2011; Wemheuer *et al*., 2014). It appears to carry out a life style well adapted to the nutrient- and energy-poor marine pelagic realm. The entire population of this organism, and possibly even single cells, simultaneously transcribe almost the entire genome during phytoplankton bloom situations and convert complementary energy by harvesting light. This versatile physiology together with the wealth of high-affinity transporter proteins, despite the absence of distinct genomic and metabolic traits typical for the *Roseobacter* clade, and a streamlined genome relative to other *Roseobacter* clade organisms, helps to explain the success of this abundant taxon in temperate and presumably also (sub)polar marine systems. Our findings illustrate how important it is to dissect the role of individual taxa for better understanding their participation in processing organic and inorganic matter in pelagic marine systems.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Renate Gahl-Janssen for growing *P. temperata* RCA23 and harvesting biomass for DNA extraction, Melissa Büngener for technical assistance in sequence analyses, the crew of RV Heincke and M Wolterink for

helping with sample collection and processing during the North Sea cruise and the prophage induction experiment. This work was supported by grants from the Volkswagen Foundation in Lower Saxony, Germany, and by TRR51 of Deutsche Forschungsgemeinschaft.

References

Béjà O, Suzuki MT, Heidelberg JF, Nelson WC, Preston CM, Hamada T *et al.* (2002). Unsuspected diversity among marine aerobic anoxygenic phototrophs. *Nature* **415**: 630–633.

Biers EJ, Wang K, Pennington C, Belas R, Chen F, Moran MA. (2008). Occurrence and expression of gene transfer agent (GTA) genes in marine bacterioplankton. *Appl Environ Microbiol* **74**: 2933–2939.

Bolger AM, Lohse M, Usadel B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics;* e-pub ahead of print 28 April 2014; doi:10.1093/bioinformatics/btu170.

Brinkhoff T, Giebel HA, Simon M. (2008). Diversity, ecology, and genomics of the Roseobacter clade: a short overview. *Arch Microbiol* **189**: 531–539.

Cunliffe M. (2011). Correlating carbon monoxide oxidation with cox genes in the abundant Marine Roseobacter Clade. *ISME J* **5**: 685–691.

Delcher AL, Bratke KA, Powers EC, Salzberg SL. (2007). Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**: 673–679.

Embree M, Nagarajan H, Movahedi N, Chitsaz H, Zengler K. (2014). Single-cell genome and metatranscriptome sequencing reveal metabolic interactions of an alkane-degrading methanogenic community. *ISME J.* **8**: 757–767.

Fuerch T, Preusse M, Tomasch J, Zech H, Wagner-Döbler I, Rabus R *et al.* (2009). Metabolic fluxes in the central carbon metabolism of Dinoroseobacter shibae and Phaeobacter gallaeciensis, two members of the marine Roseobacter clade. *BMC Microbiol* **9**: 209.

Giebel HA, Brinkhoff T, Zwisler W, Selje N, Simon M. (2009). Distribution of Roseobacter RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environ Microbiol* **11**: 2164–2178.

Giebel HA, Kalhoefer D, Lemke A, Thole S, Gahl-Janssen R, Simon M *et al.* (2011). Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J* **5**: 8–19.

Giebel HA, Kalhoefer D, Gahl-Janssen R, Choo YJ, Lee K, Cho JC *et al.* (2013). Planktomarina temperata gen. nov., sp. nov., belonging to the globally distributed RCA cluster of the marine Roseobacter clade, isolated from the German Wadden Sea. *Int J Syst Evol Microbiol* **63**: 4207–4217.

Gilbert JA, Field D, Huang Y, Edwards R, Weizhong L, Gilna R *et al.* (2008). Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. *PLoS One* **3**: e3042.

Gilbert JA, Field D, Swift P, Thomas S, Cummings D, Temperton B *et al.* (2010). The taxonomic and functional diversity of microbes at a temperate coastal site: a 'multi-omic' study of seasonal and diel temporal variation. *PLoS One* **5**: e15545.

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D *et al.* (2005). Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242–1245.

Giovannoni SJ, Stingl U. (2007). The importance of culturing bacterioplankton in the 'omics' age. *Nat Rev Microbiol* **5**: 820–826.

Giovannoni SJ, Stingl U. (2005). Molecular diversity and ecology of microbial plankton. *Nature* **437**: 343–348.

Giovannoni SJ, Cameron Thrash J, Temperton B. (2014). Implications of streamlining theory for microbial ecology. *ISME J;* e-pub ahead of print 17 April 2014; doi:10.1038/ismej.2014.60.

Kopylova E, Noé L, Touzet H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.

Langmead B, Salzberg S. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.

Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S *et al.* (2009). The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci USA* **106**: 15527–15533.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

Luo H, Csu M, Hughes A, Moran MA. (2013). Evolution of divergent life history strategies in marine Alphaproteobacteria. *mbio* **4**: e00373-13.

Luo H, Swan BK, Stepanauskas R, Hughes A, Moran MA. (2014). Evolutionary analysis of a streamlined lineage of surface ocean Roseobacters. *ISME J* **8**: 1428–1439.

Mayali X, Franks PS, Azam F. (2008). Cultivation and ecosystem role of a marine Roseobacter clade-affiliated cluster bacterium. *Appl Environ Microbiol* **74**: 2595–2603.

Mogk A, Deuerling E, Vorderwülbecke S, Vierling E, Bukau B. (2003). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol Microbiol* **50**: 585–595.

Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE *et al.* (2010). Genome characteristics of a generalist marine bacterial lineage. *ISME J* **4**: 784–798.

Nicolas P, Mäder U, Dervyn E, Rochat T, Léduc A *et al.* (2012). Condition-dependent transcriptome reveals high-level regulatory architecture in Bacillus subtilis. *Science* **335**: 1103–1106.

Passalacqua KD, Varadarajan A, Ondov BD, Okou DT, Zwick ME, Bergman NH. (2009). Structure and complexity of a bacterial transcriptome. *J Bacteriol* **191**: 3203–3211.

Pernthaler A, Pernthaler J, Schattenhofer M, Amann R. (2002). Identification of DNA-synthesizing bacterial cells in coastal North Sea plankton. *Appl Environ Microbiol* **68**: 5728–5736.

Pradella S, Päuker O, Petersen J. (2010). Genome organization of the marine Roseobacter clade member Marinovum algicola. *Arch Microbiol* **192**: 115–126.

Rich VI, Pham VD, Eppley JM, Shi Y, DeLong EF. (2011). Time-series analyses of Monterey Bay coastal microbial picoplankton using a 'genome proxy' microarray. *Environ Microbiol* **13**: 116–134.

Rusch DB, Halpern A, Sutton G, Heidelberg KB, Williamson S, Yooseph S *et al.* (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: e77.

Selje N, Simon M, Brinkhoff T. (2004). A newly discovered Roseobacter cluster in temperate and polar oceans. *Nature* **427**: 445–448.

Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A *et al.* (2010). The primary transcriptome of the major human pathogen Helicobacter pylori. *Nature* **460**: 250–254.

Sowell SM, Wilhelm LJ, Norbeck AD, Lipton LS, Nicora CD, Barofsky DF *et al.* (2009). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **3**: 93–105.

Staden R. (1996). The Staden sequence analysis package. *Mol Biotech* **5**: 233–241.

Stolyar S, He Q, Joachimiak MP, He ZL, Yang ZK, Borglin SE *et al.* (2007). Response of Desulfovibrio vulgaris to alkaline stress. *J Bacteriol* **189**: 8944–8952.

Swan BK, Tupper B, Sczyrba A, Lauro FM, Martinez-Garcia M, González JM *et al.* (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc Natl Acad Sci USA* **110**: 11463–11468.

Tech M, Merkl R. (2003). YACOP: Enhanced gene prediction obtained by a combination of existing methods. *In Silico Biology* **3**: 441–451.

Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM *et al.* (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608–611.

Thomas T, Rusch D, DeMaere MZ, Yung PY, Lewis M, Halpern A *et al.* (2010). Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *ISME J* **4**: 1557–1567.

Thompson AW, Huang K, Saito MA, Chisholm SW. (2011). Transcriptome response of high- and low-light-adapted Prochlorococcus strains to changing iron availability. *ISME J* **5**: 1580–1594.

Ting L, Williams TJ, Cowley MJ, Lauro FM, Guilhaus M, Raftery MJ *et al.* (2010). Cold adaptation in the marine bacterium, Sphingopyxis alaskensis, assessed using quantitative proteomics. *Environ Microbiol* **12**: 2658–2676.

Tripp HJ, Kitner JB, Schwalbach MS, Dacey JWH, Wilhelm LJ, Giovannoni SJ. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **452**: 741–744.

Vollmers J, Voget S, Dietrich S, Gollnow K, Smits M, Meyer K *et al.* (2013). Poles apart: extreme genome plasticity and a new xanthorhodopsin-like gene family in the genomes of Octadecabacter arcticus 238 and Octadecabacter antarcticus 307. *PLoS One* **8**: e63422.

Wagner-Döbler I, Biebl H. (2006). Environmental biology of the marine Roseobacter lineage. *Ann Rev Microbiol* **60**: 255–280.

Wang B, Lu L, Lv H, Jiang H, Qu G, Tian C *et al.* (2014). The transcriptome landscape of Prochlorococcus MED4 and the factors for stabilizing the core genome. *BMC Microbiol* **14**: 11.

Wemheuer B, Güllert S, Billerbeck S, Giebel H-A, Voget S, Simon M *et al.* (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol Ecol* **87**: 378–389.

West NJ, Obernosterer I, Zemb Lebaron P. (2008). Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. *Environ Microbiol* **10**: 738–756.

Wiegand S, Dietrich S, Hertel R, Bongaerts J, Evers S, Volland S *et al.* (2013). RNA-Seq of Bacillus licheniformis: active regulatory RNA features expressed within a productive fermentation. *BMC Genomics* **14**: 667.

Yooseph S, Nealson KH, Rusch DB, McCrown JP, Dupont CL, Kim M *et al.* (2010). Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* **468**: 60–67.

Yutin N, Béjà O. (2005). Putative novel photosynthetic reaction centre organizations in marine aerobic anoxygenic photosynthetic bacteria: insights from metagenomics and environmental genomics. *Environ Microbiol* **7**: 2027–2033.

Yutin N, Suzuki MT, Teeling H, Weber M, Venter JC, Rusch DB *et al.* (2007). Assessing diversity and biogeography of aerobic anoxygenic phototrophic bacteria in surface waters of the Atlantic and Pacific Oceans using the Global Ocean Sampling expedition metagenomes. *Environ Microbiol* **9**: 1464–1475.

Zhao Y, Wang K, Budinoff C, Buchan A, Lang A, Jiao N *et al.* (2009). Gene transfer agent (GTA) genes reveal diverse and dynamic Roseobacter and Rhodobacter populations in the Chesapeake Bay. *ISME J* **3**: 364–373.

Zong R, Zhao N. (2012). Proteomic responses of Roseobacter litoralis OCh149 to starvation and light regimen. *Microbes Environ* **27**: 430–442.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)

Supplementary Information to:

Adaptation of an abundant *Roseobacter* RCA organism to pelagic systems revealed by genomic and transcriptomic analyses

Sonja Voget [1], Bernd Wemheuer [1], Thorsten Brinkhoff [2], John Vollmers [1], Sascha Dietrich[1], Helge-A. Giebel [2], Christine Beardsley [2], Carla Sardemann [2], Insa Bakenhus [2], Sara Billerbeck [2], Rolf Daniel [1], Meinhard Simon [2] *

[1] Institute of Microbiology & Genetics, Genomic & Applied Microbiology and Göttingen Genomics Laboratory, University of Göttingen, D-37077 Göttingen, Germany.
[2] Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, D-26111 Oldenburg, Germany.

* corresponding author. m.simon@icbm.de

Supplementary information includes 4 Figures with the circular representation of the genome of *Planktomarina temperata* RCA23 (S1), the flagellar genes (S2), the photosynthesis gene cluster of *P. temperata* RCA23 and other organisms affiliated to the *Roseobacter* clade (S3), expression patterns of the photosynthesis gene cluster of *P. temperata* RCA23 during day and night during a phytoplankton spring bloom in the North Sea (S4) and 7 Tables with genomic and metagenomic data on this organism and others of the *Roseobacter* clade. For Tables S4, S6 and S7 see extra files.

Figure legends

Figure S1: Circular representation of the genome of *Planktomarina temperata* RCA23. From the outer to the innermost circle: 1. Genomic islands (GIs, red). 2 and 3: Leading and lagging strand colored according to the assigned Clusters of Orthologous Groups (COG) categories. 4: transposases/integrases (red), rRNA genes (purple). 5: tRNAs. 6-22: Homologs to genes in 17 organisms of the *Roseobacter* clade as listed in Table S3, colored according to subclades like in Fig. 1. ORFs of *P. temperata* were determined by reciprocal BLAST-analysis with global alignments (see Methods) against the proteome datasets of the 17 organisms. Query organisms: *Rhodobacterales bacterium* HTCC2150, *Maritimibacter alkaliphilus* HTCC2654, *Jannaschia* sp. CCS1, *Dinoroseobacter shibae* DFL12, *Loktanella vestfoldensis* SKA53, *Octadecabacter antarcticus* 307, *Octadecabacter arcticus* 238, *Roseovarius* sp. TM1035, *Roseovarius* sp. 217, *Sagittula stellata* E-37, *Oceanicola batsensis* HTCC2597, *Sulfitobacter* NAS-14.1, *Sulfitobacter* sp. GAI101, *Roseobacter litoralis* Och149, *Rhodobacterales bacterium* HTCC2083, *Phaeobacter inhibens* DSM17395, *Ruegeria pomeroyi* DSS-3. 23: G+C-content of the chromosome of *P. temperata* with violet areas below average and olive areas above average.

Figure S2 Neighbor-joining phylogeny of flagellar gene sets in *Roseobacter* clade members (marked in red) and related gene sets in other *Alpha-* and *Gammaproteobacteria* based on concatenated protein sequences. The subtree containing flagellar protein sequences of Gammaproteobacteria was collapsed. Bootstrap values are given for major nodes. The resulting phylogenetic tree shows that the flagellar gene sets of the *Roseobacter* clade are divided into two distinct groups (group I and II) which are more closely related to gene sets of Gammaproteobacteria than to each other. The flagellar genes of *P. temperata* (marked in bold) fall into group II which contains relatively few sequences of *Roseobacter* clade members but includes strains of the SAR116 clade (Cand. *Puniceispirillum marinum* IMCC1322, Alphaproteobacterium HIMB 100) and of *Rhodobacter sphaeroides*. In contrast, the majority of flagella gene sets in the *Roseobacter* clade belong to group I.
For the phylogenetic relationship of the flagellar gene sets, genes belonging to the flagellar cluster were identified in the genomes of 38 *Roseobacter* clade members based on existing annotations and on reciprocal BLASTp analyses. Related gene sets from other bacterial taxa (displayed in black) were obtained by BLASTp analyses of the NCBI nt database using multiple flagellar proteins as query. When multiple flagellar gene sets were present in the same genome, numbers were assigned to the individual sets (indicated by square parentheses). For each gene set, alignments of 15 shared flagellar proteins (FlgA, FlgB, FlgC, FlgD, FlgF, FlgG, FlgH, FlgI, FlhA, FlhB, FliE, FliI, FliL, FliQ, and FliR) were produced using clustalW v1.83. The alignments were concatenated and imported into ARB v5.1 where a filter was created to remove gapped sites and the remaining 3,673 alignment positions per concatenated sequence were subjected to neighbor-joining analysis.

Figure S3 Photosynthesis gene cluster (PGC) of *P. temperata* RCA23, a BAC clone and other organisms affiliated to the *Roseobacter* clade. Shown is the different organization of the PGC of members of the *Roseobacter* clade and environmental clone BAC-60D04 (NCBI Accession AE008921) derived from Californian coastal waters (Béjà *et al*., 2002). Homologous regions are indicated by blue shaded vertical areas and black lines, respectively. Re-arrangements are marked by red shaded areas. Genes are colored according to biological categories: green,

bacteriochlorophyll biosynthesis (bch); orange, carotenoid biosynthesis (crt); red, light harvesting and photosynthesis reaction center (puf); blue, cytochrome c2; grey, accessory genes.

Figure S4 Transporter proteins of *P. temperata* RCA23 in comparison to those of other organisms of the *Roseobacter* clade. Relation between total numbers of transporters and ABC transporters (A). Box-Whisker plot of ABC transporters (B), ABC-/total transporters (C) and TRAP transporters (D) of 14 members of the *Roseobacter* clade associated to other organisms, surfaces or sediment and of 24 members with a pelagic life style listed in supplementary Table S2. Data include 12 closed and 26 draft genomes. Blue circle and +: *P. temperata* RCA23. The boxes show the median (solid line), mean (dashed line), the 25^{th} and 75^{th} percentile and the whiskers the 10^{th} and 90^{th} percentiles. •: outlayers. For further details see supplementary Tables S2 and S3.

Figure S5 Expression patterns of the photosynthesis gene cluster of *P. temperata* RCA23 at station 13 during day and station 9 at night during a phytoplankton spring bloom in the southern North Sea in May 2010. Genes are colored according to biological categories: green, bacteriochlorophyll biosynthesis (bch); orange, carotenoid biosynthesis (crt); red, light harvesting and photosynthesis reaction center (puf).

Figure S6 Representation of the genome of *Planktomarina temperata* RCA23 in metagenomic data sets of the north-western Atlantic Ocean. (A) Overview of the stations visited by the Global Ocean Sampling (GOS) expedition in the western coastal Atlantic Ocean (Rusch *et al*., 2007). (B) Salinity, temperature, chlorophyll and mapped reads (bar) of the genome of *P. temperata* RCA23 at the stations shown above.

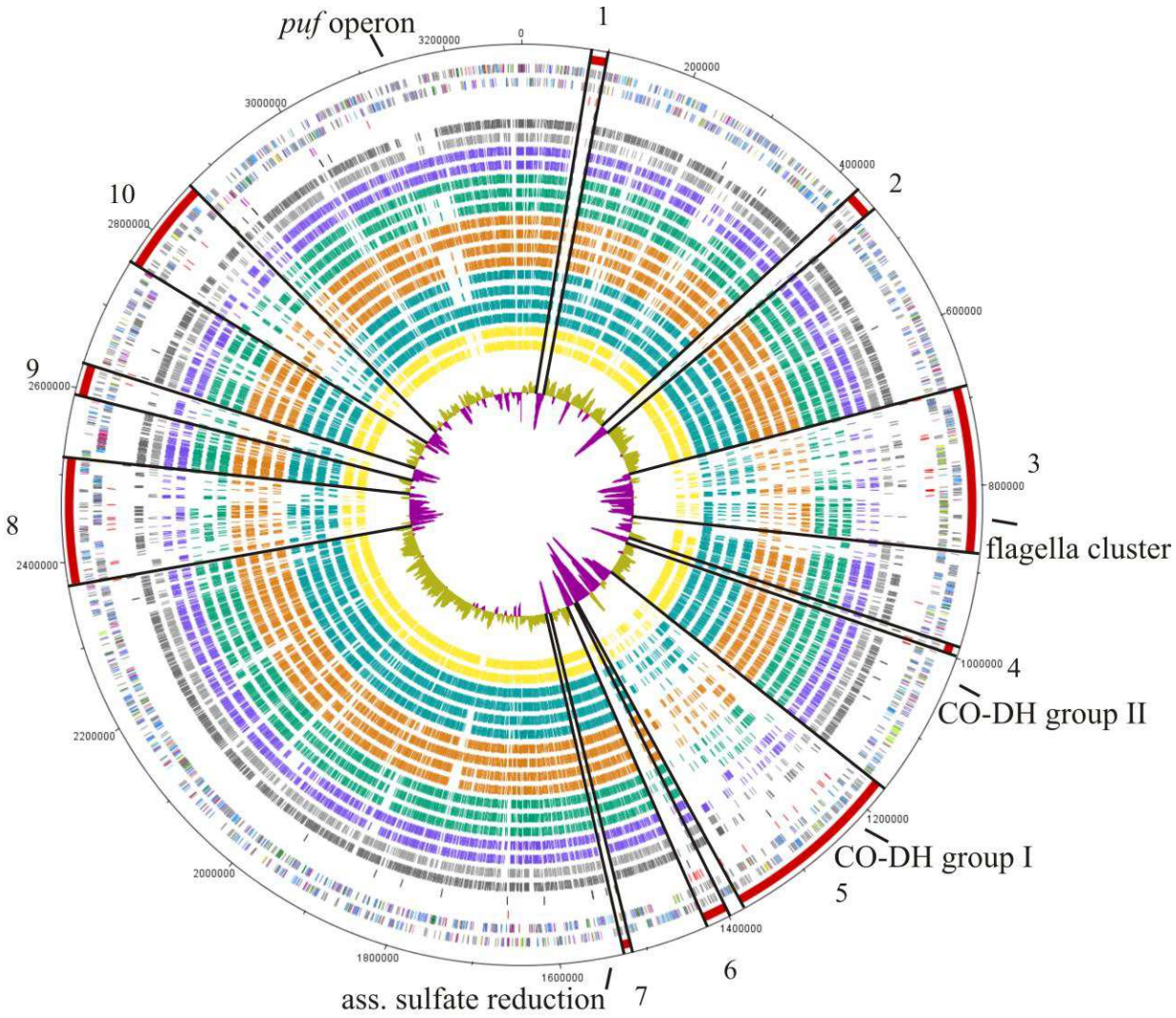

puf operon
1
2
3
flagella cluster
4
CO-DH group II
CO-DH group I
5
6
7
ass. sulfate reduction
8
9
10
0
200000
400000
600000
800000
1000000
1200000
1400000
1600000
1800000
2000000
2200000
2400000
2600000
2800000
3000000
3200000

Figure S1.

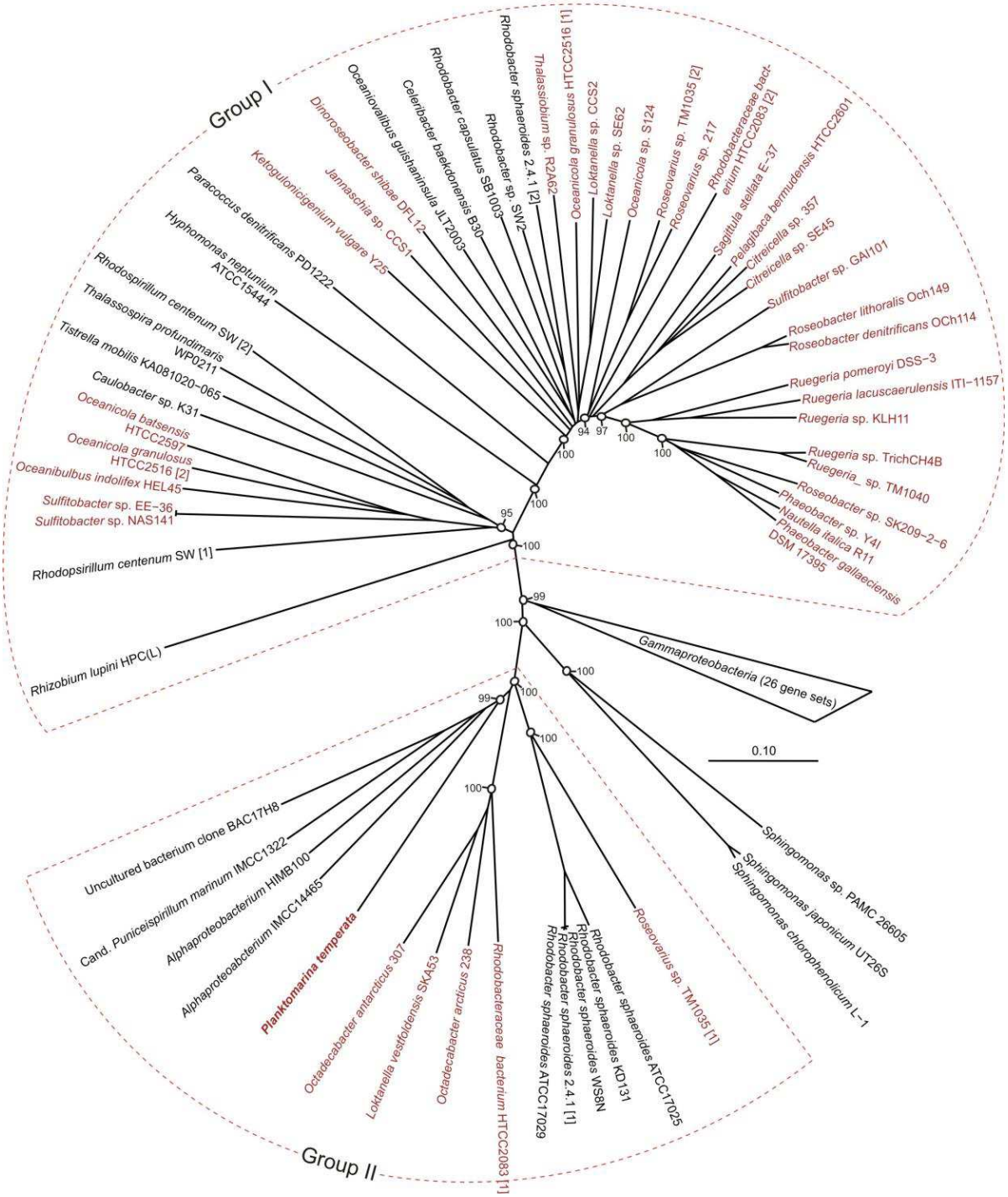

Group I
Group II
Rhodobacter sphaeroides 2.4.1 [2]
Thalassiobium sp. R2A62
Oceanicola granulosus HTCC2516 [1]
Loktanella sp. CCS2
Loktanella sp. SE62
Oceanicola sp. S124
Roseovarius sp. TM1035 [2]
Roseovarius sp. 217
Rhodobacteraceae bacterium HTCC2083 [2]
Sagittula stellata E-37
Pelagibaca bermudensis HTCC2601
Citreicella sp. 357
Citreicella sp. SE45
Sulfitobacter sp. GAI101
Roseobacter litholaralis Och149
Roseobacter denitrificans OCh114
Ruegeria pomeroyi DSS-3
Ruegeria lacuscaerulensis ITI-1157
Ruegeria sp. KLH11
Ruegeria sp. TrichCH4B
Ruegeria_ sp. TM1040
Roseobacter sp. SK209-2-6
Phaeobacter sp. Y4I
Nautella italica R11
Phaeobacter gallaeciensis DSM 17395
Rhodobacter capsulatus SB1003
Rhodobacter sp. SW2
Oceaniovalibus guishaninsula JLT2003
Celeribacter baekdonensis B30
Dinoroseobacter shibae DFL12
Jannaschia sp. CCS1
Ketogulonicigenium vulgare Y25
Paracoccus denitrificans PD1222
Hyphomonas neptunium ATCC15444
Rhodospirillum centenum SW [2]
Thalassospira profundimaris WP0211
Tistrella mobilis KA081020-065
Caulobacter sp. K31
Oceanicola batsensis HTCC2597
Oceanicola granulosus HTCC2516 [2]
Oceanibulbus indolifex HEL45
Sulfitobacter sp. EE-36
Sulfitobacter sp. NAS141
Rhodopsirillum centenum SW [1]
Rhizobium lupini HPC(L)
Gammaproteobacteria (26 gene sets)
0.10
Sphingomonas sp. PAMC 26605
Sphingomonas japonicum UT26S
Sphingomonas chlorophenolicum L-1
Rhodobacter sphaeroides ATCC17025
Roseovarius sp. TM1035 [1]
Rhodobacter sphaeroides KD131
Rhodobacter sphaeroides WS8N
Rhodobacter sphaeroides 2.4.1 [1]
Rhodobacter sphaeroides ATCC17029
Rhodobacteraceae bacterium HTCC2083 [1]
Octadecabacter arcticus 238
Loktanella vestfoldensis SKA53
Octadecabacter antarcticus 307
Planktomarina temperata
Alphaproteobacterium IMCC14465
Alphaproteobacterium HIMB100
Cand. Puniceispirillum marinum IMCC1322
Uncultured bacterium clone BAC17H8
94
97
100
95
99

Figure S2

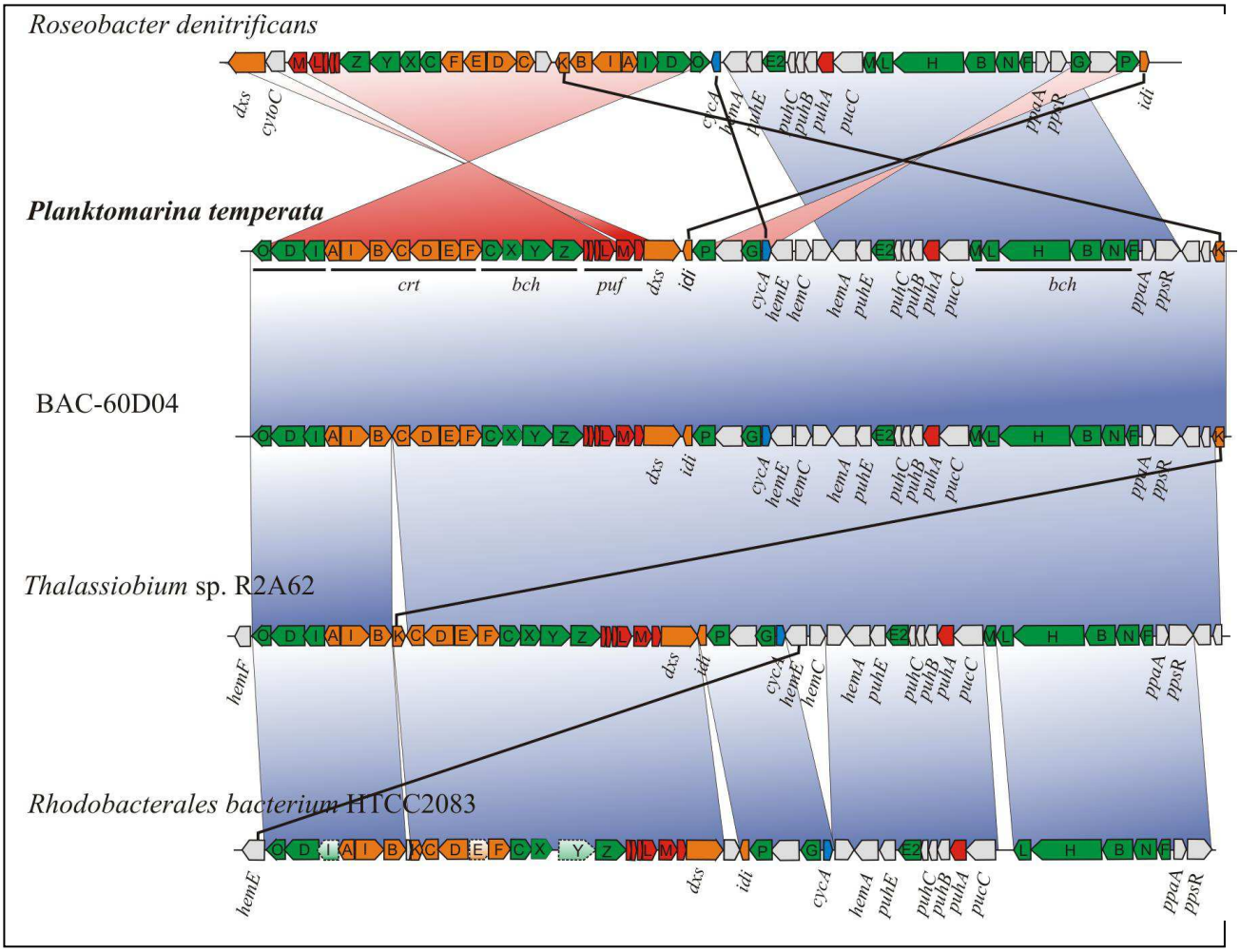
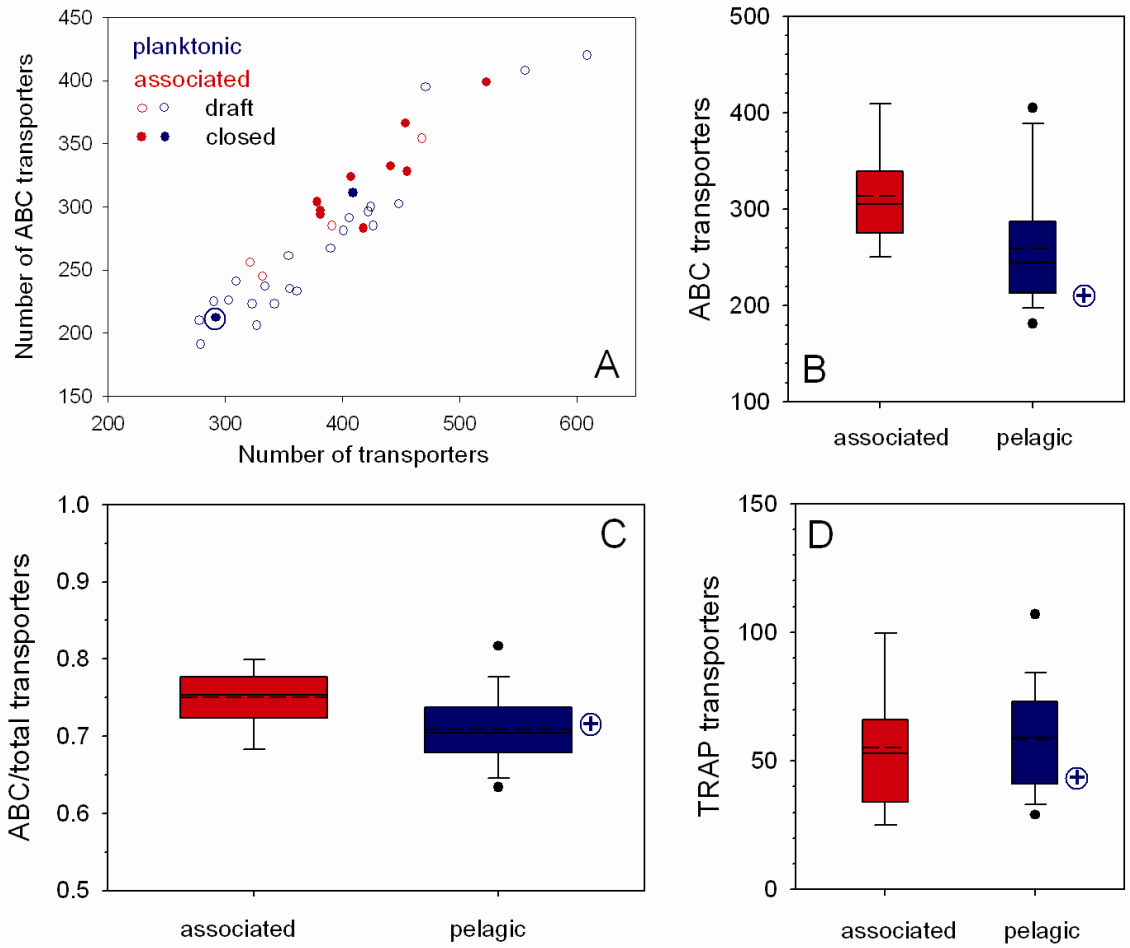

Roseobacter denitrificans
Planktomarina temperata
BAC-60D04
Thalassiobium sp. R2A62
Rhodobacterales bacterium HTCC2083
crt
bch
puf
bch

Figure S3


planktonic
associated
draft
closed
Number of ABC transporters
Number of transporters
A
ABC transporters
B
associated
pelagic
ABC/total transporters
C
associated
pelagic
TRAP transporters
D
associated
pelagic

Fig

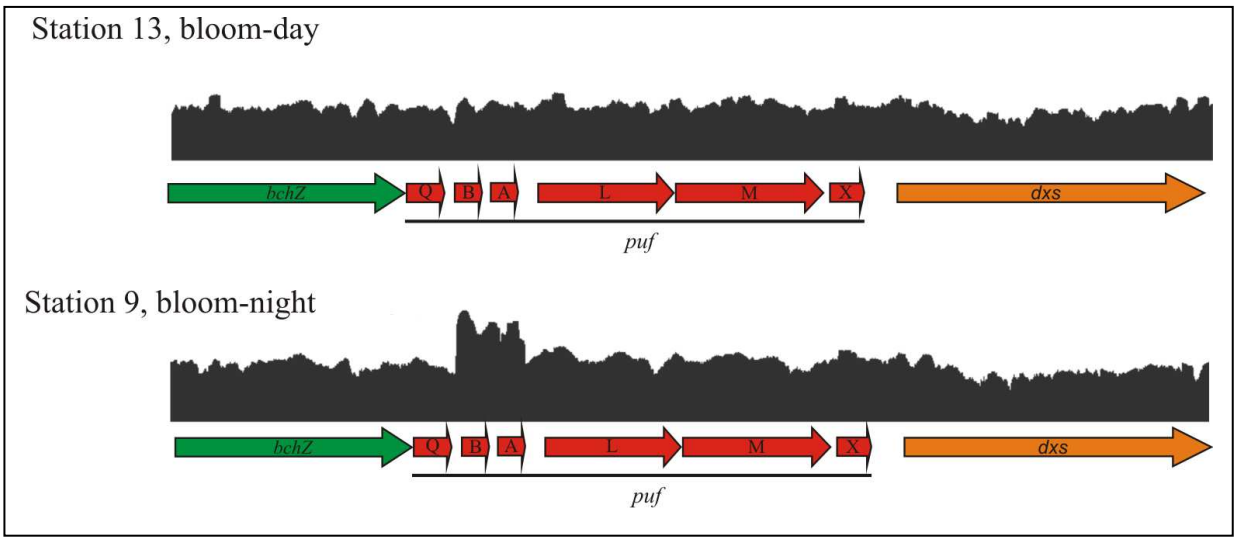


Figure S5 Voget et al

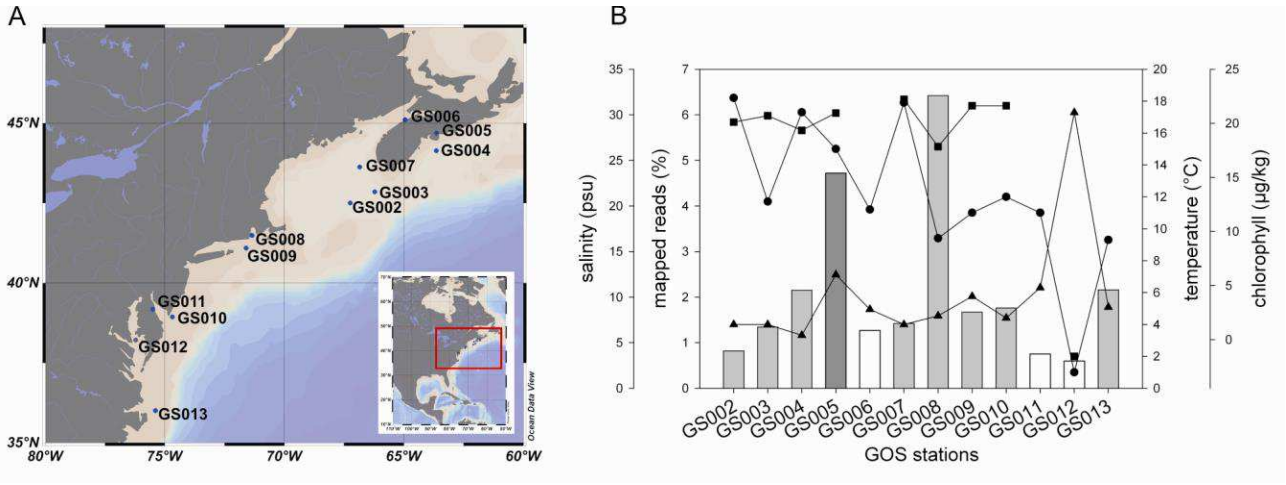

A
45°N
40°N
35°N
80°W
75°W
70°W
65°W
60°W
GS006
GS005
GS004
GS007
GS003
GS002
GS008
GS009
GS011
GS010
GS012
GS013
Ocean Data View
B
salinity (psu)
mapped reads (%)
temperature (°C)
chlorophyll (µg/kg)
GS002
GS003
GS004
GS005
GS006
GS007
GS008
GS009
GS010
GS011
GS012
GS013
GOS stations

Figure S6

Table S1. Overview of metagenomic and metatranscriptomic datasets used in this study.

Data sets used in this study	type	Sequencing technology	Raw-reads	No. of reads (trimmed, rRNA free)	Reference
Station 3, non-bloom	DNA	Illumina	24 331 052	23 381 812	This study
	RNA	Illumina	24 879 579	2 223 804	
Station 9, bloom-night	DNA	Illumina	13 125 978	12 671 944	This study
	RNA	Illumina	26 176 832	8 101 537	
Station 13, bloom-day	DNA	Illumina	16 877 252	16 156 091	This study
	RNA	Illumina	26 985 711	15 349 574	
northwestern Atlantic (stations GOS02-13)	DNA	Sanger	1 083 836	-	Rusch *et al*.
Norwegian fjord	DNA	454	863 687	-	Gilbert *et al*
	RNA	454	256 537	-	

Rusch DB, *et al.* (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. PLoS Biology 5:e77.

Gilbert JA, *et al.* (2008). Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. PLoS One 3:e3042.

Table S2. Overview of genome characteristics and general genome comparisons of organisms of the *Roseobacter* clade isolated from associations (assoc) with various organisms, surfaces or sediment and from pelagic environments and ordered according to genome size. For comparison respective data of Cand. *Pelagibacter ubique* HTCC1062 are also given. The number of plasmids was derived from plasmid replication systems. Data were taken from the Integrated Microbial Genomes and Metagenomes (IMG; http://img.jgi.doe.gov/) platform.
F: finished genomes; D: draft genomes. nd: not determined.
CDS: Coding DNA sequences; COG: Cluster of orthologous groups; ABC: ATP-binding cassette;
TRAP: Tripartite ATP-independent periplasmic.

Sta-tus	Genome Name	type	Scaffold Count	GC	genome size	% coding	Gene Count	CDS Count	COG	Homolog CDSs (alignment min. 30% and e-value <1e-20)		trans-port proteins		ABC-type transport proteins		TRAP-type transport proteins	amino acid transport proteins	Plas-mids
			total	%	Mb	%	total	total	%	total	%	total	per Mb	total	per Mb	total	total	total
F	Cand. *Pelagibacter ubique* HTCC1062	pelagic	1	30	1.31	96.1	1 394	1 354	81.13	847	27.7	95	73	67	51.1	9	31	0
F	*Nautella italica* R11	assoc	2	60	3.82	88.7	3 725	3 656		1 860	60.9	321	84	256	67.0	27	63	1
F	*Ruegeria* sp. TM1040	assoc	3	60	4.15	89.0	3 964	3 870	75.45	1 825	59.7	375	90	289	69.6	42	51	4
F	*Phaeobacter inhibens* 2.10	assoc	4	60	4.16	88.0	3 798	3 729	82.52	1 894	62.0	371	89	289	69.5	27	74	3
D	*Roseovarius* sp. TM1035	assoc	15	61	4.21	91.1	4 158	4 102	73.42	1 847	60.4	380	90	275	65.3	53	100	1
F	*Phaeobacter inhibens* DSM17395	assoc	4	60	4.23	88.8	3 960	3 891	80.03	1 891	61.9	369	87	297	70.2	22	69	3
F	*Roseobacter denitrificans* OCh 114	assoc	5	59	4.33	89.4	4 201	4 146	73.22	1 923	62.9	432	100	324	74.8	64	72	4
F	*Dinoroseobacter shibae* DFL-12, DSM 16493	assoc	6	66	4.42	89.9	4 271	4 219	74.29	1 861	60.9	407	92	273	61.8	78	55	5
D	*Ruegeria* sp. KLH11	assoc	6	58	4.49	86.4	4 338	4 274	68.88	1 797	58.8	327	73	241	53.7	34	51	2
F	*Ruegeria pomeroyi* DSS-3	assoc	2	64	4.60	90.0	4 355	4 283	78.05	1 941	63.5	447	97	321	69.8	72	114	1
D	*Silicibacter* sp. TrichCH4B	assoc	8	59	4.69	89.1	4 814	4 735	68.92	1 821	59.6	461	98	348	74.2	55	45	nd
F	*Roseobacter litoralis* Och 149	assoc	4	57	4.74	89.1	4 577	4 537	77.91	1 982	64.9	509	107	387	81.6	66	95	3
F	*Octadecabacter antarcticus* 307	assoc	18	55	4.91	83.4	5 544	5 495	58.75	1 906	62.4	428	87	339	69.0	48	58	1
F	*Octadecabacter arcticus* 238	assoc	8	55	5.39	81.8	5 883	5 834	60.58	1 888	61.8	395	73	305	56.6	50	80	2
D	*Citreicella* sp. SE45	assoc	9	67	5.52	87.9	5 499	5 427	71.65	1 721	56.3	644	117	443	80.3	132	103	3

Sta-tus	Genome Name	type	Scaffold Count	GC	genome size	% coding	Gene Count	CDS Count	COG	Homolog CDSs (alignment min. 30% and e-value <1e-20)		trans-port proteins		ABC-type transport proteins		TRAP-type transport proteins	amino acid transport proteins	Plas-mids
			total	%	Mb	%	total	total	%	total	%	total	per Mb	total	per Mb	total	total	total
D	*Rhodobacterales* sp.HTCC2255	pelagic	12	37	2.54	92.7	2 240	2 177	87.46	1 565	51,2	247	111	179	47.6	38	54	0
D	*Loktanella vestfoldensis* SKA53	pelagic	14	60	3.06	91.9	3 117	3 068	75.01	1 843	60.3	272	89	205	67.0	34	45	0
F	***Planktomarina temperata* RCA23**	**pelagic**	**1**	**54**	**3.29**	**89.8**	**3 101**	**3 056**	**83.23**	**/**	**/**	**287**	**87**	**208**	**63.2**	**43**	**53**	**0**
D	*Loktanella* sp. CCS2	pelagic	11	55	3.49	92.2	3 703	3 660	69.46	1 841	60.2	300	86	234	67.0	35	46	0
D	*Thalassiobium* sp. R2A62	pelagic	1	55	3.49	90.1	3 744	3 696	67.68	1 913	62.6	285	82	221	63.3	31	48	0
D	*Ruegeriar lacuscaerulensis* ITI-1157	pelagic	2	63	3.52	90.9	3 677	3 611	72.97	1 828	59.8	298	85	222	63.1	28	44	1
D	*Sulfitobacter* sp. EE-36	pelagic	15	60	3.55	91.0	3 542	3 474	76.31	1 764	57.7	333	94	215	60.6	73	54	2
D	*Rhodobacterales* sp. HTCC2150	pelagic	25	49	3.58	91.6	3 713	3 667	70.11	1 943	63.6	273	76	186	52.0	53	59	3
D	*Roseovarius nubinhibens* ISM	pelagic	10	64	3.67	89.8	3 605	3 547	74.84	1 737	56.8	322	88	202	55.0	76	70	1
D	*Sulfitobacter* sp. NAS-14.1	pelagic	27	60	4.00	90.1	4 026	3 962	73.7	1 788	58.5	348	87	229	57.3	65	53	5
D	*Rhodobacterales* sp. HTCC2083	pelagic	5	53	4.02	87.6	4 226	4 179	69.12	2 009	65.7	347	86	255	63.4	66	76	1
D	*Oceanicola granulosus* HTCC2516	pelagic	85	70	4.04	91.5	3 855	3 792	77.12	1 736	56.8	459	114	384	95.0	35	52	2
D	*Oceanibulbus indolifex* HEL-45	pelagic	105	60	4.11	89.4	4 208	4 153	73.31	1 780	58.2	411	100	286	69.6	68	79	1
D	*Roseobacter* sp. AzwK-3b	pelagic	31	62	4.18	88.8	4 197	4 145	69.67	1 825	59.7	325	78	229	54.8	59	78	0
F	*Jannaschia* sp. CCS1	pelagic	2	62	4.40	90.8	4 339	4 283	73.17	1 917	62.7	400	91	303	68.9	59	39	1
D	*Rhodobacterales* sp. Y4I	pelagic	5	64	4.34	86.2	4 206	4 133	73.3	1 782	58.3	317	73	218	50.2	42	73	2
D	*Oceanicola batsensis* HTCC2597	pelagic	23	66	4.44	89.2	4 261	4 212	75.1	1 739	56.9	437	98	293	66.0	82	108	4
D	*Maritimibacter alkaliphilus* HTCC2654	pelagic	46	64	4.53	90.1	4 763	4 712	67.86	1 751	57.3	381	84	260	57.4	64	88	3

Sta-tus	Genome Name	type	Scaffold Count	GC	genome size	% coding	Gene Count	CDS Count	COG	Homolog CDSs (alignment min. 30% and e-value <1e-20)		trans-port proteins		ABC-type transport proteins		TRAP-type transport proteins	amino acid transport proteins	Plas-mids
			total	%	Mb	%	total	total	%	total	%	total	per Mb	total	per Mb	total	total	total
D	*Sulfitobacter* sp. GAI101	pelagic	9	59	4.53	87.2	4 258	4 203	76.33	1 857	60.8	414	91	291	64.2	74	59	3
D	*Roseobacter* sp. SK209-2-6	pelagic	29	57	4.55	88.8	4 610	4 537	71.13	1 851	60.6	394	87	276	60.7	59	86	3
D	*Roseobacter* sp. MED193	pelagic	19	57	4.65	89.1	4 605	4 535	71.9	1 901	62.2	395	85	281	60.4	51	80	2
D	*Roseovarius* sp. 217	pelagic	37	61	4.76	90.2	4 823	4 772	72.53	1 842	60.3	417	88	277	58.2	73	100	2
D	*Sagittula stellata* E-37	pelagic	39	65	5.26	88.3	5 121	5 067	72.66	1 908	62.4	546	104	399	75.9	90	89	3
D	*Pelagibaca bermudensis* HTCC2601	pelagic	103	66	5.42	88.6	5 519	5 452	70.95	1 830	59.9	596	110	408	75.3	116	87	3

Table S3. Mean values±standard deviation of genomic features and COG categories indicated of the members of the *Roseobacter* clade associated to other organisms, surfaces and sediment and with a pelagic life style listed in supplementary Table S4, *Planktomarina temperata* RCA23 and of *Photobacterium angustum* and *Sphingopyxis alaskensis*. *: significantly different means as tested by Student's t-test ($P<0.03$; genome size, % coding genes, ABC/total transporters, COG V) or Mann Whitney Rank Sum test ($P<0.015$; ABC transporters, COG I). N: number of organism of the subgroup; CDS: Coding DNA Sequences. Data of *P. angustum* and *S. alaskensis* are from Lauro *et al*. (2009) and the Integrated Microbial Genomes and Metagenomes (IMG; http://img.jgi.doe.gov/) platform.

Feature	**Associated**	**Pelagic**	***P. temperata***	***P. angustum***	***S. alaskensis***
N	14	24	1	1	1
GC content (%)	60.07±3.58	59.28±6.82	54	39.69	65.46
Genome size (Mb)	4 547±478	4 059±681 *	3.29	5.10	3.37
% coding genes	88.05±2.57	89.82±1.62 *	89.75	85.19	90.63
CDS	4 442.7±689.2	4 003.9±699.8	3 056	4 558	3 208
Transporter proteins	419.0±82.5	366.8±86.0	287	293	116
ABC transporters	313.4±53.6	260.9±63.5 *	208	182	50
ABC/total	0.75±0.04	0.71±0.05 *	0.72	0.62	0.43
ABC/Mb	68.8±7.9	63.2±9.8	63.2	36.7	14.9
TRAP transporters	55.0±28.1	58.9+21.3	43	4	0
COG N (cell motility; %)	1.53±0.44	1.22±0.54	1.71	3.31	1.01
COG K (transcription; %)	8.03±1.01	7.53±0.99	6.15	7.53	6.62
COG V (defense; %)	1.37±0.20	1.14±0.18 *	1.01	1.47	1.24
COG T (signal transduction ; %)	4.45±0.79	4.16±0.85	2.72	7.07	3.63
COG I (lipid transport and metabolism; %)	4.58±0.47	5.34±0.86 *	6.18	2.96	4.41
COG Q (secondary metabolites, biosynthesis, transport and catabolism, %)	3.52±0.44	3.87±0.65	4.63	2.17	3.54

Table S4. Complete CDS list of *P. temperata* RCA23 and combined results of the reciprocal blast comparisons against 39 *Roseobacter* genomes. Orthologous genes with an e-value of the corresponding blast hit <1e-20 and min. global alignment of 30% are colored.

See extra file!

Table S5. Percent of COG categories N (cell motility), K (transcription), V (defense), T (signal transduction), I (lipid transport and metabolism) and Q (secondary metabolites, biosynthesis, transport and catabolism) of organisms of the *Roseobacter* clade isolated from associations with various organisms, surfaces or sediment and from pelagic environments. For comparison respective data of Cand. *Pelagibacter ubique* are also given. Data were taken from the Integrated Microbial Genomes and Metagenomes (IMG; http://img.jgi.doe.gov/) platform.

Genome Name	**N**	**K**	**V**	**T**	**I**	**Q**
	%	%	%	%	%	%
Associated						
Dinoroseobacter shibae DFL-12, DSM 16493	1.14	6.86	1.26	3.89	4.37	3.65
Nautella italica R11	1.82	8.07	1.46	5.17	4.68	3.39
Octadecabacter antarcticus 307	1.32	7.11	1.23	3.51	4.71	3.45
Octadecabacter arcticus 238, DSM 13978	1.21	6.25	0.83	2.90	3.60	2.63
Phaeobacter inhibens 2.10	1.69	8.62	1.40	4.98	4.75	3.73
Phaeobacter inhibens DSM 17395	1.64	8.30	1.33	4.67	4.86	3.63
Roseobacter denitrificans OCh 114	1.43	7.15	1.72	4.52	4.52	3.38
Roseobacter litoralis Och 149	1.35	7.63	1.40	4.37	4.21	3.56
Ruegeria pomeroyi DSS-3	0.85	9.89	1.44	4.15	5.74	4.50
Ruegeria sp. TM1040	2.17	9.29	1.34	6.08	4.61	3.28
Citreicella sp. SE45	1.27	8.60	1.37	4.19	4.29	3.86
Roseovarius sp. TM1035	2.46	7.73	1.51	4.59	4.65	3.93
Ruegeria sp. KLH11	1.17	7.97	1.54	4.05	4.75	3.38
Ruegeria sp. TrichCH4B	1.90	9.07	1.33	5.21	4.31	2.95

	N	K	V	T	I	Q
Pelagic						
***Planktomarina temperata* RCA23**	**1.71**	**6.15**	**1.01**	**2.72**	**6.18**	**4.63**
Jannaschia sp. CCS1	1.32	8.31	0.94	4.19	5.01	4.03
Loktanella sp. CCS2	1.17	7.23	1.17	4.04	4.94	2.92
Loktanella vestfoldensis SKA53	1.97	6.46	1.20	4.15	5.09	3.55
Maritimibacter alkaliphilus HTCC2654	0.53	7.74	0.93	4.02	6.96	4.36
Oceanibulbus indolifex HEL-45	1.04	7.52	0.91	5.38	4.93	3.57
Oceanicola batsensis HTCC2597	1.13	7.06	1.03	3.47	8.00	4.94
Oceanicola granulosus HTCC2516	2.25	7.27	1.21	3.77	4.57	3.80
Pelagibaca bermudensis HTCC2601	1.86	7.99	1.02	4.16	4.75	3.80
Rhodobacterales bacterium HTCC2083	1.64	6.95	1.06	2.91	5.92	4.62
Rhodobacterales bacterium HTCC2150	0.31	6.80	1.23	3.73	5.38	4.80
Rhodobacterales bacterium Y4I	1.78	8.27	1.39	6.23	4.57	3.60
Rhodobacterales bacterium HTCC2255	0.21	5.80	0.93	2.48	4.35	2.69
Roseobacter sp. AzwK-3b	0.75	6.94	1.57	4.58	4.75	3.08
Roseobacter sp. MED193	0.97	9.21	1.27	4.92	5.83	4.08
Roseobacter sp. SK209-2-6	1.62	10.09	1.04	5.15	4.79	3.54
Roseovarius nubinhibens ISM	0.26	8.67	0.93	3.82	5.00	4.04
Roseovarius sp. 217	1.17	8.66	1.03	4.43	5.32	3.72
Sagittula stellata E-37	1.45	7.87	1.10	4.97	5.29	4.27
Ruegeria lacuscaerulensis ITI-1157	1.16	7.38	1.30	4.21	5.03	3.24
Sulfitobacter sp. EE-36	1.33	7.33	1.26	4.00	6.10	4.14
Sulfitobacter sp. GAI101	1.51	6.68	1.45	4.86	5.97	4.95
Sulfitobacter sp. NAS-14.1	1.01	7.48	1.15	3.81	4.95	3.67
Thalassiobium sp. R2A62	1.10	6.91	1.26	3.75	4.38	2.88

Table S6. Normalized read counts (NPKM) of the transcripts at stations 9 (bloom, night) and 13 (bloom, day) and differences between night and day

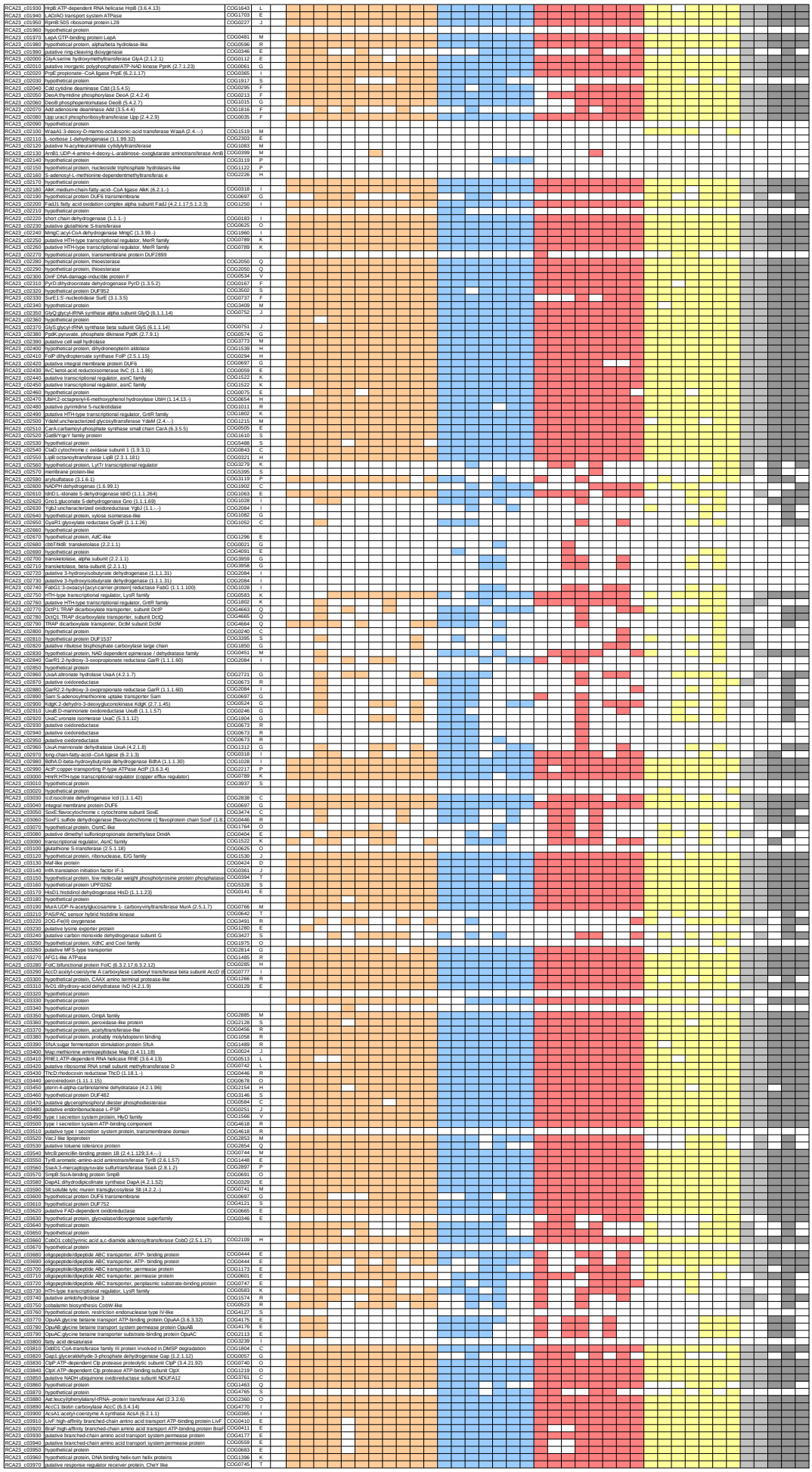
See extra file!

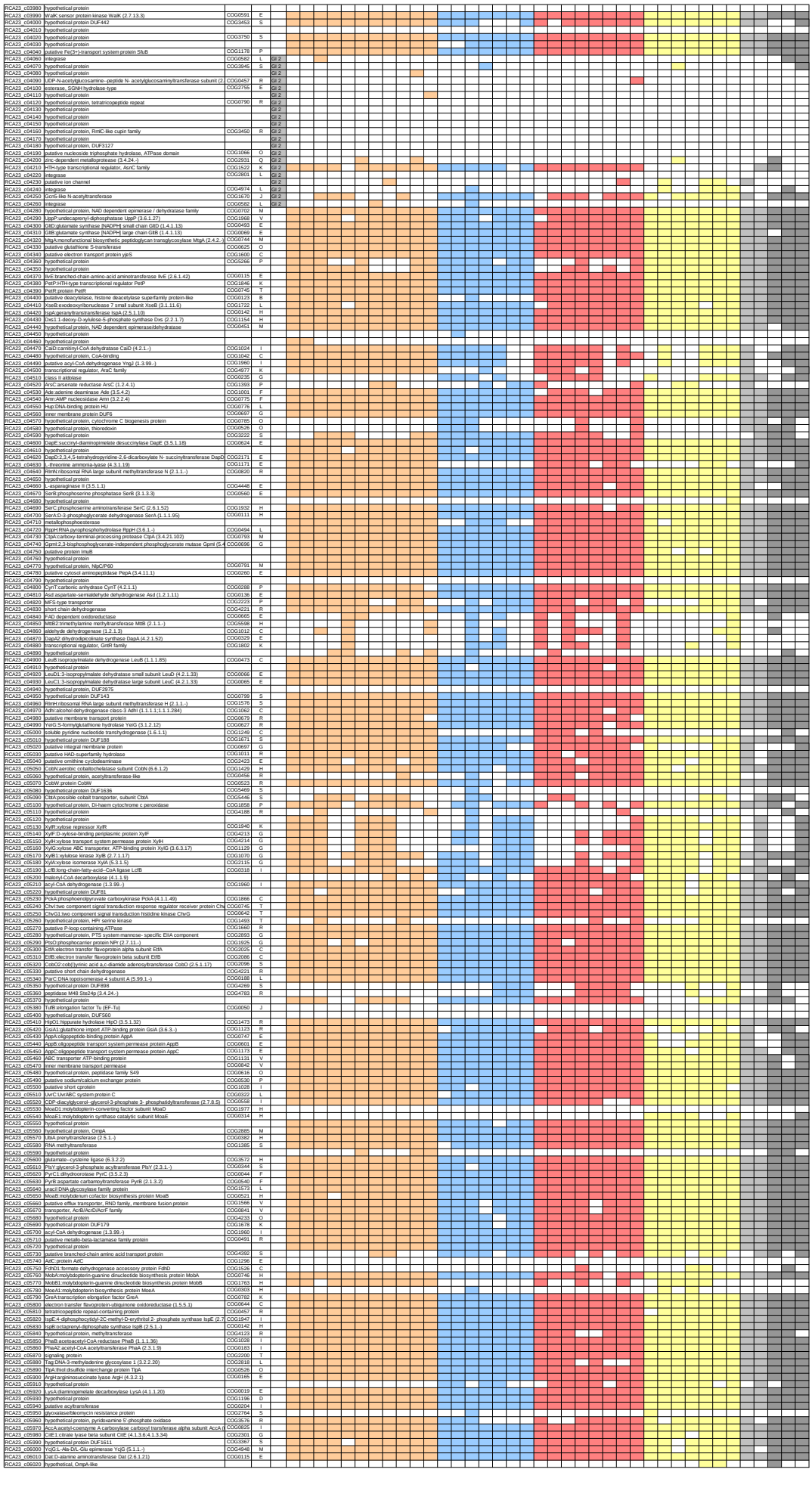
Table S7 Coverage of *P. temperata* RC23 genes from combined stations GS012-13 of the GOS data set and the Norwegian fjord metagenome.

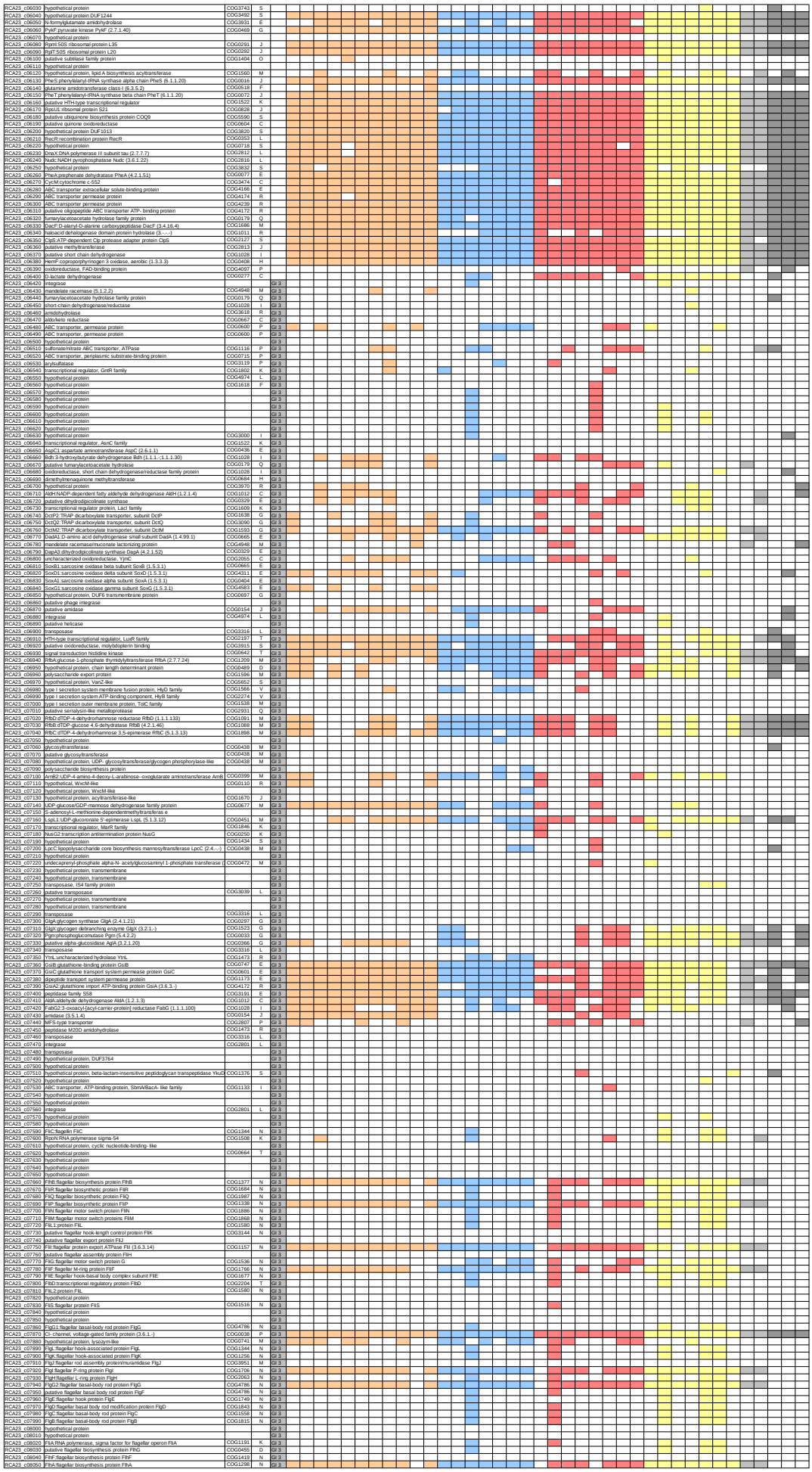
See extra file!

Table S4: combined results of the reciprocal blast comparisons against 38 Roseobacter genomes. Orthologous genes with an e-value of the corresponding blast hit <1e-20 and and min. global alignment of 30% are coloured.

locus tag	Annotation		COG category	genomic island
RCA23_c00010	DnaA:chromosomal replication initiatior protein DnaA	COG0593	L	
RCA23_c00020	DnaN:DNA polymerase III beta subunit DnaN (2.7.7.7)	COG0592	L	
RCA23_c00030	RecF:DNA replication and repair protein RecF	COG1195	L	
RCA23_c00040	hypothetical protein, LysE type translocator	COG1280	E	
RCA23_c00050	GyrB:DNA gyrase subunit B (5.99.1.3)	COG0187	L	
RCA23_c00060	FAD dependent oxidoreductase (1.4.3.-)	COG0665	E	
RCA23_c00070	hypothetical protein, rhodanese-like sulphurtransferase	COG1054	R	
RCA23_c00080	PncA:pyrazinamidase/nicotinamidase PncA (3.5.1.-;3.5.1.19)	COG1335	Q	
RCA23_c00090	PncB:nicotinate phosphoribosyltransferase PncB (2.4.2.11)	COG1488	H	
RCA23_c00100	aminoglycoside phosphotransferase	COG2334	R	
RCA23_c00110	putative aminotransferase class III	COG0160	E	
RCA23_c00120	putative peptide chain release factor	COG1186	J	
RCA23_c00130	hypothetical integral membrane protein	COG1738	S	
RCA23_c00140	MepA:penicillin-insensitive murein endopeptidase MepA (3.4.24.-)	COG3770	M	
RCA23_c00150	putative MFS-type transporter	COG2270	R	
RCA23_c00160	hypothetical protein	COG0824	R	
RCA23_c00170	hypothetical protein, YGGT family	COG0762	S	
RCA23_c00180	RecQ:ATP-dependent DNA helicase RecQ (3.6.4.12)	COG0514	L	
RCA23_c00190	hypothetical protein DUF328	COG3022	S	
RCA23_c00200	Idi1:isopentenyl-diphosphate delta-isomerase Idi (5.3.3.2)	COG1443	I	
RCA23_c00210	HemA1:5-aminolevulinate synthase HemA (2.3.1.37)	COG0156	H	
RCA23_c00220	CycA:cytochrome c2	COG3474	C	
RCA23_c00230	PhaA1:acetyl-CoA acetyltransferase PhaA (2.3.1.9)	COG0183	I	
RCA23_c00240	DnaE1:DNA polymerase III 2 alpha subunit DnaE (2.7.7.7)	COG0587	L	
RCA23_c00250	NusG1:transcription antitermination protein NusG	COG0250	K	
RCA23_c00260	putative lytic murein transglycosylase	COG2951	M	
RCA23_c00270	putative ribonuclease R	COG0557	K	
RCA23_c00280	hypothetical protein DUF461	COG2847	S	
RCA23_c00290	AroC:chorismate synthase AroC (4.2.3.5)	COG0082	E	
RCA23_c00300	hypothetical protein			
RCA23_c00310	ThiB:thiamine-binding periplasmic protein ThiB	COG4143	H	
RCA23_c00320	ThiP:thiamine transport system permease protein ThiP	COG1178	P	
RCA23_c00330	ThiQ:thiamine import ATP-binding protein ThiQ (3.6.3.-)	COG3840	H	
RCA23_c00340	PetC:cytochrome c1	COG2857	C	
RCA23_c00350	PetB:cytochrome b	COG1290	C	
RCA23_c00360	PetA:ubiquinol-cytochrome c reductase iron-sulfur subunit PetA (1.10.2.2)	COG0723	C	
RCA23_c00370	hypothetical protein, glutathione S-transferase	COG0625	O	
RCA23_c00380	IdhA:inositol 2-dehydrogenase IdhA (1.1.1.18)	COG0673	R	
RCA23_c00390	PmbA:protein PmbA	COG0312	R	
RCA23_c00400	putative inositol monophosphatase	COG0483	G	
RCA23_c00410	putative 3-deoxy-D-manno-octulosonic-acid transferase (2.4.-.-)	COG1519	M	
RCA23_c00420	LpxK:tetraacyldisaccharide 4'-kinase LpxK (2.7.1.130)	COG1663	M	
RCA23_c00430	hypothetical protein, thioredoxin	COG1651	O	
RCA23_c00440	hypothetical protein DUF721	COG5389	S	
RCA23_c00450	YfhQ:A/G-specific adenine glycosylase YfhQ (3.2.2.-)	COG1194	L	
RCA23_c00460	AlkB:alkane 1-monooxygenase AlkB (1.14.15.3)			
RCA23_c00470	CcrM:modification methylase CcrM (2.1.1.72)	COG0863	L	
RCA23_c00480	RnhB:ribonuclease HII (3.1.26.4)	COG0164	L	
RCA23_c00490	putative exodeoxyribonuclease III	COG0708	L	
RCA23_c00500	hypothetical protein, thioredoxin	COG3118	O	
RCA23_c00510	putative ATP-dependent protease La (LON)	COG2802	R	
RCA23_c00520	hypothetical protein, Trm112p-like	COG2835	S	
RCA23_c00530	putative 2-octaprenyl-6-methoxyphenol hydroxylase	COG0654	H	
RCA23_c00540	GatA1:glutamyl-tRNA(Gln) amidotransferase subunit A (6.3.5.-)	COG0154	J	
RCA23_c00550	DapL:LL-diaminopimelate aminotransferase DapL (2.6.1.83)	COG0436	E	
RCA23_c00560	putative DNA transclocase FtsK	COG1674	D	
RCA23_c00570	putative outer membrane lipoprotein carrier protein LolA	COG2834	M	
RCA23_c00580	hypothetical protein	COG4764	S	
RCA23_c00590	hemimethylated DNA-binding protein, YccV like	COG3785	S	
RCA23_c00600	putative gamma-glutamyltransferase ywrD (2.3.2.2)	COG0405	E	
RCA23_c00610	ThlA:acetyl-CoA acetyltransferase ThlA (2.3.1.9)	COG0183	I	
RCA23_c00620	hypothetical protein	COG1956	T	
RCA23_c00630	putative HTH-type transcriptional regulator	COG1396	K	
RCA23_c00640	hypothetical protein, homoserine/homoserine lactone efflux protein	COG1280	E	
RCA23_c00650	LysE-type translocator	COG1280	E	
RCA23_c00660	dimethylglycine dehydrogenase (1.5.99.2)	COG0404	E	
RCA23_c00670	putative homocysteine S-methyltransferase	COG2040	E	
RCA23_c00680	pyridoxamine 5'-phosphate oxidase-like protein, FMN-binding	COG3576	R	
RCA23_c00690	Sardh:sarcosine dehydrogenase (1.5.99.1)	COG0404	E	
RCA23_c00700	hypothetical protein	COG0790	R	
RCA23_c00710	putative agmatine deiminase AguA (3.5.3.12)	COG2957	E	
RCA23_c00720	Cl- channel, voltage-gated family protein (3.6.1.-)	COG0038	P	
RCA23_c00730	Cl- channel, voltage-gated family protein (3.6.1.-)	COG0038	P	
RCA23_c00750	DapF:diaminopimelate epimerase DapF (5.1.1.7)	COG0253	E	GI 1
RCA23_c00760	MiaB:(dimethylallyl)adenosine tRNA methylthiotransferase MiaB (2.-.-.-)	COG0621	J	GI 1
RCA23_c00770	hypothetical protein, glutathione S- transferotein	COG0625	O	GI 1
RCA23_c00780	HTH-type transcriptional regulator, LacI family	COG1609	K	GI 1
RCA23_c00790	hypothetical protein			GI 1
RCA23_c00800	hypothetical protein	COG4274	S	GI 1
RCA23_c00810	hypothetical protein, phosphoglycerate mutase- like	COG0406	G	GI 1
RCA23_c00820	hypothetical protein	COG0790	R	GI 1
RCA23_c00830	hypothetical protein, glyoxalase/bleomycin resistance protein/dihydroxybiphenyl	COG0346	E	GI 1
RCA23_c00840	putative phage integrase	COG4974	L	GI 1
RCA23_c00850	hypothetical protein			GI 1
RCA23_c00860	hypothetical protein			GI 1
RCA23_c00870	putative endonuclease			GI 1
RCA23_c00880	DNA integration/recombination/invertion protein	COG1961	L	GI 1
RCA23_c00890	hypothetical protein			GI 1
RCA23_c00900	hypothetical protein, peptidoglycan binding- like			GI 1
RCA23_c00910	hypothetical protein			GI 1
RCA23_c00920	putative phage integrase	COG4974	L	GI 1
RCA23_c00940	putative HTH-type transcriptional repressor, ArsR family	COG0640	K	
RCA23_c00950	AtpI:ATP synthase protein I			
RCA23_c00960	AtpB:ATP synthase subunit AtpB	COG0356	C	
RCA23_c00970	AtpE:ATP synthase subunit c			
RCA23_c00980	ATP synthase B' chain AtpX	COG0711	C	
RCA23_c00990	ATP synthase B chain AtpF			
RCA23_c01000	PdhR:pyruvate dehydrogenase complex repressor	COG2186	K	
RCA23_c01010	YiaD:inner membrane lipoprotein YiaD	COG2885	M	
RCA23_c01020	Nth:endonuclease III (4.2.99.18)	COG0177	L	
RCA23_c01030	putative pfkB family carbohydrate kinase	COG0524	G	
RCA23_c01040	hypothetical protein			
RCA23_c01050	glycosyl transferase family 14			
RCA23_c01060	hypothetical protein			
RCA23_c01070	HIT-like protein	COG0537	F	
RCA23_c01080	ABC transporter ATP binding protein	COG1131	V	
RCA23_c01090	hypothetical protein	COG4391	S	
RCA23_c01100	hypothetical protein			
RCA23_c01110	PolA:DNA polymerase I (2.7.7.7)	COG0749	L	
RCA23_c01120	MetB:cystathionine gamma-synthase MetB (2.5.1.48)	COG0626	E	
RCA23_c01130	RluE:ribosomal large subunit pseudouridine synthase RluE (5.4.99.-)	COG1187	J	
RCA23_c01140	hypothetical protein DUF1285	COG3816	S	
RCA23_c01150	ATPase, MoxR type	COG0714	R	
RCA23_c01160	hypothetical protein	COG1721	R	
RCA23_c01170	hypothetical protein			
RCA23_c01180	hypothetical protein DUF1355	COG5426	S	
RCA23_c01190	GlcD:glycolate oxidase subunit GlcD	COG0277	C	
RCA23_c01200	GlcE:glycolate oxidase subunit GlcE	COG0277	C	
RCA23_c01210	GlcF:glycolate oxidase iron-sulfur subunit GlcF	COG0247	C	
RCA23_c01220	hypothetical protein, trypsin	COG3591	E	
RCA23_c01230	putative chitinase (3.2.1.14)	COG3325	G	
RCA23_c01240	IbpA1:small heat shock protein IbpA	COG0071	O	
RCA23_c01250	GabD:succinate-semialdehyde dehydrogenase GabD (1.2.1.16)	COG1012	C	
RCA23_c01260	hypothetical protein	COG0657	I	
RCA23_c01270	hypothetical protein	COG5617	S	
RCA23_c01280	putative cyclopentanol dehydrogenase CpnA (1.1.1.163)	COG1028	I	
RCA23_c01290	hypothetical protein			
RCA23_c01300	hypothetical protein DUF1523			
RCA23_c01310	aldehyde dehydrogenase (1.2.1.3)	COG1012	C	
RCA23_c01320	DeoC:deoxyribose-phosphate aldolase DeoC (4.1.2.4)	COG0274	F	
RCA23_c01330	CfxE:ribulose-phosphate 3-epimerase, chromosomal (5.1.3.1)	COG0036	G	
RCA23_c01340	hemolysin-type calcium-binding region			
RCA23_c01350	5-oxoprolinase (ATP-hydrolyzing) (3.5.2.9)	COG0145	E	
RCA23_c01360	aminotransferase class-III (2.6.1.-;2.6.1.18)	COG0161	H	
RCA23_c01370	AllA:ureidoglycolate hydrolase AllA (3.5.3.19)	COG3194	F	
RCA23_c01380	GalM:aldose 1-epimerase GalM (5.1.3.3)	COG2017	G	
RCA23_c01390	BgaB:beta-galactosidase BgaB (3.2.1.23)	COG1874	G	
RCA23_c01400	putative gluconolactonase	COG3386	G	
RCA23_c01410	DgoA:2-dehydro-3-deoxy-6-phosphogalactonate aldolase DgoA (4.1.2.21)	COG0800	G	
RCA23_c01420	DgoK:2-dehydro-3-deoxygalactonokinase DgoK (2.7.1.58)	COG3734	G	
RCA23_c01430	short chain dehydrogenase	COG1028	I	
RCA23_c01440	RafA:alpha-galactosidase RafA (3.2.1.22)	COG3345	G	
RCA23_c01450	putative ABC transporter inner membrane component	COG0395	G	
RCA23_c01460	putative ABC transporter inner membrane component	COG1175	G	
RCA23_c01470	putative extracellular solute-binding protein	COG1653	G	
RCA23_c01480	HTH-type transcriptional regulator, IclR family	COG1414	K	
RCA23_c01490	sugar ABC transporter ATP-binding protein	COG3839	G	
RCA23_c01500	hypothetical protein, DUF81 family			
RCA23_c01510	putative transcriptional regulator, gntR family	COG1167	K	
RCA23_c01520	Tpa:taurine--pyruvate aminotransferase Tpa (2.6.1.77)	COG0161	H	
RCA23_c01530	putative taurine-binding periplasmic protein	COG4521	P	
RCA23_c01540	TauB:taurine transport system ATP-binding protein TauB (3.6.3.36)	COG1116	P	
RCA23_c01550	TauC:taurine transport system permease protein TauC	COG0600	P	
RCA23_c01560	dimethylglycine dehydrogenase (1.5.99.2)	COG0404	E	
RCA23_c01570	MttB1:trimethylamine methyltransferase MttB (2.1.1.-)	COG5598	H	
RCA23_c01580	FadN1:3-hydroxyacyl-CoA dehydrogenase FadN (1.1.1.35)	COG1250	I	
RCA23_c01590	hypothetical protein	COG0596	R	
RCA23_c01600	CysD:O-acetylhomoserine (thiol)-lyase CysD (2.5.1.49)	COG2873	E	
RCA23_c01610	putative signaling protein	COG2200	T	
RCA23_c01620	Xsc:sulfoacetaldehyde acetyltransferase Xsc (2.3.3.15)	COG0028	E	
RCA23_c01630	Pta:phosphate acetyltransferase Pta (2.3.1.8)	COG0280	C	
RCA23_c01640	DmsA:DMSO reductase chain A (1.2.7.-)	COG0243	C	
RCA23_c01650	alpha/beta hydrolase	COG0596	R	
RCA23_c01660	DmsB:DMSO reductase chain B	COG0437	C	
RCA23_c01670	DmsC:DMSO reductase chain C	COG3302	R	
RCA23_c01680	sodium:alanine symporter	COG1115	E	
RCA23_c01690	Tdk:thymidine kinase Tdk (2.7.1.21)	COG1435	F	
RCA23_c01700	NADPH-dependent FMN reductase	COG0431	R	
RCA23_c01710	CarB:carbamoyl-phosphate synthase large chain CarB (6.3.5.5)	COG0458	E	
RCA23_c01720	HTH-type transcriptional regulator, AsnC family	COG1522	K	
RCA23_c01730	hypothetical protein			
RCA23_c01740	AspS:aspartyl-tRNA synthase AspS (6.1.1.12)	COG0173	J	
RCA23_c01750	hypothetical protein			
RCA23_c01760	putative signal transduction response regulator receiver protein	COG2197	T	
RCA23_c01770	methylmalonyl-CoA epimerase (5.1.99.1)	COG0346	E	
RCA23_c01780	hypothetical protein DUF1467	COG5454	S	
RCA23_c01790	hypothetical protein DUF540	COG2981	E	
RCA23_c01800	putative nitroreductase	COG0778	C	
RCA23_c01810	HdaH:histone deacetylase-like amidohydrolase HdaH (3.5.1.-)	COG0123	B	
RCA23_c01820	hypothetical protein			
RCA23_c01830	hypothetical protein, peptidase family M48	COG0501	O	
RCA23_c01840	hypothetical protein			
RCA23_c01850	RimO:ribosomal protein S12 methylthiotransferase RimO (2.-.-.-)	COG0621	J	
RCA23_c01860	hypothetical protein, transmembrane			
RCA23_c01870	GcrA cell cycle regulator	COG5352	S	
RCA23_c01880	ABC-2 type transport system membrane protein	COG0842	V	
RCA23_c01890	ArgD:acetylornithine aminotransferase ArgD (2.6.1.11)	COG4992	E	
RCA23_c01900	ArgF:ornithine carbamoyltransferase ArgF (2.1.3.3)	COG0078	E	
RCA23_c01910	hypothetical protein			
RCA23_c01920	GntZ:6-phosphogluconate dehydrogenase GntZ (1.1.1.44)	COG0362	G	

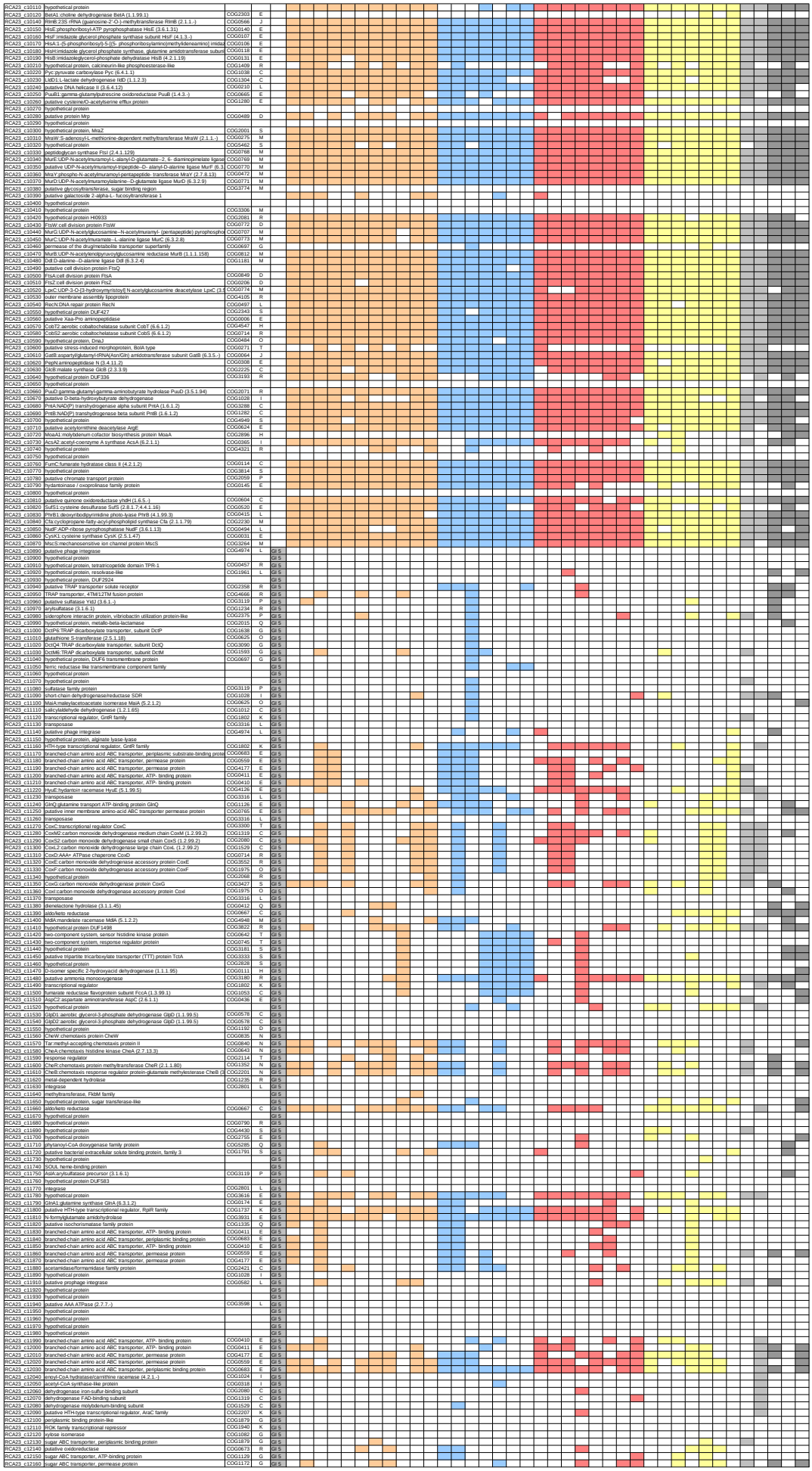

RCA23_c01930 HrpB ATP-dependent RNA helicase HrpB (3.6.4.13) COG1643 L
RCA23_c01940 LAO/AO transport system ATPase COG1703 E
RCA23_c01950 RpmB 50S ribosomal protein L28 COG0227 J
RCA23_c01960 hypothetical protein
RCA23_c01970 LepA GTP-binding protein LepA COG0481 M
RCA23_c01980 hypothetical protein, alpha/beta hydrolase-like COG0596 R
RCA23_c01990 putative ring-cleaving dioxygenase COG0346 E
RCA23_c02000 GlyA serine hydroxymethyltransferase GlyA (2.1.2.1) COG0112 E
RCA23_c02010 putative inorganic polyphosphate/ATP-NAD kinase PpnK (2.7.1.23) COG0061 G
RCA23_c02020 PrpE propionate--CoA ligase PrpE (6.2.1.17) COG0365 I
RCA23_c02030 hypothetical protein COG1917 S
RCA23_c02040 CddI cytidine deaminase CddI (3.5.4.5) COG0295 F
RCA23_c02050 DeoA thymidine phosphorylase DeoA (2.4.2.4) COG0213 F
RCA23_c02060 DeoB phosphopentomutase DeoB (5.4.2.7) COG1015 G
RCA23_c02070 AddI adenosine deaminase AddI (3.5.4.4) COG1816 F
RCA23_c02080 Upp uracil phosphoribosyltransferase Upp (2.4.2.9) COG0035 F
RCA23_c02090 hypothetical protein
RCA23_c02100 WaaA3 3-deoxy-D-manno-octulosonic-acid transferase WaaA (2.4.--) COG1519 M
RCA23_c02110 L-sorbose 1-dehydrogenase (1.1.99.32) COG2303 E
RCA23_c02120 putative N-acylneuraminate cytidylyltransferase COG1083 M
RCA23_c02130 AmtB1 UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase AmtB COG0399 M
RCA23_c02140 hypothetical protein COG3179 P
RCA23_c02150 hypothetical protein, nucleoside triphosphate hydrolases-like COG1122 P
RCA23_c02160 S-adenosyl-L-methionine-dependentmethyltransferas e COG2226 H
RCA23_c02170 hypothetical protein
RCA23_c02180 AlkK medium-chain fatty-acid--CoA ligase AlkK (6.2.1.-) COG0318 I
RCA23_c02190 hypothetical protein DUF6 transmembrane COG0697 G
RCA23_c02200 FadJ1 fatty acid oxidation complex alpha subunit FadJ (4.2.1.17,5.1.2.3) COG1250 I
RCA23_c02210 hypothetical protein
RCA23_c02220 short chain dehydrogenase (1.1.1.-) COG0183 I
RCA23_c02230 putative glutathione S-transferase COG0625 O
RCA23_c02240 MmgC acyl-CoA dehydrogenase MmgC (1.3.99.-) COG1960 I
RCA23_c02250 putative HTH-type transcriptional regulator, MerR family COG0789 K
RCA23_c02260 putative HTH-type transcriptional regulator, MerR family COG0789 K
RCA23_c02270 hypothetical protein, transmembrane protein DUF2899
RCA23_c02280 hypothetical protein, thioesterase COG2050 Q
RCA23_c02290 hypothetical protein, thioesterase COG2050 Q
RCA23_c02300 DinF DNA-damage-inducible protein F COG0534 V
RCA23_c02310 PyrD dihydroorotate dehydrogenase PyrD (1.3.5.2) COG0167 F
RCA23_c02320 hypothetical protein DUF952 COG3502 S
RCA23_c02330 SurE1 5'-nucleotidase SurE (3.1.3.5) COG0737 F
RCA23_c02340 hypothetical protein COG3409 M
RCA23_c02350 GlyQ glycyl-tRNA synthase alpha subunit GlyQ (6.1.1.14) COG0752 J
RCA23_c02360 hypothetical protein
RCA23_c02370 GlyS glycyl-tRNA synthase beta subunit GlyS (6.1.1.14) COG0751 J
RCA23_c02380 PpdK pyruvate, phosphate dikinase PpdK (2.7.9.1) COG0574 G
RCA23_c02390 putative cell wall hydrolase COG3773 M
RCA23_c02400 hypothetical protein, dihydroneopterin aldolase COG1539 H
RCA23_c02410 FolP dihydropteroate synthase FolP (2.5.1.15) COG0294 H
RCA23_c02420 putative integral membrane protein DUF6 COG0697 G
RCA23_c02430 IlvC ketol-acid reductoisomerase IlvC (1.1.1.86) COG0059 E
RCA23_c02440 putative transcriptional regulator, asnC family COG1522 K
RCA23_c02450 putative transcriptional regulator, asnC family COG1522 K
RCA23_c02460 hypothetical protein COG0075 E
RCA23_c02470 UbiH 2-octaprenyl-6-methoxyphenol hydroxylase UbiH (1.14.13.-) COG0654 H
RCA23_c02480 putative pyrimidine 5-nucleotidase COG1011 R
RCA23_c02490 putative HTH-type transcriptional regulator, GntR family COG1802 K
RCA23_c02500 YdaM uncharacterized glycosyltransferase YdaM (2.4.--) COG1215 M
RCA23_c02510 CarA carbamoyl-phosphate synthase small chain CarA (6.3.5.5) COG0505 E
RCA23_c02520 GatB/YqeY family protein COG1610 S
RCA23_c02530 hypothetical protein COG5488 S
RCA23_c02540 CtaD cytochrome c oxidase subunit 1 (1.9.3.1) COG0843 C
RCA23_c02550 LipB octanoyltransferase LipB (2.3.1.181) COG0321 H
RCA23_c02560 hypothetical protein, LytTr transcriptional regulator COG3279 K
RCA23_c02570 membrane protein-like COG5395 S
RCA23_c02580 acylsulfatase (3.1.6.1) COG3119 P
RCA23_c02600 NADPH dehydrogenas (1.6.99.1) COG1902 C
RCA23_c02610 IdnD1 -idonate 5-dehydrogenase IdnD (1.1.1.264) COG1063 E
RCA23_c02620 Gno1 gluconate 5-dehydrogenase Gno (1.1.1.69) COG1028 I
RCA23_c02630 YjbJ uncharacterized oxidoreductase YjbJ (1.1.--) COG2084 I
RCA23_c02640 hypothetical protein, xylose isomerase-like COG1082 G
RCA23_c02650 GyaR1 glyoxylate reductase GyaR (1.1.1.26) COG1052 C
RCA23_c02660 hypothetical protein
RCA23_c02670 hypothetical protein, AdlC-like COG1296 E
RCA23_c02680 cbbT/tktB transketolase (2.2.1.1) COG0021 G
RCA23_c02690 hypothetical protein COG4091 E
RCA23_c02700 transketolase, alpha subunit (2.2.1.1) COG3959 G
RCA23_c02710 transketolase, beta-subunit (2.2.1.1) COG3958 G
RCA23_c02720 putative 3-hydroxyisobutyrate dehydrogenase (1.1.1.31) COG2084 I
RCA23_c02730 putative 3-hydroxyisobutyrate dehydrogenase (1.1.1.31) COG2084 I
RCA23_c02740 FabG1 3-oxoacyl-[acyl-carrier-protein] reductase FabG (1.1.1.100) COG1028 I
RCA23_c02750 HTH-type transcriptional regulator, LysR family COG0583 K
RCA23_c02760 putative HTH-type transcriptional regulator, GntR family COG1802 K
RCA23_c02770 DctP1 TRAP dicarboxylate transporter, subunit DctP COG4663 Q
RCA23_c02780 DctQ1 TRAP dicarboxylate transporter, subunit DctQ COG4665 Q
RCA23_c02790 TRAP dicarboxylate transporter, DctM subunit DctM COG4664 Q
RCA23_c02800 hypothetical protein COG0240 C
RCA23_c02810 hypothetical protein DUF1537 COG3395 S
RCA23_c02820 putative ribulose bisphosphate carboxylase large chain COG1850 G
RCA23_c02830 hypothetical protein, NAD dependent epimerase / dehydratase family COG0451 M
RCA23_c02840 GarR1 2-hydroxy-3-oxopropionate reductase GarR (1.1.1.60) COG2084 I
RCA23_c02850 hypothetical protein
RCA23_c02860 UxaA altronate hydrolase UxaA (4.2.1.7) COG2721 G
RCA23_c02870 putative oxidoreductase COG0673 R
RCA23_c02880 GarR2 2-hydroxy-3-oxopropionate reductase GarR (1.1.1.60) COG2084 I
RCA23_c02890 Sam S-adenosylmethionine uptake transporter Sam COG0697 G
RCA23_c02900 KdgK 2-dehydro-3-deoxygluconokinase KdgK (2.7.1.45) COG0524 G
RCA23_c02910 UxuB D-mannonate oxidoreductase UxuB (1.1.1.57) COG0246 G
RCA23_c02920 UxaC uronate isomerase UxaC (5.3.1.12) COG1904 G
RCA23_c02930 putative oxidoreductase COG0673 R
RCA23_c02940 putative oxidoreductase COG0673 R
RCA23_c02950 putative oxidoreductase COG0673 R
RCA23_c02960 UxuA mannonate dehydratase UxuA (4.2.1.8) COG1312 G
RCA23_c02970 long-chain-fatty-acid--CoA ligase (6.2.1.3) COG0318 I
RCA23_c02980 BdhA D-beta-hydroxybutyrate dehydrogenase BdhA (1.1.1.30) COG1028 I
RCA23_c02990 ActP copper-transporting P-type ATPase ActP (3.6.3.4) COG2217 P
RCA23_c03000 HmrR HTH-type transcriptional regulator (copper efflux regulator) COG0789 K
RCA23_c03010 hypothetical protein COG3937 S
RCA23_c03020 hypothetical protein
RCA23_c03030 Icd isocitrate dehydrogenase Icd (1.1.1.42) COG2838 C
RCA23_c03040 integral membrane protein DUF6 COG0697 G
RCA23_c03050 SoxE flavocytochrome c cytochrome subunit SoxE COG3474 C
RCA23_c03060 SoxF1 sulfide dehydrogenase (flavocytochrome c) flavoprotein chain SoxF (1.8 COG3446 R
RCA23_c03070 hypothetical protein, CsnK-like COG1764 O
RCA23_c03080 putative dimethyl sulfoniopropionate demethylase DmdA COG0404 E
RCA23_c03090 transcriptional regulator, AsnC family COG1522 K
RCA23_c03100 glutathione S-transferase (2.5.1.18) COG0625 O
RCA23_c03120 hypothetical protein, ribonuclease, E/G family COG1530 J
RCA23_c03130 Maf-like protein COG0424 D
RCA23_c03140 InfA translation initiation factor IF-1 COG0361 J
RCA23_c03150 hypothetical protein, low molecular weight phosphotyrosine protein phosphatase COG0394 T
RCA23_c03160 hypothetical protein UPF0262 COG5328 S
RCA23_c03170 HisD1 histidinol dehydrogenase HisD (1.1.1.23) COG0141 E
RCA23_c03180 hypothetical protein
RCA23_c03190 MurA UDP-N-acetylglucosamine 1- carboxyvinyltransferase MurA (2.5.1.7) COG0766 M
RCA23_c03210 PASPHC sensor hybrid histidine kinase COG0642 T
RCA23_c03220 2OG-Fe(II) oxygenase COG3491 R
RCA23_c03230 putative lysine exporter protein COG1280 E
RCA23_c03240 putative carbon monoxide dehydrogenase subunit G COG1427 S
RCA23_c03250 hypothetical protein, XdhC and CoxI family COG1975 O
RCA23_c03260 putative MFS-type transporter COG2814 G
RCA23_c03270 AFG1-like ATPase COG1485 R
RCA23_c03280 FolC bifunctional protein FolC (6.3.2.17,6.3.2.12) COG0285 H
RCA23_c03290 AccD acetyl-coenzyme A carboxylase carboxyl transferase beta subunit AccD (6 COG0777 I
RCA23_c03300 hypothetical protein, CAAX amino terminal protease-like COG1266 R
RCA23_c03310 IlvD1 dihydroxy-acid dehydratase IlvD (4.2.1.9) COG0129 E
RCA23_c03320 hypothetical protein
RCA23_c03330 hypothetical protein
RCA23_c03340 hypothetical protein
RCA23_c03350 hypothetical protein, OmpA family COG2885 M
RCA23_c03360 hypothetical protein, peroxidase-like protein COG2128 S
RCA23_c03370 hypothetical protein, acetyltransferase-like COG0456 R
RCA23_c03380 hypothetical protein, probably molybdopterin binding COG1058 R
RCA23_c03390 StsA sugar fermentation stimulation protein StsA COG1489 R
RCA23_c03400 Map methionine aminopeptidase Map (3.4.11.18) COG0024 J
RCA23_c03410 RhlE1 ATP-dependent RNA helicase RhlE (3.6.4.13) COG0513 L
RCA23_c03420 putative ribosomal RNA small subunit methyltransferase D COG0742 L
RCA23_c03430 ThcD rhodocoxin reductase ThcD (1.18.1.-) COG0446 R
RCA23_c03440 ferredoxin (1.11.1.15) COG0678 O
RCA23_c03450 pterin-4-alpha-carbinolamine dehydratase (4.2.1.96) COG2154 H
RCA23_c03460 hypothetical protein DUF482 COG3146 S
RCA23_c03470 putative glycerophosphoryl diester phosphodiesterase COG0584 C
RCA23_c03480 putative endoribonuclease L-PSP COG0251 J
RCA23_c03490 type I secretion system protein, HlyD family COG1566 V
RCA23_c03500 type I secretion system ATP-binding component COG4618 R
RCA23_c03510 putative type I secretion system protein, transmembrane domain COG4618 R
RCA23_c03520 VacJ like lipoprotein COG2853 M
RCA23_c03530 putative toluene tolerance protein COG2854 Q
RCA23_c03540 MecB penicillin-binding protein 1B (2.4.1.129,3.4.--) COG0744 M
RCA23_c03550 TyrB aromatic-amino-acid aminotransferase TyrB (2.6.1.57) COG1448 E
RCA23_c03560 SsaA 3-mercaptopyruvate sulfurtransferase SsaA (2.8.1.2) COG2897 P
RCA23_c03570 SmpB SsrA-binding protein SmpB COG0691 O
RCA23_c03580 DapA1 dihydrodipicolinate synthase DapA (4.2.1.52) COG0329 E
RCA23_c03590 Slt soluble lytic murein transglycosylase Slt (4.2.2.-) COG0741 M
RCA23_c03600 hypothetical protein DUF6 transmembrane COG0697 G
RCA23_c03610 hypothetical protein DUF752 COG4121 S
RCA23_c03620 putative FAD-dependent oxidoreductase COG0665 E
RCA23_c03630 hypothetical protein, glyoxalase/dioxygenase superfamily COG0346 E
RCA23_c03640 hypothetical protein
RCA23_c03650 hypothetical protein
RCA23_c03660 CobO1 cob(I)yrinic acid a,c-diamide adenosyltransferase CobO (2.5.1.17) COG2109 H
RCA23_c03670 hypothetical protein
RCA23_c03680 oligopeptide/dipeptide ABC transporter, ATP- binding protein COG0444 E
RCA23_c03690 oligopeptide/dipeptide ABC transporter, ATP- binding protein COG0444 E
RCA23_c03700 oligopeptide/dipeptide ABC transporter, permease protein COG1173 E
RCA23_c03710 oligopeptide/dipeptide ABC transporter, permease protein COG0601 E
RCA23_c03720 oligopeptide/dipeptide ABC transporter, periplasmic substrate-binding protein COG0747 E
RCA23_c03730 HTH-type transcriptional regulator, LysR family COG0583 K
RCA23_c03740 putative amidohydrolase 3 COG1574 R
RCA23_c03750 cobalamin biosynthesis CobW-like COG0523 R
RCA23_c03760 hypothetical protein, restriction endonuclease type IV-like COG4127 S
RCA23_c03770 OpuAA glycine betaine transport ATP-binding protein OpuAA (3.6.3.32) COG4175 E
RCA23_c03780 OpuAB glycine betaine transport system permease protein OpuAB COG4176 E
RCA23_c03790 OpuAC glycine betaine transporter substrate-binding protein OpuAC COG2113 E
RCA23_c03800 fatty acid desaturase COG3239 I
RCA23_c03810 DdtO1 CoA-transferase family III protein involved in DMSP degradation COG1804 C
RCA23_c03820 Gap1 glyceraldehyde-3-phosphate dehydrogenase Gap (1.2.1.12) COG0057 G
RCA23_c03830 ClpP ATP-dependent Clp protease proteolytic subunit ClpP (3.4.21.92) COG0740 O
RCA23_c03840 ClpX ATP-dependent Clp protease ATP-binding subunit ClpX COG1219 O
RCA23_c03850 putative NADH ubiquinone oxidoreductase subunit NDUFA12 COG3761 C
RCA23_c03860 hypothetical protein COG1463 Q
RCA23_c03870 hypothetical protein COG4765 S
RCA23_c03880 Aat leucyl/phenylalanyl-tRNA--protein transferase Aat (2.3.2.6) COG3760 O
RCA23_c03890 AccC1 biotin carboxylase AccC (6.3.4.14) COG4770 I
RCA23_c03900 AcsA1 acetyl-coenzyme A synthase AcsA (6.2.1.1) COG0365 I
RCA23_c03910 LivF high-affinity branched-chain amino acid transport ATP-binding protein LivF COG0410 E
RCA23_c03920 BraF high-affinity branched-chain amino acid transport ATP-binding protein BraF COG0411 E
RCA23_c03930 putative branched-chain amino acid transport system permease protein COG4177 E
RCA23_c03940 putative branched-chain amino acid transport system permease protein COG0559 E
RCA23_c03950 hypothetical protein COG0683 E
RCA23_c03960 hypothetical protein, DNA binding helix-turn helix proteins COG1396 K
RCA23_c03970 putative response regulator receiver protein, CheY like COG0745 T



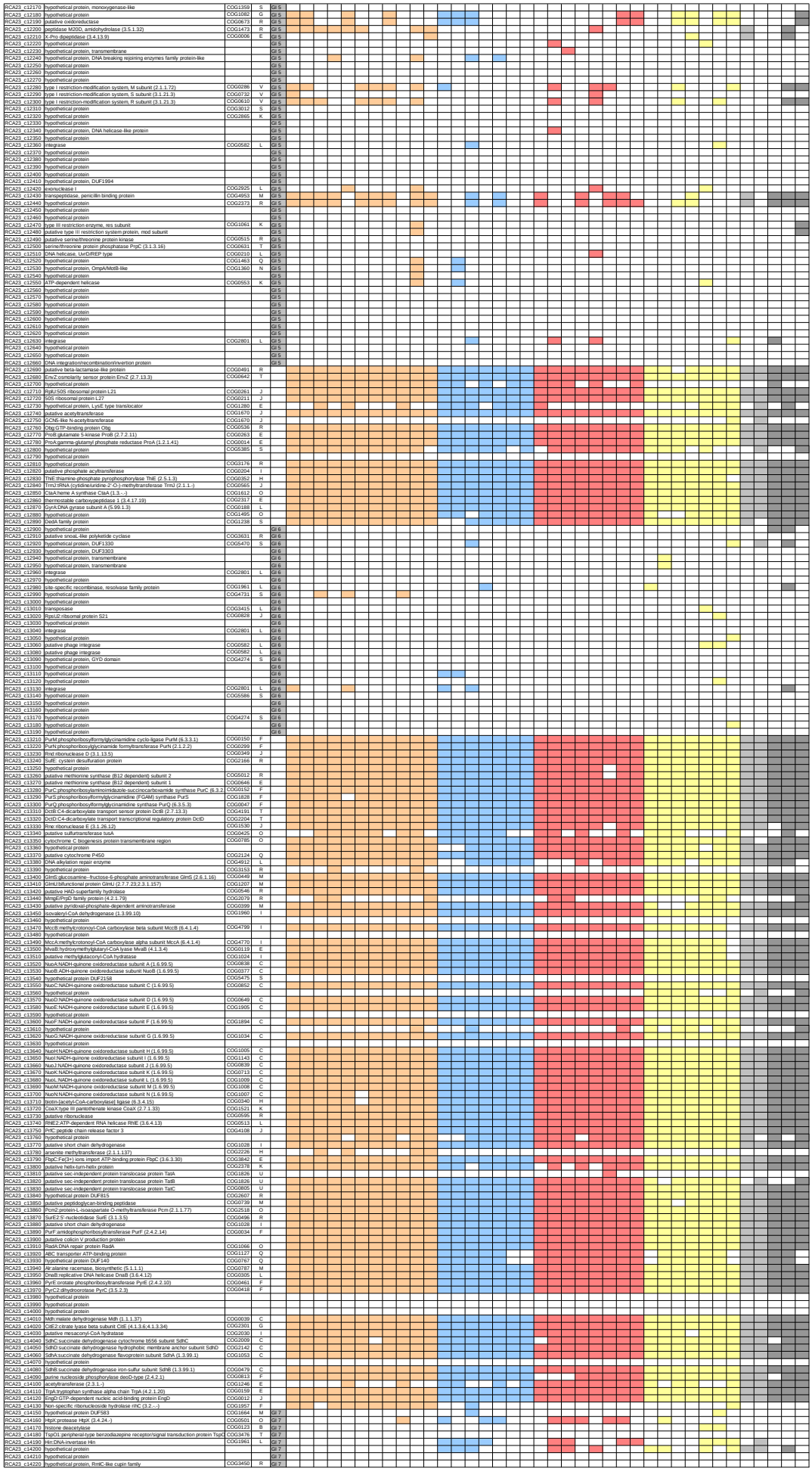


RCA23_c06030	hypothetical protein	COG3743	S
RCA23_c06040	hypothetical protein DUF1244	COG3492	S
RCA23_c06050	N-formylglutamate amidohydrolase	COG3931	E
RCA23_c06060	PykF:pyruvate kinase PykF (2.7.1.40)	COG0469	G
RCA23_c06070	hypothetical protein		
RCA23_c06080	RpmI:50S ribosomal protein L35	COG0291	J
RCA23_c06090	RplT:50S ribosomal protein L20	COG0292	J
RCA23_c06100	putative subtilase family protein	COG1404	O
RCA23_c06110	hypothetical protein		
RCA23_c06120	hypothetical protein, lipid A biosynthesis acyltransferase	COG1560	M
RCA23_c06130	PheS:phenylalanyl-tRNA synthase alpha chain PheS (6.1.1.20)	COG0016	J
RCA23_c06140	glutamine amidotransferase class-I (6.3.5.2)	COG0518	F
RCA23_c06150	PheT:phenylalanyl-tRNA synthase beta chain PheT (6.1.1.20)	COG0072	J
RCA23_c06160	putative HTH-type transcriptional regulator	COG1522	K
RCA23_c06170	RpsU:ribosomal protein S21	COG0828	J
RCA23_c06180	putative ubiquinone biosynthesis protein COQ9	COG5590	S
RCA23_c06190	putative quinone oxidoreductase	COG0604	C
RCA23_c06200	hypothetical protein DUF1013	COG3820	S
RCA23_c06210	RecR:recombination protein RecR	COG0353	L
RCA23_c06220	hypothetical protein	COG0718	S
RCA23_c06230	DnaX:DNA polymerase III subunit tau (2.7.7.7)	COG2812	L
RCA23_c06240	NudC:NADH pyrophosphatase NudC (3.6.1.22)	COG2816	L
RCA23_c06250	hypothetical protein	COG3832	S
RCA23_c06260	PheA:prephenate dehydratase PheA (4.2.1.51)	COG0077	E
RCA23_c06270	CycM:cytochrome c-552	COG3474	C
RCA23_c06280	ABC transporter extracellular solute-binding protein	COG4166	E
RCA23_c06290	ABC transporter permease protein	COG4174	R
RCA23_c06300	ABC transporter permease protein	COG4239	R
RCA23_c06310	putative oligopeptide ABC transporter ATP- binding protein	COG4172	R
RCA23_c06320	fumarylacetoacetate hydrolase family protein	COG0179	Q
RCA23_c06330	DacF:D-alanyl-D-alanine carboxypeptidase DacF (3.4.16.4)	COG1686	M
RCA23_c06340	haloacid dehalogenase domain protein hydrolase (3.-.-.-)	COG1011	R
RCA23_c06350	ClpS:ATP-dependent Clp protease adapter protein ClpS	COG2127	S
RCA23_c06360	putative methyltransferase	COG2813	J
RCA23_c06370	putative short chain dehydrogenase	COG1028	I
RCA23_c06380	HemF:coproporphyrinogen 3 oxidase, aerobic (1.3.3.3)	COG0408	H
RCA23_c06390	oxidoreductase, FAD-binding protein	COG4097	P
RCA23_c06400	D-lactate dehydrogenase	COG0277	C
RCA23_c06420	integrase		GI 3
RCA23_c06430	mandelate racemase (5.1.2.2)	COG4948	M
RCA23_c06440	fumarylacetoacetate hydrolase family protein	COG0179	Q
RCA23_c06450	short-chain dehydrogenase/reductase	COG1028	I
RCA23_c06460	amidohydrolase	COG3618	R
RCA23_c06470	aldo/keto reductase	COG0667	C
RCA23_c06480	ABC transporter, permease protein	COG0600	P
RCA23_c06490	ABC transporter, permease protein	COG0600	P
RCA23_c06500	hypothetical protein		
RCA23_c06510	sulfonate/nitrate ABC transporter, ATPase	COG1116	P
RCA23_c06520	ABC transporter, periplasmic substrate-binding protein	COG0715	P
RCA23_c06530	arylsulfatase	COG3119	P
RCA23_c06540	transcriptional regulator, GntR family	COG1802	K
RCA23_c06550	hypothetical protein	COG4974	L
RCA23_c06560	hypothetical protein	COG1618	F
RCA23_c06570	hypothetical protein		
RCA23_c06580	hypothetical protein		
RCA23_c06590	hypothetical protein		
RCA23_c06600	hypothetical protein		
RCA23_c06610	hypothetical protein		
RCA23_c06620	hypothetical protein		
RCA23_c06630	hypothetical protein	COG3000	I
RCA23_c06640	transcriptional regulator, AsnC family	COG1522	K
RCA23_c06650	AspC1:aspartate aminotransferase AspC (2.6.1.1)	COG0436	E
RCA23_c06660	Bdh:3-hydroxybutyrate dehydrogenase Bdh (1.1.1.-;1.1.1.30)	COG1028	I
RCA23_c06670	putative fumarylacetoacetate hydrolase	COG0179	Q
RCA23_c06680	oxidoreductase, short chain dehydrogenase/reductase family protein	COG1028	I
RCA23_c06690	dimethylmenaquinone methyltransferase	COG0684	H
RCA23_c06700	hypothetical protein	COG3970	R
RCA23_c06710	AldH:NADP-dependent fatty aldehyde dehydrogenase AldH (1.2.1.4)	COG1012	C
RCA23_c06720	putative dihydrodipicolinate synthase	COG0329	E
RCA23_c06730	transcriptional regulator protein, LacI family	COG1609	K
RCA23_c06740	DctP2:TRAP dicarboxylate transporter, subunit DctP	COG1638	G
RCA23_c06750	DctQ2:TRAP dicarboxylate transporter, subunit DctQ	COG3090	G
RCA23_c06760	DctM2:TRAP dicarboxylate transporter, subunit DctM	COG1593	G
RCA23_c06770	DadA1,D-amino acid dehydrogenase small subunit DadA (1.4.99.1)	COG0665	E
RCA23_c06780	mandelate racemase/muconate lactonizing protein	COG4948	M
RCA23_c06790	DapA3:dihydrodipicolinate synthase DapA (4.2.1.52)	COG0329	E
RCA23_c06800	uncharacterized oxidoreductase, YjmC	COG2055	C
RCA23_c06810	SoxB1:sarcosine oxidase beta subunit SoxB (1.5.3.1)	COG0665	E
RCA23_c06820	SoxD1:sarcosine oxidase delta subunit SoxD (1.5.3.1)	COG4311	E
RCA23_c06830	SoxA1:sarcosine oxidase alpha subunit SoxA (1.5.3.1)	COG0404	E
RCA23_c06840	SoxG1:sarcosine oxidase gamma subunit SoxG (1.5.3.1)	COG4583	E
RCA23_c06850	hypothetical protein, DUF6 transmembrane protein	COG0697	G
RCA23_c06860	putative phage integrase		GI 3
RCA23_c06870	putative amidase	COG0154	J
RCA23_c06880	integrase	COG4974	L
RCA23_c06890	putative helicase		GI 3
RCA23_c06900	transposase	COG3316	L
RCA23_c06910	HTH-type transcriptional regulator, LuxR family	COG2197	T
RCA23_c06920	putative oxidoreductase, molybdopterin binding	COG3915	S
RCA23_c06930	signal transduction histidine kinase	COG0642	T
RCA23_c06940	RlbA:glucose-1-phosphate thymidylyltransferase RlbA (2.7.7.24)	COG1209	M
RCA23_c06950	hypothetical protein, chain length determinant protein	COG0489	D
RCA23_c06960	polysaccharide export protein	COG1596	M
RCA23_c06970	hypothetical protein, VanZ-like	COG5652	S
RCA23_c06980	type I secretion system membrane fusion protein, HlyD family	COG1566	V
RCA23_c06990	type I secretion system ATP-binding component, HlyB family	COG2274	V
RCA23_c07000	type I secretion outer membrane protein, TolC family	COG1538	M
RCA23_c07010	putative serralysin-like metalloprotease	COG2931	Q
RCA23_c07020	RlbD:dTDP-4-dehydrorhamnose reductase RlbD (1.1.1.133)	COG1091	M
RCA23_c07030	RlbB:dTDP-glucose 4,6-dehydratase RlbB (4.2.1.46)	COG1088	M
RCA23_c07040	RlbC:dTDP-4-dehydrorhamnose 3,5-epimerase RlbC (5.1.3.13)	COG1898	M
RCA23_c07050	hypothetical protein		
RCA23_c07060	glycosyltransferase	COG0438	M
RCA23_c07070	putative glycosyltransferase	COG0438	M
RCA23_c07080	hypothetical protein, UDP- glycosyltransferase/glycogen phosphorylase-like	COG0438	M
RCA23_c07090	polysaccharide biosynthesis protein		
RCA23_c07100	AmlB2:UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase AmlB	COG0399	M
RCA23_c07110	hypothetical, WxcM-like	COG0110	R
RCA23_c07120	hypothetical protein, WxcM-like		
RCA23_c07130	hypothetical protein, acyltransferase-like	COG1670	J
RCA23_c07140	UDP-glucose/GDP-mannose dehydrogenase family protein	COG0677	M
RCA23_c07150	S-adenosyl-L-methionine-dependentmethyltransferas e		
RCA23_c07160	LspL1:UDP-glucuronate 5'-epimerase LspL (5.1.3.12)	COG0451	M
RCA23_c07170	transcriptional regulator, MarR family	COG1846	K
RCA23_c07180	NusG2:transcription antitermination protein NusG	COG0250	K
RCA23_c07190	hypothetical protein	COG1434	S
RCA23_c07200	LpcC:lipopolysaccharide core biosynthesis mannosyltransferase LpcC (2.4.-.-)	COG0438	M
RCA23_c07210	hypothetical protein		
RCA23_c07220	undecaprenyl-phosphate alpha-N- acetylglucosaminyl 1-phosphate transferase (	COG0472	M
RCA23_c07230	hypothetical protein, transmembrane		
RCA23_c07240	hypothetical protein, transmembrane		
RCA23_c07250	transposase, IS4 family protein		
RCA23_c07260	putative transposase	COG3039	L
RCA23_c07270	hypothetical protein, transmembrane		
RCA23_c07280	hypothetical protein, transmembrane		
RCA23_c07290	transposase	COG3316	L
RCA23_c07300	GlgA:glycogen synthase GlgA (2.4.1.21)	COG0297	G
RCA23_c07310	GlgX:glycogen debranching enzyme GlgX (3.2.1.-)	COG1523	G
RCA23_c07320	Pgm:phosphoglucomutase Pgm (5.4.2.2)	COG0033	G
RCA23_c07330	putative alpha-glucosidase AglA (3.2.1.20)	COG0366	G
RCA23_c07340	transposase	COG3316	L
RCA23_c07350	YtnL:uncharacterized hydrolase YtnL	COG1473	R
RCA23_c07360	GsiB:glutathione-binding protein GsiB	COG0747	E
RCA23_c07370	GsiC:glutathione transport system permease protein GsiC	COG0601	E
RCA23_c07380	dpeptide transport system permease protein	COG1173	E
RCA23_c07390	GsiA2:glutathione import ATP-binding protein GsiA (3.6.3.-)	COG4172	R
RCA23_c07400	peptidase family S58	COG3191	E
RCA23_c07410	AldA:aldehyde dehydrogenase AldA (1.2.1.3)	COG1012	C
RCA23_c07420	FabG2:3-oxoacyl-[acyl-carrier-protein] reductase FabG (1.1.1.100)	COG1028	I
RCA23_c07430	amidase (3.5.1.4)	COG0154	J
RCA23_c07440	MFS-type transporter	COG2807	P
RCA23_c07450	peptidase M20/D amidohydrolase	COG1473	R
RCA23_c07460	transposase	COG3316	L
RCA23_c07470	integrase	COG2801	L
RCA23_c07480	transposase		GI 3
RCA23_c07490	hypothetical protein, DUF3764		
RCA23_c07500	hypothetical protein		
RCA23_c07510	hypothetical protein, beta-lactam-insensitive peptidoglycan transpeptidase YkuD	COG1376	S
RCA23_c07520	hypothetical protein		
RCA23_c07530	ABC transporter, ATP-binding protein, SbmA/BacA- like family	COG1133	I
RCA23_c07540	hypothetical protein		
RCA23_c07550	hypothetical protein		
RCA23_c07560	integrase	COG2801	L
RCA23_c07570	hypothetical protein		
RCA23_c07580	hypothetical protein		
RCA23_c07590	FlC:flagellin FlC	COG1344	N
RCA23_c07600	RpoN:RNA polymerase sigma-54	COG1508	K
RCA23_c07610	hypothetical protein, cyclic nucleotide-binding- like		
RCA23_c07620	hypothetical protein	COG0664	T
RCA23_c07630	hypothetical protein		
RCA23_c07640	hypothetical protein		
RCA23_c07650	hypothetical protein		
RCA23_c07660	FlhB:flagellar biosynthesis protein FlhB	COG1377	N
RCA23_c07670	FlIR:flagellar biosynthetic protein FlIR	COG1684	N
RCA23_c07680	FlIQ:flagellar biosynthetic protein FlIQ	COG1987	N
RCA23_c07690	FlIP:flagellar biosynthetic protein FlIP	COG1338	N
RCA23_c07700	FlIN:flagellar motor switch protein FlIN	COG1886	N
RCA23_c07710	FlIM:flagellar motor switch proteins FlIM	COG1868	N
RCA23_c07720	FlIL1 protein FlIL	COG1580	N
RCA23_c07730	putative flagellar hook-length control protein FlIK	COG3144	N
RCA23_c07740	putative flagellar export protein FlIJ		
RCA23_c07750	FlI:flagellar protein export ATPase FlI (3.6.3.14)	COG1157	N
RCA23_c07760	putative flagellar assembly protein FlIH		
RCA23_c07770	FlG:flagellar motor switch protein G	COG1536	N
RCA23_c07780	FlF:flagellar M-ring protein FlF	COG1766	N
RCA23_c07790	FlE:flagellar hook-basal body complex subunit FlE	COG1677	N
RCA23_c07800	FlD:transcriptional regulatory protein FlD	COG2204	T
RCA23_c07810	FlL2 protein FlL	COG1580	N
RCA23_c07820	hypothetical protein		
RCA23_c07830	FlS:flagellar protein FlS	COG1516	N
RCA23_c07840	hypothetical protein		
RCA23_c07850	hypothetical protein		
RCA23_c07860	FlgG1:flagellar basal-body rod protein FlgG	COG4786	N
RCA23_c07870	Ch- channel, voltage-gated family protein (3.6.1.-)	COG0038	P
RCA23_c07880	hypothetical protein, lysozym-like	COG0741	M
RCA23_c07890	FlgL:flagellar hook-associated protein FlgL	COG1344	N
RCA23_c07900	FlgK:flagellar hook-associated protein FlgK	COG1256	N
RCA23_c07910	FlgJ:flagellar rod assembly protein/muramidase FlgJ	COG3951	M
RCA23_c07920	FlgI:flagellar P-ring protein FlgI	COG1706	N
RCA23_c07930	FlgH:flagellar L-ring protein FlgH	COG2063	N
RCA23_c07940	FlgG2:flagellar basal-body rod protein FlgG	COG4786	N
RCA23_c07950	putative flagellar basal body rod protein FlgF	COG4786	N
RCA23_c07960	FlgE:flagellar hook protein FlgE	COG1749	N
RCA23_c07970	FlgD:flagellar basal body rod modification protein FlgD	COG1843	N
RCA23_c07980	FlgC:flagellar basal-body rod protein FlgC	COG1558	N
RCA23_c07990	FlgB:flagellar basal-body rod protein FlgB	COG1815	N
RCA23_c08000	hypothetical protein		
RCA23_c08010	hypothetical protein		
RCA23_c08020	FlA:RNA polymerase, sigma factor for flagellar operon FlA	COG1191	K
RCA23_c08030	putative flagellar biosynthesis protein FlhG	COG0455	D
RCA23_c08040	FlhF:flagellar biosynthesis protein FlhF	COG1419	N
RCA23_c08050	FlhA:flagellar biosynthesis protein FlhA	COG1298	N

RCA23_c08060	hypothetical membrane lipoprotein, DUF400		GI 3
RCA23_c08070	hypothetical protein		GI 3
RCA23_c08080	FlgA-flagellar basal body P-ring biosynthesis protein FlgA	COG1261	N GI 3
RCA23_c08090	putative negative regulator of flagellin synthesis FlgM		GI 3
RCA23_c08100	hypothetical protein		GI 3
RCA23_c08110	hypothetical protein, HCP-like	COG0790	R GI 3
RCA23_c08120	hypothetical protein		GI 3
RCA23_c08130	sigma54 specific transcriptional regulator, Fis family	COG2204	T GI 3
RCA23_c08140	FliD;flagellar hook-associated protein FliD	COG1345	N GI 3
RCA23_c08150	MotB;chemotaxis protein MotB	COG1360	N GI 3
RCA23_c08160	MotA;chemotaxis protein MotA	COG1291	N GI 3
RCA23_c08170	hypothetical protein, DUF1566		GI 3
RCA23_c08180	putative transmembrane protein, DUF6		GI 3
RCA23_c08190	putative polyketide hydroxylase SchC (1.14.13.-)	COG0654	H GI 3
RCA23_c08200	aldehyde dehydrogenase (1.2.1.3)	COG1012	C
RCA23_c08210	Acmsd;2-amino-3-carboxymuconate-6-semialdehyde decarboxylase (4.1.1.45)	COG2159	R
RCA23_c08220	transcriptional regulator, IclR family	COG1414	K
RCA23_c08230	TRAP transporter, 4TM/12TM fusion protein	COG4666	R
RCA23_c08240	TRAP transporter solute receptor, TAXI family	COG2358	R
RCA23_c08250	thiamine pyrophosphate enzyme-like TPP-binding	COG0028	E
RCA23_c08260	hypothetical protein	COG2268	S
RCA23_c08270	short chain dehydrogenase (1.1.1.-)	COG1028	I
RCA23_c08280	putative aldehyde dehydrogenase yfmT (1.2.1.3)	COG1012	C
RCA23_c08290	branched-chain amino acid ABC transporter, ATP- binding protein	COG0410	E
RCA23_c08300	branched-chain amino acid ABC transporter, ATP- binding protein	COG0411	E
RCA23_c08310	putative transporter, permease protein	COG4177	E
RCA23_c08320	putative transporter, permease protein	COG0559	E
RCA23_c08330	putative transporter, periplasmic binding protein	COG0683	E
RCA23_c08340	hypothetical protein, PrpF protein-like	COG2828	S
RCA23_c08350	4-carboxy-4-hydroxy-2-oxoadipate aldolase/oxaloacetate decarboxylase (4.1.1.-)	COG0684	H
RCA23_c08360	putative N-acetylglucosaminyl- phosphatidylinositol de-N-acetylase (3.5.1.89)	COG2120	S
RCA23_c08370	hypothetical protein		
RCA23_c08380	transcriptional regulator	COG1802	K
RCA23_c08390	GcvT1;aminomethyltransferase GcvT (2.1.2.10)	COG0404	E
RCA23_c08400	DctP3 TRAP dicarboxylate transporter, subunit DctP	COG1638	G
RCA23_c08410	putative TRAP transporter, DctQ component	COG3090	G
RCA23_c08420	DctM3 TRAP dicarboxylate transporter, subunit DctM	COG1593	G
RCA23_c08430	XylE;metapyrocatechase XylE (1.13.11.2)	COG2514	R
RCA23_c08440	MFS-type transporter	COG2814	G
RCA23_c08450	hypothetical protein, cytochrome c		
RCA23_c08460	hypothetical protein, copper resistance protein C	COG2372	R
RCA23_c08470	hypothetical protein, copper resistance protein D	COG1276	P
RCA23_c08480	hypothetical protein	COG3613	F
RCA23_c08490	hypothetical protein		
RCA23_c08500	hypothetical protein		
RCA23_c08510	fumarate reductase flavoprotein subunit (1.3.99.1)	COG1053	C
RCA23_c08520	PrpB;methylisocitrate lyase PrpB (4.1.3.30)	COG2513	G
RCA23_c08530	putative isochorismatase family protein	COG1335	Q
RCA23_c08540	HyuA1 hydantoin utilization protein A (3.5.2.-)	COG0145	E
RCA23_c08550	HyuB1 hydantoin utilization protein B (3.5.2.14)	COG0146	E
RCA23_c08560	hypothetical protein	COG1942	R
RCA23_c08570	HTH-type transcriptional regulator, GntR family	COG2188	K
RCA23_c08580	LeuC2 3-isopropylmalate dehydratase large subunit LeuC (4.2.1.33)	COG0065	E
RCA23_c08590	LeuD2 3-isopropylmalate dehydratase small subunit LeuD (4.2.1.33)	COG0066	E
RCA23_c08600	DctP4 TRAP dicarboxylate transporter, subunit DctP	COG1638	G
RCA23_c08610	putative TRAP transporter, DctQ component	COG3090	G
RCA23_c08620	DctM4 TRAP dicarboxylate transporter, subunit DctM	COG1593	G
RCA23_c08630	putative glutamate synthase [NADPH] small chain	COG0493	E
RCA23_c08640	hypothetical protein, dihydrooorotate dehydrogenase	COG0167	F
RCA23_c08650	putative HTH-type transcriptional regulator	COG1309	K
RCA23_c08660	AmaB;N-carbamoyl-L-amino acid hydrolase AmaB (3.5.1.87)	COG0624	E
RCA23_c08670	Dht;D-hydantoinase/dihydropyrimidinase Dht (3.5.2.2)	COG0044	F
RCA23_c08680	ABC transporter ATP-binding protein	COG1116	P
RCA23_c08690	putative ABC transporter permease protein	COG0600	P
RCA23_c08700	putative ABC transporter permease protein	COG0600	P
RCA23_c08710	putative thiamine biosynthesis protein	COG0715	P
RCA23_c08720	putative integral membrane protein DUF6	COG0697	G
RCA23_c08730	hypothetical protein DUF6	COG0697	G
RCA23_c08740	PhoR;alkaline phosphatase synthesis sensor protein PhoR (2.7.13.3)	COG5002	T
RCA23_c08750	putative phosphate binding protein PstS	COG0226	P
RCA23_c08760	putative phosphate transport system permease protein PstC	COG0573	P
RCA23_c08770	putative phosphate transport system permease protein PstA	COG0581	P
RCA23_c08780	PstB;phosphate import ATP-binding protein PstB (3.6.3.27)	COG1117	P
RCA23_c08790	PhoU;phosphate transport system regulatory protein PhoU	COG0704	P
RCA23_c08800	PhoB;phosphate regulon transl protein PhoB	COG0745	T
RCA23_c08810	putative hippurate hydrolase protei (3.5.1.32)	COG1473	R
RCA23_c08820	UreD;urease accessory protein UreD	COG0829	O
RCA23_c08830	UreA;urease gamma subunit UreA (3.5.1.5)	COG0831	E
RCA23_c08840	UreB;urease beta subunit UreB (3.5.1.5)	COG0832	E
RCA23_c08850	UreC;urease alpha subunit UreC (3.5.1.5)	COG0804	E
RCA23_c08860	UreE;urease accessory protein UreE	COG2371	O
RCA23_c08870	UreF;urease accessory protein UreF	COG0830	O
RCA23_c08880	UreG;urease accessory protein UreG	COG0378	O
RCA23_c08890	integrase	COG2801	L
RCA23_c08900	hypothetical protein	COG4585	T
RCA23_c08910	Alkj1;glucose-methanol-choline oxidoreductase Alkj (1.1.99.-)	COG2303	E
RCA23_c08920	hypothetical protein, YaeQ family protein	COG4681	S
RCA23_c08930	hypothetical protein, DUF6 transmembrane protein	COG0697	G
RCA23_c08940	Ugd;UDP-glucose 6-dehydrogenase Ugd (1.1.1.22)	COG1004	M
RCA23_c08950	WaaA2;3-deoxy-D-manno-octulosonic-acid transferase WaaA (2.4.-.-)	COG1519	M
RCA23_c08960	WecB;UDP-N-acetylglucosamine 2-epimerase WecB (5.1.3.14)	COG0381	M
RCA23_c08970	putative N-acetylneuramic acid synthase	COG2089	M
RCA23_c08980	hypothetical protein		
RCA23_c08990	hypothetical protein	COG1651	O
RCA23_c09000	BetC;choline-sulfatase BetC (3.1.6.6)	COG3119	P
RCA23_c09010	permease	COG2233	F
RCA23_c09020	hypothetical protein		
RCA23_c09030	membrane dipeptidase (3.4.13.19)	COG2355	E
RCA23_c09040	hypothetical protein		
RCA23_c09050	hypothetical protein		
RCA23_c09060	TruA;tRNA pseudouridine synthase A (5.4.99.12)	COG0101	J
RCA23_c09070	YqiX-like protein	COG3106	R
RCA23_c09080	hypothetical protein, acetyltransferase-like	COG0456	R
RCA23_c09090	hypothetical protein	COG3768	S
RCA23_c09100	IleS;isoleucyl-tRNA synthase IleS (6.1.1.5)	COG0060	J
RCA23_c09110	putative integral membrane protein DUF6		
RCA23_c09120	Pcs;phosphatidylcholine synthase Pcs (2.7.8.24)	COG1183	I
RCA23_c09130	putative tyrosine recombinase xerC	COG4974	L
RCA23_c09140	hypothetical protein	COG3159	S
RCA23_c09150	transaldolase (2.2.1.2)	COG0176	G
RCA23_c09160	PriA;primosomal protein N' (3.6.1.-)	COG1198	L
RCA23_c09170	McmA1;methylmalonyl-CoA mutase McmA (5.4.99.2)	COG1884	I
RCA23_c09180	PlsB;glycerol-3-phosphate acyltransferase PlsB (2.3.1.15)	COG2937	I
RCA23_c09190	Ccr;crotonyl-CoA reductase Ccr	COG0604	C
RCA23_c09200	putative ATPase	COG0507	L
RCA23_c09210	MtbB3;trimethylamine methyltransferase MtbB (2.1.1.-)	COG5598	H
RCA23_c09220	hypothetical protein DUF1052	COG5321	S
RCA23_c09240	hypothetical protein, DUF411	COG3019	R GI 4
RCA23_c09250	hypothetical protein		GI 4
RCA23_c09260	hypothetical protein		GI 4
RCA23_c09270	hypothetical protein	COG4642	S GI 4
RCA23_c09280	hypothetical protein	COG3904	S GI 4
RCA23_c09290	putative DNA-binding protein, transposase-like	COG3415	L GI 4
RCA23_c09300	HTH-type transcriptional regulator, LysR family	COG0583	K GI 4
RCA23_c09310	hypothetical protein		GI 4
RCA23_c09320	hypothetical protein, hydrolase-like	COG0627	R GI 4
RCA23_c09330	putative integrase	COG2801	L GI 4
RCA23_c09340	putative transposase		GI 4
RCA23_c09350	hypothetical protein, ornithine cyclodeaminase- like	COG2423	E
RCA23_c09360	LrgB-like protein	COG1346	M
RCA23_c09370	hypothetical protein	COG1380	R
RCA23_c09380	putative branched-chain-amino-acid aminotransferase IlvE (2.6.1.42)	COG0115	E
RCA23_c09390	GarR1;2-hydroxy-3-oxopropionate reductase GarR (1.1.1.60)	COG0084	I
RCA23_c09400	Bcr;bicyclomycin resistance protein Bcr	COG2814	G
RCA23_c09410	hypothetical protein		
RCA23_c09420	hypothetical protein, photolyase	COG3046	R
RCA23_c09430	DNA photolyase, FAD-binding/cryptochrome	COG0415	L
RCA23_c09440	CoxS1;carbon monoxide dehydrogenase small chain CoxS (1.2.99.2)	COG2080	C
RCA23_c09450	CoxL1;carbon monoxide dehydrogenase large chain CoxL (1.2.99.2)	COG1529	C
RCA23_c09460	CoxM1;carbon monoxide dehydrogenase medium chain CoxM (1.2.99.2)	COG1319	C
RCA23_c09470	putative chaperone coxD, AAA ATPase	COG0714	R
RCA23_c09480	CoxE-like protein	COG3552	R
RCA23_c09490	IolB1 5-dkexy-glucuronate isomerase IolB (5.3.1.-)	COG3718	G
RCA23_c09500	branched-chain amino acid transport protein, AzlD-like	COG4392	S
RCA23_c09510	branched-chain amino acid transport protein, AzlC-like	COG1296	E
RCA23_c09520	alpha-IPM synthase/homocitrate synthase	COG0119	E
RCA23_c09530	cytochrome P450	COG2124	Q
RCA23_c09540	DctP5 TRAP dicarboxylate transporter, subunit DctP	COG1638	G
RCA23_c09550	DctQ3 TRAP dicarboxylate transporter, subunit DctQ	COG3090	G
RCA23_c09560	DctM5 TRAP dicarboxylate transporter, subunit DctM	COG1593	G
RCA23_c09570	putative citrate transporter	COG0471	P
RCA23_c09580	TpiA;triosephosphate isomerase TpiA (5.3.1.1)	COG0149	G
RCA23_c09590	putative regulator protein of competence- specific genes TfoX	COG3070	K
RCA23_c09600	SufA; iron-sulfur cluster assembly protein	COG0316	S
RCA23_c09610	hypothetical protein DUF59	COG2151	R
RCA23_c09620	Tgt;queuine tRNA-ribosyltransferase Tgt (2.4.2.29)	COG0343	J
RCA23_c09630	hypothetical protein		
RCA23_c09640	LonA;ATP-dependent protease La (3.4.21.53)	COG0466	O
RCA23_c09650	putative phosphoglycerate mutase family protein	COG0406	G
RCA23_c09670	PcmI_protein-L-isoaspartate O-methyltransferase Pcm (2.1.1.77)	COG2518	O
RCA23_c09680	outer membrane efflux protein	COG1538	M
RCA23_c09690	hypothetical protein		
RCA23_c09700	CobQ;cobyric acid synthase CobQ	COG1492	H
RCA23_c09710	hypothetical protein		
RCA23_c09720	Efp;translation elongation factor P	COG0231	J
RCA23_c09730	YgfZ;hypothetical protein tRNA modifying protein YgfZ	COG0354	R
RCA23_c09740	putative glycosyltransferase, family 2	COG0463	M
RCA23_c09750	MsbA;lipid A export ATP-binding/permease protein MsbA (3.6.3.-)	COG1132	V
RCA23_c09760	SgaA1;serine--glyoxylate aminotransferase SgaA (2.6.1.45)	COG0075	E
RCA23_c09770	HisC1;histidinol-phosphate aminotransferase HisC (2.6.1.9)	COG0079	E
RCA23_c09780	ValS;valyl-tRNA synthase ValS (6.1.1.9)	COG0525	J
RCA23_c09790	hypothetical protein, metal-dependent phosphohydrolase	COG4341	R
RCA23_c09800	putative phytanoyl-CoA dioxygenase family protein	COG5285	Q
RCA23_c09810	molybdenum-containing hydroxylase	COG1529	C
RCA23_c09820	hypothetical protein DUF2235	COG3673	S
RCA23_c09830	MetF;5,10-methylenetetrahydrofolate reductase MetF (1.5.1.20)	COG0685	E
RCA23_c09840	MetR;HTH-type transcriptional regulator MetR	COG0583	K
RCA23_c09850	SuhB;inositol-1-monophosphatase SuhB (3.1.3.25)	COG0483	G
RCA23_c09860	hypothetical protein	COG3063	N
RCA23_c09880	transcriptional regulator, HxlR family	COG1733	K
RCA23_c09890	hypothetical protein, DUF3764		
RCA23_c09900	hypothetical protein		
RCA23_c09910	hypothetical protein, DsrEIF-like	COG1553	P
RCA23_c09920	Gpo;glutathione peroxidase Gpo (1.11.1.9)	COG0386	O
RCA23_c09930	putative ion channel		
RCA23_c09950	hypothetical protein		
RCA23_c09940	RluD ribosomal large subunit pseudouridine synthase D (5.4.99.-)	COG0564	J
RCA23_c09960	RpoH1;RNA polymerase sigma-32 factor RpoH	COG0568	K
RCA23_c09970	putative oxidoreductase	COG0673	R
RCA23_c09980	YbiC;oligoendopeptidase F (3.4.24.-)	COG1164	E
RCA23_c09990	RhmA;2-keto-3-deoxy-L-rhamnonate aldolase RhmA (4.1.2.-)	COG3836	G
RCA23_c10000	hypothetical protein, lysophospholipase L2	COG2267	I
RCA23_c10010	putative sterol-binding protein	COG3255	I
RCA23_c10020	hypothetical protein		
RCA23_c10030	RhlB;ATP-dependent RNA helicase RhlB	COG4581	L
RCA23_c10040	putative heat shock protein	COG1188	J
RCA23_c10050	FdxA;ferredoxin FdxA	COG1146	C
RCA23_c10060	putative transcriptional regulator, CarD family	COG1329	K
RCA23_c10070	CobS1;cobalamin-5-phosphate synthase CobS (2.-.-.-)	COG0368	H
RCA23_c10080	CobT1;nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	COG2038	H
RCA23_c10090	hypothetical protein		
RCA23_c10100	iphyoalase/bleomycin resistance protein	COG3565	R

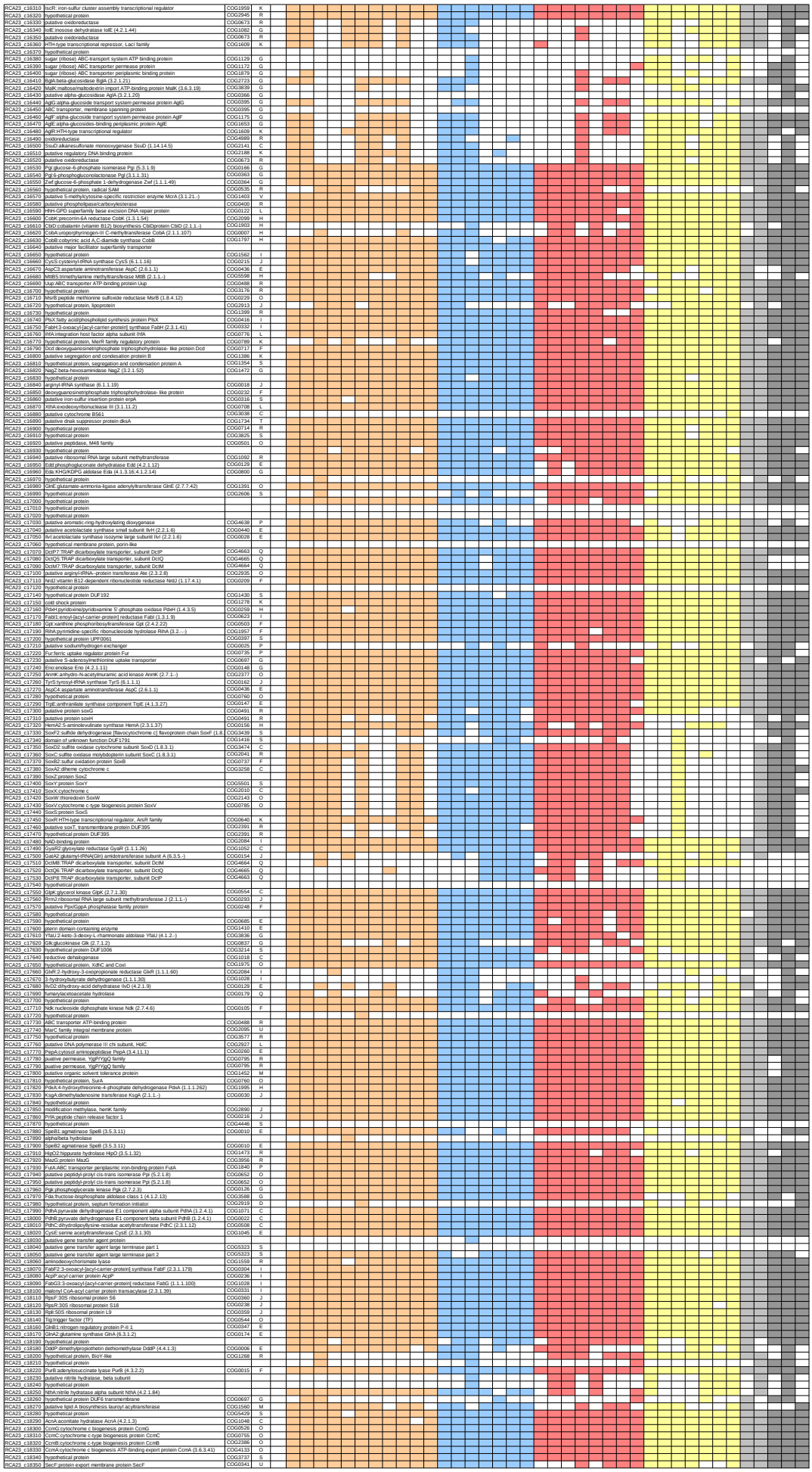


RCA23_c10110	hypothetical protein		
RCA23_c10120	BetA1 choline dehydrogenase BetA (1.1.99.1)	COG2303	E
RCA23_c10140	RlmB:23S rRNA (guanosine-2'-O-)-methyltransferase RlmB (2.1.1-)	COG0566	J
RCA23_c10150	HisE phosphoribosyl-ATP pyrophosphatase HisE (3.6.1.31)	COG0140	E
RCA23_c10160	HisF:imidazole glycerol phosphate synthase subunit HisF (4.1.3-)	COG0107	E
RCA23_c10170	HisA:1-(5-phosphoribosyl)-5-[(5- phosphoribosylamino)methylideneamino] imidaz	COG0106	E
RCA23_c10180	HisH:imidazole glycerol phosphate synthase, glutamine amidotransferase subunit	COG0118	E
RCA23_c10190	HisB:imidazoleglycerol-phosphate dehydratase HisB (4.2.1.19)	COG0131	E
RCA23_c10210	hypothetical protein, calcineurin-like phosphoesterase-like	COG1409	R
RCA23_c10220	Pyc pyruvate carboxylase Pyc (6.4.1.1)	COG1038	C
RCA23_c10230	LldD1:L-lactate dehydrogenase lldD (1.1.2.3)	COG1304	C
RCA23_c10240	putative DNA helicase II (3.6.4.12)	COG0210	L
RCA23_c10250	PuuB1:gamma-glutamylputrescine oxidoreductase PuuB (1.4.3.-)	COG0665	E
RCA23_c10260	putative cysteine/O-acetylserine efflux protein	COG1280	E
RCA23_c10270	hypothetical protein		
RCA23_c10280	putative protein Mrp	COG0489	D
RCA23_c10290	hypothetical protein		
RCA23_c10300	hypothetical protein, MraZ	COG2001	S
RCA23_c10310	MraW:S-adenosyl-L-methionine-dependent methyltransferase MraW (2.1.1.-)	COG0275	M
RCA23_c10320	hypothetical protein	COG5462	S
RCA23_c10330	peptidoglycan synthase FtsI (2.4.1.129)	COG0768	M
RCA23_c10340	MurE:UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2, 6- diaminopimelate ligase	COG0769	M
RCA23_c10350	putative UDP-N-acetylmuramoyl-tripeptide--D- alanyl-D-alanine ligase MurF (6.3	COG0770	M
RCA23_c10360	MraY:phospho-N-acetylmuramoyl-pentapeptide- transferase MraY (2.7.8.13)	COG0472	M
RCA23_c10370	MurD:UDP-N-acetylmuramoylalanine--D-glutamate ligase MurD (6.3.2.9)	COG0771	M
RCA23_c10380	putative glycosyltransferase, sugar binding region	COG3774	M
RCA23_c10390	putative galactoside 2-alpha-L- fucosyltransferase 1		
RCA23_c10400	hypothetical protein		
RCA23_c10410	hypothetical protein	COG3306	M
RCA23_c10420	hypothetical protein H0933	COG2081	R
RCA23_c10430	FtsW:cell division protein FtsW	COG0772	D
RCA23_c10440	MurG:UDP-N-acetylglucosamine--N-acetylmuramyl- (pentapeptide) pyrophospho	COG0707	M
RCA23_c10450	MurC:UDP-N-acetylmuramate--L-alanine ligase MurC (6.3.2.8)	COG0773	M
RCA23_c10460	permease of the drug/metabolite transporter superfamily	COG0697	G
RCA23_c10470	MurB:UDP-N-acetylenolpyruvoylglucosamine reductase MurB (1.1.1.158)	COG0812	M
RCA23_c10480	DdlD-alanine--D-alanine ligase Ddl (6.3.2.4)	COG1181	M
RCA23_c10490	putative cell division protein FtsQ		
RCA23_c10500	FtsA:cell division protein FtsA	COG0849	D
RCA23_c10510	FtsZ:cell division protein FtsZ	COG0206	D
RCA23_c10520	LpxC:UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase LpxC (3.5	COG0774	M
RCA23_c10530	outer membrane assembly lipoprotein	COG4105	R
RCA23_c10540	RecN:DNA repair protein RecN	COG0497	L
RCA23_c10550	hypothetical protein DUF427	COG2343	S
RCA23_c10560	putative Xaa-Pro aminopeptidase	COG0006	E
RCA23_c10570	CobT2:aerobic cobaltochelatase subunit CobT (6.6.1.2)	COG4547	H
RCA23_c10580	CobS2:aerobic cobaltochelatase subunit CobS (6.6.1.2)	COG0714	R
RCA23_c10590	hypothetical protein, DnaJ	COG0484	O
RCA23_c10600	putative stress-induced morphoprotein, BolA type	COG0271	T
RCA23_c10610	GatB:aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit GatB (6.3.5.-)	COG0064	J
RCA23_c10620	PepN:aminopeptidase N (3.4.11.2)	COG0308	E
RCA23_c10630	GlcB:malate synthase GlcB (2.3.3.9)	COG2225	C
RCA23_c10640	hypothetical protein DUF336	COG3193	R
RCA23_c10650	hypothetical protein		
RCA23_c10660	PuuD:gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD (3.5.1.94)	COG2071	R
RCA23_c10670	putative D-beta-hydroxybutyrate dehydrogenase	COG1028	I
RCA23_c10680	PntA:NAD(P) transhydrogenase alpha subunit PntA (1.6.1.2)	COG3288	C
RCA23_c10690	PntB:NAD(P) transhydrogenase beta subunit PntB (1.6.1.2)	COG1282	C
RCA23_c10700	hypothetical protein	COG4949	S
RCA23_c10710	putative acetylornithine deacetylase ArgE	COG0624	E
RCA23_c10720	MoaA1:molybdenum cofactor biosynthesis protein MoaA	COG2896	H
RCA23_c10730	AcsA2:acetyl-coenzyme A synthase AcsA (6.2.1.1)	COG0365	I
RCA23_c10740	hypothetical protein	COG4321	R
RCA23_c10750	hypothetical protein		
RCA23_c10760	FumC:fumarate hydratase class II (4.2.1.2)	COG0114	C
RCA23_c10770	hypothetical protein	COG3814	S
RCA23_c10780	putative chromate transport protein	COG0059	P
RCA23_c10790	hydantoinase / oxoprolinase family protein	COG0145	E
RCA23_c10800	hypothetical protein		
RCA23_c10810	putative quinone oxidoreductase yhdH (1.6.5.-)	COG0604	C
RCA23_c10820	SufS1:cysteine desulfurase SufS (2.8.1.7;4.4.1.16)	COG0520	E
RCA23_c10830	PhrB1:deoxyribodipyrimidine photo-lyase PhrB (4.1.99.3)	COG0415	L
RCA23_c10840	CfaCyclopropane-fatty-acyl-phospholipid synthase Cfa (2.1.1.79)	COG2230	M
RCA23_c10850	NudF:ADP-ribose pyrophosphatase NudF (3.6.1.13)	COG0494	L
RCA23_c10860	CysK1:cysteine synthase CysK (2.5.1.47)	COG0031	E
RCA23_c10870	MscS:mechanosensitive ion channel protein MscS	COG3264	M
RCA23_c10890	putative phage integrase	COG4974	L
RCA23_c10900	hypothetical protein		
RCA23_c10910	hypothetical protein, tetratricopeptide domain TPR-1	COG0457	R
RCA23_c10920	hypothetical protein, resolvase-like	COG1961	L
RCA23_c10930	hypothetical protein, DUF2924		
RCA23_c10940	putative TRAP transporter solute receptor	COG2358	R
RCA23_c10950	TRAP transporter, 4TM/12TM fusion protein	COG4666	R
RCA23_c10960	putative sulfatase Yid3 (3.6.1.-)	COG3119	P
RCA23_c10970	arylsulfatase (3.1.6.1)	COG1234	R
RCA23_c10980	siderophore interactin protein, vibriobactin utilization protein-like	COG2375	P
RCA23_c10990	hypothetical protein, metallo-beta-lactamase	COG2015	Q
RCA23_c11000	DctP6 TRAP dicarboxylate transporter, subunit DctP	COG1638	G
RCA23_c11010	glutathione S-transferase (2.5.1.18)	COG0625	O
RCA23_c11020	DctQ4 TRAP dicarboxylate transporter, subunit DctQ	COG3090	G
RCA23_c11030	DctM6 TRAP dicarboxylate transporter, subunit DctM	COG1593	G
RCA23_c11040	hypothetical protein, DUF6 transmembrane protein	COG0697	G
RCA23_c11050	ferric reductase like transmembrane component family		
RCA23_c11060	hypothetical protein		
RCA23_c11070	hypothetical protein		
RCA23_c11080	sulfatase family protein	COG3119	P
RCA23_c11090	short-chain dehydrogenase/reductase SDR	COG1028	I
RCA23_c11100	MaiA:maleylacetoacetate isomerase MaiA (5.2.1.2)	COG0625	O
RCA23_c11110	salicylaldehyde dehydrogenase (1.2.1.65)	CO10012	C
RCA23_c11120	transcriptional regulator, GntR family	COG1802	K
RCA23_c11130	transposase	COG3316	L
RCA23_c11140	putative phage integrase	COG4974	L
RCA23_c11150	hypothetical protein, alginate lyase-lyase		
RCA23_c11160	HTH-type transcriptional regulator, GntR family	COG1802	K
RCA23_c11170	branched-chain amino acid ABC transporter, periplasmic substrate-binding prote	COG0683	E
RCA23_c11180	branched-chain amino acid ABC transporter, permease protein	COG0559	E
RCA23_c11190	branched-chain amino acid ABC transporter, permease protein	COG4177	E
RCA23_c11200	branched-chain amino acid ABC transporter, ATP- binding protein	COG0411	E
RCA23_c11210	branched-chain amino acid ABC transporter, ATP- binding protein	COG0410	E
RCA23_c11220	HyuE:hydantoin racemase HyuE (5.1.99.5)	COG4126	E
RCA23_c11230	transposase	COG3316	L
RCA23_c11240	GlnQ:glutamine transport ATP-binding protein GlnQ	COG1126	E
RCA23_c11250	putative inner membrane amino-acid ABC transporter permease protein	COG0765	E
RCA23_c11260	transposase	COG3316	L
RCA23_c11270	CoxC:transcriptional regulator CoxC	COG1300	T
RCA23_c11280	CoxM2:carbon monoxide dehydrogenase medium chain CoxM (1.2.99.2)	COG1319	C
RCA23_c11290	CoxS2:carbon monoxide dehydrogenase small chain CoxS (1.2.99.2)	COG2080	C
RCA23_c11300	CoxL2:carbon monoxide dehydrogenase large chain CoxL (1.2.99.2)	COG1529	C
RCA23_c11310	CoxD:AAA+ ATPase chaperone CoxD	COG0714	R
RCA23_c11320	CoxE:carbon monoxide dehydrogenase accessory protein CoxE	COG3552	R
RCA23_c11330	CoxF:carbon monoxide dehydrogenase accessory protein CoxF	COG1975	O
RCA23_c11340	hypothetical protein	COG2068	R
RCA23_c11350	CoxG:carbon monoxide dehydrogenase protein CoxG	COG3427	S
RCA23_c11360	CoxI:carbon monoxide dehydrogenase accessory protein CoxI	COG1975	O
RCA23_c11370	transposase	COG3316	L
RCA23_c11380	dienelactone hydrolase (3.1.1.45)	COG0412	Q
RCA23_c11390	aldo/keto reductase	COG0667	C
RCA23_c11400	MdlA:mandelate racemase MdlA (5.1.2.2)	COG4948	M
RCA23_c11410	hypothetical protein DUF1498	COG3822	R
RCA23_c11420	two-component system, sensor histidine kinase protein	COG0642	T
RCA23_c11430	two-component system, response regulator protein	COG0745	T
RCA23_c11440	hypothetical protein	COG3181	S
RCA23_c11450	putative tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S
RCA23_c11460	hypothetical protein	COG2828	S
RCA23_c11470	D-isomer specific 2-hydroxyacid dehydrogenase (1.1.1.95)	COG0111	H
RCA23_c11480	putative ammonia monooxygenase	COG3180	R
RCA23_c11490	transcriptional regulator	COG1802	K
RCA23_c11500	fumarate reductase flavoprotein subunit FccA (1.3.99.1)	COG1053	C
RCA23_c11510	AspC2:aspartate aminotransferase AspC (2.6.1.1)	COG0436	E
RCA23_c11520	hypothetical protein		
RCA23_c11530	GlpD1:aerobic glycerol-3-phosphate dehydrogenase GlpD (1.1.99.5)	COG0578	C
RCA23_c11540	GlpD2:aerobic glycerol-3-phosphate dehydrogenase GlpD (1.1.99.5)	COG0578	C
RCA23_c11550	hypothetical protein	COG1192	D
RCA23_c11560	CheW:chemotaxis protein CheW	COG0835	N
RCA23_c11570	Tar:methyl-accepting chemotaxis protein II	COG0840	N
RCA23_c11580	CheA:chemotaxis histidine kinase CheA (2.7.13.3)	COG0643	N
RCA23_c11590	response regulator	COG2114	T
RCA23_c11600	CheR:chemotaxis protein methyltransferase CheR (2.1.1.80)	COG1352	N
RCA23_c11610	CheB:chemotaxis response regulator protein-glutamate methylesterase CheB (3	COG2201	N
RCA23_c11620	metal-dependent hydrolase	COG1235	R
RCA23_c11630	integrase	COG2801	L
RCA23_c11640	methyltransferase, FkbM family		
RCA23_c11650	hypothetical protein, sugar transferase-like		
RCA23_c11660	aldo/keto reductase	COG0667	C
RCA23_c11670	hypothetical protein		
RCA23_c11680	hypothetical protein	COG0790	R
RCA23_c11690	hypothetical protein	COG4430	S
RCA23_c11700	hypothetical protein	COG2755	E
RCA23_c11710	phytanoyl-CoA dioxygenase family protein	COG5285	Q
RCA23_c11720	putative bacterial extracellular solute binding protein, family 3	COG1791	S
RCA23_c11730	hypothetical protein		
RCA23_c11740	SOUL heme-binding protein		
RCA23_c11750	AslA:arylsulfatase precursor (3.1.6.1)	COG3119	P
RCA23_c11760	hypothetical protein DUF583		
RCA23_c11770	integrase	COG2801	L
RCA23_c11780	hypothetical protein	COG3616	E
RCA23_c11790	GlnA1:glutamine synthase GlnA (6.3.1.2)	COG0174	E
RCA23_c11800	putative HTH-type transcriptional regulator, RpiR family	COG1737	K
RCA23_c11810	N-formylglutamate amidohydrolase	COG3931	E
RCA23_c11820	putative isochorismatase family protein	COG1335	Q
RCA23_c11830	branched-chain amino acid ABC transporter, ATP- binding protein	COG0411	E
RCA23_c11840	branched-chain amino acid ABC transporter, periplasmic binding protein	COG0683	E
RCA23_c11850	branched-chain amino acid ABC transporter, ATP- binding protein	COG0410	E
RCA23_c11860	branched-chain amino acid ABC transporter, permease protein	COG0559	E
RCA23_c11870	branched-chain amino acid ABC transporter, permease protein	COG4177	E
RCA23_c11880	acetamidase/formamidase family protein	COG2421	C
RCA23_c11890	hypothetical protein	CO01028	I
RCA23_c11910	putative prophage integrase	COG0582	L
RCA23_c11920	hypothetical protein		
RCA23_c11930	hypothetical protein		
RCA23_c11940	putative AAA ATPase (2.7.7.-)	COG3598	L
RCA23_c11950	hypothetical protein		
RCA23_c11960	hypothetical protein		
RCA23_c11970	hypothetical protein		
RCA23_c11980	hypothetical protein		
RCA23_c11990	branched-chain amino acid ABC transporter, ATP- binding protein	COG0410	E
RCA23_c12000	branched-chain amino acid ABC transporter, ATP- binding protein	COG0411	E
RCA23_c12010	branched-chain amino acid ABC transporter, permease protein	COG4177	E
RCA23_c12020	branched-chain amino acid ABC transporter, permease protein	COG0559	E
RCA23_c12030	branched-chain amino acid ABC transporter, periplasmic binding protein	COG0683	E
RCA23_c12040	4-hroxyl-CoA hydratase/carnithine racemase (4.2.1.-)	COG1024	I
RCA23_c12050	acetyl-CoA synthase-like protein	COG0318	I
RCA23_c12060	dehydrogenase iron-sulfur-binding subunit	COG2080	C
RCA23_c12070	dehydrogenase FAD-binding subunit	COG1319	C
RCA23_c12080	dehydrogenase molybdenum-binding subunit	COG1529	C
RCA23_c12090	putative HTH-type transcriptional regulator, AraC family	COG2207	K
RCA23_c12100	periplasmic binding protein-like	COG1879	G
RCA23_c12110	ROK family transcriptional repressor	COG1940	K
RCA23_c12120	xylose isomerase	COG1082	G
RCA23_c12130	sugar ABC transporter, periplasmic binding protein	COG1879	G
RCA23_c12140	putative oxidoreductase	COG0673	R
RCA23_c12150	sugar ABC transporter, ATP-binding protein	COG1129	G
RCA23_c12160	sugar ABC transporter, permease protein	COG1172	G

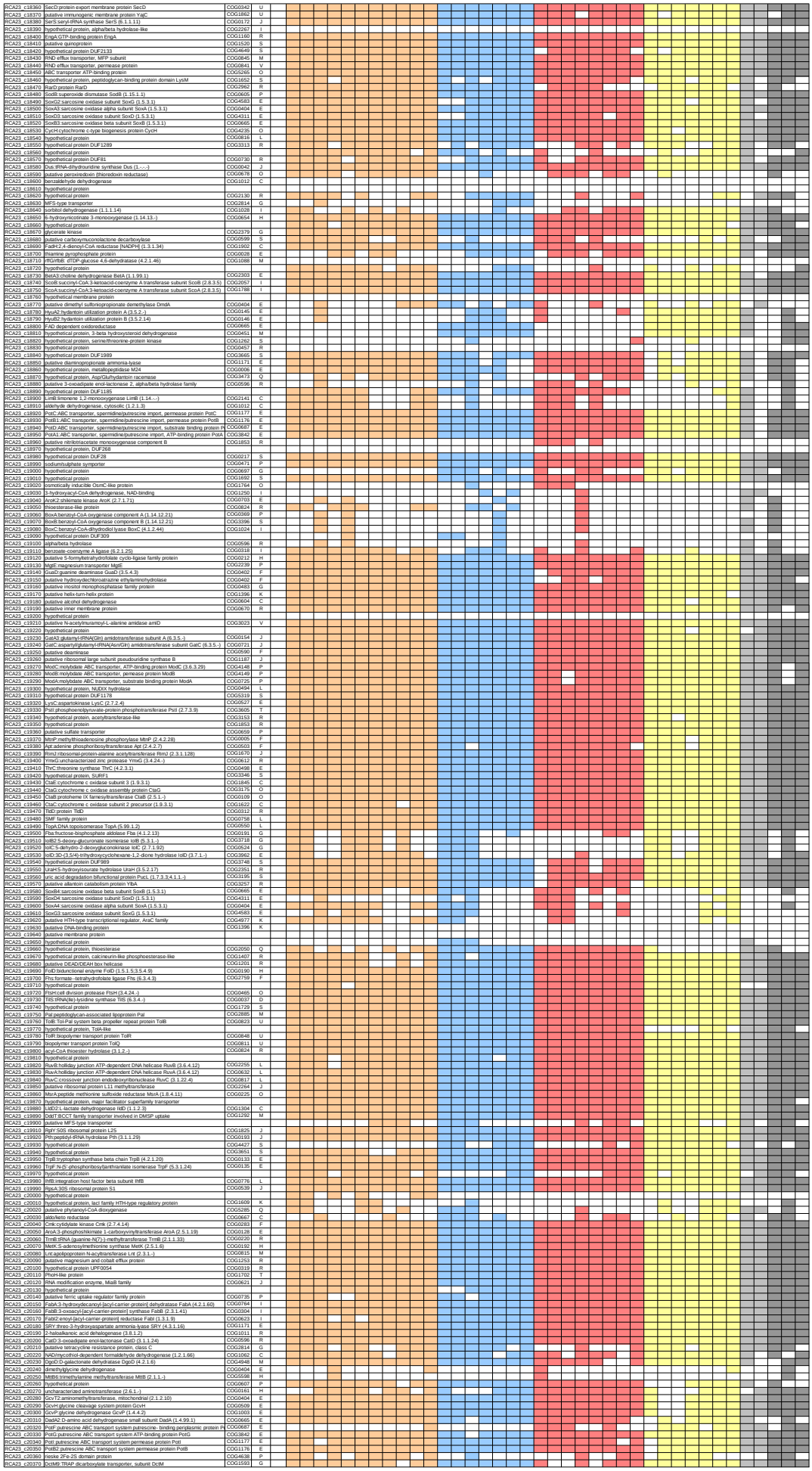


RCA23_c12170	hypothetical protein, monooxygenase-like	COG1359	S
RCA23_c12180	hypothetical protein	COG1082	G
RCA23_c12190	putative oxidoreductase	COG0673	R
RCA23_c12200	peptidase M20D, amidohydrolase (3.5.1.32)	COG1473	R
RCA23_c12210	X-Pro dipeptidase (3.4.13.9)	COG0006	E
RCA23_c12220	hypothetical protein		
RCA23_c12230	hypothetical protein, transmembrane		
RCA23_c12240	hypothetical protein, DNA breaking rejoining enzymes family protein-like		
RCA23_c12250	hypothetical protein		
RCA23_c12260	hypothetical protein		
RCA23_c12270	hypothetical protein		
RCA23_c12280	type I restriction-modification system, M subunit (2.1.1.72)	COG0286	V
RCA23_c12290	type I restriction-modification system, S subunit (3.1.21.3)	COG0732	V
RCA23_c12300	type I restriction-modification system, R subunit (3.1.21.3)	COG0610	V
RCA23_c12310	hypothetical protein	COG3012	S
RCA23_c12320	hypothetical protein	COG2865	K
RCA23_c12330	hypothetical protein		
RCA23_c12340	hypothetical protein, DNA helicase-like protein		
RCA23_c12350	hypothetical protein		
RCA23_c12360	integrase	COG0582	L
RCA23_c12370	hypothetical protein		
RCA23_c12380	hypothetical protein		
RCA23_c12390	hypothetical protein		
RCA23_c12400	hypothetical protein		
RCA23_c12410	hypothetical protein, DUF1994		
RCA23_c12420	exonuclease I	COG2925	L
RCA23_c12430	transpeptidase, penicillin binding protein	COG4953	M
RCA23_c12440	hypothetical protein	COG2373	R
RCA23_c12450	hypothetical protein		
RCA23_c12460	hypothetical protein		
RCA23_c12470	type III restriction enzyme, res subunit	COG1061	K
RCA23_c12480	putative type III restriction system protein, mod subunit		
RCA23_c12490	putative serine/threonine protein kinase	COG0515	R
RCA23_c12500	serine/threonine protein phosphatase PrpC (3.1.3.16)	COG0631	T
RCA23_c12510	DNA helicase, UvrD/REP type	COG0210	L
RCA23_c12520	hypothetical protein	COG1463	Q
RCA23_c12530	hypothetical protein, OmpA/MotB-like	COG1360	N
RCA23_c12540	hypothetical protein		
RCA23_c12550	ATP-dependent helicase	COG0553	K
RCA23_c12560	hypothetical protein		
RCA23_c12570	hypothetical protein		
RCA23_c12580	hypothetical protein		
RCA23_c12590	hypothetical protein		
RCA23_c12600	hypothetical protein		
RCA23_c12610	hypothetical protein		
RCA23_c12620	hypothetical protein		
RCA23_c12630	integrase	COG2801	L
RCA23_c12640	hypothetical protein		
RCA23_c12650	hypothetical protein		
RCA23_c12660	DNA integration/recombination/inversion protein		
RCA23_c12690	putative beta-lactamase-like protein	COG0491	R
RCA23_c12680	EnvZ-osmolarity sensor protein EnvZ (2.7.13.3)	COG0642	T
RCA23_c12700	hypothetical protein		
RCA23_c12710	RplU:50S ribosomal protein L21	COG0261	J
RCA23_c12720	50S ribosomal protein L27	COG0211	J
RCA23_c12730	hypothetical protein, LysE type translocator	COG1280	E
RCA23_c12740	putative acetyltransferase	COG1670	J
RCA23_c12750	GCN5-like N-acetyltransferase	COG1670	J
RCA23_c12760	Obg:GTP-binding protein Obg	COG0536	R
RCA23_c12770	ProB:glutamate 5-kinase ProB (2.7.2.11)	COG0263	E
RCA23_c12780	ProA:gamma-glutamyl phosphate reductase ProA (1.2.1.41)	COG0014	E
RCA23_c12800	hypothetical protein	COG5385	S
RCA23_c12790	hypothetical protein		
RCA23_c12810	hypothetical protein	COG3176	R
RCA23_c12820	putative phosphate acyltransferase	COG0204	I
RCA23_c12830	ThiE:thiamine-phosphate pyrophosphorylase ThiE (2.5.1.3)	COG0352	H
RCA23_c12840	TrmJ:tRNA (cytidine/uridine-2'-O-)-methyltransferase TrmJ (2.1.1.-)	COG0565	J
RCA23_c12850	CtaA:heme A synthase CtaA (1.3.-.-)	COG1612	O
RCA23_c12860	thermostable carboxypeptidase 1 (3.4.17.19)	COG2317	E
RCA23_c12870	GyrA:DNA gyrase subunit A (5.99.1.3)	COG0188	L
RCA23_c12880	hypothetical protein	COG1495	O
RCA23_c12890	DedA family protein	COG1238	S
RCA23_c12900	hypothetical protein		
RCA23_c12910	putative snoaL-like polyketide cyclase	COG3631	R
RCA23_c12920	hypothetical protein, DUF1330	COG5470	S
RCA23_c12930	hypothetical protein, DUF3303		
RCA23_c12940	hypothetical protein, transmembrane		
RCA23_c12950	hypothetical protein, transmembrane		
RCA23_c12960	integrase	COG2801	L
RCA23_c12970	hypothetical protein		
RCA23_c12980	site-specific recombinase, resolvase family protein	COG1961	L
RCA23_c12990	hypothetical protein	COG4731	S
RCA23_c13000	hypothetical protein		
RCA23_c13010	transposase	COG3415	L
RCA23_c13020	RpsU:ribosomal protein S21	COG0828	J
RCA23_c13030	hypothetical protein		
RCA23_c13040	integrase	COG2801	L
RCA23_c13050	hypothetical protein		
RCA23_c13060	putative phage integrase	COG0582	L
RCA23_c13080	putative phage integrase	COG0582	L
RCA23_c13090	hypothetical protein, GYD domain	COG4274	S
RCA23_c13100	hypothetical protein		
RCA23_c13110	hypothetical protein		
RCA23_c13120	hypothetical protein		
RCA23_c13130	integrase	COG2801	L
RCA23_c13140	hypothetical protein	COG5586	S
RCA23_c13150	hypothetical protein		
RCA23_c13160	hypothetical protein		
RCA23_c13170	hypothetical protein	COG4274	S
RCA23_c13180	hypothetical protein		
RCA23_c13190	hypothetical protein		
RCA23_c13210	PurM:phosphoribosylformylglycinemidine cyclo-ligase PurM (6.3.3.1)	COG0150	F
RCA23_c13220	PurN:phosphoribosylglycinamide formyltransferase PurN (2.1.2.2)	COG0299	F
RCA23_c13230	RnsI:ribonuclease D (3.1.13.5)	COG0349	J
RCA23_c13240	SufE: cystein desulfuration protein	COG2166	R
RCA23_c13250	hypothetical protein		
RCA23_c13260	putative methionine synthase (B12 dependent) subunit 2	COG5012	R
RCA23_c13270	putative methionine synthase (B12 dependent) subunit 1	COG0646	E
RCA23_c13280	PurC:phosphoribosylaminoimidazole-succinocarboxamide synthase PurC (6.3.2	COG0152	F
RCA23_c13290	PurS:phosphoribosylformylglycinamidine (FGAM) synthase PurS	COG1828	F
RCA23_c13300	PurQ:phosphoribosylformylglycinamidine synthase PurQ (6.3.5.3)	COG0047	F
RCA23_c13310	DctB:C4-dicarboxylate transport sensor protein DctB (2.7.13.3)	COG4191	T
RCA23_c13320	DctD:C4-dicarboxylate transport transcriptional regulatory protein DctD	COG2204	T
RCA23_c13330	Rnw:ribonuclease E (3.1.26.12)	COG1530	J
RCA23_c13340	putative sulfurtransferase tusA	COG0425	O
RCA23_c13350	cytochrome C biogenesis protein transmembrane region	COG0785	O
RCA23_c13360	hypothetical protein		
RCA23_c13370	putative cytochrome P450	COG2124	Q
RCA23_c13380	DNA alkylation repair enzyme	COG4912	L
RCA23_c13390	hypothetical protein	COG3153	R
RCA23_c13400	GlmS:glucosamine--fructose-6-phosphate aminotransferase GlmS (2.6.1.16)	COG0449	M
RCA23_c13410	GlmU:bifunctional protein GlmU (2.7.7.23;2.3.1.157)	COG1207	M
RCA23_c13420	putative HAD-superfamily hydrolase	COG0546	R
RCA23_c13440	MngE/PnpD family protein (4.2.1.79)	COG2079	R
RCA23_c13430	putative pyridoxal-phosphate-dependent aminotransferase	COG0399	M
RCA23_c13450	Isovaleryl-CoA dehydrogenase (1.3.99.10)	COG1960	I
RCA23_c13460	hypothetical protein		
RCA23_c13470	MccB:methylcrotonoyl-CoA carboxylase beta subunit MccB (6.4.1.4)	COG4799	I
RCA23_c13480	hypothetical protein		
RCA23_c13490	MccA:methylcrotonoyl-CoA carboxylase alpha subunit MccA (6.4.1.4)	COG4770	I
RCA23_c13500	MvaB:hydroxymethylglutaryl-CoA lyase MvaB (4.1.3.4)	COG0119	E
RCA23_c13510	putative methylglutaconyl-CoA hydratase	COG1024	I
RCA23_c13520	NuoA:NADH-quinone oxidoreductase subunit A (1.6.99.5)	COG0838	C
RCA23_c13530	NuoB:ADH-quinone oxidoreductase subunit NuoB (1.6.99.5)	COG0377	C
RCA23_c13540	hypothetical protein DUF2158	COG5475	S
RCA23_c13550	NuoC:NADH-quinone oxidoreductase subunit C (1.6.99.5)	COG0852	C
RCA23_c13560	hypothetical protein		
RCA23_c13570	NuoD:NADH-quinone oxidoreductase subunit D (1.6.99.5)	COG0649	C
RCA23_c13580	NuoE:NADH-quinone oxidoreductase subunit E (1.6.99.5)	COG1905	C
RCA23_c13590	hypothetical protein		
RCA23_c13600	NuoF:NADH-quinone oxidoreductase subunit F (1.6.99.5)	COG1894	C
RCA23_c13610	hypothetical protein		
RCA23_c13620	NuoG:NADH-quinone oxidoreductase subunit G (1.6.99.5)	COG1034	C
RCA23_c13630	hypothetical protein		
RCA23_c13640	NuoH:NADH-quinone oxidoreductase subunit H (1.6.99.5)	COG1005	C
RCA23_c13650	NuoI:NADH-quinone oxidoreductase subunit I (1.6.99.5)	COG1143	C
RCA23_c13660	NuoJ:NADH-quinone oxidoreductase subunit J (1.6.99.5)	COG0839	C
RCA23_c13670	NuoK:NADH-quinone oxidoreductase subunit K (1.6.99.5)	COG0713	C
RCA23_c13680	NuoL:NADH-quinone oxidoreductase subunit L (1.6.99.5)	COG1009	C
RCA23_c13690	NuoM:NADH-quinone oxidoreductase subunit M (1.6.99.5)	COG1008	C
RCA23_c13700	NuoN:NADH-quinone oxidoreductase subunit N (1.6.99.5)	COG1007	C
RCA23_c13710	biotin-[acetyl-CoA-carboxylase] ligase (6.3.4.15)	COG0340	H
RCA23_c13720	CoaX:type III pantothenate kinase CoaX (2.7.1.33)	COG1521	K
RCA23_c13730	putative ribonuclease	COG0595	R
RCA23_c13740	RhlE2:ATP-dependent RNA helicase RhlE (3.6.4.13)	COG0513	L
RCA23_c13750	PrfC:peptide chain release factor 3	COG4108	J
RCA23_c13760	hypothetical protein		
RCA23_c13770	putative short chain dehydrogenase	COG1028	I
RCA23_c13780	arsenite methyltransferase (2.1.1.137)	COG2226	H
RCA23_c13790	FbpC:Fe(3+) ions import ATP-binding protein FbpC (3.6.3.30)	COG3842	E
RCA23_c13800	putative helix-turn-helix protein	COG2378	K
RCA23_c13810	putative sec-independent protein translocase protein TatA	COG1826	U
RCA23_c13820	putative sec-independent protein translocase protein TatB	COG1826	U
RCA23_c13830	putative sec-independent protein translocase protein TatC	COG0805	U
RCA23_c13840	hypothetical protein DUF815	COG2607	R
RCA23_c13850	putative peptidoglycan-binding peptidase	COG0739	M
RCA23_c13860	PcmI:protein-L-isoaspartate O-methyltransferase PcmI (2.1.1.77)	COG1518	O
RCA23_c13870	SurE2:5'-nucleotidase SurE (3.1.3.5)	COG0496	R
RCA23_c13880	putative short chain dehydrogenase	COG1028	I
RCA23_c13890	PurF:amidophosphoribosyltransferase PurF (2.4.2.14)	COG0034	F
RCA23_c13900	putative colicin V production protein		
RCA23_c13910	RadA:DNA repair protein RadA	COG1066	O
RCA23_c13920	ABC transporter ATP-binding protein	COG1127	Q
RCA23_c13930	hypothetical protein DUF140	COG0767	Q
RCA23_c13940	Alr:alanine racemase, biosynthetic (5.1.1.1)	COG0787	M
RCA23_c13950	DnaB:replicative DNA helicase DnaB (3.6.4.12)	COG0305	L
RCA23_c13960	PyrE:orotate phosphoribosyltransferase PyrE (2.4.2.10)	COG0461	F
RCA23_c13970	PyrC:2-dihydroorotase PyrC (3.5.2.3)	COG0418	F
RCA23_c13980	hypothetical protein		
RCA23_c13990	hypothetical protein		
RCA23_c14000	hypothetical protein		
RCA23_c14010	Mdh:malate dehydrogenase Mdh (1.1.1.37)	COG0039	C
RCA23_c14020	CitE2:citrate lyase beta subunit CitE (4.1.3.6;4.1.3.34)	COG2301	G
RCA23_c14030	putative mesaconyl-CoA hydratase	COG2030	I
RCA23_c14040	SdhC:succinate dehydrogenase cytochrome b556 subunit SdhC	COG2009	C
RCA23_c14050	SdhD:succinate dehydrogenase hydrophobic membrane anchor subunit SdhD	COG2142	C
RCA23_c14060	SdhA:succinate dehydrogenase flavoprotein subunit SdhA (1.3.99.1)	COG1053	C
RCA23_c14070	hypothetical protein		
RCA23_c14080	SdhB:succinate dehydrogenase iron-sulfur subunit SdhB (1.3.99.1)	COG0479	C
RCA23_c14090	purine nucleoside phosphorylase deoD-type (2.4.2.1)	COG0813	F
RCA23_c14100	acetyltransferase (2.3.1.-)	COG1246	E
RCA23_c14110	TrpA:tryptophan synthase alpha chain TrpA (4.2.1.20)	COG0159	E
RCA23_c14120	EngD:GTP-dependent nucleic acid-binding protein EngD	COG0012	J
RCA23_c14130	Non-specific ribonucleoside hydrolase rihC (3.2.-.-)	COG1957	F
RCA23_c14150	hypothetical protein DUF583	COG1664	M
RCA23_c14160	HtpX:protease HtpX (3.4.24.-)	COG0501	O
RCA23_c14170	histone deacetylase	COG0123	B
RCA23_c14180	TspO1:peripheral-type benzodiazepine receptor/signal transduction protein TspO	COG3476	T
RCA23_c14190	Hin:DNA-invertase Hin	COG1961	L
RCA23_c14200	hypothetical protein		
RCA23_c14210	hypothetical protein		
RCA23_c14220	hypothetical protein, RmlC-like cupin family	COG3450	R

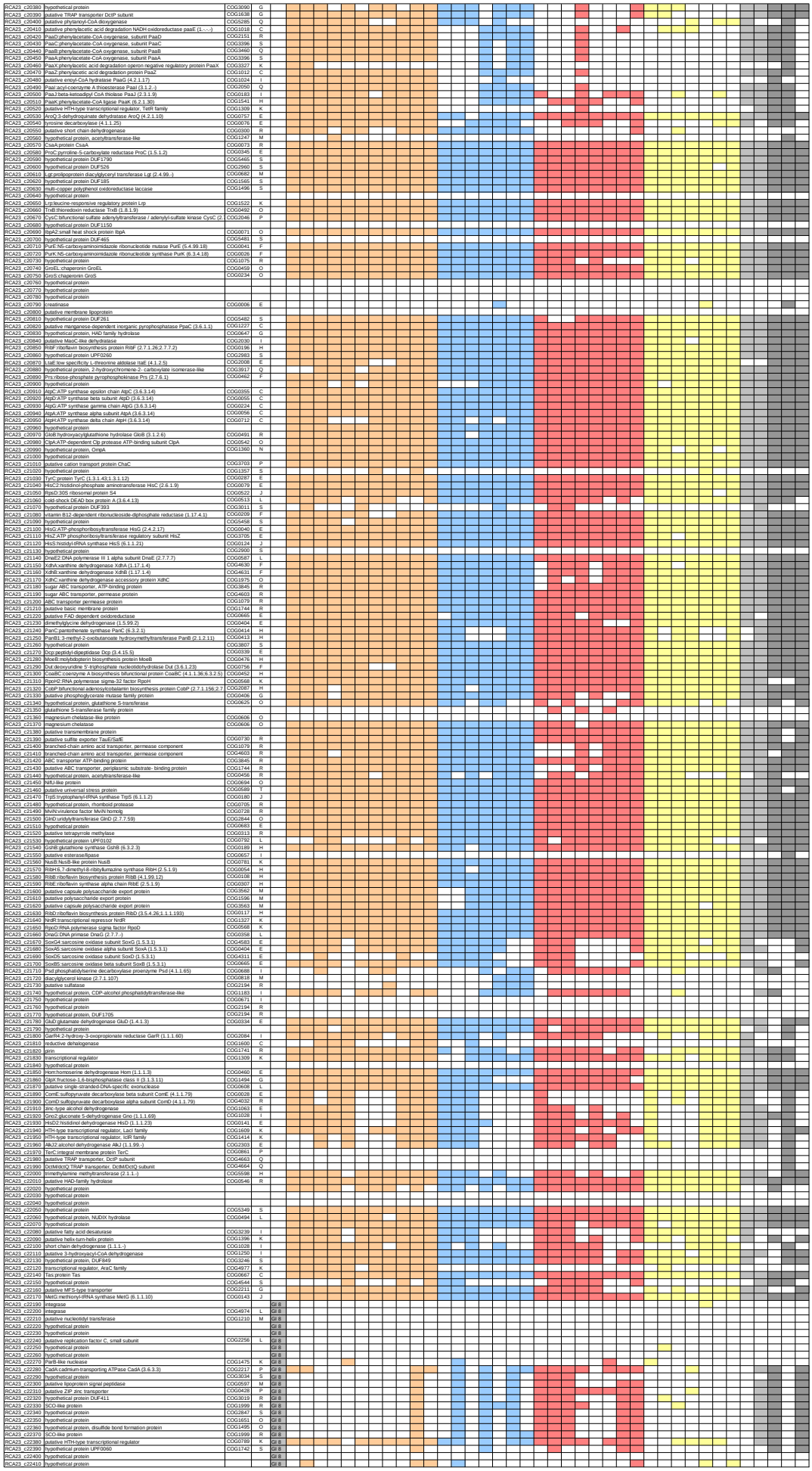

RCA23_c14230 integrase COG4974 L GI 7
RCA23_c14240 hypothetical membrane protein GI 7
RCA23_c14260 hypothetical protein COG4530 S
RCA23_c14270 hypothetical protein DUF45 COG1451 R
RCA23_c14280 HTH-type transcriptional regulator, GntR family COG1802 K
RCA23_c14290 Lpd1 dihydrolipoyl dehydrogenase Lpd (1.8.1.4) COG1249 C
RCA23_c14300 SucB:dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutara COG0508 C
RCA23_c14310 SucA:2-oxoglutarate dehydrogenase E1 component SucA (1.2.4.2) COG0567 C
RCA23_c14320 SucD1: succinyl-CoA ligase [ADP-forming] alpha subunit SucD (6.2.1.5) COG0074 C
RCA23_c14330 SucC:succinyl-CoA ligase [ADP-forming] beta subunit SucC (6.2.1.5) COG0045 C
RCA23_c14340 butyryl-CoA dehydrogenase (1.2.99.2) COG1960 I
RCA23_c14350 InfC:translation initiation factor IF-3 COG0290 J
RCA23_c14360 Fpr:ferredoxin--NADP reductase Fpr (1.18.1.2) COG1018 C
RCA23_c14370 hypothetical protein COG3749 S
RCA23_c14380 hypothetical protein COG0175 E
RCA23_c14390 Cysi:cysi/sir: sulfite reductase (ferredoxin) (1.8.7.1) COG0155 P
RCA23_c14400 hypothetical protein
RCA23_c14410 CysG:siroheme synthase CysG (2.1.1.107:4.99.1.4) COG0007 H
RCA23_c14420 HTH-type transcriptional regulator, AsnC family COG1522 K
RCA23_c14450 peptidase COG0624 E
RCA23_c14460 hypothetical protein COG1426 S
RCA23_c14470 IspG:4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase IspG (1.17.7.1) COG0821 I
RCA23_c14480 TPR-repeat containing protein COG4783 R
RCA23_c14490 aspartate aminotransferase (2.6.1.1) COG0436 E
RCA23_c14500 MrcA:penicillin-binding protein 1A (2.4.2.-:3.4.-.-) COG0009 M
RCA23_c14510 PrfB:peptide chain release factor 2 COG1186 J
RCA23_c14520 hypothetical protein DUF583 COG1664 M
RCA23_c14530 putative peptidase, M23 family COG0739 M
RCA23_c14540 hypothetical protein DUF455 COG2833 S
RCA23_c14550 Bcp:peroxiredoxin Bcp (1.11.1.15) COG1225 O
RCA23_c14560 hypothetical protein
RCA23_c14570 QueA:S-adenosylmethionine:tRNA ribosyltransferase- isomerase QueA (5.-.-.-) COG0809 J
RCA23_c14580 MFS-type transporter
RCA23_c14590 hypothetical protein DUF924 COG3803 S
RCA23_c14600 Lpd2 dihydrolipoyl dehydrogenase Lpd (1.8.1.4) COG1249 C
RCA23_c14610 UvrA:uvrABC system protein A COG0178 L
RCA23_c14620 MmgE/PrpD family protein COG2079 R
RCA23_c14630 MmsB1:3-hydroxyisobutyrate dehydrogenase MmsB (1.1.1.31) COG1084 I
RCA23_c14640 3-hydroxyisobutyryl-CoA hydrolase (3.1.2.4) COG1024 I
RCA23_c14650 Icd:butyryl-CoA dehydrogenase (1.3.99.-) COG1960 I
RCA23_c14660 MmsA:methylmalonate-semialdehyde dehydrogenase MmsA (1.2.1.27:1.2.1.18) COG1012 C
RCA23_c14670 putative HTH-type transcriptional regulator, LysR family COG0583 K
RCA23_c14680 CoaD:phosphopantetheine adenylyltransferase CoaD (2.7.7.3) COG0669 H
RCA23_c14690 Gap2:glyceraldehyde-3-phosphate dehydrogenase Gap (1.2.1.12) COG0057 G
RCA23_c14700 Gap3:glyceraldehyde-3-phosphate dehydrogenase Gap (1.2.1.12) COG0057 G
RCA23_c14710 TktA:transketolase TktA (2.2.1.1) COG0021 G
RCA23_c14720 hypothetical protein
RCA23_c14730 ZapA:cell division protein ZapA COG3027 S
RCA23_c14740 putative glutaredoxin COG0278 O
RCA23_c14750 hypothetical protein
RCA23_c14760 hypothetical protein, BolA-like COG0271 T
RCA23_c14770 PurL:phosphoribosylformylglycinamidine synthase PurL (6.3.5.3) COG0046 F
RCA23_c14780 HTH-type transcriptional regulator, LysR family COG0583 K
RCA23_c14790 hypothetical protein, pyruvate ferredoxin/flavodoxin oxidoreductase COG4231 C
RCA23_c14800 MurI:glutamate racemase MurI (5.1.1.3) COG0796 M
RCA23_c14810 ArgC:N-acetyl-gamma-glutamyl-phosphate reductase ArgC (1.2.1.38) COG0002 E
RCA23_c14820 CcmE:cytochrome c-type biogenesis protein CcmE COG2332 O
RCA23_c14830 CcmF:cytochrome c-type biogenesis protein CcmF COG1138 O
RCA23_c14840 CcmH:cytochrome c-type biogenesis protein CcmH COG3088 O
RCA23_c14850 putative enoyl-CoA hydratase FadB (4.2.1.17) COG1024 I
RCA23_c14860 hypothetical protein
RCA23_c14870 GltA:citrate synthase GltA (2.3.3.1) COG0372 C
RCA23_c14880 GltX2:glutamyl-tRNA synthase 2 (6.1.1.17) COG0008 J
RCA23_c14890 hypothetical protein competence protein E COG0658 R
RCA23_c14900 LexA:LexA repressor (3.4.21.88) COG1974 K
RCA23_c14910 MoaA2:molybdopterin biosynthesis protein MoaA COG0303 H
RCA23_c14920 MoaC:molybdenum cofactor biosynthesis protein MoaC COG0315 H
RCA23_c14930 TrpC:indole-3-glycerol phosphate synthase TrpC (4.1.1.48) COG0134 E
RCA23_c14940 TrpD:anthranilate phosphoribosyltransferase TrpD (2.4.2.18) COG0547 E
RCA23_c14950 TrpG:anthranilate synthase component TrpG (4.1.3.27) COG0512 E
RCA23_c14960 hypothetical protein, divergent polysaccharide deacetylase
RCA23_c14970 hypothetical protein
RCA23_c14990 hypothetical protein COG3108 S
RCA23_c15000 putative L,D-transpeptidase YcbB (2.-.-.-) COG2989 S
RCA23_c15010 LpxD:UDP-3-O-[3-hydroxymyristoyl] glucosamine N- acyltransferase LpxD (2.3.1 COG1044 M
RCA23_c15020 acyl carrier protein COG0236 I
RCA23_c15030 FabF1:3-oxoacyl-[acyl-carrier-protein] synthase FabF (2.3.1.179) COG0304 I
RCA23_c15040 hypothetical protein, invasion protein B COG5342 R
RCA23_c15050 inner membrane protein COG4536 P
RCA23_c15060 XerD:tyrosine recombinase XerD COG4974 L
RCA23_c15070 hypothetical protein
RCA23_c15080 hypothetical protein
RCA23_c15090 AroK1:shikimate kinase AroK (2.7.1.71) COG0703 E
RCA23_c15100 AroB:3-dehydroquinate synthase AroB (4.2.3.4) COG0337 E
RCA23_c15110 single-stranded DNA-binding protein COG0629 L
RCA23_c15120 GsiA3:glutathione import ATP-binding protein GsiA (3.6.3.-) COG0444 E
RCA23_c15130 putative ABC transporter inner membrane component COG1173 E
RCA23_c15140 putative ABC transporter permease protein COG0601 E
RCA23_c15150 ABC transporter extracellular solute-binding protein COG0747 E
RCA23_c15160 MiaA:tRNA delta(2)-isopentenylpyrophosphate transferase MiaA (2.5.1.8) COG0324 J
RCA23_c15170 PyrH:uridylate kinase PyrH (2.7.4.22) COG0528 F
RCA23_c15180 Frr:ribosome recycling factor Frr COG0233 J
RCA23_c15190 UppS:undecaprenyl pyrophosphate synthase UppS (2.5.1.31) COG0020 I
RCA23_c15200 putative cytidyltransferase COG0575 I
RCA23_c15210 Dxr:1-deoxy-D-xylulose 5-phosphate reductoisomerase Dxr (1.1.1.267) COG0743 I
RCA23_c15220 RseP:RIP metalloprotease RseP COG0750 M
RCA23_c15230 putative outer membrane assembly factor COG4775 M
RCA23_c15240 putative outer lein
RCA23_c15250 FabZ-(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase FabZ (4.2.1.-) COG0764 I
RCA23_c15260 LpxA:acyl-[acyl-carrier-protein]--UDP-N- acetylglucosamine O-acyltransferase L COG1043 M
RCA23_c15270 hypothetical protein DUF1009 COG3494 S
RCA23_c15280 LpxB:lipid-A-disaccharide synthase LpxB (2.4.1.182) COG0763 M
RCA23_c15290 TrmU:tRNA (5-methylaminomethyl-2-thiouridylate)- methyltransferase TrmU (2.1. COG0482 J
RCA23_c15300 hypothetical protein
RCA23_c15310 CtrA:cell cycle transcriptional regulator COG0745 T
RCA23_c15320 LigA:DNA ligase LigA (6.5.1.2) COG0272 L
RCA23_c15330 RecG:ATP-dependent DNA helicase RecG (3.6.4.12) COG1200 L
RCA23_c15340 hypothetical protein
RCA23_c15350 hypothetical protein COG0822 C
RCA23_c15360 HisI:phosphoribosyl-AMP cyclohydrolase HisI (3.5.4.19) COG0139 E
RCA23_c15370 GluQ:glutamyl-Q tRNA(Asp) synthase GluQ (6.1.1.-) COG0008 J
RCA23_c15380 hypothetical protein, methyltransferase COG4976 R
RCA23_c15390 Gid:tRNA uridine 5-carboxymethylaminomethyl modification enzyme Gid COG1206 J
RCA23_c15400 putative crotonase COG1024 I
RCA23_c15410 putative thioesterase COG2050 Q
RCA23_c15420 RplM:50S ribosomal protein L13 COG0102 J
RCA23_c15430 RpsI:30S ribosomal protein S9 COG0103 J
RCA23_c15440 integrase
RCA23_c15460 BetA2:choline dehydrogenase BetA (1.1.99.1) COG2303 E
RCA23_c15470 ABC transporter, permease protein COG4662 H
RCA23_c15480 ABC transporter, ATP-binding protein COG1131 V
RCA23_c15490 ABC transporter, periplasmic substrate-binding protein COG2998 H
RCA23_c15500 hypothetical membrane protein COG1238 S
RCA23_c15510 hypothetical protein DUF6 family, transmembrane COG2962 R
RCA23_c15520 hypothetical protein COG2068 R
RCA23_c15530 dimethylglycine dehydrogenase (1.5.99.2) COG0404 E
RCA23_c15540 dimethylglycine dehydrogenase (1.5.99.2) COG0404 E
RCA23_c15550 short-chain dehydrogenase/reductase SDR COG1028 I
RCA23_c15560 hypothetical transmembrane protein COG1788 R
RCA23_c15570 hypothetical protein, tetR regulator COG1309 K
RCA23_c15580 MetA:homoserine O-succinyltransferase MetA (2.3.1.46) COG1897 E
RCA23_c15590 hypothetical protein
RCA23_c15600 putative integral membrane protein COG2510 S
RCA23_c15610 Gcn5-like N-acetyltransferase COG1670 J
RCA23_c15620 GuaA:GMP synthase GuaA (6.3.5.2) COG0519 F
RCA23_c15630 MtbB4:trimethylamine methyltransferase MttB (2.1.1.-) COG5598 H
RCA23_c15640 hypothetical protein
RCA23_c15650 LipA:lipoyl synthase LipA (2.8.1.8) COG0320 H
RCA23_c15660 hypothetical protein
RCA23_c15670 Hpt:hypoxanthine phosphoribosyltransferase Hpt (2.4.2.8) COG0634 F
RCA23_c15680 hypothetical protein COG2867 I
RCA23_c15690 AmtB:ammonium transporter AmtB COG0004 P
RCA23_c15700 putative competence-damaged protein COG1546 R
RCA23_c15710 PgsA:phosphatidylglycerophosphatase A (3.1.3.27) COG1267 I
RCA23_c15720 IspF:2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase IspF (4.6.1.12) COG0245 I
RCA23_c15730 DusB:tRNA-dihydrouridine synthase B (1.-.-.-) COG0042 J
RCA23_c15740 NtrB:histidine kinase, nitrogen regulation protein NtrB (2.7.13.3) COG3852 T
RCA23_c15750 NtrC:nitrogen regulation protein NtrC COG2204 T
RCA23_c15760 NtrY:histidine kinase, nitrogen regulation protein NtrY (2.7.13.3) COG5000 T
RCA23_c15770 NtrX:nitrogen assimilation regulatory protein NtrX COG2204 T
RCA23_c15780 TrkA:Trk system potassium uptake protein TrkA COG0569 P
RCA23_c15790 putative Trk system potassium uptake protein trkH COG0168 P
RCA23_c15800 Hfq:RNA-binding protein Hfq COG1923 R
RCA23_c15810 HflX:GTP-binding protein HflX COG2262 R
RCA23_c15820 QuiP:acyl-homoserine lactone acylase QuiP (3.5.1.97) COG2366 R
RCA23_c15830 hypothetical protein COG2930 S
RCA23_c15840 HemB:delta-aminolevulinic acid dehydratase HemB (4.2.1.24) COG0113 H
RCA23_c15850 hypothetical protein
RCA23_c15860 Mfd:transcription-repair-coupling factor Mfd (3.6.1.-) COG1197 L
RCA23_c15870 putative DSBA-like thioredoxin family protein COG2761 O
RCA23_c15880 long-chain-fatty-acid-CoA ligase (6.2.1.3) COG0318 I
RCA23_c15890 HTH-type transcriptional regulator COG1396 K
RCA23_c15900 AgpZ:aquaporin AgpZ COG0580 G
RCA23_c15910 putative extracellular solute-binding protein COG4166 E
RCA23_c15920 LysE:lysine exporter protein LysE COG1279 R
RCA23_c15930 hypothetical protein DUF502 COG2928 S
RCA23_c15940 PsuG:pseudouridine-5'-phosphate glycosidase PsuG (3.2.-.-) COG2313 Q
RCA23_c15950 hypothetical protein, pfkB family carbohydrate kinase COG0524 G
RCA23_c15960 CspA1:cold shock protein CspA COG1278 K
RCA23_c15970 RpsB:30S ribosomal protein S2 COG0052 J
RCA23_c15980 Tsf:elongation factor Ts COG0264 J
RCA23_c15990 hypothetical protein
RCA23_c16000 hypothetical protein, acetyltransferase-like COG1670 J
RCA23_c16010 hypothetical protein, transcription factor NusA like
RCA23_c16020 putative phenylacetic acid degradation protein COG0663 R
RCA23_c16030 Gmk:guanylate kinase Gmk (2.7.4.8) COG0194 F
RCA23_c16040 hypothetical protein, YicC-like COG1561 S
RCA23_c16050 hypothetical protein, DUF1457 COG5388 S
RCA23_c16060 phospho-2-dehydro-3-deoxyheptonate aldolase (2.5.1.54) COG3200 E
RCA23_c16070 putative HTH-type transcriptional regulator, AraC family COG4977 K
RCA23_c16080 putative amino-acid binding protein COG0683 E
RCA23_c16090 putative branched-chain amino acid transport ATP- binding protein COG0411 E
RCA23_c16100 putative branched-chain amino acid transport ATP- binding protein COG0410 E
RCA23_c16110 putative branched-chain amino acid transport system permease protein COG0559 E
RCA23_c16120 putative branched-chain amino acid transport system permease protein COG4177 E
RCA23_c16130 TypA:GTP-binding protein TypA COG1217 T
RCA23_c16140 hypothetical protein DUF1330 COG5470 S
RCA23_c16150 AlaS alanyl-tRNA synthase AlaS (6.1.1.7) COG0013 J
RCA23_c16160 RecA protein RecA COG0468 L
RCA23_c16170 hypothetical protein
RCA23_c16180 sensor transduction histidine kianse COG0642 T
RCA23_c16190 hypothetical protein
RCA23_c16200 hypothetical protein, NOL1/NOP2/sun family COG0144 J
RCA23_c16210 GuaB:inosine-5'-monophosphate dehydrogenase GuaB (1.1.1.205) COG0516 F
RCA23_c16220 hypothetical protein COG3894 R
RCA23_c16240 SufS2:cysteine desulfurase SufS (2.8.1.7:4.4.1.16) COG0520 E
RCA23_c16250 hypothetical protein
RCA23_c16260 hypothetical protein
RCA23_c16270 SufD:FeS assembly protein SufD COG0719 O
RCA23_c16280 SufC:FeS assembly ATPase SufC COG0396 O
RCA23_c16290 SufB:FeS assembly protein SufB COG0719 O
RCA23_c16300 putative aminotransferase COG1104 E

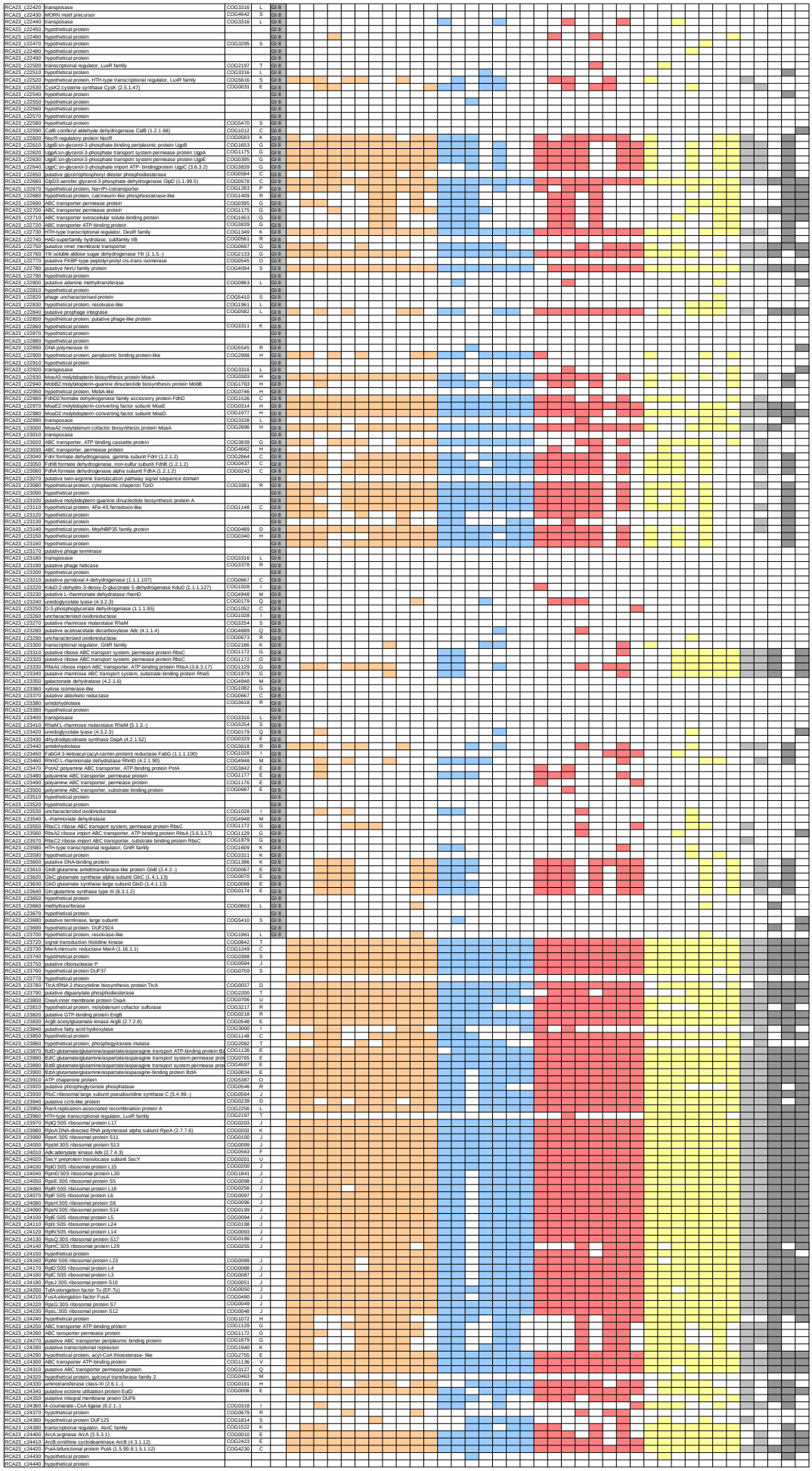


RCA23_c16310	SicR: iron-sulfur cluster assembly transcriptional regulator	COG1959	K
RCA23_c16320	hypothetical protein	COG2945	R
RCA23_c16330	putative oxidoreductase	COG0673	R
RCA23_c16340	IolE:inosose dehydratase IolE (4.2.1.44)	COG1082	G
RCA23_c16350	putative oxidoreductase	COG0673	R
RCA23_c16360	HTH-type transcriptional repressor, LacI family	COG1609	K
RCA23_c16370	hypothetical protein		
RCA23_c16380	sugar (ribose) ABC-transport system ATP binding protein	COG1129	G
RCA23_c16390	sugar (ribose) ABC transporter permeasa protein	COG1172	G
RCA23_c16400	sugar (ribose) ABC transporter periplasmic binding protein	COG1879	G
RCA23_c16410	BglA:beta-glucosidase BglA (3.2.1.21)	COG2723	G
RCA23_c16420	MalK:maltose/maltodextrin import ATP-binding protein MalK (3.6.3.19)	COG3839	G
RCA23_c16430	putative alpha-glucosidase AglA (3.2.1.20)	COG0366	G
RCA23_c16440	AglG:alpha-glucoside transport system permease protein AglG	COG0395	G
RCA23_c16450	ABC transporter, membrane spanning protein	COG0395	G
RCA23_c16460	AglF:alpha-glucoside transport system permease protein AglF	COG1175	G
RCA23_c16470	AglE:alpha-glucosides-binding periplasmic protein AglE	COG1653	G
RCA23_c16480	AglR:HTH-type transcriptional regulator	COG1609	K
RCA23_c16490	oxidoreductase	COG4989	R
RCA23_c16500	SsuD:alkanesulfonate monooxygenase SsuD (1.14.14.5)	COG2141	C
RCA23_c16510	putative regulatory DNA binding protein	COG2188	K
RCA23_c16520	putative oxidoreductase	COG0673	R
RCA23_c16530	Pgi:glucose-6-phosphate isomerase Pgi (5.3.1.9)	COG0166	G
RCA23_c16540	Pgl:6-phosphogluconolactonase Pgl (3.1.1.31)	COG0363	G
RCA23_c16550	Zwf:glucose-6-phosphate 1-dehydrogenase Zwf (1.1.1.49)	COG0364	G
RCA23_c16560	hypothetical protein, radical SAM	COG0535	R
RCA23_c16570	putative 5-methylcytosine-specific restriction enzyme McrA (3.1.21..-)	COG1403	V
RCA23_c16580	putative phospholipase/carboxylesterase	COG0400	R
RCA23_c16590	HNH-GPD superfamily base excision DNA repair protein	COG0122	L
RCA23_c16600	CobK:precorrin-6A reductase CobK (1.3.1.54)	COG2099	H
RCA23_c16610	CbiD:cobalamin (vitamin B12) biosynthesis CbiDprotein CbiD (2.1.1..)	COG1903	H
RCA23_c16620	CobA:uroporphyrinogen-III C-methyltransferase CobA (2.1.1.107)	COG0007	H
RCA23_c16630	CobB:cobyrinic acid A,C-diamide synthase CobB	COG1797	H
RCA23_c16640	putative major facilitator superfamily transporter		
RCA23_c16650	hypothetical protein	COG1562	I
RCA23_c16660	CysS:cysteinyl-tRNA synthase CysS (6.1.1.16)	COG0215	J
RCA23_c16670	AspC2:aspartate aminotransferase AspC (2.6.1.)	COG0436	E
RCA23_c16680	MtbS:trimethylamine methyltransferase MtbS (2.1.1..)	COG5598	H
RCA23_c16690	Uup:ABC transporter ATP-binding protein Uup	COG0488	R
RCA23_c16700	hypothetical protein	COG3176	R
RCA23_c16710	MsrB:peptide methionine sulfoxide reductase MsrB (1.8.4.12)	COG0229	O
RCA23_c16720	hypothetical protein, lipoprotein	COG2913	J
RCA23_c16730	hypothetical protein	COG1399	R
RCA23_c16740	PlsX:fatty acid/phospholipid synthesis protein PlsX	COG0416	I
RCA23_c16750	FabH:3-oxoacyl-[acyl-carrier-protein] synthase FabH (2.3.1.41)	COG0332	I
RCA23_c16760	IhfA:integration host factor alpha subunit IhfA	COG0776	L
RCA23_c16770	hypothetical protein, MerR family regulatory protein	COG0789	K
RCA23_c16790	Dcd:deoxyguanosinetriphosphate triphosphohydrolase- like protein Dcd	COG0717	F
RCA23_c16800	putative segregation and condesation protein B	COG1386	K
RCA23_c16810	hypothetical protein, segregation and condensation protein A	COG1354	S
RCA23_c16820	NagZ:beta-hexosaminidase NagZ (3.2.1.52)	COG1472	G
RCA23_c16830	hypothetical protein		
RCA23_c16840	arginyl-tRNA synthase (6.1.1.19)	COG0018	J
RCA23_c16850	deoxyguanosinetriphosphate triphosphohydrolase- like protein	COG0232	F
RCA23_c16860	putative iron-sulfur insertion protein erpA	COG0316	S
RCA23_c16870	XthA:exodeoxyribonuclease III (3.1.11.2)	COG0708	L
RCA23_c16880	putative cytochrome B561	COG3038	C
RCA23_c16890	putative dnak suppressor protein dksA	COG1734	T
RCA23_c16900	hypothetical protein	COG0714	R
RCA23_c16910	hypothetical protein	COG3825	S
RCA23_c16920	putative peptidase, M48 family	COG0501	O
RCA23_c16930	hypothetical protein		
RCA23_c16940	putative ribosomal RNA large subunit methyltransferase	COG1092	R
RCA23_c16950	Edd:phosphogluconate dehydratase Edd (4.2.1.12)	COG0129	E
RCA23_c16960	Eda:KHG/KDPG aldolase Eda (4.1.3.16,4.1.2.14)	COG0800	G
RCA23_c16970	hypothetical protein		
RCA23_c16980	GlnE:glutamate-ammonia-ligase adenylyltransferase GlnE (2.7.7.42)	COG1391	O
RCA23_c16990	hypothetical protein	COG2606	S
RCA23_c17000	hypothetical protein		
RCA23_c17010	hypothetical protein		
RCA23_c17020	hypothetical protein		
RCA23_c17030	putative aromatic-ring-hydroxylating dioxygenase	COG4638	P
RCA23_c17040	putative acetolactate synthase small subunit IlvH (2.2.1.6)	COG0440	E
RCA23_c17050	IlvI:acetolactate synthase isozyme large subunit IlvI (2.2.1.6)	COG0028	E
RCA23_c17060	hypothetical membrane protein, porin-like	COG1470	
RCA23_c17070	DctP7:TRAP dicarboxylate transporter, subunit DctP	COG4663	Q
RCA23_c17080	DctQ5:TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q
RCA23_c17090	DctM7:TRAP dicarboxylate transporter, subunit DctM	COG4664	Q
RCA23_c17100	putative arginyl-tRNA--protein transferase Ate (2.3.2.8)	COG2935	O
RCA23_c17110	NrdJ:vitamin B12-dependent ribonucleotide reductase NrdJ (1.17.4.1)	COG0209	F
RCA23_c17120	hypothetical protein		
RCA23_c17140	hypothetical protein DUF192	COG1430	S
RCA23_c17150	cold shock protein	COG1278	K
RCA23_c17160	PdxH:pyridoxine/pyridoxamine 5'-phosphate oxidase PdxH (1.4.3.5)	COG0259	H
RCA23_c17170	FabI1:enoyl-[acyl-carrier-protein] reductase FabI (1.3.1.9)	COG0623	I
RCA23_c17180	Gpt:xanthine phosphoribosyltransferase Gpt (2.4.2.22)	COG0503	F
RCA23_c17190	RihA:pyrimidine-specific ribonucleoside hydrolase RihA (3.2.-.-)	COG1957	F
RCA23_c17200	hypothetical protein LPF0061	COG0397	S
RCA23_c17210	putative sodium/hydrogen exchanger	COG0025	P
RCA23_c17220	Fur:ferric uptake regulator protein Fur	COG0735	P
RCA23_c17230	putative S-adenosylmethionine uptake transporter	COG0697	G
RCA23_c17240	Eno:enolase Eno (4.2.1.11)	COG0148	G
RCA23_c17250	AmmK:anhydro-N-acetylmuramic acid kinase AmmK (2.7.1.-)	COG2377	O
RCA23_c17260	TyrS:tyrosyl-tRNA synthase TyrS (6.1.1.1)	COG0162	J
RCA23_c17270	AspC4:aspartate aminotransferase AspC (2.6.1.)	COG0436	E
RCA23_c17280	hypothetical protein	COG0760	O
RCA23_c17290	TrpE:anthranilate synthase component TrpE (4.1.3.27)	COG0147	E
RCA23_c17300	putative protein soxG	COG0491	R
RCA23_c17310	putative protein soxH	COG0491	R
RCA23_c17320	HemA2:5-aminolevulinate synthase HemA (2.3.1.37)	COG0156	H
RCA23_c17330	SoxF2:sulfide dehydrogenase [flavocytochrome c] flavoprotein chain SoxF (1.8.)	COG3439	S
RCA23_c17340	domain of unknown function DUF1791	COG1416	S
RCA23_c17350	SoxD2:sulfite oxidase cytochrome subunit SoxD (1.8.3.1)	COG3474	C
RCA23_c17360	SoxC:sulfite oxidase molybdopterin subunit SoxC (1.8.3.1)	COG2041	R
RCA23_c17370	SoxB2:sulfur oxidation protein SoxB	COG0737	F
RCA23_c17380	SoxA2:diheme cytochrome c	COG3258	C
RCA23_c17390	SoxZ:protein SoxZ		
RCA23_c17400	SoxY:protein SoxY	COG5501	S
RCA23_c17410	SoxX:cytochrome c	COG2010	C
RCA23_c17420	SoxW:thioredoxin SoxW	COG2143	O
RCA23_c17430	SoxV:cytochrome c-type biogenesis protein SoxV	COG0785	O
RCA23_c17440	SoxS:protein SoxS		
RCA23_c17450	SoxR:HTH-type transcriptional regulator, ArsR family	COG0640	K
RCA23_c17460	putative soxT, transmembrane protein DUF395	COG2391	R
RCA23_c17470	hypothetical protein DUF395	COG2391	R
RCA23_c17480	NAD-binding protein	COG2084	I
RCA23_c17490	GyaR2:glyoxylate reductase GyaR (1.1.1.26)	COG1052	C
RCA23_c17500	GatA2:glutamyl-tRNA(Gln) amidotransferase subunit A (6.3.5..)	COG0154	J
RCA23_c17510	DctM8:TRAP dicarboxylate transporter, subunit DctM	COG4664	Q
RCA23_c17520	DctQ6:TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q
RCA23_c17530	DctP8:TRAP dicarboxylate transporter, subunit DctP	COG4663	Q
RCA23_c17540	hypothetical protein		
RCA23_c17550	GlpK:glycerol kinase GlpK (2.7.1.30)	COG0554	C
RCA23_c17560	RrmJ:ribosomal RNA large subunit methyltransferase J (2.1.1..)	COG0293	J
RCA23_c17570	putative PpxGppA phosphatase family protein	COG0248	F
RCA23_c17580	hypothetical protein		
RCA23_c17590	hypothetical protein	COG0685	E
RCA23_c17600	sterin domain containing enzyme	COG1410	E
RCA23_c17610	YfaU:2-keto-3-deoxy-L-rhamnonate aldolase YfaU (4.1.2..)	COG3836	G
RCA23_c17620	Glk:glucokinase Glk (2.7.1.2)	COG0837	G
RCA23_c17630	hypothetical protein DUF1006	COG3214	S
RCA23_c17640	reductive dehalogenase	COG1018	C
RCA23_c17650	hypothetical protein, XdhC and CoxI	COG1975	O
RCA23_c17660	GlxR:2-hydroxy-3-oxopropionate reductase GlxR (1.1.1.60)	COG2084	I
RCA23_c17670	3-hydroxybutyrate dehydrogenase (1.1.1.30)	COG1028	I
RCA23_c17680	IlvD2:dihydroxy-acid dehydratase IlvD (4.2.1.9)	COG0129	E
RCA23_c17690	fumarylacetoacetate hydrolase	COG0179	Q
RCA23_c17700	hypothetical protein		
RCA23_c17710	Ndk:nucleoside diphosphate kinase Ndk (2.7.4.6)	COG0105	F
RCA23_c17720	hypothetical protein		
RCA23_c17730	ABC transporter ATP-binding protein	COG0488	R
RCA23_c17740	MatC family integral membrane protein	COG2095	U
RCA23_c17750	hypothetical protein	COG3577	R
RCA23_c17760	putative DNA polymerase III chi subunit, HolC	COG2927	L
RCA23_c17770	PepA:cytosol aminopeptidase PepA (3.4.11.1)	COG0260	E
RCA23_c17780	putative permease, YjgP/YjgQ family	COG0795	R
RCA23_c17790	putative permease, YjgP/YjgQ family	COG0795	R
RCA23_c17800	putative organic solvent tolerance protein	COG1452	M
RCA23_c17810	hypothetical protein, SurA	COG0760	O
RCA23_c17820	PdxA:4-hydroxythreonine-4-phosphate dehydrogenase PdxA (1.1.1.262)	COG1995	H
RCA23_c17830	KsgA:dimethyladenosine transferase KsgA (2.1.1..)	COG0030	J
RCA23_c17840	hypothetical protein		
RCA23_c17850	modification methylase, hemK family	COG2890	J
RCA23_c17860	PrfA:peptide chain release factor 1	COG0216	J
RCA23_c17870	hypothetical protein	COG4446	S
RCA23_c17880	SpeB1:agmatinase SpeB (3.5.3.11)	COG0010	E
RCA23_c17890	alpha/beta hydrolase		
RCA23_c17900	SpeB2:agmatinase SpeB (3.5.3.11)	COG0010	E
RCA23_c17910	HipO2:hippurate hydrolase HipO (3.5.1.32)	COG1473	R
RCA23_c17920	MaoG:protein MaoG	COG3956	R
RCA23_c17930	FutA:ABC transporter periplasmic iron-binding protein FutA	COG1840	P
RCA23_c17940	putative peptidyl-prolyl cis-trans isomerase Ppi (5.2.1.8)	COG0652	O
RCA23_c17950	putative peptidyl-prolyl cis-trans isomerase Ppi (5.2.1.8)	COG0652	O
RCA23_c17960	Pgk:phosphoglycerate kinase Pgk (2.7.2.3)	COG0126	G
RCA23_c17970	Fda:fructose-bisphosphate aldolase class 1 (4.1.2.13)	COG3588	G
RCA23_c17980	hypothetical protein, septum formation initiator	COG2919	D
RCA23_c17990	PdhA:pyruvate dehydrogenase E1 component alpha subunit PdhA (1.2.4.1)	COG1071	C
RCA23_c18000	PdhB:pyruvate dehydrogenase E1 component beta subunit PdhB (1.2.4.1)	COG0022	C
RCA23_c18010	PdhC:dihydrolipoyllysine-residue acetyltransferase PdhC (2.3.1.12)	COG0508	C
RCA23_c18020	CysE:serine acetyltransferase CysE (2.3.1.30)	COG1045	E
RCA23_c18030	putative gene transfer agent protein		
RCA23_c18040	putative gene transfer agent large terminase part 1	COG5323	S
RCA23_c18050	putative gene transfer agent large terminase part 2	COG5323	S
RCA23_c18060	aminodeoxychorismate lyase	COG1559	R
RCA23_c18070	FabF2:3-oxoacyl-[acyl-carrier-protein] synthase FabF (2.3.1.179)	COG0304	I
RCA23_c18080	AcpP:acyl carrier protein AcpP	COG0236	I
RCA23_c18090	FabG3:3-oxoacyl-[acyl-carrier-protein] reductase FabG (1.1.1.100)	COG1028	I
RCA23_c18100	malonyl CoA-acyl carrier protein transacylase (2.3.1.39)	COG0331	I
RCA23_c18110	RpsF:30S ribosomal protein S6	COG0360	J
RCA23_c18120	RpsR:30S ribosomal protein S18	COG0238	J
RCA23_c18130	RplI:50S ribosomal protein L9	COG0359	J
RCA23_c18140	Tig:trigger factor (TF)	COG0544	O
RCA23_c18160	GlnB1:nitrogen regulatory protein P-II 1	COG0347	E
RCA23_c18170	GlnA2:glutamine synthase GlnA (6.3.1.2)	COG0174	E
RCA23_c18190	hypothetical protein		
RCA23_c18180	DddP:dimethylpropiothetin dethiomethylase DddP (4.4.1.3)	COG0006	E
RCA23_c18200	hypothetical protein, BioY-like	COG1268	R
RCA23_c18210	hypothetical protein		
RCA23_c18220	PurB:adenylosuccinate lyase PurB (4.3.2.2)	COG0015	F
RCA23_c18230	putative nitrile hydratase, beta subunit		
RCA23_c18240	hypothetical protein		
RCA23_c18250	NthA:nitrile hydratase alpha subunit NthA (4.2.1.84)		
RCA23_c18260	hypothetical protein DUF6 transmembrane	COG0697	G
RCA23_c18270	putative lipid A biosynthesis lauroyl acyltransferase	COG1560	M
RCA23_c18280	hypothetical protein	COG5429	S
RCA23_c18290	AcnA:aconitate hydratase AcnA (4.2.1.3)	COG1048	C
RCA23_c18300	CcmG:cytochrome c biogenesis protein CcmG	COG0526	O
RCA23_c18310	CcmC:cytochrome c-type biogenesis protein CcmC	COG0755	O
RCA23_c18320	CcmB:cytochrome c-type biogenesis protein CcmB	COG2386	O
RCA23_c18330	CcmA:cytochrome c biogenesis ATP-binding export protein CcmA (3.6.3.41)	COG4133	O
RCA23_c18340	hypothetical protein	COG3737	S
RCA23_c18350	SecF:protein export membrane protein SecF	COG0341	U



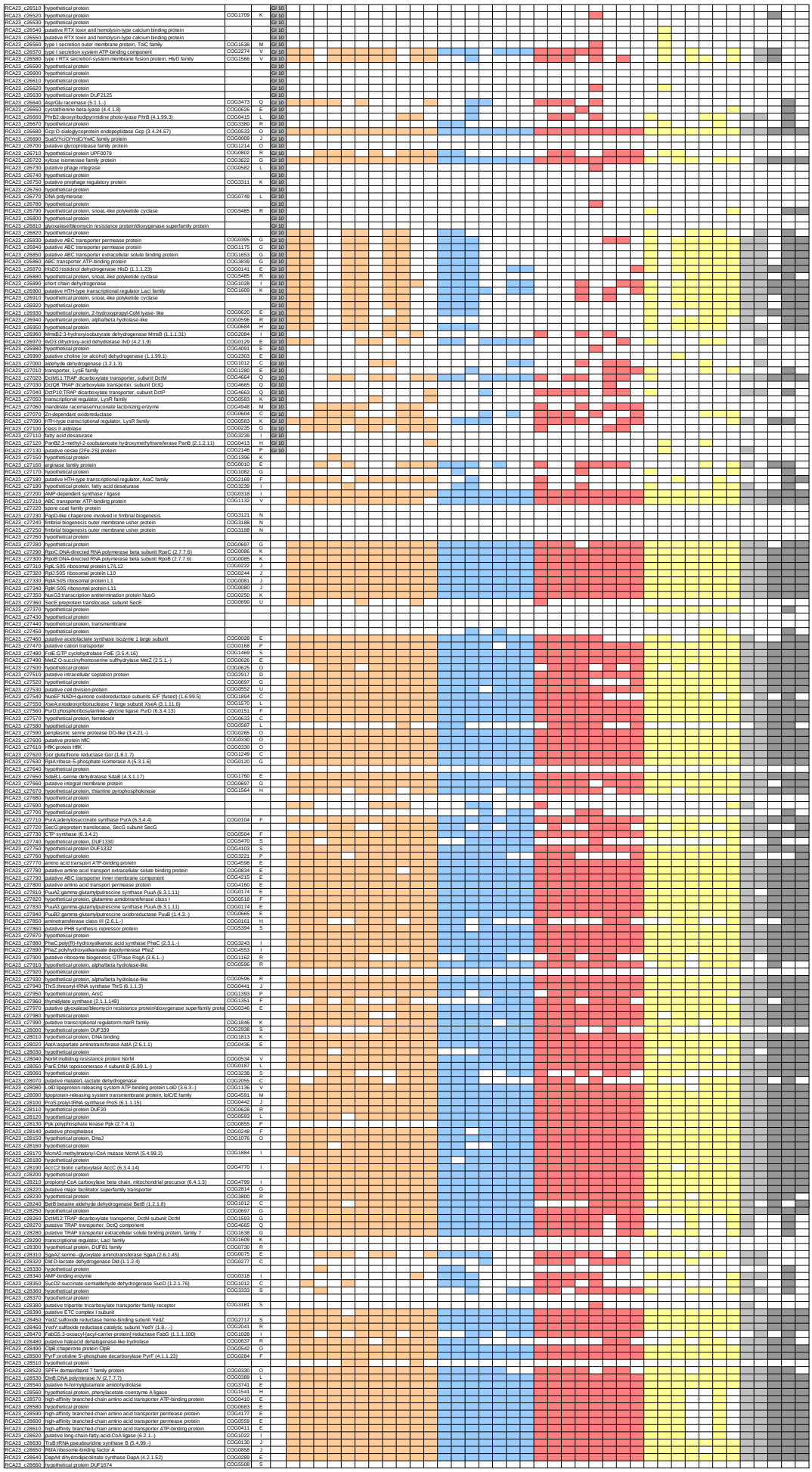
RCA23_c18360	SecD protein export membrane protein SecD	COG0342	U
RCA23_c18370	putative immunogenic membrane protein YajC	COG1862	U
RCA23_c18380	SerS seryl-tRNA synthase SerS (6.1.1.11)	COG0172	J
RCA23_c18390	hypothetical protein, alpha/beta hydrolase-like	COG2267	I
RCA23_c18400	EngA;GTP-binding protein EngA	COG1160	R
RCA23_c18410	putative quinoprotein	COG1520	S
RCA23_c18420	hypothetical protein DUF2133	COG4649	S
RCA23_c18430	RND efflux transporter, MFP subunit	COG0845	M
RCA23_c18440	RND efflux transporter, permease protein	COG0841	V
RCA23_c18450	ABC transporter ATP-binding protein	COG5265	O
RCA23_c18460	hypothetical protein, peptidoglycan-binding protein domain LysM	COG1652	S
RCA23_c18470	RarD protein RarD	COG2962	R
RCA23_c18480	SodB superoxide dismutase SodB (1.15.1.1)	COG0605	P
RCA23_c18490	SoxG2 sarcosine oxidase subunit SoxG (1.5.3.1)	COG4583	E
RCA23_c18500	SoxA3 sarcosine oxidase alpha subunit SoxA (1.5.3.1)	COG0404	E
RCA23_c18510	SoxD3 sarcosine oxidase subunit SoxD (1.5.3.1)	COG4311	E
RCA23_c18520	SoxB3 sarcosine oxidase beta subunit SoxB (1.5.3.1)	COG0665	E
RCA23_c18530	CycH;cytochrome c-type biogenesis protein CycH	COG4235	O
RCA23_c18540	hypothetical protein	COG0816	L
RCA23_c18550	hypothetical protein DUF1289	COG3313	R
RCA23_c18560	hypothetical protein		
RCA23_c18570	hypothetical protein DUF81	COG0730	R
RCA23_c18580	Dus;tRNA-dihydrouridine synthase Dus (1.-.-.-)	COG0042	J
RCA23_c18590	putative peroxiredoxin (thioredoxin reductase)	COG0678	O
RCA23_c18600	benzaldehyde dehydrogenase	COG1012	C
RCA23_c18610	hypothetical protein		
RCA23_c18620	hypothetical protein	COG2130	R
RCA23_c18630	MFS-type transporter	COG2814	G
RCA23_c18640	sorbitol dehydrogenase (1.1.1.14)	COG1028	I
RCA23_c18650	6-hydroxynicotinate 3-monooxygenase (1.14.13.-)	COG0654	H
RCA23_c18660	hypothetical protein		
RCA23_c18670	glycerate kinase	COG2379	G
RCA23_c18680	putative carboxymuconolactone decarboxylase	COG0599	S
RCA23_c18690	FadH;2,4-dienoyl-CoA reductase [NADPH] (1.3.1.34)	COG1902	C
RCA23_c18700	thiamine pyrophosphate protein	COG0028	E
RCA23_c18710	tfdG/rfbB; dTDP-glucose 4,6-dehydratase (4.2.1.46)	COG1088	M
RCA23_c18720	hypothetical protein		
RCA23_c18730	BetA3;choline dehydrogenase BetA (1.1.99.1)	COG2303	E
RCA23_c18740	ScoB;succinyl-CoA:3-ketoacid-coenzyme A transferase subunit ScoB (2.8.3.5)	COG2057	I
RCA23_c18750	ScoA;succinyl-CoA:3-ketoacid-coenzyme A transferase subunit ScoA (2.8.3.5)	COG1788	I
RCA23_c18760	hypothetical membrane protein		
RCA23_c18770	putative dimethyl sulfoniopropionate demethylase DmdA	COG0404	E
RCA23_c18780	HyuA2 hydantoin utilization protein A (3.5.2.-)	COG0145	E
RCA23_c18790	HyuB2 hydantoin utilization protein B (3.5.2.14)	COG0146	E
RCA23_c18800	FAD dependent oxidoreductase	COG0665	E
RCA23_c18810	hypothetical protein, 3-beta hydroxysteroid dehydrogenase	COG0451	M
RCA23_c18820	hypothetical protein, serine/threonine-protein kinase	COG1262	S
RCA23_c18830	hypothetical protein	COG0457	R
RCA23_c18840	hypothetical protein DUF1989	COG3665	S
RCA23_c18850	putative diaminopropionate ammonia-lyase	COG1171	E
RCA23_c18860	hypothetical protein, metallopeptidase M24	COG0006	E
RCA23_c18870	hypothetical protein, Asp/Glu/hydantoin racemase	COG3473	Q
RCA23_c18880	putative 3-oxoadipate enol-lactonase 2, alpha/beta hydrolase family	COG0596	R
RCA23_c18890	hypothetical protein DUF1185		
RCA23_c18900	LimB limonene 1,2-monooxygenase LimB (1.14.-.-)	COG2141	C
RCA23_c18910	aldehyde dehydrogenase, cytosolic (1.2.1.3)	COG1012	C
RCA23_c18920	PotC ABC transporter, spermidine/putrescine import, permease protein PotC	COG1177	E
RCA23_c18930	PotB1 ABC transporter, spermidine/putrescine import, permease protein PotB	COG1176	E
RCA23_c18940	PotD ABC transporter, spermidine/putrescine import, substrate binding protein P	COG0687	E
RCA23_c18950	PotA1 ABC transporter, spermidine/putrescine import, ATP-binding protein PotA	COG3842	E
RCA23_c18960	putative nitrilotriacetate monooxygenase component B	COG1853	R
RCA23_c18970	hypothetical protein, DUF268		
RCA23_c18980	hypothetical protein DUF28	COG0217	S
RCA23_c18990	sodium/sulphate symporter	COG0471	P
RCA23_c19000	hypothetical protein	COG0697	G
RCA23_c19010	hypothetical protein	COG1692	S
RCA23_c19020	osmotically inducible OsmC-like protein	COG1764	O
RCA23_c19030	3-hydroxyacyl-CoA dehydrogenase, NAD-binding	COG1250	I
RCA23_c19040	AroK2 shikimate kinase AroK (2.7.1.71)	COG0703	E
RCA23_c19050	thioesterase-like protein	COG0824	R
RCA23_c19060	BoxA;benzoyl-CoA oxygenase component A (1.14.12.21)	COG0369	P
RCA23_c19070	BoxB;benzoyl-CoA oxygenase component B (1.14.12.21)	COG3396	S
RCA23_c19080	BoxC;benzoyl-CoA-dihydrodiol lyase BoxC (4.1.2.44)	COG1024	I
RCA23_c19090	hypothetical protein DUF309		
RCA23_c19100	alpha/beta hydrolase	COG0596	R
RCA23_c19110	benzoate-coenzyme A ligase (6.2.1.25)	COG0318	I
RCA23_c19120	putative 5-formyltetrahydrofolate cyclo-ligase family protein	COG0212	H
RCA23_c19130	MgtE magnesium transporter MgtE	COG2239	P
RCA23_c19140	GuaD guanine deaminase GuaD (3.5.4.3)	COG0402	F
RCA23_c19150	putative hydroxydechloroatrazine ethylaminohydrolase	COG0402	F
RCA23_c19160	putative inositol monophosphatase family protein	COG0483	G
RCA23_c19170	putative helix-turn-helix protein	COG1396	K
RCA23_c19180	putative alcohol dehydrogenase	COG0604	C
RCA23_c19190	putative inner membrane protein	COG0670	R
RCA23_c19200	hypothetical protein		
RCA23_c19210	putative N-acetylmuramoyl-L-alanine amidase amiD	COG3023	V
RCA23_c19220	hypothetical protein		
RCA23_c19230	GatA3 glutamyl-tRNA(Gln) amidotransferase subunit A (6.3.5.-)	COG0154	J
RCA23_c19240	GatC;aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit GatC (6.3.5.-)	COG0721	J
RCA23_c19250	putative deaminase	COG0590	F
RCA23_c19260	putative ribosomal large subunit pseudouridine synthase B	COG1187	J
RCA23_c19270	ModC;molybdate ABC transporter, ATP-binding protein ModC (3.6.3.29)	COG4148	P
RCA23_c19280	ModB;molybdate ABC transporter, permease protein ModB	COG4149	P
RCA23_c19290	ModA;molybdate ABC transporter, substrate binding protein ModA	COG0725	P
RCA23_c19300	hypothetical protein, NUDIX hydrolase	COG4994	L
RCA23_c19310	hypothetical protein DUF1178	COG5319	S
RCA23_c19320	LysC;aspartokinase LysC (2.7.2.4)	COG0527	E
RCA23_c19330	PstI phosphoenolpyruvate-protein phosphotransferase PstI (2.7.3.9)	COG3605	T
RCA23_c19340	hypothetical protein, acetyltransferase-like	COG3153	R
RCA23_c19350	hypothetical protein	COG1853	R
RCA23_c19360	putative sulfate transporter	COG0659	P
RCA23_c19370	MtnP;methylthioadenosine phosphorylase MtnP (2.4.2.28)	COG0005	F
RCA23_c19380	Apt;adenine phosphoribosyltransferase Apt (2.4.2.7)	COG0503	F
RCA23_c19390	RimJ;ribosomal-protein-alanine acetyltransferase RimJ (2.3.1.128)	COG1670	J
RCA23_c19400	YmxG;uncharacterized zinc protease YmxG (3.4.24.-)	COG0612	R
RCA23_c19410	ThrC;threonine synthase ThrC (4.2.3.1)	COG0498	E
RCA23_c19420	hypothetical protein, SURF1	COG1346	S
RCA23_c19430	CtaE;cytochrome c oxidase subunit 3 (1.9.3.1)	COG1845	C
RCA23_c19440	CtaG;cytochrome c oxidase assembly protein CtaG	COG3175	O
RCA23_c19450	CtaB;protoheme IX farnesyltransferase CtaB (2.5.1.-)	COG0109	O
RCA23_c19460	CtaC;cytochrome c oxidase subunit 2 precursor (1.9.3.1)	COG1622	C
RCA23_c19470	TldD protein TldD	COG0312	R
RCA23_c19480	SMF family protein	COG0758	L
RCA23_c19490	TopA;DNA topoisomerase TopA (5.99.1.2)	COG0550	L
RCA23_c19500	Fba;fructose-bisphosphate aldolase Fba (4.1.2.13)	COG0191	G
RCA23_c19510	IolB2 5-deoxy-glucuronate isomerase IolB (5.3.1.-)	COG3718	G
RCA23_c19520	IolC;5-dehydro-2-deoxygluconokinase IolC (2.7.1.92)	COG0524	G
RCA23_c19530	IolD;3D-(3,5/4)-trihydroxycyclohexane-1,2-dione hydrolase IolD (3.7.1.-)	COG3962	E
RCA23_c19540	hypothetical protein DUF989	COG3748	S
RCA23_c19550	UraH;5-hydroxyisourate hydrolase UraH (3.5.2.17)	COG2351	R
RCA23_c19560	uric acid degradation bifunctional protein PucL (1.7.3.3;4.1.1.-)	COG3195	S
RCA23_c19570	putative allantoin catabolism protein YibA	COG3257	R
RCA23_c19580	SoxB4;sarcosine oxidase beta subunit SoxB (1.5.3.1)	COG0665	E
RCA23_c19590	SoxD4;sarcosine oxidase subunit SoxD (1.5.3.1)	COG4311	E
RCA23_c19600	SoxA4;sarcosine oxidase alpha subunit SoxA (1.5.3.1)	COG0404	E
RCA23_c19610	SoxG3;sarcosine oxidase subunit SoxG (1.5.3.1)	COG4583	E
RCA23_c19620	putative HTH-type transcriptional regulator, AraC family	COG4977	K
RCA23_c19630	putative DNA-binding protein	COG1396	K
RCA23_c19640	putative membrane protein		
RCA23_c19650	hypothetical protein		
RCA23_c19660	hypothetical protein, thioesterase	COG2050	Q
RCA23_c19670	hypothetical protein, calcineurin-like phosphoesterase-like	COG1407	R
RCA23_c19680	putative DEAD/DEAH box helicase	COG1201	R
RCA23_c19690	FolD;bifunctional enzyme FolD (1.5.1.5;3.5.4.9)	COG0190	H
RCA23_c19700	Fhs;formate--tetrahydrofolate ligase Fhs (6.3.4.3)	COG2759	F
RCA23_c19710	hypothetical protein		
RCA23_c19720	FtsH;cell division protease FtsH (3.4.24.-)	COG0465	O
RCA23_c19730	TilS;tRNA(Ile)-lysidine synthase TilS (6.3.4.-)	COG0037	D
RCA23_c19740	hypothetical protein	COG1729	S
RCA23_c19750	Pal;peptidoglycan-associated lipoprotein Pal	COG2885	M
RCA23_c19760	TolB;Tol-Pal system beta propeller repeat protein TolB	COG0823	U
RCA23_c19770	hypothetical protein, TolA-like		
RCA23_c19780	TolR;biopolymer transport protein TolR	COG0848	U
RCA23_c19790	biopolymer transport protein TolQ	COG0811	U
RCA23_c19800	acyl-CoA thioester hydrolase (3.1.2.-)	COG0824	R
RCA23_c19810	hypothetical protein		
RCA23_c19820	RuvB;holliday junction ATP-dependent DNA helicase RuvB (3.6.4.12)	COG2255	L
RCA23_c19830	RuvA;holliday junction ATP-dependent DNA helicase RuvA (3.6.4.12)	COG0632	L
RCA23_c19840	RuvC;crossover junction endodeoxyribonuclease RuvC (3.1.22.4)	COG0817	L
RCA23_c19850	putative ribosomal protein L11 methyltransferase	COG2264	J
RCA23_c19860	MsrA;peptide methionine sulfoxide reductase MsrA (1.8.4.11)	COG0225	O
RCA23_c19870	hypothetical protein, major facilitator superfamily transporter		
RCA23_c19880	LldD2;L-lactate dehydrogenase lldD (1.1.2.3)	COG1304	C
RCA23_c19890	DdtT BCCT family transporter involved in DMSP uptake	COG1292	M
RCA23_c19900	putative MFS-type transporter		
RCA23_c19910	RplY 50S ribosomal protein L25	COG1825	J
RCA23_c19920	Pth;peptidyl-tRNA hydrolase Pth (3.1.1.29)	COG0193	J
RCA23_c19930	hypothetical protein	COG4427	S
RCA23_c19940	hypothetical protein	COG3651	S
RCA23_c19950	TrpB tryptophan synthase beta chain TrpB (4.2.1.20)	COG0133	E
RCA23_c19960	TrpF N-5'-phosphoribosyl)anthranilate isomerase TrpF (5.3.1.24)	COG0135	E
RCA23_c19970	hypothetical protein		
RCA23_c19980	IhfB;integration host factor beta subunit IhfB	COG0776	L
RCA23_c19990	RpsA;30S ribosomal protein S1	COG0539	J
RCA23_c20000	hypothetical protein		
RCA23_c20010	hypothetical protein, lacI family HTH-type regulatory protein	COG1609	K
RCA23_c20020	putative phytanoyl-CoA dioxygenase	COG5285	Q
RCA23_c20030	aldo/keto reductase	COG0667	C
RCA23_c20040	Cmk;cytidylate kinase Cmk (2.7.4.14)	COG0283	F
RCA23_c20050	AroA;3-phosphoshikimate 1-carboxyvinyltransferase AroA (2.5.1.19)	COG0128	E
RCA23_c20060	TrmB;tRNA (guanine-N(7)-)-methyltransferase TrmB (2.1.1.33)	COG0220	R
RCA23_c20070	MetK S-adenosylmethionine synthase MetK (2.5.1.6)	COG0192	H
RCA23_c20080	Lnt;apolipoprotein N-acyltransferase Lnt (2.3.1.-)	COG0815	M
RCA23_c20090	putative magnesium and cobalt efflux protein	COG1253	R
RCA23_c20100	hypothetical protein UPF0054	COG0319	R
RCA23_c20110	PhoH-like protein	COG1702	T
RCA23_c20120	RNA modification enzyme, MiaB family	COG0621	J
RCA23_c20130	hypothetical protein		
RCA23_c20140	putative ferric uptake regulator family protein	COG0735	P
RCA23_c20150	FabA;3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase FabA (4.2.1.60)	COG0764	I
RCA23_c20160	FatB3 3-oxoacyl-[acyl-carrier-protein] synthase FatB (2.3.1.41)	COG0304	I
RCA23_c20170	FabI2 enoyl-[acyl-carrier-protein] reductase FabI (1.3.1.9)	COG0623	I
RCA23_c20180	SRY threo-3-hydroxyaspartate ammonia-lyase SRY (4.3.1.16)	COG1171	E
RCA23_c20190	2-haloalkanoic acid dehalogenase (3.8.1.2)	COG1011	R
RCA23_c20200	CatD;3-oxoadipate enol-lactonase CatD (3.1.1.24)	COG0596	R
RCA23_c20210	putative tetracycline resistance protein, class C	COG2814	G
RCA23_c20220	NAD/mycothiol-dependent formaldehyde dehydrogenase (1.2.1.66)	COG1062	C
RCA23_c20230	DgoD;D-galactonate dehydratase DgoD (4.2.1.6)	COG4948	M
RCA23_c20240	dimethylglycine dehydrogenase	COG0404	E
RCA23_c20250	MtbB6 trimethylamine methyltransferase MttB (2.1.1.-)	COG5598	H
RCA23_c20260	hypothetical protein	COG0607	P
RCA23_c20270	uncharacterized aminotransferase (2.6.1.-)	COG0161	H
RCA23_c20280	GcvT2;aminomethyltransferase, mitochondrial (2.1.2.10)	COG0404	E
RCA23_c20290	GcvH;glycine cleavage system protein GcvH	COG0509	E
RCA23_c20300	GcvP;glycine dehydrogenase GcvP (1.4.4.2)	COG1003	E
RCA23_c20310	DadA2;D-amino acid dehydrogenase small subunit DadA (1.4.99.1)	COG0665	E
RCA23_c20320	PotF;putrescine ABC transport system putrescine- binding periplasmic protein P	COG0687	E
RCA23_c20330	PotG;putrescine ABC transport system ATP-binding protein PotG	COG3842	E
RCA23_c20340	PotI;putrescine ABC transport system permease protein PotI	COG1177	E
RCA23_c20350	PotB2;putrescine ABC transport system permease protein PotB	COG1176	E
RCA23_c20360	Ileske 2Fe-2S domain protein	COG4638	P
RCA23_c20370	DctM9;TRAP dicarboxylate transporter, subunit DctM	COG1593	G





RCA23_c22420	transposase	COG3316	L
RCA23_c22430	MORN motif precursor	COG4642	S
RCA23_c22440	transposase	COG3316	L
RCA23_c22450	hypothetical protein		
RCA23_c22460	hypothetical protein		
RCA23_c22470	hypothetical protein	COG3295	S
RCA23_c22480	hypothetical protein		
RCA23_c22490	hypothetical protein		
RCA23_c22500	transcriptional regulator, LuxR family	COG2197	T
RCA23_c22510	hypothetical protein	COG3316	L
RCA23_c22520	hypothetical protein, HTH-type transcriptional regulator, LuxR family	COG5616	S
RCA23_c22530	CysK2 cysteine synthase CysK (2.5.1.47)	COG0031	E
RCA23_c22540	hypothetical protein		
RCA23_c22550	hypothetical protein		
RCA23_c22560	hypothetical protein		
RCA23_c22570	hypothetical protein		
RCA23_c22580	hypothetical protein	COG5470	S
RCA23_c22590	CalB coniferyl aldehyde dehydrogenase CalB (1.2.1.68)	COG1012	C
RCA23_c22600	NocR regulatory protein NocR	COG0583	K
RCA23_c22610	UgpB sn-glycerol-3-phosphate-binding periplasmic protein UgpB	COG1653	G
RCA23_c22620	UgpA sn-glycerol-3-phosphate transport system permease protein UgpA	COG1175	G
RCA23_c22630	UgpE sn-glycerol-3-phosphate transport system permease protein UgpE	COG0395	G
RCA23_c22640	UgpC sn-glycerol-3-phosphate import ATP- bindingprotein UgpC (3.6.3.2)	COG3839	G
RCA23_c22650	putative glycerophosphoryl diester phosphodiesterase	COG0584	C
RCA23_c22660	GlpD3 aerobic glycerol-3-phosphate dehydrogenase GlpD (1.1.99.5)	COG0578	C
RCA23_c22670	hypothetical protein, Na+/Pi-cotransporter	COG1283	P
RCA23_c22680	hypothetical protein, calcineurin-like phosphoesterase-like	COG1409	R
RCA23_c22690	ABC transporter permease protein	COG0395	G
RCA23_c22700	ABC transporter permease protein	COG1175	G
RCA23_c22710	ABC transporter extracellular solute-binding protein	COG1653	G
RCA23_c22720	ABC transporter ATP-binding protein	COG3839	G
RCA23_c22730	HTH-type transcriptional regulator, DeoR family	COG1349	K
RCA23_c22740	HAD-superfamily hydrolase, subfamily IIB	COG0561	R
RCA23_c22750	putative inner membrane transporter	COG0697	G
RCA23_c22760	YIi soluble aldose sugar dehydrogenase YIi (1.1.5..)	COG2133	G
RCA23_c22770	putative FKBP-type peptidyl-prolyl cis-trans isomerase	COG0545	O
RCA23_c22780	putative NirU family protein	COG4094	S
RCA23_c22790	hypothetical protein		
RCA23_c22800	putative adenine methyltransferase	COG0863	L
RCA23_c22810	hypothetical protein		
RCA23_c22820	phage uncharacterised protein	COG5410	S
RCA23_c22830	hypothetical protein, resolvase-like	COG1961	L
RCA23_c22840	putative prophage integrase	COG0582	L
RCA23_c22850	hypothetical protein, putative phage-like protein		
RCA23_c22860	hypothetical protein	COG3311	K
RCA23_c22870	hypothetical protein		
RCA23_c22880	hypothetical protein		
RCA23_c22890	DNA polymerase III	COG5545	R
RCA23_c22900	hypothetical protein, periplasmic binding protein-like	COG2998	H
RCA23_c22910	hypothetical protein		
RCA23_c22920	transposase	COG3316	L
RCA23_c22930	MoeA3 molybdopterin biosynthesis protein MoeA	COG0303	H
RCA23_c22940	MobB2 molybdopterin-guanine dinucleotide biosynthesis protein MobB	COG1763	H
RCA23_c22950	hypothetical protein, MobA-like	COG0746	H
RCA23_c22960	FdhD2 formate dehydrogenase family accessory protein FdhD	COG1526	C
RCA23_c22970	MoaE2 molybdopterin-converting factor subunit MoaE	COG0314	H
RCA23_c22980	MoaD2 molybdopterin-converting factor subunit MoaD	COG1977	H
RCA23_c22990	transposase	COG3328	L
RCA23_c23000	MoaA2 molybdenum cofactor biosynthesis protein MoaA	COG2896	H
RCA23_c23010	transposase		
RCA23_c23020	ABC transporter, ATP-binding cassette protein	COG3839	G
RCA23_c23030	ABC transporter, permease protein	COG4662	H
RCA23_c23040	FdhI formate dehydrogenase, gamma subunit FdhI (1.2.1.2)	COG2864	C
RCA23_c23050	FdhB formate dehydrogenase, iron-sulfur subunit FdhB (1.2.1.2)	COG0437	C
RCA23_c23060	FdhA formate dehydrogenase alpha subunit FdhA (1.2.1.2)	COG0243	C
RCA23_c23070	putative twin-arginine translocation pathway signal sequence domain		
RCA23_c23080	hypothetical protein, cytoplasmic chaperon TorD	COG3381	R
RCA23_c23090	hypothetical protein		
RCA23_c23100	putative molybdopterin-guanine dinucleotide biosynthesis protein A		
RCA23_c23110	hypothetical protein, 4Fe-4S ferredoxin-like	COG1148	C
RCA23_c23120	hypothetical protein		
RCA23_c23130	hypothetical protein		
RCA23_c23140	hypothetical protein, MrpNBP35 family protein	COG0489	D
RCA23_c23150	hypothetical protein	COG0340	H
RCA23_c23160	hypothetical protein		
RCA23_c23170	putative phage terminase		
RCA23_c23180	transposase	COG3316	L
RCA23_c23190	putative phage helicase	COG3378	R
RCA23_c23200	hypothetical protein		
RCA23_c23210	putative pyridoxal 4-dehydrogenase (1.1.1.107)	COG0667	C
RCA23_c23220	KduD 2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase KduD (1.1.1.127)	COG1028	I
RCA23_c23230	putative L-rhamnonate dehydratase rhamD	COG4948	M
RCA23_c23240	ureidoglycolate lyase (4.3.2.3)	COG0179	Q
RCA23_c23250	D-3-phosphoglycerate dehydrogenase (1.1.1.95)	COG1052	C
RCA23_c23260	uncharacterized oxidoreductase	COG1028	I
RCA23_c23270	putative rhamnose mutarotase RhaM	COG3254	S
RCA23_c23280	putative acetoacetate decarboxylase Adc (4.1.1.4)	COG4689	Q
RCA23_c23290	uncharacterized oxidoreductase	COG0673	R
RCA23_c23300	transcriptional regulator, GntR family	COG2186	K
RCA23_c23310	putative ribose ABC transport system, permease protein RbsC	COG1172	G
RCA23_c23320	putative ribose ABC transport system, permease protein RbsC	COG1172	G
RCA23_c23330	RbsA1 ribose import ABC transporter, ATP-binding protein RbsA (3.6.3.17)	COG1129	G
RCA23_c23340	putative rhamnose ABC transport system, substrate-binding protein RhaS	COG1879	G
RCA23_c23350	galactonate dehydratase (4.2.1.6)	COG4948	M
RCA23_c23360	lyase isomerase-like	COG1082	G
RCA23_c23370	putative aldo/keto reductase	COG0667	C
RCA23_c23380	amidohydrolase	COG3618	R
RCA23_c23390	hypothetical protein		
RCA23_c23400	transposase	COG3316	L
RCA23_c23410	RhaM L-rhamnose mutarotase RhaM (5.1.3..)	COG3254	S
RCA23_c23420	ureidoglycolate lyase (4.3.2.3)	COG0179	Q
RCA23_c23430	dihydrodipicolinate synthase DapA (4.2.1.52)	COG0329	E
RCA23_c23440	amidohydrolase	COG3618	R
RCA23_c23450	FabG4 3-ketoacyl-(acyl-carrier-protein) reductase FabG (1.1.1.100)	COG1028	I
RCA23_c23460	RhmD L-rhamnonate dehydratase RhmD (4.2.1.90)	COG4948	M
RCA23_c23470	PotA2 polyamine ABC transporter, ATP-binding protein PotA	COG3842	E
RCA23_c23480	polyamine ABC transporter, permease protein	COG1177	E
RCA23_c23490	polyamine ABC transporter, permease protein	COG1176	E
RCA23_c23500	polyamine ABC transporter, substrate binding protein	COG0687	E
RCA23_c23510	hypothetical protein		
RCA23_c23520	hypothetical protein		
RCA23_c23530	uncharacterized oxidoreductase	COG1028	I
RCA23_c23540	L-rhamnonate dehydratase	COG4948	M
RCA23_c23550	RbsC1 ribose ABC transport system, permease protein RbsC	COG1172	G
RCA23_c23560	RbsA2 ribose import ABC transporter, ATP-binding protein RbsA (3.6.3.17)	COG1129	G
RCA23_c23570	RbsC2 ribose import ABC transporter, substrate binding protein RbsC	COG1879	G
RCA23_c23580	HTH-type transcriptional regulator, GntR family	COG1609	K
RCA23_c23590	hypothetical protein	COG3311	K
RCA23_c23600	putative DNA-binding protein	COG1396	K
RCA23_c23610	GlxB glutamine amidotransferase-like protein GlxB (2.4.2..)	COG0067	E
RCA23_c23620	GlxC glutamate synthase alpha subunit GlxC (1.4.1.13)	COG0070	E
RCA23_c23630	GlxD glutamate synthase large subunit GlxD (1.4.1.13)	COG0069	E
RCA23_c23640	Glnglutamine synthase type III (6.3.1.2)	COG0174	E
RCA23_c23650	hypothetical protein		
RCA23_c23660	methyltransferase	COG0863	L
RCA23_c23670	hypothetical protein		
RCA23_c23680	putative terminase, large subunit	COG5410	S
RCA23_c23690	hypothetical protein, DUF2924		
RCA23_c23700	hypothetical protein, resolvase-like	COG1961	L
RCA23_c23720	signal transduction histidine kinase	COG0642	T
RCA23_c23730	MerA mercuric reductase MerA (1.16.1.1)	COG1249	C
RCA23_c23740	hypothetical protein	COG0398	S
RCA23_c23750	putative ribonuclease P	COG0594	J
RCA23_c23760	hypothetical protein DUF37	COG0759	S
RCA23_c23770	hypothetical protein		
RCA23_c23780	TtcA tRNA 2-thiocytidine biosynthesis protein TtcA	COG0037	D
RCA23_c23790	putative diguanylate phosphodiesterase	COG2200	T
RCA23_c23800	OxaA inner membrane protein OxaA	COG0706	U
RCA23_c23810	hypothetical protein, molybdenum cofactor sulfurase	COG3217	R
RCA23_c23820	putative GTP-binding protein EngB	COG0218	R
RCA23_c23830	ArgB acetylglutamate kinase ArgB (2.7.2.8)	COG0548	E
RCA23_c23840	putative fatty acid hydroxylase	COG3000	I
RCA23_c23850	hypothetical protein	COG1148	C
RCA23_c23860	hypothetical protein, phosphoglycerate mutase	COG2062	T
RCA23_c23870	BztD glutamate/glutamine/aspartate/asparagine transport ATP-binding protein Bzt	COG1126	E
RCA23_c23880	BztC glutamate/glutamine/aspartate/asparagine transport system permease prot	COG0765	E
RCA23_c23890	BztB glutamate/glutamine/aspartate/asparagine transport system permease prot	COG4597	E
RCA23_c23900	BztA glutamate/glutamine/aspartate/asparagine-binding protein BztA	COG0834	E
RCA23_c23910	ATP chaperone protein	COG5387	O
RCA23_c23920	putative phosphoglycerate phosphatase	COG0546	R
RCA23_c23930	RluC ribosomal large subunit pseudouridine synthase C (5.4.99..)	COG0564	J
RCA23_c23940	putative ccrb-like protein	COG0239	D
RCA23_c23950	RarA replication-associated recombination protein A	COG2256	L
RCA23_c23960	HTH-type transcriptional regulator, LuxR family	COG2197	T
RCA23_c23970	RplQ 50S ribosomal protein L17	COG0203	J
RCA23_c23980	RpoA DNA-directed RNA polymerase alpha subunit RpoA (2.7.7.6)	COG0202	K
RCA23_c23990	RpsK 30S ribosomal protein S11	COG0100	J
RCA23_c24000	RpsM 30S ribosomal protein S13	COG0099	J
RCA23_c24010	Adk adenylate kinase Adk (2.7.4.3)	COG0563	F
RCA23_c24020	SecY preprotein translocase subunit SecY	COG0201	U
RCA23_c24030	RplO 50S ribosomal protein L15	COG0200	J
RCA23_c24040	RpmD 50S ribosomal protein L30	COG1841	J
RCA23_c24050	RpsE 30S ribosomal protein S5	COG0098	J
RCA23_c24060	RplR 50S ribosomal protein L18	COG0256	J
RCA23_c24070	RplF 50S ribosomal protein L6	COG0097	J
RCA23_c24080	RpsH 30S ribosomal protein S8	COG0096	J
RCA23_c24090	RpsN 30S ribosomal protein S14	COG0199	J
RCA23_c24100	RplE 50S ribosomal protein L5	COG0094	J
RCA23_c24110	RplX 50S ribosomal protein L24	COG0198	J
RCA23_c24120	RplN 50S ribosomal protein L14	COG0093	J
RCA23_c24130	RpsQ 30S ribosomal protein S17	COG0186	J
RCA23_c24140	RpmC 50S ribosomal protein L29	COG0255	J
RCA23_c24150	hypothetical protein		
RCA23_c24160	RplW 50S ribosomal protein L23	COG0089	J
RCA23_c24170	RplD 50S ribosomal protein L4	COG0088	J
RCA23_c24180	RplC 50S ribosomal protein L3	COG0087	J
RCA23_c24190	RpsJ 30S ribosomal protein S10	COG0051	J
RCA23_c24200	TufA elongation factor Tu (EF-Tu)	COG0050	J
RCA23_c24210	FusA elongation factor FusA	COG0480	J
RCA23_c24220	RpsG 30S ribosomal protein S7	COG0049	J
RCA23_c24230	RpsL 30S ribosomal protein S12	COG0048	J
RCA23_c24240	hypothetical protein	COG1072	H
RCA23_c24250	ABC transporter ATP-binding protein	COG1129	G
RCA23_c24260	ABC tansporter permease protein	COG1172	G
RCA23_c24270	putative ABC transporter periplasmic binding protein	COG1879	G
RCA23_c24280	putative transcriptional repressor	COG1940	K
RCA23_c24290	hypothetical protein, acyl-CoA thioesterase- like	COG2755	E
RCA23_c24300	ABC transporter ATP-binding protein	COG1136	V
RCA23_c24310	putative ABC transporter permease protein	COG3127	Q
RCA23_c24320	hypothetical protein, gylcosyl transferase family 2	COG0463	M
RCA23_c24330	aminotransferase class-III (2.6.1..)	COG0161	H
RCA23_c24340	putative ectoine utilization protein EutD	COG0006	E
RCA23_c24350	putative integral membrane protein DUF5		
RCA23_c24360	4-coumarate--CoA ligase (6.2.1..)	COG0318	I
RCA23_c24370	hypothetical protein	COG0679	R
RCA23_c24380	hypothetical protein DUF125	COG1814	S
RCA23_c24390	transcriptional regulator, AsnC family	COG1522	K
RCA23_c24400	ArcA arginase ArcA (3.5.3.1)	COG0010	E
RCA23_c24410	ArcB ornithine cyclodeaminase ArcB (4.3.1.12)	COG2423	E
RCA23_c24420	PutA bifunctional protein PutA (1.5.99.8;1.5.1.12)	COG4230	C
RCA23_c24430	hypothetical protein		
RCA23_c24440	hypothetical protein		

RCA23_c24460	hypothetical protein		
RCA23_c24450	creatinase (3.5.3.3)	COG0006	E
RCA23_c24470	putative hydrolase	COG0596	R
RCA23_c24480	ProW:glycine betaine/L-proline transport system permease protein ProW	COG4176	E
RCA23_c24490	ProV:glycine betaine/L-proline transport ATP-binding protein ProV	COG4175	E
RCA23_c24500	ProX:glycine betaine-binding periplasmic protein ProX	COG2113	E
RCA23_c24510	putative HTH-type transcriptional regluator, LysR family	COG0583	K
RCA23_c24520	OcyD:D-cysteine desulfhydrase DcyD (4.4.1.15)	COG2515	E
RCA23_c24530	hypothetical protein		
RCA23_c24540	ZnuA:high-affinity zinc uptake system protein ZnuA	COG4531	P
RCA23_c24550	Zur:zinc uptake regulator	COG0735	P
RCA23_c24560	ZnuC:zinc import ATP-binding protein ZnuC (3.6.3.-)	COG1121	P
RCA23_c24570	ZnuB:high-affinity zinc uptake system membrane protein ZnuB	COG1108	P
RCA23_c24580	Hpd:4-hydroxyphenylpyruvate dioxygenase Hpd (1.13.11.27)	COG3185	E
RCA23_c24590	hypothetical protein		GI 9
RCA23_c24600	hypothetical protein		GI 9
RCA23_c24610	hypothetical protein	COG1028	I GI 9
RCA23_c24620	hypothetical protein	COG1402	R GI 9
RCA23_c24630	putative gluconolactonase	COG3386	G GI 9
RCA23_c24640	RbsK:ribokinase RbsK (2.7.1.15)	COG0524	G GI 9
RCA23_c24650	RbsD:ribose ABC transporter protein RbsD	COG4154	G GI 9
RCA23_c24660	putative ribose ABC transporter, ATP-binding protein RbsA	COG3839	G GI 9
RCA23_c24670	putative ribose ABC transporter, substrate binding protein RbsB	COG1653	G GI 9
RCA23_c24680	putative ribose ABC transporter, permease protein RbsC	COG1175	G GI 9
RCA23_c24690	putative ribose ABC transporter, permease protein RbsC	COG0395	G GI 9
RCA23_c24700	hypothetical protein, L-fucose isomerase-like		GI 9
RCA23_c24710	sugar isomerase	COG0794	M GI 9
RCA23_c24720	XylB:2-xylulose kinase XylB (2.7.1.17)	COG1070	G GI 9
RCA23_c24730	SgcX:aminopeptidase SgcX (3.4.11.-)	COG1363	G GI 9
RCA23_c24740	BtpA:photosystem I biogenesis protein BtpA	COG0434	R GI 9
RCA23_c24750	PtdJ:pyridoxal 4-dehydrogenase PtdJ (1.1.1.107)	COG0667	C GI 9
RCA23_c24760	putative amidohydrolase	COG3618	R GI 9
RCA23_c24770	putative carbohydrate kinase, pfkB family	COG0524	G GI 9
RCA23_c24780	xylose isomerase family protein	COG4952	M GI 9
RCA23_c24790	putative HTH-type transcriptional regulator, GntR family	COG1802	K GI 9
RCA23_c24800	mandelate racemase/muconate lactonizing enzyme	COG4948	M GI 9
RCA23_c24810	fatty acid desaturase	COG3239	I GI 9
RCA23_c24820	hypothetical protein		GI 9
RCA23_c24830	transposase	COG3316	L GI 9
RCA23_c24840	transporter, LysE family	COG1280	E GI 9
RCA23_c24850	hypothetical protein		GI 9
RCA23_c24860	putative transporter, periplasmic binding protein	COG2358	R GI 9
RCA23_c24870	RbtD:ribitol 2-dehydrogenase RbtD (1.1.1.56)	COG4221	R GI 9
RCA23_c24880	hypothetical protein		GI 9
RCA23_c24890	UDP-N-acetylglucosamine--peptide N- acetylglucosaminyltransferase 110 kDa s	COG0457	R GI 9
RCA23_c24900	putative phage integrase	COG4974	L GI 9
RCA23_c24920	DacC:D-alanyl-D-alanine carboxypeptidase DacC (3.4.16.4)	COG1686	M
RCA23_c24930	Tmk:thymidylate kinase Tmk (2.7.4.9)	COG0125	F
RCA23_c24940	hypothetical protein, DNA polymerase III delta subunit	COG2812	L
RCA23_c24950	TatD family deoxyribonuclease	COG0084	L
RCA23_c24960	hypothetical protein, metallo-beta-lactamase	COG1235	R
RCA23_c24970	membrane transport protein	COG0679	R
RCA23_c24980	FAD dependent oxidoreductase	COG0665	E
RCA23_c24990	UvrB UvrABC system protein B	COG0556	L
RCA23_c25000	hypothetical protein		
RCA23_c25010	hypothetical protein		
RCA23_c25020	hypothetical protein		
RCA23_c25030	LysS:lysyl-tRNA synthase LysS (6.1.1.6)	COG1384	J
RCA23_c25040	hypothetical protein		
RCA23_c25050	putative D-Ala-D-Ala carboxypeptidase 3 (S13) family protein	COG2027	M
RCA23_c25060	hypothetical protein		
RCA23_c25070	NadD:nicotinic acid mononucleotide adenyltransferase NadD (2.7.7.18)	COG1057	H
RCA23_c25080	ABC transporter, ATP-binding cassette protein, ChvD family	COG0488	R
RCA23_c25090	CspA2:cold shock protein CspA	COG1278	K
RCA23_c25100	GltX1:glutamyl-tRNA synthase 1 (6.1.1.17)	COG0008	J
RCA23_c25110	NadE:glutamine-dependent NAD(+) synthase NadE (6.3.5.1)	COG0171	H
RCA23_c25120	putative phosphatidylinositol-4-phosphate 5- kinase	COG4642	S
RCA23_c25130	hypothetical protein	COG1028	I
RCA23_c25140	LeuA:2-isopropylmalate synthase LeuA (2.3.3.13)	COG0119	E
RCA23_c25150	MreB:rod shape-determining protein MreB	COG1077	D
RCA23_c25160	MreC:rod shape-determining protein MreC	COG1792	M
RCA23_c25170	hypothetical protein		
RCA23_c25180	MrdA:penicillin-binding protein 2	COG0768	M
RCA23_c25190	MrdB:rod shape-determining protein RodA	COG0772	D
RCA23_c25200	GhrA:glyoxylate/hydroxypyruvate reductase GhrA (1.1.1.79;1.1.1.81)	COG0111	H
RCA23_c25210	hypothetical protein		
RCA23_c25220	hypothetical protein	COG1573	L
RCA23_c25230	Nat/H+ antiporter NhaA	COG3004	P
RCA23_c25240	KpsT:capsule polysaccharide export protein KpsT (3.6.3.38)	COG1134	G
RCA23_c25250	KpsE:capsule polysaccharide export protein KpsE	COG3524	M
RCA23_c25270	hypothetical protein	COG1832	R
RCA23_c25280	PetF:ferredoxin PetF	COG0633	C
RCA23_c25290	selenium binding protein		
RCA23_c25300	aldehyde dehydrogenase (1.2.1.3)	COG1012	C
RCA23_c25310	DdtD2:CoA-transferase family III protein involved in DMSP degradation	COG1804	C
RCA23_c25320	GlnB2:nitrogen regulatory protein P-II 2	COG0347	E
RCA23_c25330	ammonium transporter	COG0004	P
RCA23_c25340	HpaR:homoprotocatechuate degradative operon repressor	COG1846	K
RCA23_c25350	HpcC:5-carboxymethyl-2-hydroxymuconic semialdehyde dehydrogenase HpcC (	COG1012	C
RCA23_c25360	HpaD:3,4-dihydroxyphenylacetate 2,3-dioxygenase HpaD (1.13.11.2)	COG0346	E
RCA23_c25370	putative lyase	COG0015	F
RCA23_c25380	hypothetical protein DUF849	COG3246	S
RCA23_c25390	PcaG:protocatechuate 3,4-dioxygenase alpha chain PcaG (1.13.11.3)	COG3485	Q
RCA23_c25400	PcaH:protocatechuate 3,4-dioxygenase beta chain PcaH (1.13.11.3)	COG3485	Q
RCA23_c25410	PcaC:4-carboxymuconolactone decarboxylase PcaC (4.1.1.44)	COG0599	S
RCA23_c25420	hypothetical protein		
RCA23_c25430	transcriptional regulatory protein	COG0745	T
RCA23_c25440	ChvG2:two component signal transduction histidine kinase ChvG	COG0642	T
RCA23_c25450	hypothetical protein		
RCA23_c25460	phosphate transporter	COG0306	P
RCA23_c25470	hypothetical protein		
RCA23_c25480	KdsA2:dehydro-3-deoxyphosphooctonate aldolase KdsA (2.5.1.55)	COG2877	M
RCA23_c25490	KdsD:arabinose 5-phosphate isomerase KdsD (5.3.1.13)	COG0794	M
RCA23_c25500	KdsB:3-deoxy -manno-octulosonate cytidylyltransferase KdsB (2.7.7.38)	COG1212	M
RCA23_c25510	hypothetical protein		
RCA23_c25520	bacterial surface protein		
RCA23_c25530	KatG:catalase-peroxidase KatG (1.11.1.6;1.11.1.7)	COG0376	P
RCA23_c25540	hypothetical protein	COG2067	I
RCA23_c25550	thioesterase/thiol ester dehydrase-isomerase	COG0824	R
RCA23_c25560	iron-containing alcohol dehydrogenase	COG1454	C
RCA23_c25570	hypothetical protein	COG1960	I
RCA23_c25580	nodulation protein N-like protein	COG2030	I
RCA23_c25590	putative phosphotransferase, eukaryotic acyl-CoA dehydrogenase	COG3173	R
RCA23_c25600	putative HTH-type transcriptional regulator, TetR family	COG1309	K
RCA23_c25610	FadJ2:fatty acid oxidation complex alpha subunit FadJ (4.2.1.17;1.1.1.35)	COG1250	I
RCA23_c25620	acyl-CoA dehydrogenase (1.3.99.-)	COG1960	I
RCA23_c25630	acyl-CoA dehydrogenase (1.3.99.-)	COG1960	I
RCA23_c25640	hypothetical protein, MarR family		
RCA23_c25650	acyl-CoA dehydrogenase (1.3.99.-)	COG1960	I
RCA23_c25660	Qor:quinone oxidoreductase (1.6.5.5.)	COG0604	C
RCA23_c25670	hypothetical protein		
RCA23_c25680	putative DNA repair protein RecO	COG1381	L
RCA23_c25690	hypothetical protein	COG5447	S
RCA23_c25700	Era:GTP-binding protein Era	COG1159	R
RCA23_c25710	Rnc:ribonuclease 3 (3.1.26.3)	COG0571	K
RCA23_c25720	LepB:signal peptidase I (3.4.21.89)	COG0681	U
RCA23_c25730	AcpS:holo-[acyl-carrier-protein] synthase AcpS (2.7.8.7)	COG0736	I
RCA23_c25740	PdxJ:pyridoxine 5'-phosphate synthase PdxJ (2.6.99.2)	COG0854	H
RCA23_c25750	hypothetical protein DUF2062	COG3216	S
RCA23_c25760	SpoT:guanosine-3',5'-bis(diphosphate) 3'- pyrophosphohydrolase SpoT (3.1.7.2)	COG0317	T
RCA23_c25770	RpoZ:DNA-directed RNA polymerase subunit omega (2.7.7.6)	COG1758	K
RCA23_c25780	putative 2-amino-4-hydroxy-6- hydroxymethyldihydropteridine pyrophosphokinas	COG0801	H
RCA23_c25790	hypothetical protein DUF88	COG1432	S
RCA23_c25800	IspH:4-hydroxy-3-methylbut-2-enyl diphosphate reductase IspH (1.17.1.2)	COG0761	I
RCA23_c25810	hypothetical protein, methyltransferase	COG2227	H
RCA23_c25820	RnhA:ribonuclease H (3.1.26.4)	COG0328	L
RCA23_c25830	hypothetical protein		
RCA23_c25840	Fmt:methionyl-tRNA formyltransferase Fmt (2.1.2.9)	COG0223	J
RCA23_c25850	Def1:peptide deformylase Def (3.5.1.88)	COG0242	J
RCA23_c25860	Def2:peptide deformylase Def (3.5.1.88)	COG0242	J
RCA23_c25870	aminotransferase	COG1168	E
RCA23_c25880	hypothetical protein		
RCA23_c25890	CobM:precorrin-4 C(11)-methyltransferase CobM (2.1.1.133)	COG2875	H
RCA23_c25900	CobJ:precorrin-3B C(17)-methyltransferase CobJ (2.1.1.131)	COG1010	H
RCA23_c25910	CobI:precorrin-2 C(20)-methyltransferase CobI (2.1.1.130)	COG2243	H
RCA23_c25920	CobL:precorrin-6Y C(5,15)-methyltransferase CobL (2.1.1.132)	COG2242	H
RCA23_c25930	CobH:precorrin-8X methylmutase CobH (5.4.1.2)	COG2082	H
RCA23_c25940	CbiX:sirohydrochlorin cobalt/chelatase CbiX (4.99.1.3)	COG2138	S
RCA23_c25950	methyltransferase, FkbM family	COG3774	M
RCA23_c25960	hypothetical protein, glutathione S-transferase	COG0435	O
RCA23_c25970	hypothetical protein DUF1636	COG5469	S
RCA23_c25980	CobC:threonine-phosphate decarboxylase CobC (4.1.1.81)	COG0079	E
RCA23_c25990	CobD:cobalamin biosynthesis protein CobD	COG1270	H
RCA23_c26000	putative peptidoglycan-binding lytic murein transglycosylase	COG2951	M
RCA23_c26010	putative chromosome partition protein smc	COG1196	D
RCA23_c26020	hypothetical protein		
RCA23_c26030	hypothetical protein	COG1028	I
RCA23_c26040	hypothetical protein DUF81	COG0730	R
RCA23_c26050	RpsB:50S ribosomal protein L2	COG0090	J
RCA23_c26060	RpsS:30S ribosomal protein S19	COG0185	J
RCA23_c26070	RplV:50S ribosomal protein L22	COG0091	J
RCA23_c26080	RpsC:30S ribosomal protein S3	COG0092	J
RCA23_c26090	RplP:50S ribosomal protein L16	COG0197	J
RCA23_c26100	hypothetical protein		GI 10
RCA23_c26120	hypothetical protein	COG0666	R GI 10
RCA23_c26130	transcriptional regulator		GI 10
RCA23_c26140	hypothetical protein		GI 10
RCA23_c26150	hypothetical protein, UPF0311		GI 10
RCA23_c26160	transcriptional regulator, MarR family	COG1846	K GI 10
RCA23_c26170	long-chain-fatty-acid-CoA ligase	COG0318	I GI 10
RCA23_c26180	DctM1:0.TRAP dicarboxylate transporter, subunit DctM	COG1593	G GI 10
RCA23_c26190	DctQ7:TRAP dicarboxylate transporter, subunit DctQ		GI 10
RCA23_c26200	DctP9:TRAP dicarboxylate transporter, subunit DctP	COG1638	G GI 10
RCA23_c26210	Crt:3-hydroxybutyryl-CoA dehydratase Crt (4.2.1.55)	COG1024	I GI 10
RCA23_c26220	metal-dependent hydrolase	COG2159	R GI 10
RCA23_c26230	FAD dependent monooxygenase	COG0654	H GI 10
RCA23_c26240	arylsulfatase (3.1.6.1)	COG3119	P GI 10
RCA23_c26250	transcriptional regulator, MarR family	COG1846	K GI 10
RCA23_c26260	putative transporter, periplasmic binding protein	COG2358	R GI 10
RCA23_c26270	TRAP transporter, 4TM/12TM fusion protein	COG4666	R GI 10
RCA23_c26280	hypothetical protein, UPF0261	COG5441	S GI 10
RCA23_c26290	PuuA1:gamma-glutamylputrescine synthase PuuA (6.3.1.11)	COG0174	E GI 10
RCA23_c26300	cytochrome P450	COG2124	Q GI 10
RCA23_c26310	glutamine amidotransferase class-I	COG0518	F GI 10
RCA23_c26320	TctC1:tripartite tricarboxylate transporter (TTT) protein TctC	COG3181	S GI 10
RCA23_c26330	TctA1:tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S GI 10
RCA23_c26340	hypothetical protein, transmembrane		GI 10
RCA23_c26350	PuuC:gamma-glutamyl-gamma-aminobutyraldehyde dehydrogenase PuuC (1.2.1.	COG1012	C GI 10
RCA23_c26360	hypothetical protein		GI 10
RCA23_c26370	hypothetical protein		GI 10
RCA23_c26380	hypothetical protein		GI 10
RCA23_c26390	putative fucosyltransferase		GI 10
RCA23_c26400	integrase	COG2801	L GI 10
RCA23_c26410	putative extracellular solute-binding protein	COG1840	P GI 10
RCA23_c26420	two-component system, sensor histidine kinase protein	COG0642	T GI 10
RCA23_c26430	two-component system, response regulator protein	COG0745	T GI 10
RCA23_c26440	TctC2:tripartite tricarboxylate transporter (TTT) protein TctC	COG3181	S GI 10
RCA23_c26450	TctB:tripartite tricarboxylate transporter (TTT) protein TctB		GI 10
RCA23_c26460	TctA2:tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S GI 10
RCA23_c26470	hypothetical protein		GI 10
RCA23_c26480	hypothetical protein		GI 10
RCA23_c26490	transposase A		GI 10



RCA23_c28670	putative ribosomal RNA small subunit methyltransferase B	COG0144	J	
RCA23_c28680	hypothetical protein, heparinase I/III	COG5360	S	
RCA23_c28690	PurH bifunctional purine biosynthesis protein PurH (2.1.2.3;3.5.4.10)	COG0138	F	
RCA23_c28700	signal peptidase II (3.4.23.36)	COG0597	M	
RCA23_c28710	uncharacterized zinc protease yAwA (3.4.24.-)	COG0612	R	
RCA23_c28720	uncharacterized zinc protease yAwB (3.4.24.-)	COG0612	R	
RCA23_c28730	MutL DNA mismatch repair protein MutL	COG0323	L	
RCA23_c28740	putative RmuC family protein	COG1322	S	
RCA23_c28750	ChrR transcriptional activator ChrR	COG1806	T	
RCA23_c28760	SigK RNA polymerase sigma factor SigK	COG1595	K	
RCA23_c28770	hypothetical protein	COG2907	R	
RCA23_c28780	hypothetical protein DUF1365	COG3496	S	
RCA23_c28790	hypothetical protein	COG2211	G	
RCA23_c28800	hypothetical protein			
RCA23_c28810	putative short chain dehydrogenase	COG4221	R	
RCA23_c28820	saccharopine dehydrogenase (NAD+,L-lysine- forming) (1.5.1.7)	COG3288	C	
RCA23_c28830	putative glutathione S-transferase	COG0625	O	
RCA23_c28840	putative phosphoglycerate mutase family protein	COG0406	G	
RCA23_c28850	Fdh formate dehydrogenase Fdh (1.2.1.2)	COG1383	R	
RCA23_c28860	hypothetical protein	COG1357	S	
RCA23_c28870	hypothetical protein DUF1643	COG4333	S	
RCA23_c28880	DehH haloacetate dehalogenase DehH (3.8.1.3)	COG0596	R	
RCA23_c28890	RpsO:30S ribosomal protein S15	COG0184	J	
RCA23_c28900	hypothetical protein, DUF6 transmembrane protein	COG0697	G	
RCA23_c28910	Pnp:polyribonucleotide nucleotidyltransferase Pnp (2.7.7.8)	COG1185	J	
RCA23_c28920	aldehyde dehydrogenase (1.2.1.3)	COG1012	C	
RCA23_c28930	RluA:ribosomal large subunit pseudouridine synthase A (5.4.99.-)	COG0564	J	
RCA23_c28940	OtnG protein OtnG			
RCA23_c28950	hypothetical protein, DUF940 putative lipoprotein			
RCA23_c28960	hypothetical protein			
RCA23_c28970	GalE:UDP-glucose 4-epimerase GalE (5.1.3.2)	COG1087	M	
RCA23_c28980	putative FAD linked oxidase	COG0277	C	
RCA23_c28990	Ffh:signal recognition particle protein Ffh	COG0541	U	
RCA23_c29000	putative chorismate mutase type II	COG1605	E	
RCA23_c29010	RpsP:30S ribosomal protein S16	COG0228	J	
RCA23_c29020	RimM:ribosome maturation factor RimM	COG0806	J	
RCA23_c29030	TrmD:tRNA (guanine-N(1)-)-methyltransferase TrmD (2.1.1.31)	COG0336	J	
RCA23_c29040	RplS:50S ribosomal protein L19	COG0335	J	
RCA23_c29050	RpmE:50S ribosomal protein L31	COG0254	J	
RCA23_c29060	MjlZ:ATPase MjlZ	COG1192	D	
RCA23_c29070	fructokinase (2.7.1.4)	COG0524	G	
RCA23_c29080	PobA:P-hydroxybenzoate hydroxylase PobA (1.14.13.2)	COG0654	H	
RCA23_c29090	KynU:kynureninase KynU (3.7.1.3)	COG3844	E	
RCA23_c29100	HTH-type transcriptional regulator, GntR family	COG1802	K	
RCA23_c29110	hypothetical protein, glutathione S-transferase	COG0625	O	
RCA23_c29120	L-threonine 3-dehydrogenase (1.1.1.103)	COG0451	M	
RCA23_c29130	ainyl-CoA hydratase (4.2.1.17)	COG1024	I	
RCA23_c29140	ainyl-CoA hydratase (4.2.1.17)	COG1250	I	
RCA23_c29150	FadH2:3-hydroxyacyl-CoA dehydrogenase FadH (1.1.1.35)	COG1028	I	
RCA23_c29160	hypothetical protein			
RCA23_c29170	integrase	COG4974	L	
RCA23_c29180	hypothetical protein, porin-like			
RCA23_c29190	LeuS:leucyl-tRNA synthase LeuS (6.1.1.4)	COG0495	J	
RCA23_c29200	hypothetical protein DUF2159			
RCA23_c29210	hypothetical protein			
RCA23_c29220	hypothetical protein, glutathione S-transferase	COG0625	O	
RCA23_c29230	hypothetical protein	COG2081	R	
RCA23_c29240	hypothetical protein, porin			
RCA23_c29250	hypothetical protein, alanine racemase-like	COG0325	R	
RCA23_c29260	hypothetical protein, YkuD	COG3786	S	
RCA23_c29270	RibA:GTP cyclohydrolase II RibA (3.5.4.25)	COG0807	H	
RCA23_c29280	response regulator receiver protein	COG0745	T	
RCA23_c29290	hydratase/decarboxylase (4.1.1.-)	COG3971	Q	
RCA23_c29300	5-carboxymethyl-2-hydroxymuconate delta- isomerase (5.3.3.10)	COG0179	Q	
RCA23_c29310	5-carboxymethyl-2-hydroxymuconate delta isomerase (5.3.3.10)	COG3232	E	
RCA23_c29320	hypothetical protein, HTH-type transcriptional regulator, MarR family	COG1846	K	
RCA23_c29330	HpaH:4-hydroxyphenylacetate 3-monooxygenase oxygenase component HpaH (	COG2368	Q	
RCA23_c29340	ABC sugar transporter, ATPase subunit	COG3839	G	
RCA23_c29350	MtlK:mannitol 2-dehydrogenase MtlK (1.1.1.67)	COG0246	G	
RCA23_c29360	transcriptional regulator, AraC family	COG2207	K	
RCA23_c29370	ABC sugar transporter, permease protein	COG0395	G	
RCA23_c29380	ABC sugar transporter, permease protein	COG1175	G	
RCA23_c29390	ABC transporter, periplasmic substrate-binding protein	COG1653	G	
RCA23_c29400	putative flavohemoglobin / bacterial hemoglobin	COG1017	C	
RCA23_c29410	hypothetical protein			
RCA23_c29420	NhaA:Na+/H+ antiporter NhaA	COG3004	P	
RCA23_c29430	hypothetical protein			
RCA23_c29440	hypothetical protein DUF55	COG2947	S	
RCA23_c29450	YciI protein YciI	COG2350	S	
RCA23_c29460	GpsA:glycerol-3-phosphate dehydrogenase GpsA (1.1.1.94)	COG0240	C	
RCA23_c29470	HemD:hypothetical protein, uroporphyrinogen-III synthase HemD	COG1587	H	
RCA23_c29480	hypothetical protein	COG3264	M	
RCA23_c29490	hypothetical protein, HemY	COG3898	S	
RCA23_c29500	transcriptional regulatory protein, Ars family	COG0640	K	
RCA23_c29510	BchO:magnesium-chelatase subunit BchO (6.6.1.1)	COG0596	R	
RCA23_c29520	BchD:magnesium-chelatase subunit BchD (6.6.1.1)	COG1240	H	
RCA23_c29530	BchI:magnesium-chelatase subunit BchI (6.6.1.1)	COG1239	H	
RCA23_c29540	CrtA:spheroidene monooxygenase CrtA (1.-.-.-)			
RCA23_c29550	CrtI:phytoene dehydrogenase CrtI (1.14.99.-)	COG1233	Q	
RCA23_c29560	CrtB:phytoene synthase CrtB (2.5.1.32)	COG1562	I	
RCA23_c29570	CrtC:hydroxyneurosporene dehydrogenase CrtC (1.-.-.-)			
RCA23_c29580	CrtD:methoxyneurosporene dehydrogenase CrtD (1.14.99.-)	COG1233	Q	
RCA23_c29590	CrtE:geranylgeranyl pyrophosphate synthase CrtE (2.5.1.29)	COG0142	H	
RCA23_c29600	CrtF:hydroxyneurosporene methyltransferase CrtF (2.1.1.-)			
RCA23_c29610	BchC:2-desacetyl-2-hydroxyethyl bacteriochlorophyllide A dehydrogenase BchC	COG1063	E	
RCA23_c29620	BchX:chlorophyllide reductase BchX (1.3.1.-)	COG1348	P	
RCA23_c29630	BchY:chlorophyllide reductase BchY (1.18.1.-)			
RCA23_c29640	BchZ:chlorophyllide reductase subunit BchZ (1.18.-.-)	COG2710	C	
RCA23_c29650	PufQ:protein PufQ			
RCA23_c29660	PufB:light-harvesting protein B-870 beta chain PufB			
RCA23_c29670	PufA:light-harvesting protein B-870 alpha chain PufA			
RCA23_c29680	PufL:reaction center protein L chain PufL			
RCA23_c29690	PufM:reaction center protein M chain PufM			
RCA23_c29700	protein PufX			
RCA23_c29710	Dxs:2.1-deoxy-D-xylulose-5-phosphate synthase Dxs (2.2.1.7)	COG1154	H	
RCA23_c29720	Idi2:isopentenyl-diphosphate delta-isomerase Idi (5.3.3.2)	COG1443	I	
RCA23_c29730	BchP:geranylgeranyl reductase BchP	COG0644	C	
RCA23_c29740	bacteriochlorophyll synthase 44.5 kDa chain			
RCA23_c29750	BchG:bacteriochlorophyll synthase BchG	COG0382	H	
RCA23_c29760	cytochrome c-551	COG3474	C	
RCA23_c29770	HemE:uroporphyrinogen decarboxylase HemE (4.1.1.37)	COG0407	H	
RCA23_c29780	HemC:porphobilinogen deaminase HemC (2.5.1.61)	COG0181	H	
RCA23_c29790	hypothetical protein, NmrA-like	COG0702	M	
RCA23_c29800	HemA:5-aminolevulinate synthase HemA (2.3.1.37)	COG0156	H	
RCA23_c29810	hypothetical protein			
RCA23_c29820	AcsF:aerobic Mg-protoporphyrin IX monomethyl ester oxidative cyclase AcsF (1.14.13.81)			
RCA23_c29830	hypothetical protein			
RCA23_c29840	hypothetical protein			
RCA23_c29850	hypothetical protein			
RCA23_c29860	PuhA:reaction center protein PuhA			
RCA23_c29870	PucC:protein PucC			
RCA23_c29880	BchM:magnesium-protoporphyrin O-methyltransferase BchM (2.1.1.11)	COG2227	H	
RCA23_c29890	BchL:light-independent protochlorophyllide reductase iron-sulfur ATP-binding prot	COG1348	P	
RCA23_c29900	BchH:magnesium-chelatase subunit BchH (6.6.1.1)	COG1429	H	
RCA23_c29910	BchB:light-independent protochlorophyllide reductase subunit BchB (1.18.-.-)	COG2710	C	
RCA23_c29920	BchN:light-independent protochlorophyllide reductase subunit BchN	COG2710	C	
RCA23_c29930	BchF:2-vinyl bacteriochlorophyllide hydratase BchF (4.2.1.-)			
RCA23_c29940	putative transcriptional regulator PpaA			
RCA23_c29950	PpsR:transcriptional regulator PpsR	COG3829	K	
RCA23_c29960	hypothetical protein	COG5012	R	
RCA23_c29970	FixJ:transcriptional regulator protein FixJ	COG4566	T	
RCA23_c29980	TspO2:peripheral-type benzodiazepine receptor/signal transduction protein TspO	COG3476	T	
RCA23_c29990	hypothetical protein, integral membrane proteins YeeE/YedE	COG2391	R	
RCA23_c30000	hypothetical protein, integral membrane proteins YeeE/YedE	COG2391	R	
RCA23_c30010	putative beta-lactamase hydrolase-like protein	COG0491	R	
RCA23_c30020	hypothetical protein	COG3453	S	
RCA23_c30030	sulphate transporter	COG0659	P	
RCA23_c30040	amidase (3.5.1.4)	COG0154	J	
RCA23_c30050	hypothetical protein			
RCA23_c30060	Aph:aminoglycoside phosphotransferase Aph (2.7.1.95)			
RCA23_c30070	hypothetical protein	COG2885	M	
RCA23_c30080	DnaJ:chaperone protein DnaJ	COG0484	O	
RCA23_c30090	DnaK:chaperone protein DnaK	COG0443	O	
RCA23_c30100	putative alpha-ketoglutarate-dependent dioxygenase AlkB	COG3145	L	
RCA23_c30110	putative ABC-2 type transporter	COG1682	G	
RCA23_c30120	CysQ:3'(2'),5'-bisphosphate nucleotidase CysQ (3.1.3.7)	COG1218	P	
RCA23_c30130	GalU:UTP-glucose-1-phosphate uridylyltransferase GalU (2.7.7.9)	COG1210	M	
RCA23_c30140	hypothetical protein	COG0463	M	
RCA23_c30150	hypothetical protein			
RCA23_c30160	hypothetical protein beta-1,6-N- acetylglucosaminyltransferases			
RCA23_c30170	hypothetical protein			
RCA23_c30180	PtsN:nitrogen regulatory protein PtsN (2.7.1.-)	COG1762	G	
RCA23_c30190	putative sigma 54 modulation protein	COG1544	J	
RCA23_c30200	LptB:lipopolysaccharide export system ATP-binding protein LptB (3.6.3.-)	COG1137	R	
RCA23_c30210	putative lipopolysaccharide export system protein LptA	COG1934	S	
RCA23_c30220	hypothetical protein			
RCA23_c30230	putative 3'-5'-exonuclease	COG0349	J	
RCA23_c30240	glycine amidinotransferase (2.1.4.1)	COG1834	E	
RCA23_c30250	hypothetical protein	COG4123	R	
RCA23_c30260	hypothetical protein, metal-dependent phosphohydrolase	COG1896	R	
RCA23_c30270	AhcY:S-adenosyl-L-homocysteine hydrolase AhcY (3.3.1.1)	COG0499	H	
RCA23_c30280	RegA:photosynthetic apparatus regulatory protein RegA	COG4567	T	
RCA23_c30290	SenC:protein SenC	COG1999	R	
RCA23_c30300	RegB:sensor histidine kinase RegB (2.7.13.3)	COG0642	T	
RCA23_c30310	hypothetical protein			
RCA23_c30320	hypothetical protein, aminoglycoside phosphotransferase	COG3178	R	
RCA23_c30330	hypothetical protein, nucleotidyl transferase	COG1208	M	
RCA23_c30340	AddB:double-strand break repair protein AddB	COG3893	L	
RCA23_c30350	AddA:double-strand break repair helicase AddA	COG1074	L	
RCA23_c30360	TrxA:thioredoxin TrxA	COG3118	O	
RCA23_c30370	HslV:ATP-dependent protease HslV (3.4.25.-)	COG5405	O	
RCA23_c30380	HslU:ATP-dependent hsl protease ATP-binding subunit HslU	COG1220	O	
RCA23_c30390	MFS-type transporter	COG2814	G	
RCA23_c30400	hypothetical protein, Smr protein/MutS2	COG2840	S	
RCA23_c30410	putative lytic murein transglycosylase	COG2821	M	
RCA23_c30420	hypothetical protein, TIM44	COG4395	S	
RCA23_c30430	putative cytoplasmic membrane protein FxsA	COG3030	R	
RCA23_c30440	SecB:protein-export protein SecB	COG1952	U	
RCA23_c30450	DnaQ:DNA polymerase III subunit epsilon (2.7.7.7)	COG0847	L	
RCA23_c30460	CoaE:dephospho-CoA kinase CoaE (2.7.1.24)	COG0237	H	
RCA23_c30470	AroE:shikimate 5-dehydrogenase AroE (1.1.1.25)	COG0169	E	
RCA23_c30480	Rho:transcription termination factor Rho (3.6.1.-)	COG1158	K	
RCA23_c30490	MnmE:tRNA modification GTPase MnmE (3.6.-.-)	COG0486	R	
RCA23_c30500	MnmG:tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	COG0445	D	
RCA23_c30510	RsmG:ribosomal RNA small subunit methyltransferase G (2.1.1.-)	COG0357	M	
RCA23_c30520	ParA:chromosome-partitioning protein ParA	COG1192	D	
RCA23_c30530	ParB:chromosome-partitioning protein ParB	COG1475	K	
RCA23_c30540	HemN:oxygen-independent coproporphyrinogen III oxidase HemN (1.3.99.22)	COG0635	H	
RCA23_c30550	RdgB:nucleoside-triphosphatase RdgB (3.6.1.15)	COG0127	F	
RCA23_c30560	Rph:ribonuclease PH (2.7.7.56)	COG0689	J	
RCA23_c30570	HrcA:heat-inducible transcription repressor HrcA	COG1420	K	
RCA23_c30580	GrpE:protein GrpE	COG0576	O	
RCA23_c30590	MutS:DNA mismatch repair protein MutS	COG0249	L	
RCA23_c30600	MaeB:NADP-dependent malic enzyme MaeB (1.1.1.40)	COG0281	C	
RCA23_c30610	BNB family carbohydrate kinase	COG0524	G	
RCA23_c30620	putative aminotransferase class IV	COG0115	E	
RCA23_c30630	hypothetical protein, branched-chain-amino-acid aminotransferase-like			
RCA23_c30640	ArgG:argininosuccinate synthase ArgG (6.3.4.5)	COG0137	E	
RCA23_c30650	threonine dehydratase, biosynthetic (4.3.1.19)	COG1171	E	
RCA23_c30660	putative NUDIX hydrolase	COG0494	L	
RCA23_c30670	HslO:heat shock protein 33	COG1281	O	
RCA23_c30680	putative NUDIX hydrolase	COG0494	L	

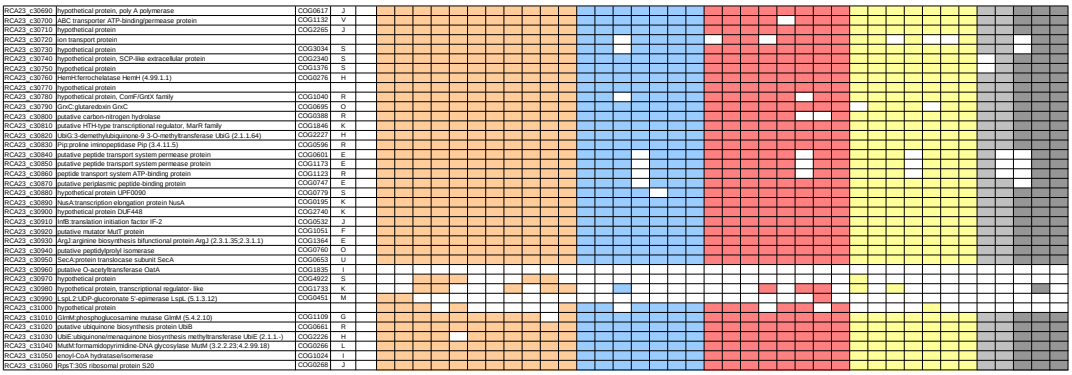

RCA23_c30690 hypothetical protein, poly A polymerase COG0617 J
RCA23_c30700 ABC transporter ATP-binding/permease protein COG1132 V
RCA23_c30710 hypothetical protein COG2265 J
RCA23_c30720 ion transport protein
RCA23_c30730 hypothetical protein COG3034 S
RCA23_c30740 hypothetical protein, SCP-like extracellular protein COG2340 S
RCA23_c30750 hypothetical protein COG1376 S
RCA23_c30760 HemH:ferrochelatase HemH (4.99.1.1) COG0276 H
RCA23_c30770 hypothetical protein
RCA23_c30780 hypothetical protein, ComF/GntX family COG1040 R
RCA23_c30790 GrxC:glutaredoxin GrxC COG0695 O
RCA23_c30800 putative carbon-nitrogen hydrolase COG0388 R
RCA23_c30810 putative HTH-type transcriptional regulator, MarR family COG1846 K
RCA23_c30820 UbiG:3-demethylubiquinone-9 3-O-methyltransferase UbiG (2.1.1.64) COG2227 H
RCA23_c30830 Pip:proline iminopeptidase Pip (3.4.11.5) COG0596 R
RCA23_c30840 putative peptide transport system permease protein COG0601 E
RCA23_c30850 putative peptide transport system permease protein COG1173 E
RCA23_c30860 peptide transport system ATP-binding protein COG1123 R
RCA23_c30870 putative periplasmic peptide-binding protein COG0747 E
RCA23_c30880 hypothetical protein UPF0090 COG0779 S
RCA23_c30890 NusA:transcription elongation protein NusA COG0195 K
RCA23_c30900 hypothetical protein DUF448 COG2740 K
RCA23_c30910 InfB:translation initiation factor IF-2 COG0532 J
RCA23_c30920 putative mutator MutT protein COG1051 F
RCA23_c30930 ArgJ:arginine biosynthesis bifunctional protein ArgJ (2.3.1.35;2.3.1.1) COG1364 E
RCA23_c30940 putative peptidylprolyl isomerase COG0760 O
RCA23_c30950 SecA:protein translocase subunit SecA COG0653 U
RCA23_c30960 putative O-acetyltransferase OatA COG1835 I
RCA23_c30970 hypothetical protein COG4922 S
RCA23_c30980 hypothetical protein, transcriptional regulator- like COG1733 K
RCA23_c30990 LgsL2:UDP-glucuronate 5'-epimerase LsgL (5.1.3.12) COG0451 M
RCA23_c31000 hypothetical protein
RCA23_c31010 GlmM:phosphoglucosamine mutase GlmM (5.4.2.10) COG1109 G
RCA23_c31020 putative ubiquinone biosynthesis protein UbiB COG0661 R
RCA23_c31030 UbiE:ubiquinone/menaquinone biosynthesis methyltransferase UbiE (2.1.1.-) COG2226 H
RCA23_c31040 MutM:formamidopyrimidine-DNA glycosylase MutM (3.2.2.23;4.2.99.18) COG0266 L
RCA23_c31050 Enoyl-CoA hydratase/isomerase COG1024 I
RCA23_c31060 RpsT:30S ribosomal protein S20 COG0268 J

Table S6: COG categories, genomic islands, normalized read counts (NPKM) of the transcripts at stations 9 (bloom,night) and 13 (bloom, day) and differences between night and day

locus_tag	start	stop	annotation	COG ID	COG category	genomic island	NPKM station 9 bloom-night	NPKM station 13 bloom-day	NPKM difference night-day
RCA23_c00010	101	1447	chromosomal replication initiatior protein DnaA	COG0593	L		617	502	115
RCA23_c00020	1550	2668	DNA polymerase III beta subunit DnaN	COG0592	L		605	575	30
RCA23_c00030	2778	3800	DNA replication and repair protein RecF	COG1195	L		358	419	-61
RCA23_c00040	3797	4417	hypothetical protein, LysE type translocator	COG1280	E		455	635	-180
RCA23_c00050	4567	6984	DNA gyrase subunit B	COG0187	L		599	596	3
RCA23_c00060	7002	8384	FAD dependent oxidoreductase	COG0665	E		339	421	-82
RCA23_c00070	9282	8392	hypothetical protein, rhodanese-like sulphurtransferase	COG1054	R		414	392	22
RCA23_c00080	9413	10003	pyrazinamidase/nicotinamidase PncA	COG1335	Q		456	502	-46
RCA23_c00090	10003	11295	nicotinate phosphoribosyltransferase PncB	COG1488	H		608	548	60
RCA23_c00100	12224	11292	aminoglycoside phosphotransferase	COG2334	R		392	456	-64
RCA23_c00110	13477	12221	putative aminotransferase class III	COG0160	E		255	339	-84
RCA23_c00120	13889	13467	putative peptide chain release factor	COG1186	J		432	659	-227
RCA23_c00130	14772	14128	hypothetical integral membrane protein	COG1738	S		378	472	-94
RCA23_c00140	15799	14879	penicillin-insensitive murein endopeptidase MepA	COG3770	M		424	459	-35
RCA23_c00150	17181	15796	putative MFS-type transporter	COG2270	R		588	535	53
RCA23_c00160	17672	17241	hypothetical protein	COG0824	R		354	357	-3
RCA23_c00170	17836	18201	hypothetical protein, YGGT family	COG0762	S		446	469	-23
RCA23_c00180	18217	20277	ATP-dependent DNA helicase RecQ	COG0514	L		362	427	-65
RCA23_c00190	20300	21073	hypothetical protein DUF328	COG3022	S		375	423	-48
RCA23_c00200	21576	21070	isopentenyl-diphosphate delta-isomerase Idi	COG1443	I		495	470	25
RCA23_c00210	21812	23056	5-aminolevulinate synthase HemA	COG0156	H		577	477	100
RCA23_c00220	23097	23540	cytochrome c2	COG3474	C		489	500	-11
RCA23_c00230	24777	23602	acetyl-CoA acetyltransferase PhaA	COG0183	I		276	293	-17
RCA23_c00240	24942	27992	DNA polymerase III 2 alpha subunit DnaE	COG0587	L		333	404	-71

RCA23_c00250	28122	28625	transcription antitermination protein NusG	COG0250	K	338	431	-93
RCA23_c00260	29748	28636	putative lytic murein transglycosylase	COG2951	M	373	308	65
RCA23_c00270	31991	29745	putative ribonuclease R	COG0557	K	415	447	-32
RCA23_c00280	32084	32503	hypothetical protein DUF461	COG2847	S	758	537	221
RCA23_c00290	33615	32512	chorismate synthase AroC	COG0082	E	389	392	-3
RCA23_c00300	34180	34422	hypothetical protein			461	357	104
RCA23_c00310	34649	35638	thiamine-binding periplasmic protein ThiB	COG4143	H	445	484	-39
RCA23_c00320	35638	37182	thiamine transport system permease protein ThiP	COG1178	P	287	490	-203
RCA23_c00330	37151	37867	thiamine import ATP-binding protein ThiQ	COG3840	H	337	496	-159
RCA23_c00340	38720	37920	cytochrome c1	COG2857	C	615	381	234
RCA23_c00350	40072	38735	cytochrome b	COG1290	C	894	546	348
RCA23_c00360	40644	40072	ubiquinol-cytochrome c reductase iron-sulfur subunit PetA	COG0723	C	566	383	183
RCA23_c00370	41451	40837	hypothetical protein, glutathione S-transferase	COG0625	O	244	366	-122
RCA23_c00380	42523	41495	inositol 2-dehydrogenase IdhA	COG0673	R	282	357	-75
RCA23_c00390	42704	44050	protein PmbA	COG0312	R	317	402	-85
RCA23_c00400	44043	44810	putative inositol monophosphatase	COG0483	G	442	524	-82
RCA23_c00410	44819	45955	putative 3-deoxy-D-manno-octulosonic-acid transferase	COG1519	M	381	447	-66
RCA23_c00420	45952	46938	tetraacyldisaccharide 4'-kinase LpxK	COG1663	M	299	391	-92
RCA23_c00430	47600	46941	hypothetical protein, thioredoxin	COG1651	O	615	604	11
RCA23_c00440	48123	47602	hypothetical protein DUF721	COG5389	S	395	653	-258
RCA23_c00450	48220	49254	A/G-specific adenine glycosylase YfhQ	COG1194	L	216	265	-49
RCA23_c00460	49425	50567	alkane 1-monooxygenase AlkB			705	622	83
RCA23_c00470	51766	50654	modification methylase CcrM	COG0863	L	547	559	-12
RCA23_c00480	52492	51857	ribonuclease HII	COG0164	L	826	659	167
RCA23_c00490	52617	53405	putative exodeoxyribonuclease III	COG0708	L	412	409	3
RCA23_c00500	53466	54383	hypothetical protein, thioredoxin	COG3118	O	360	421	-61
RCA23_c00510	54410	55051	putative ATP-dependent protease La (LON)	COG2802	R	461	381	80
RCA23_c00520	55048	55233	hypothetical protein, Trm112p-like	COG2835	S	557	439	118
RCA23_c00530	56447	55230	putative 2-octaprenyl-6-methoxyphenol hydroxylase	COG0654	H	336	427	-91
RCA23_c00540	56547	57857	glutamyl-tRNA(Gln) amidotransferase subunit A	COG0154	J	218	476	-258

RCA23_c00550	57922	59085	LL-diaminopimelate aminotransferase DapL	COG0436	E		518	488	30
RCA23_c00560	59099	61888	putative DNA transclocase FtsK	COG1674	D		506	510	-4
RCA23_c00570	61992	62549	putative outer membrane lipoprotein carrier protein LolA	COG2834	M		382	413	-31
RCA23_c00580	63223	62720	hypothetical protein	COG4764	S		478	353	125
RCA23_c00590	63513	63836	hemimethylated DNA-binding protein, YccV like	COG3785	S		588	541	47
RCA23_c00600	65410	63833	putative gamma-glutamyltransferase ywrD	COG0405	E		253	306	-53
RCA23_c00610	65571	66749	acetyl-CoA acetyltransferase ThlA	COG0183	I		289	358	-69
RCA23_c00620	66761	67216	hypothetical protein	COG1956	T		801	697	104
RCA23_c00630	67302	67871	putative HTH-type transcriptional regulator	COG1396	K		624	670	-46
RCA23_c00640	67871	68488	hypothetical protein, homoserine/homoserine lactone efflux	COG1280	E		417	485	-68
RCA23_c00650	68502	69128	LysE-type translocator	COG1280	E		274	391	-117
RCA23_c00660	69140	71587	dimethylglycine dehydrogenase	COG0404	E		513	451	62
RCA23_c00670	71603	72469	putative homocysteine S-methyltransferase	COG2040	E		334	377	-43
RCA23_c00680	72474	72929	pyridoxamine 5'-phosphate oxidase-like protein, FMN-bindir	COG3576	R		400	404	-4
RCA23_c00690	72966	75419	sarcosine dehydrogenase	COG0404	E		503	505	-2
RCA23_c00700	75596	76150	hypothetical protein	COG0790	R		229	273	-44
RCA23_c00710	77326	76202	putative agmatine deiminase AguA	COG2957	E		244	302	-58
RCA23_c00720	78340	77450	Cl- channel, voltage-gated family protein	COG0038	P		169	179	-10
RCA23_c00730	79210	78368	Cl- channel, voltage-gated family protein	COG0038	P		284	229	55
RCA23_c00750	79638	80468	diaminopimelate epimerase DapF	COG0253	E	GI 1	356	453	-97
RCA23_c00760	80465	81721	(dimethylallyl)adenosine tRNA methylthiotransferase MiaB	COG0621	J	GI 1	300	392	-92
RCA23_c00770	81775	82368	hypothetical protein, glutathione S-transferotein	COG0625	O	GI 1	340	473	-133
RCA23_c00780	83403	82369	HTH-type transcriptional regulator, LacI family	COG1609	K	GI 1	392	377	15
RCA23_c00790	84698	84859	hypothetical protein			GI 1	512	328	184
RCA23_c00800	85627	85304	hypothetical protein	COG4274	S	GI 1	503	298	205
RCA23_c00810	86298	86852	hypothetical protein, phosphoglycerate mutase-like	COG0406	G	GI 1	577	360	217
RCA23_c00820	86971	87357	hypothetical protein	COG0790	R	GI 1	602	405	197
RCA23_c00830	87971	87558	hypothetical protein, glyoxalase/bleomycin resistance protei	COG0346	E	GI 1	350	319	31
RCA23_c00840	88760	88251	putative phage integrase	COG4974	L	GI 1	284	247	37
RCA23_c00850	89246	89632	hypothetical protein			GI 1	12	10	2

RCA23_c00860	89849	90679	hypothetical protein			GI 1	44	43	1
RCA23_c00870	92517	91066	putative endonuclease			GI 1	1	0	1
RCA23_c00880	93431	93994	DNA integration/recombination/invertion protein	COG1961	L	GI 1	253	243	10
RCA23_c00890	94002	94223	hypothetical protein			GI 1	529	399	130
RCA23_c00900	94400	95290	hypothetical protein, peptidoglycan binding-like			GI 1	0	0	0
RCA23_c00910	95706	96575	hypothetical protein			GI 1	28	29	-1
RCA23_c00920	98235	96589	putative phage integrase	COG4974	L	GI 1	54	35	19
RCA23_c00940	98894	98478	putative HTH-type transcriptional repressor, ArsR family	COG0640	K		463	367	96
RCA23_c00950	99032	99355	ATP synthase protein I				599	344	255
RCA23_c00960	99360	100142	ATP synthase subunit AtpB	COG0356	C		361	248	113
RCA23_c00970	100189	100425	ATP synthase subunit c				766	349	417
RCA23_c01000	102506	101736	pyruvate dehydrogenase complex repressor	COG2186	K		376	382	-6
RCA23_c01010	103165	102623	inner membrane lipoprotein YiaD	COG2885	M		265	319	-54
RCA23_c01020	103461	104105	endonuclease III	COG0177	L		351	403	-52
RCA23_c01030	104102	105094	putative pfkB family carbohydrate kinase	COG0524	G		636	568	68
RCA23_c01040	105925	105104	hypothetical protein				439	403	36
RCA23_c01050	106040	106831	glycosyl transferase family 14				610	421	189
RCA23_c01060	106838	107617	hypothetical protein				588	434	154
RCA23_c01070	107628	108008	HIT-like protein	COG0537	F		461	501	-40
RCA23_c01080	108944	108024	ABC transporter ATP binding protein	COG1131	V		655	495	160
RCA23_c01090	109004	109222	hypothetical protein	COG4391	S		754	472	282
RCA23_c01100	109664	109233	hypothetical protein				564	410	154
RCA23_c01110	109753	112548	DNA polymerase I	COG0749	L		456	441	15
RCA23_c01120	112575	113720	cystathionine gamma-synthase MetB	COG0626	E		208	301	-93
RCA23_c01130	113714	114265	ribosomal large subunit pseudouridine synthase RluE	COG1187	J		276	499	-223
RCA23_c01140	114879	114286	hypothetical protein DUF1285	COG3816	S		371	533	-162
RCA23_c01150	114939	115946	ATPase, MoxR type	COG0714	R		192	243	-51
RCA23_c01160	115943	116806	hypothetical protein	COG1721	R		276	277	-1
RCA23_c01170	116803	119550	hypothetical protein				224	357	-133
RCA23_c01180	119595	121607	hypothetical protein DUF1355	COG5426	S		287	442	-155

RCA23_c01190	122292	123830	glycolate oxidase subunit GlcD	COG0277	C	492	613	-121
RCA23_c01200	123827	124996	glycolate oxidase subunit GlcE	COG0277	C	449	505	-56
RCA23_c01210	124996	126312	glycolate oxidase iron-sulfur subunit GlcF	COG0247	C	438	414	24
RCA23_c01220	126476	127198	hypothetical protein, trypsin	COG3591	E	379	640	-261
RCA23_c01230	128473	127643	putative chitinase	COG3325	G	43	18	25
RCA23_c01240	129159	129614	small heat shock protein IbpA	COG0071	O	1.541	759	782
RCA23_c01250	131155	129680	succinate-semialdehyde dehydrogenase GabD	COG1012	C	335	398	-63
RCA23_c01260	131221	132024	hypothetical protein	COG0657	I	241	333	-92
RCA23_c01270	132344	132021	hypothetical protein	COG5617	S	233	327	-94
RCA23_c01280	133081	132341	putative cyclopentanol dehydrogenase CpnA	COG1028	I	381	387	-6
RCA23_c01290	133264	133566	hypothetical protein			287	356	-69
RCA23_c01300	133661	134278	hypothetical protein DUF1523			469	370	99
RCA23_c01310	136598	134286	aldehyde dehydrogenase	COG1012	C	256	353	-97
RCA23_c01320	137589	136603	deoxyribose-phosphate aldolase DeoC	COG0274	F	287	273	14
RCA23_c01330	137692	138366	ribulose-phosphate 3-epimerase, chromosomal	COG0036	G	290	386	-96
RCA23_c01340	140190	138385	hemolysin-type calcium-binding region			322	456	-134
RCA23_c01350	140316	143915	5-oxoprolinase (ATP-hydrolyzing)	COG0145	E	257	381	-124
RCA23_c01360	145374	144022	aminotransferase class-III	COG0161	H	314	322	-8
RCA23_c01370	145865	145371	ureidoglycolate hydrolase AllA	COG3194	F	551	501	50
RCA23_c01380	146898	145891	aldose 1-epimerase GalM	COG2017	G	273	397	-124
RCA23_c01390	148778	146898	beta-galactosidase BgaB	COG1874	G	225	271	-46
RCA23_c01400	149622	148786	putative gluconolactonase	COG3386	G	260	321	-61
RCA23_c01410	150218	149616	2-dehydro-3-deoxy-6-phosphogalactonate aldolase DgoA	COG0800	G	205	303	-98
RCA23_c01420	151120	150215	2-dehydro-3-deoxygalactonokinase DgoK	COG3734	G	202	287	-85
RCA23_c01430	151884	151117	short chain dehydrogenase	COG1028	I	245	234	11
RCA23_c01440	154013	151929	alpha-galactosidase RafA	COG3345	G	239	275	-36
RCA23_c01450	154942	154043	putative ABC transporter inner membrane component	COG0395	G	388	341	47
RCA23_c01460	155930	154920	putative ABC transporter inner membrane component	COG1175	G	389	289	100
RCA23_c01470	157438	156014	putative extracellular solute-binding protein	COG1653	G	365	261	104
RCA23_c01480	157460	158245	HTH-type transcriptional regulator, IclR family	COG1414	K	141	131	10

RCA23_c01490	158266	159324	sugar ABC transporter ATP-binding protein	COG3839	G	231	267	-36
RCA23_c01500	160108	159344	Sulfite exporter TauE/SafE			177	299	-122
RCA23_c01510	161562	160108	putative transcriptional regulator, gntR family	COG1167	K	244	244	0
RCA23_c01520	163233	161716	taurine--pyruvate aminotransferase Tpa	COG0161	H	341	283	58
RCA23_c01530	163567	164568	taurine ABC transporter, periplasmic binding protein TauA	COG4521	P	501	386	115
RCA23_c01540	164690	165496	taurine ABC transport system ATP-binding protein TauB	COG1116	P	614	476	138
RCA23_c01550	165493	166803	taurine ABC transport system permease protein TauC	COG0600	P	618	484	134
RCA23_c01560	169252	166847	dimethylglycine dehydrogenase	COG0404	E	324	431	-107
RCA23_c01570	169351	170925	trimethylamine methyltransferase MttB	COG5598	H	239	366	-127
RCA23_c01580	173231	170928	3-hydroxyacyl-CoA dehydrogenase FadN	COG1250	I	294	374	-80
RCA23_c01590	174034	173246	hypothetical protein	COG0596	R	346	339	7
RCA23_c01600	175440	174031	O-acetylhomoserine (thiol)-lyase CysD	COG2873	E	375	470	-95
RCA23_c01610	175521	177014	putative signaling protein	COG2200	T	515	534	-19
RCA23_c01620	177159	178928	sulfoacetaldehyde acetyltransferase Xsc	COG0028	E	486	430	56
RCA23_c01630	178940	179944	phosphate acetyltransferase Pta	COG0280	C	250	314	-64
RCA23_c01640	179971	182781	DMSO reductase chain A	COG0243	C	273	340	-67
RCA23_c01650	182778	183659	alpha/beta hydrolase	COG0596	R	217	299	-82
RCA23_c01660	183656	184402	DMSO reductase chain B	COG0437	C	282	401	-119
RCA23_c01670	184402	185274	DMSO reductase chain C	COG3302	R	342	514	-172
RCA23_c01680	186741	185380	sodium:alanine symporter	COG1115	E	347	371	-24
RCA23_c01690	187519	186917	thymidine kinase Tdk	COG1435	F	259	365	-106
RCA23_c01700	187650	188366	NADPH-dependent FMN reductase	COG0431	R	400	533	-133
RCA23_c01710	191838	188422	carbamoyl-phosphate synthase large chain CarB	COG0458	E	371	400	-29
RCA23_c01720	192331	191858	HTH-type transcriptional regulator, AsnC family	COG1522	K	562	484	78
RCA23_c01730	192395	192673	hypothetical protein			876	934	-58
RCA23_c01740	192769	194544	aspartyl-tRNA synthase AspS	COG0173	J	459	475	-16
RCA23_c01750	194967	194602	hypothetical protein			335	405	-70
RCA23_c01760	195297	195013	putative signal transduction response regulator receiver pro	COG2197	T	358	337	21
RCA23_c01770	195361	195822	methylmalonyl-CoA epimerase	COG0346	E	605	483	122
RCA23_c01780	195850	196143	hypothetical protein DUF1467	COG5454	S	593	437	156

RCA23_c01790	196798	196100	hypothetical protein DUF540	COG2981	E	452	400	52
RCA23_c01800	197310	196795	putative nitroreductase	COG0778	C	264	302	-38
RCA23_c01810	198509	197412	histone deacetylase-like amidohydrolase HdaH	COG0123	B	254	420	-166
RCA23_c01820	198760	198527	hypothetical protein			331	417	-86
RCA23_c01830	198898	199593	hypothetical protein, peptidase family M48	COG0501	O	218	295	-77
RCA23_c01840	199628	199873	hypothetical protein			216	359	-143
RCA23_c01850	201592	200219	ribosomal protein S12 methylthiotransferase RimO	COG0621	J	203	311	-108
RCA23_c01860	203057	201645	hypothetical protein, transmembrane			183	253	-70
RCA23_c01870	203697	203125	GcrA cell cycle regulator	COG5352	S	261	323	-62
RCA23_c01880	203799	204602	ABC-2 type transport system membrane protein	COG0842	V	589	598	-9
RCA23_c01890	204812	205990	acetylornithine aminotransferase ArgD	COG4992	E	327	501	-174
RCA23_c01900	206062	206985	ornithine carbamoyltransferase ArgF	COG0078	E	299	408	-109
RCA23_c01910	207091	207885	hypothetical protein			277	502	-225
RCA23_c01920	207908	209323	6-phosphogluconate dehydrogenase GntZ	COG0362	G	254	365	-111
RCA23_c01930	211949	209328	ATP-dependent RNA helicase HrpB	COG1643	L	228	291	-63
RCA23_c01940	211993	212991	LAO/AO transport system ATPase	COG1703	E	204	314	-110
RCA23_c01950	213208	213495	50S ribosomal protein L28	COG0227	J	729	469	260
RCA23_c01960	213684	214121	hypothetical protein			684	509	175
RCA23_c01970	214253	216049	GTP-binding protein LepA	COG0481	M	566	504	62
RCA23_c01980	216818	216063	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	269	374	-105
RCA23_c01990	216901	217473	putative ring-cleaving dioxygenase	COG0346	E	435	594	-159
RCA23_c02000	218824	217544	serine hydroxymethyltransferase GlyA	COG0112	E	425	364	61
RCA23_c02010	218982	219755	putative inorganic polyphosphate/ATP-NAD kinase PpnK	COG0061	G	432	438	-6
RCA23_c02020	222034	220145	propionate--CoA ligase PrpE	COG0365	I	297	361	-64
RCA23_c02030	222198	223292	hypothetical protein	COG1917	S	290	425	-135
RCA23_c02040	223384	223809	cytidine deaminase Cdd	COG0295	F	257	367	-110
RCA23_c02050	223809	225131	thymidine phosphorylase DeoA	COG0213	F	209	351	-142
RCA23_c02060	225142	226296	phosphopentomutase DeoB	COG1015	G	241	350	-109
RCA23_c02070	226293	227261	adenosine deaminase Add	COG1816	F	392	461	-69
RCA23_c02080	227588	228214	uracil phosphoribosyltransferase Upp	COG0035	F	717	465	252

RCA23_c02090	228449	228868	hypothetical protein			647	366	281
RCA23_c02100	230234	228900	3-deoxy-D-manno-octulosonic-acid transferase WaaA	COG1519	M	739	512	227
RCA23_c02110	231721	230234	L-sorbose 1-dehydrogenase	COG2303	E	814	536	278
RCA23_c02120	232388	231708	putative N-acylneuraminate cytidylyltransferase	COG1083	M	721	442	279
RCA23_c02130	233528	232392	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotrar	COG0399	M	732	488	244
RCA23_c02140	236514	234343	hypothetical protein	COG3119	P	188	160	28
RCA23_c02150	237421	236696	hypothetical protein, nucleoside triphosphate hydrolases-lik	COG1122	P	241	189	52
RCA23_c02160	238069	237422	S-adenosyl-L-methionine-dependent methyltransferase	COG2226	H	312	194	118
RCA23_c02170	239306	238389	hypothetical protein			275	268	7
RCA23_c02180	241023	239398	medium-chain-fatty-acid--CoA ligase AlkK	COG0318	I	311	345	-34
RCA23_c02190	241093	241995	hypothetical protein DUF6 transmembrane	COG0697	G	423	493	-70
RCA23_c02200	244207	242024	fatty acid oxidation complex alpha subunit FadJ	COG1250	I	352	406	-54
RCA23_c02210	244632	244219	hypothetical protein			532	651	-119
RCA23_c02220	245865	244654	short chain dehydrogenase	COG0183	I	230	341	-111
RCA23_c02230	246509	245877	putative glutathione S-transferase	COG0625	O	519	394	125
RCA23_c02240	248334	246553	acyl-CoA dehydrogenase MmgC	COG1960	I	389	367	22
RCA23_c02250	248746	248357	putative HTH-type transcriptional regulator, MerR family	COG0789	K	557	407	150
RCA23_c02260	249251	248883	putative HTH-type transcriptional regulator, MerR family	COG0789	K	553	505	48
RCA23_c02270	250441	249299	hypothetical protein, transmembrane protein DUF2899			442	566	-124
RCA23_c02280	250528	250956	hypothetical protein, thioesterase	COG2050	Q	148	209	-61
RCA23_c02290	250953	251453	hypothetical protein, thioesterase	COG2050	Q	252	263	-11
RCA23_c02300	251446	252768	DNA-damage-inducible protein F	COG0534	V	309	520	-211
RCA23_c02310	253850	252789	dihydroorotate dehydrogenase PyrD	COG0167	F	240	322	-82
RCA23_c02320	254185	253847	hypothetical protein DUF952	COG3502	S	527	496	31
RCA23_c02330	254360	255892	5'-nucleotidase SurE	COG0737	F	403	477	-74
RCA23_c02340	257694	255964	hypothetical protein	COG3409	M	265	336	-71
RCA23_c02350	257885	258799	glycyl-tRNA synthase alpha subunit GlyQ	COG0752	J	609	605	4
RCA23_c02360	258799	259317	hypothetical protein			393	433	-40
RCA23_c02370	259317	261374	glycyl-tRNA synthase beta subunit GlyS	COG0751	J	339	366	-27
RCA23_c02380	261437	263977	pyruvate, phosphate dikinase PpdK	COG0574	G	240	338	-98

RCA23_c02390	264078	264761	putative cell wall hydrolase	COG3773	M	429	521	-92
RCA23_c02400	264811	265737	hypothetical protein, dihydroneopterin aldolase	COG1539	H	326	427	-101
RCA23_c02410	265734	266720	dihydropteroate synthase FolP	COG0294	H	218	368	-150
RCA23_c02420	267675	266692	putative integral membrane protein DUF6	COG0697	G	277	365	-88
RCA23_c02430	268729	267707	ketol-acid reductoisomerase IlvC	COG0059	E	455	347	108
RCA23_c02440	268864	269319	putative transcriptional regulator, asnC family	COG1522	K	260	420	-160
RCA23_c02450	269316	269771	putative transcriptional regulator, asnC family	COG1522	K	255	394	-139
RCA23_c02460	269846	270979	hypothetical protein	COG0075	E	378	383	-5
RCA23_c02470	272183	270987	2-octaprenyl-6-methoxyphenol hydroxylase UbiH	COG0654	H	210	332	-122
RCA23_c02480	272931	272161	putative pyrimidine 5-nucleotidase	COG1011	R	348	434	-86
RCA23_c02490	273012	273620	putative HTH-type transcriptional regulator, GntR family	COG1802	K	242	345	-103
RCA23_c02500	275505	273580	uncharacterized glycosyltransferase YdaM	COG1215	M	317	431	-114
RCA23_c02510	276701	275577	carbamoyl-phosphate synthase small chain CarA	COG0505	E	362	363	-1
RCA23_c02520	276880	277338	GatB/YqeY family protein	COG1610	S	554	518	36
RCA23_c02530	277906	277379	hypothetical protein	COG5488	S	343	384	-41
RCA23_c02540	279835	278162	cytochrome c oxidase subunit 1	COG0843	C	599	404	195
RCA23_c02550	280700	280038	octanoyltransferase LipB	COG0321	H	284	374	-90
RCA23_c02560	281476	280700	hypothetical protein, LytTr transcriptional regulator	COG3279	K	157	207	-50
RCA23_c02570	281566	282084	membrane protein-like	COG5395	S	319	344	-25
RCA23_c02590	284544	282886	arylsulfatase	COG3119	P	331	339	-8
RCA23_c02600	284793	285923	NADPH dehydrogenas	COG1902	C	211	275	-64
RCA23_c02610	287001	285961	L-idonate 5-dehydrogenase IdnD	COG1063	E	214	271	-57
RCA23_c02620	287045	287839	gluconate 5-dehydrogenase Gno	COG1028	I	244	357	-113
RCA23_c02630	288037	288939	uncharacterized oxidoreductase YgbJ	COG2084	I	282	358	-76
RCA23_c02640	288936	289799	hypothetical protein, xylose isomerase-like	COG1082	G	259	421	-162
RCA23_c02650	290765	289839	glyoxylate reductase GyaR	COG1052	C	275	373	-98
RCA23_c02660	290938	290762	hypothetical protein			401	386	15
RCA23_c02670	290989	291273	hypothetical protein, AzlC-like	COG1296	E	119	114	5
RCA23_c02680	293832	291886	cbbT/tktB: transketolase	COG0021	G	35	55	-20
RCA23_c02690	295174	293834	hypothetical protein	COG4091	E	45	80	-35

RCA23_c02700	295374	296207	transketolase, alpha subunit	COG3959	G	64	84	-20
RCA23_c02710	296204	297247	transketolase, beta-subunit	COG3958	G	33	56	-23
RCA23_c02720	297249	297809	putative 3-hydroxyisobutyrate dehydrogenase	COG2084	I	23	47	-24
RCA23_c02730	297809	298171	putative 3-hydroxyisobutyrate dehydrogenase	COG2084	I	41	116	-75
RCA23_c02740	298168	298920	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	63	100	-37
RCA23_c02750	299907	298930	HTH-type transcriptional regulator, LysR family	COG0583	K	53	71	-18
RCA23_c02760	300726	300040	putative HTH-type transcriptional regulator, GntR family	COG1802	K	134	109	25
RCA23_c02770	300938	301951	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	153	99	54
RCA23_c02780	302000	302872	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	180	134	46
RCA23_c02800	304513	305322	hypothetical protein	COG0240	C	83	98	-15
RCA23_c02810	305336	306664	hypothetical protein DUF1537	COG3395	S	99	119	-20
RCA23_c02820	306661	307920	putative ribulose bisphosphate carboxylase large chain	COG1850	G	55	88	-33
RCA23_c02830	308882	307917	hypothetical protein, NAD dependent epimerase / dehydrata	COG0451	M	98	93	5
RCA23_c02840	309016	309939	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I	65	89	-24
RCA23_c02850	310592	310035	hypothetical protein			69	68	1
RCA23_c02860	310977	312479	altronate hydrolase UxaA	COG2721	G	287	313	-26
RCA23_c02870	312455	313480	putative oxidoreductase	COG0673	R	374	392	-18
RCA23_c02880	313516	314403	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I	370	358	12
RCA23_c02890	314406	315287	S-adenosylmethionine uptake transporter Sam	COG0697	G	714	584	130
RCA23_c02900	315284	316147	2-dehydro-3-deoxygluconokinase KdgK	COG0524	G	444	419	25
RCA23_c02910	317333	316134	D-mannonate oxidoreductase UxuB	COG0246	G	274	353	-79
RCA23_c02920	318704	317334	uronate isomerase UxaC	COG1904	G	271	302	-31
RCA23_c02930	318798	319817	putative oxidoreductase	COG0673	R	175	238	-63
RCA23_c02940	319829	320809	putative oxidoreductase	COG0673	R	304	345	-41
RCA23_c02950	320806	321861	putative oxidoreductase	COG0673	R	242	272	-30
RCA23_c02960	321849	323057	mannonate dehydratase UxuA	COG1312	G	216	284	-68
RCA23_c02970	323103	324641	long-chain-fatty-acid--CoA ligase	COG0318	I	230	340	-110
RCA23_c02980	324634	325440	D-beta-hydroxybutyrate dehydrogenase BdhA	COG1028	I	247	448	-201
RCA23_c02990	325531	327783	copper-transporting P-type ATPase ActP	COG2217	P	292	457	-165
RCA23_c03000	327811	328200	HTH-type transcriptional regulator (copper efflux regulator)	COG0789	K	384	514	-130

RCA23_c03010	328492	328253	hypothetical protein	COG3937	S	328	304	24
RCA23_c03020	328847	329161	hypothetical protein			479	454	25
RCA23_c03030	331571	329382	isocitrate dehydrogenase Icd	COG2838	C	360	383	-23
RCA23_c03040	332596	331688	integral membrane protein DUF6	COG0697	G	370	367	3
RCA23_c03050	332863	333540	flavocytochrome c cytochrome subunit SoxE	COG3474	C	492	664	-172
RCA23_c03060	333621	334832	sulfide dehydrogenase [flavocytochrome c] flavoprotein cha	COG0446	R	319	404	-85
RCA23_c03070	335365	334916	hypothetical protein, OsmC-like	COG1764	O	444	380	64
RCA23_c03080	336582	335440	putative dimethyl sulfoniopropionate demethylase DmdA	COG0404	E	508	485	23
RCA23_c03090	337231	336746	transcriptional regulator, AsnC family	COG1522	K	732	486	246
RCA23_c03100	338152	338784	glutathione S-transferase	COG0625	O	93	79	14
RCA23_c03120	340630	339608	hypothetical protein, ribonuclease, E/G family	COG1530	J	306	281	25
RCA23_c03130	341160	340627	Maf-like protein	COG0424	D	156	192	-36
RCA23_c03140	341419	341201	translation initiation factor IF-1	COG0361	J	574	601	-27
RCA23_c03150	342020	341622	hypothetical protein, low molecular weight phosphotyrosine	COG0394	T	491	374	117
RCA23_c03160	342565	342086	hypothetical protein UPF0262	COG5328	S	455	421	34
RCA23_c03170	343866	342565	histidinol dehydrogenase HisD	COG0141	E	235	274	-39
RCA23_c03180	344390	343932	hypothetical protein			233	414	-181
RCA23_c03190	345655	344387	UDP-N-acetylglucosamine 1-carboxyvinyltransferase MurA	COG0766	M	277	290	-13
RCA23_c03210	346329	347666	PAS/PAC sensor hybrid histidine kinase	COG0642	T	293	203	90
RCA23_c03220	347759	348817	2OG-Fe(II) oxygenase	COG3491	R	359	267	92
RCA23_c03230	348821	349465	putative lysine exporter protein	COG1280	E	425	357	68
RCA23_c03240	349742	350194	putative carbon monoxide dehydrogenase subunit G	COG3427	S	365	468	-103
RCA23_c03250	350184	351113	hypothetical protein, XdhC and CoxI family	COG1975	O	275	339	-64
RCA23_c03260	352364	351108	putative MFS-type transporter	COG2814	G	383	409	-26
RCA23_c03270	352457	353512	AFG1-like ATPase	COG1485	R	237	332	-95
RCA23_c03280	354792	353509	bifunctional protein FolC	COG0285	H	200	338	-138
RCA23_c03290	355709	354789	acetyl-coenzyme A carboxylase carboxyl transferase beta s	COG0777	I	408	418	-10
RCA23_c03300	356729	355773	hypothetical protein, CAAX amino terminal protease-like	COG1266	R	413	454	-41
RCA23_c03310	356791	358524	dihydroxy-acid dehydratase IlvD	COG0129	E	332	368	-36
RCA23_c03320	358660	359103	hypothetical protein			227	276	-49

RCA23_c03330	359279	360058	hypothetical protein			249	344	-95
RCA23_c03340	360076	360423	hypothetical protein			380	445	-65
RCA23_c03350	361268	360426	hypothetical protein, OmpA family	COG2885	M	332	403	-71
RCA23_c03360	361905	361339	hypothetical protein, peroxidase-like protein	COG2128	S	428	461	-33
RCA23_c03370	362627	361902	hypothetical protein, acetyltransferase-like	COG0456	R	256	347	-91
RCA23_c03380	363346	362624	hypothetical protein, probably molybdopterin binding	COG1058	R	218	322	-104
RCA23_c03390	363392	364114	sugar fermentation stimulation protein SfsA	COG1489	R	302	425	-123
RCA23_c03400	364191	365000	methionine aminopeptidase Map	COG0024	J	369	393	-24
RCA23_c03410	366577	365054	ATP-dependent RNA helicase RhlE	COG0513	L	275	300	-25
RCA23_c03420	367398	366829	putative ribosomal RNA small subunit methyltransferase D	COG0742	L	236	306	-70
RCA23_c03430	368600	367395	rhodocoxin reductase ThcD	COG0446	R	286	329	-43
RCA23_c03440	368681	369172	peroxiredoxin	COG0678	O	574	527	47
RCA23_c03450	369613	369326	pterin-4-alpha-carbinolamine dehydratase	COG2154	H	559	495	64
RCA23_c03460	370797	369610	hypothetical protein DUF482	COG3146	S	291	351	-60
RCA23_c03470	371604	370849	putative glycerophosphoryl diester phosphodiesterase	COG0584	C	398	441	-43
RCA23_c03480	372065	371601	putative endoribonuclease L-PSP	COG0251	J	345	507	-162
RCA23_c03490	373364	372132	type I secretion system protein, HlyD family	COG1566	V	257	378	-121
RCA23_c03500	374871	373411	type I secretion system ATP-binding component	COG4618	R	195	310	-115
RCA23_c03510	375137	374874	putative type I secretion system protein, transmembrane do	COG4618	R	484	480	4
RCA23_c03520	375269	376000	VacJ like lipoprotein	COG2853	M	564	440	124
RCA23_c03530	375984	376589	putative toluene tolerance protein	COG2854	Q	423	353	70
RCA23_c03540	378789	376654	penicillin-binding protein 1B	COG0744	M	283	374	-91
RCA23_c03550	380155	378971	aromatic-amino-acid aminotransferase TyrB	COG1448	E	343	390	-47
RCA23_c03560	381012	380155	3-mercaptopyruvate sulfurtransferase SseA	COG2897	P	307	316	-9
RCA23_c03570	381576	381100	SsrA-binding protein SmpB	COG0691	O	752	622	130
RCA23_c03580	382567	381680	dihydrodipicolinate synthase DapA	COG0329	E	245	271	-26
RCA23_c03590	382609	384669	soluble lytic murein transglycosylase Slt	COG0741	M	231	304	-73
RCA23_c03600	385589	384717	hypothetical protein DUF6 transmembrane	COG0697	G	244	358	-114
RCA23_c03610	386248	385586	hypothetical protein DUF752	COG4121	S	268	375	-107
RCA23_c03620	386288	387352	putative FAD-dependent oxidoreductase	COG0665	E	279	439	-160

RCA23_c03630	387390	387752	hypothetical protein, glyoxalase/dioxygenase superfamily	COG0346	E	217	367	-150
RCA23_c03640	388138	387773	hypothetical protein			350	445	-95
RCA23_c03650	388488	388135	hypothetical protein			529	416	113
RCA23_c03660	389111	388503	cob(I)yrinic acid a,c-diamide adenosyltransferase CobO	COG2109	H	496	404	92
RCA23_c03670	389202	389606	hypothetical protein			210	318	-108
RCA23_c03680	390600	389629	oligopeptide/dipeptide ABC transporter, ATP-binding protein	COG0444	E	260	314	-54
RCA23_c03690	391589	390600	oligopeptide/dipeptide ABC transporter, ATP-binding protein	COG0444	E	238	350	-112
RCA23_c03700	392467	391586	oligopeptide/dipeptide ABC transporter, permease protein	COG1173	E	262	357	-95
RCA23_c03710	393441	392464	oligopeptide/dipeptide ABC transporter, permease protein	COG0601	E	439	395	44
RCA23_c03720	395073	393502	oligopeptide/dipeptide ABC transporter, periplasmic substra	COG0747	E	379	323	56
RCA23_c03730	396231	395302	HTH-type transcriptional regulator, LysR family	COG0583	K	497	494	3
RCA23_c03740	396411	397982	putative amidohydrolase 3	COG1574	R	232	346	-114
RCA23_c03750	397979	398881	cobalamin biosynthesis CobW-like	COG0523	R	249	331	-82
RCA23_c03760	398992	400035	hypothetical protein, restriction endonuclease type IV-like	COG4127	S	241	213	28
RCA23_c03770	401165	400095	glycine betaine transport ATP-binding protein OpuAA	COG4175	E	439	469	-30
RCA23_c03780	402193	401162	glycine betaine transport system permease protein OpuAB	COG4176	E	459	546	-87
RCA23_c03790	403226	402258	glycine betaine transporter substrate-binding protein OpuAC	COG2113	E	364	379	-15
RCA23_c03800	403626	404708	fatty acid desaturase	COG3239	I	370	432	-62
RCA23_c03810	405832	404717	CoA-transferase family III protein involved in DMSP degrad	COG1804	C	286	378	-92
RCA23_c03820	405843	406949	glyceraldehyde-3-phosphate dehydrogenase Gap	COG0057	G	320	495	-175
RCA23_c03830	407067	407687	ATP-dependent Clp protease proteolytic subunit ClpP	COG0740	O	732	497	235
RCA23_c03840	407812	409077	ATP-dependent Clp protease ATP-binding subunit ClpX	COG1219	O	680	551	129
RCA23_c03850	409209	409592	putative NADH ubiquinone oxidoreductase subunit NDUFA1	COG3761	C	786	483	303
RCA23_c03860	409597	410043	ABC-type transport system involved in resistance to organic	COG1463	Q	548	559	-11
RCA23_c03870	410043	410402	Uncharacterized protein conserved in bacteria (DUF2155)	COG4765	S	308	397	-89
RCA23_c03880	411007	410366	leucyl/phenylalanyl-tRNA--protein transferase Aat	COG2360	O	466	513	-47
RCA23_c03890	412356	411004	biotin carboxylase AccC	COG4770	I	318	325	-7
RCA23_c03900	414969	413026	acetyl-coenzyme A synthase AcsA	COG0365	I	265	296	-31
RCA23_c03910	415768	415037	high-affinity branched-chain amino acid transport ATP-bindi	COG0410	E	232	342	-110
RCA23_c03920	416529	415771	high-affinity branched-chain amino acid transport ATP-bindi	COG0411	E	513	381	132

RCA23_c03930	417743	416541	putative branched-chain amino acid transport system perme	COG4177	E		708	525	183
RCA23_c03940	418779	417748	putative branched-chain amino acid transport system perme	COG0559	E		636	469	167
RCA23_c03950	420231	418885	hypothetical protein	COG0683	E		472	393	79
RCA23_c03960	420457	421773	hypothetical protein, DNA binding helix-turn helix proteins	COG1396	K		304	375	-71
RCA23_c03970	421836	422207	putative response regulator receiver protein, CheY like	COG0745	T		506	409	97
RCA23_c03980	422204	422473	hypothetical protein				308	532	-224
RCA23_c03990	422474	425128	sensor protein kinase WalK	COG0591	E		309	414	-105
RCA23_c04000	425546	425109	hypothetical protein DUF442	COG3453	S		331	346	-15
RCA23_c04010	425805	425593	hypothetical protein				528	443	85
RCA23_c04020	425917	426168	hypothetical protein	COG3750	S		925	593	332
RCA23_c04030	427065	426184	hypothetical protein				688	461	227
RCA23_c04040	427159	428790	putative Fe(3+)-transport system protein SfuB	COG1178	P		329	410	-81
RCA23_c04060	428970	430124	integrase	COG0582	L	GI 2	37	28	9
RCA23_c04070	430197	430778	hypothetical protein	COG3945	S	GI 2	17	8	9
RCA23_c04080	434845	431570	hypothetical protein			GI 2	41	29	12
RCA23_c04090	434910	437174	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltra	COG0457	R	GI 2	20	13	7
RCA23_c04100	440084	439434	esterase, SGNH hydrolase-type	COG2755	E	GI 2	0	0	0
RCA23_c04110	440656	440258	hypothetical protein			GI 2	235	150	85
RCA23_c04120	441029	441403	hypothetical protein, tetratricopeptide repeat	COG0790	R	GI 2	76	63	13
RCA23_c04130	441971	441678	hypothetical protein			GI 2	270	161	109
RCA23_c04140	442447	442121	hypothetical protein			GI 2	223	168	55
RCA23_c04150	442788	442444	hypothetical protein			GI 2	224	151	73
RCA23_c04160	443599	443222	hypothetical protein, RmlC-like cupin family	COG3450	R	GI 2	4	15	-11
RCA23_c04170	444724	444101	hypothetical protein			GI 2	45	28	17
RCA23_c04180	445304	444918	hypothetical protein, DUF3127			GI 2	56	66	-10
RCA23_c04190	446438	445509	putative nucleoside triphosphate hydrolase, ATPase domai	COG1066	O	GI 2	14	20	-6
RCA23_c04200	447532	449187	zinc-dependent metalloprotease	COG2931	Q	GI 2	119	65	54
RCA23_c04210	449574	449230	HTH-type transcriptional regulator, AsnC family	COG1522	K	GI 2	160	127	33
RCA23_c04220	450860	449781	integrase	COG2801	L	GI 2	0	0	0
RCA23_c04230	453456	452530	putative ion channel			GI 2	549	372	177

RCA23_c04240	454720	453746	integrase	COG4974	L	GI 2	326	288	38
RCA23_c04250	455125	455838	Gcn5-like N-acetyltransferase	COG1670	J	GI 2	8	1	7
RCA23_c04260	457797	456643	integrase	COG0582	L	GI 2	166	139	27
RCA23_c04280	458139	459119	hypothetical protein, NAD dependent epimerase / dehydrata	COG0702	M		246	375	-129
RCA23_c04290	459936	459124	undecaprenyl-diphosphatase UppP	COG1968	V		192	274	-82
RCA23_c04300	460124	461578	glutamate synthase [NADPH] small chain GltD	COG0493	E		382	340	42
RCA23_c04310	461639	466186	glutamate synthase [NADPH] large chain GltB	COG0069	E		343	411	-68
RCA23_c04320	466290	467033	monofunctional biosynthetic peptidoglycan transglycosylase	COG0744	M		271	373	-102
RCA23_c04330	467132	467752	putative glutathione S-transferase	COG0625	O		483	589	-106
RCA23_c04340	467733	468776	putative electron transport protein yjeS	COG1600	C		231	349	-118
RCA23_c04360	469585	468773	hypothetical protein	COG5266	P		307	349	-42
RCA23_c04350	469584	470741	hypothetical protein				351	560	-209
RCA23_c04370	471709	470840	branched-chain-amino-acid aminotransferase IlvE	COG0115	E		279	331	-52
RCA23_c04380	471872	472384	HTH-type transcriptional regulator PetP	COG1846	K		317	270	47
RCA23_c04390	472381	473094	protein PetR	COG0745	T		277	371	-94
RCA23_c04400	473091	474020	putative deacytelase, histone deacetylase superfamily prote	COG0123	B		332	484	-152
RCA23_c04410	474013	474255	exodeoxyribonuclease 7 small subunit XseB	COG1722	L		304	484	-180
RCA23_c04420	474215	475159	geranyltranstransferase IspA	COG0142	H		225	362	-137
RCA23_c04430	475200	477101	1-deoxy-D-xylulose-5-phosphate synthase Dxs	COG1154	H		238	320	-82
RCA23_c04440	478007	477105	hypothetical protein, NAD dependent epimerase/dehydratas	COG0451	M		220	302	-82
RCA23_c04450	478232	478657	hypothetical protein				298	338	-40
RCA23_c04460	478753	479079	hypothetical protein				446	435	11
RCA23_c04470	479921	479142	carnitinyl-CoA dehydratase CaiD	COG1024	I		337	370	-33
RCA23_c04480	481945	479918	hypothetical protein, CoA-binding	COG1042	C		232	293	-61
RCA23_c04490	483157	481997	putative acyl-CoA dehydrogenase YngJ	COG1960	I		288	270	18
RCA23_c04500	483234	484166	transcriptional regulator, AraC family	COG4977	K		204	299	-95
RCA23_c04510	484951	484196	class II aldolase	COG0235	G		278	347	-69
RCA23_c04520	485442	485014	arsenate reductase ArsC	COG1393	P		218	288	-70
RCA23_c04530	485529	487343	adenine deaminase Ade	COG1001	F		223	310	-87
RCA23_c04540	487330	488805	AMP nucleosidase Amn	COG0775	F		382	319	63

RCA23_c04550	488882	489214	DNA-binding protein HU	COG0776	L	749	586	163
RCA23_c04560	489360	490229	inner membrane protein DUF6	COG0697	G	402	597	-195
RCA23_c04570	490303	491016	hypothetical protein, cytochrome C biogenesis protein	COG0785	O	316	423	-107
RCA23_c04580	491039	491449	hypothetical protein, thioredoxin	COG0526	O	398	399	-1
RCA23_c04590	492021	491446	hypothetical protein	COG3222	S	181	253	-72
RCA23_c04600	493168	492038	succinyl-diaminopimelate desuccinylase DapE	COG0624	E	242	273	-31
RCA23_c04610	493554	493168	hypothetical protein			369	407	-38
RCA23_c04620	494478	493591	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltrans	COG2171	E	248	351	-103
RCA23_c04630	495411	494482	L-threonine ammonia-lyase	COG1171	E	204	305	-101
RCA23_c04640	496617	495427	ribosomal RNA large subunit methyltransferase N	COG0820	R	322	367	-45
RCA23_c04650	497212	496688	hypothetical protein			320	378	-58
RCA23_c04660	497318	498349	L-asparaginase II	COG4448	E	194	300	-106
RCA23_c04670	499288	498410	phosphoserine phosphatase SerB	COG0560	E	220	310	-90
RCA23_c04680	499417	499902	hypothetical protein			382	415	-33
RCA23_c04690	500025	501251	phosphoserine aminotransferase SerC	COG1932	H	301	441	-140
RCA23_c04700	501306	502901	D-3-phosphoglycerate dehydrogenase SerA	COG0111	H	369	392	-23
RCA23_c04710	502898	503629	metallophosphoesterase			240	328	-88
RCA23_c04720	504126	503641	RNA pyrophosphohydrolase RppH	COG0494	L	234	236	-2
RCA23_c04730	505467	504130	carboxy-terminal-processing protease CtpA	COG0793	M	363	371	-8
RCA23_c04740	507182	505665	2,3-bisphosphoglycerate-independent phosphoglycerate mu	COG0696	G	294	331	-37
RCA23_c04750	507362	508978	putative protein ImuB			218	304	-86
RCA23_c04760	509201	509545	hypothetical protein			401	555	-154
RCA23_c04770	510291	509542	hypothetical protein, NlpC/P60	COG0791	M	244	388	-144
RCA23_c04780	511667	510288	putative cytosol aminopeptidase PepA	COG0260	E	203	280	-77
RCA23_c04790	512156	511782	hypothetical protein			280	370	-90
RCA23_c04800	512302	512970	carbonic anhydrase CynT	COG0288	P	701	454	247
RCA23_c04810	514074	513052	aspartate-semialdehyde dehydrogenase Asd	COG0136	E	777	529	248
RCA23_c04820	515330	514140	MFS-type transporter	COG2223	P	321	277	44
RCA23_c04830	515444	516199	short chain dehydrogenase	COG4221	R	260	362	-102
RCA23_c04840	517557	516226	FAD dependent oxidoreductase	COG0665	E	173	210	-37

RCA23_c04850	519110	517560	trimethylamine methyltransferase MttB	COG5598	H	269	395	-126
RCA23_c04860	520642	519152	aldehyde dehydrogenase	COG1012	C	229	319	-90
RCA23_c04870	521547	520639	dihydrodipicolinate synthase DapA	COG0329	E	277	336	-59
RCA23_c04880	522241	521537	transcriptional regulator, GntR family	COG1802	K	322	303	19
RCA23_c04890	523118	522333	hypothetical protein			258	358	-100
RCA23_c04900	524284	523202	isopropylmalate dehydrogenase LeuB	COG0473	C	304	357	-53
RCA23_c04910	525496	524363	hypothetical protein			267	377	-110
RCA23_c04920	526231	525626	3-isopropylmalate dehydratase small subunit LeuD	COG0066	E	458	387	71
RCA23_c04930	527634	526231	3-isopropylmalate dehydratase large subunit LeuC	COG0065	E	288	275	13
RCA23_c04940	528605	528072	hypothetical protein, DUF2975			207	247	-40
RCA23_c04950	528844	529257	hypothetical protein DUF143	COG0799	S	450	364	86
RCA23_c04960	529257	529727	ribosomal RNA large subunit methyltransferase H	COG1576	S	245	311	-66
RCA23_c04970	529799	530911	alcohol dehydrogenase class-3 AdhI	COG1062	C	294	435	-141
RCA23_c04980	531014	531961	putative membrane transport protein	COG0679	R	233	422	-189
RCA23_c04990	532000	532896	S-formylglutathione hydrolase YeiG	COG0627	R	250	354	-104
RCA23_c05000	532900	534324	soluble pyridine nucleotide transhydrogenase	COG1249	C	327	370	-43
RCA23_c05010	534321	534755	hypothetical protein DUF188	COG1671	S	235	386	-151
RCA23_c05020	534844	535797	putative integral membrane protein	COG0697	G	439	594	-155
RCA23_c05030	535823	536443	putative HAD-superfamily hydrolase	COG1011	R	270	379	-109
RCA23_c05040	536452	537399	putative ornithine cyclodeaminase	COG2423	E	225	304	-79
RCA23_c05050	541032	537361	aerobic cobaltochelatase subunit CobN	COG1429	H	208	291	-83
RCA23_c05060	541732	541283	hypothetical protein, acetyltransferase-like	COG0456	R	273	241	32
RCA23_c05070	542827	541736	protein CobW	COG0523	R	298	352	-54
RCA23_c05080	542896	543240	hypothetical protein DUF1636	COG5469	S	269	436	-167
RCA23_c05090	543928	543248	possible cobalt transporter, subunit CbtA	COG5446	S	213	300	-87
RCA23_c05100	544412	545827	Cytochrome c peroxidase	COG1858	P	296	759	-463
RCA23_c05110	546982	545975	hypothetical protein	COG4188	R	350	299	51
RCA23_c05120	549014	547548	hypothetical protein			258	381	-123
RCA23_c05130	550285	549011	xylose repressor XylR	COG1940	K	245	284	-39
RCA23_c05140	550504	551532	D-xylose-binding periplasmic protein XylF	COG4213	G	906	544	362

RCA23_c05150	551693	553006	xylose transport system permease protein XylH	COG4214	G	453	505	-52
RCA23_c05160	553018	553776	xylose ABC transporter, ATP-binding protein XylG	COG1129	G	597	619	-22
RCA23_c05170	553780	555234	xylulose kinase XylB	COG1070	G	231	268	-37
RCA23_c05180	555316	556617	xylose isomerase XylA	COG2115	G	393	363	30
RCA23_c05190	558142	556625	long-chain-fatty-acid--CoA ligase LcfB	COG0318	I	319	296	23
RCA23_c05200	559391	558144	malonyl-CoA decarboxylase			373	355	18
RCA23_c05210	561272	559614	acyl-CoA dehydrogenase	COG1960	I	256	352	-96
RCA23_c05220	561381	562196	hypothetical protein DUF81			360	365	-5
RCA23_c05230	563871	562240	phosphoenolpyruvate carboxykinase PckA	COG1866	C	523	455	68
RCA23_c05240	564119	564820	two component signal transduction response regulator rece	COG0745	T	486	575	-89
RCA23_c05250	564856	566508	two component signal transduction histidine kinase ChvG	COG0642	T	217	329	-112
RCA23_c05260	566513	566974	hypothetical protein, HPr serine kinase	COG1493	T	220	332	-112
RCA23_c05270	566967	567878	putative P-loop containing ATPase	COG1660	R	307	374	-67
RCA23_c05280	567875	568264	hypothetical protein, PTS system mannose-specific EIIA co	COG2893	G	251	493	-242
RCA23_c05290	568274	568540	phosphocarrier protein NPr	COG1925	G	259	480	-221
RCA23_c05300	569551	568622	electron transfer flavoprotein alpha subunit EtfA	COG2025	C	357	393	-36
RCA23_c05310	570309	569551	electron transfer flavoprotein beta subunit EtfB	COG2086	C	490	366	124
RCA23_c05320	571054	570482	cob(I)yrinic acid a,c-diamide adenosyltransferase CobO	COG2096	S	330	335	-5
RCA23_c05330	572154	571324	putative short chain dehydrogenase	COG4221	R	327	465	-138
RCA23_c05340	572337	574667	DNA topoisomerase 4 subunit A	COG0188	L	319	408	-89
RCA23_c05350	574828	576000	hypothetical protein DUF898	COG4269	S	147	227	-80
RCA23_c05360	576002	577117	peptidase M48 Ste24p	COG4783	R	216	344	-128
RCA23_c05370	577197	577787	hypothetical protein			310	426	-116
RCA23_c05380	577907	579082	elongation factor Tu (EF-Tu)	COG0050	J	14	5	9
RCA23_c05400	580699	579374	hypothetical protein, DUF560			336	441	-105
RCA23_c05410	582022	580856	hippurate hydrolase HipO	COG1473	R	256	365	-109
RCA23_c05420	583851	582037	glutathione import ATP-binding protein GsiA	COG1123	R	260	312	-52
RCA23_c05430	584120	585829	oligopeptide-binding protein AppA	COG0747	E	394	382	12
RCA23_c05440	585905	586912	oligopeptide transport system permease protein AppB	COG0601	E	420	433	-13
RCA23_c05450	586905	587870	oligopeptide transport system permease protein AppC	COG1173	E	428	438	-10

RCA23_c05460	587971	588900	ABC transporter ATP-binding protein	COG1131	V	711	476	235
RCA23_c05470	588897	589658	inner membrane transport permease	COG0842	V	729	644	85
RCA23_c05480	589709	590494	hypothetical protein, peptidase family S49	COG0616	O	594	456	138
RCA23_c05490	590678	591637	putative sodium/calcium exchanger protein	COG0530	P	399	489	-90
RCA23_c05500	591649	592425	putative short cprotein	COG1028	I	366	403	-37
RCA23_c05510	592554	594575	UvrABC system protein C	COG0322	L	372	374	-2
RCA23_c05520	594704	595360	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltra	COG0558	I	788	670	118
RCA23_c05530	595366	595617	molybdopterin-converting factor subunit MoaD	COG1977	H	349	415	-66
RCA23_c05540	595619	596086	molybdopterin synthase catalytic subunit MoaE	COG0314	H	385	464	-79
RCA23_c05550	596425	596096	hypothetical protein			362	278	84
RCA23_c05560	598304	596418	hypothetical protein, OmpA	COG2885	M	263	298	-35
RCA23_c05570	599317	598358	UbiA prenyltransferase	COG0382	H	237	327	-90
RCA23_c05580	599330	600049	RNA methyltransferase	COG1385	S	221	259	-38
RCA23_c05590	600046	600558	hypothetical protein			265	322	-57
RCA23_c05600	600631	602001	glutamate--cysteine ligase	COG3572	H	440	391	49
RCA23_c05610	602642	602034	glycerol-3-phosphate acyltransferase PlsY	COG0344	S	259	318	-59
RCA23_c05620	603898	602639	dihydroorotase PyrC	COG0044	F	245	261	-16
RCA23_c05630	604984	604040	aspartate carbamoyltransferase PyrB	COG0540	F	333	548	-215
RCA23_c05640	605144	605956	uracil DNA glycosylase family protein	COG1573	L	152	228	-76
RCA23_c05650	605953	606495	molybdenum cofactor biosynthesis protein MoaB	COG0521	H	237	295	-58
RCA23_c05660	606597	608030	putative efflux transporter, RND family, membrane fusion pr	COG1566	V	265	305	-40
RCA23_c05670	608034	611426	transporter, AcrB/AcrD/AcrF family	COG0841	V	347	446	-99
RCA23_c05680	612217	611423	hypothetical protein	COG4233	O	221	319	-98
RCA23_c05690	612351	612965	hypothetical protein DUF179	COG1678	K	253	362	-109
RCA23_c05700	613054	614754	acyl-CoA dehydrogenase	COG1960	I	300	339	-39
RCA23_c05710	614764	615792	putative metallo-beta-lactamase family protein	COG0491	R	238	425	-187
RCA23_c05720	616040	616657	hypothetical protein			369	420	-51
RCA23_c05730	616959	616630	putative branched-chain amino acid transport protein	COG4392	S	293	366	-73
RCA23_c05740	617657	616956	protein AzlC	COG1296	E	276	375	-99
RCA23_c05750	617712	618581	formate dehydrogenase accessory protein FdhD	COG1526	C	297	394	-97

RCA23_c05760	618578	619183	molybdopterin-guanine dinucleotide biosynthesis protein Mo	COG0746	H	193	329	-136
RCA23_c05770	619180	619677	molybdopterin-guanine dinucleotide biosynthesis protein Mo	COG1763	H	159	338	-179
RCA23_c05780	619674	620930	molybdopterin biosynthesis protein MoeA	COG0303	H	162	269	-107
RCA23_c05790	621450	620965	transcription elongation factor GreA	COG0782	K	496	392	104
RCA23_c05800	621745	623397	electron transfer flavoprotein-ubiquinone oxidoreductase	COG0644	C	384	446	-62
RCA23_c05810	623502	625211	tetratricopeptide repeat-containing protein	COG0457	R	227	360	-133
RCA23_c05820	625201	626028	4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate synt	COG1947	I	217	334	-117
RCA23_c05830	627043	626045	octaprenyl-diphosphate synthase IspB	COG0142	H	264	336	-72
RCA23_c05840	627656	628066	hypothetical protein, methyltransferase	COG4123	R	203	250	-47
RCA23_c05850	629161	628439	acetoacetyl-CoA reductase PhaB	COG1028	I	301	347	-46
RCA23_c05860	630460	629267	acetyl-CoA acetyltransferase PhaA	COG0183	I	348	300	48
RCA23_c05870	631425	630640	signaling protein	COG2200	T	475	385	90
RCA23_c05880	631516	632085	DNA-3-methyladenine glycosylase 1	COG2818	L	457	429	28
RCA23_c05890	632617	632066	thiol:disulfide interchange protein TlpA	COG0526	O	277	392	-115
RCA23_c05900	632642	634030	argininosuccinate lyase ArgH	COG0165	E	259	303	-44
RCA23_c05910	634155	634472	hypothetical protein			386	409	-23
RCA23_c05920	634558	635823	diaminopimelate decarboxylase LysA	COG0019	E	332	435	-103
RCA23_c05930	635848	638316	hypothetical protein	COG1196	D	232	318	-86
RCA23_c05940	639228	639956	putative acyltransferase	COG0204	I	437	429	8
RCA23_c05950	640216	640605	glyoxalase/bleomycin resistance protein	COG2764	S	365	382	-17
RCA23_c05960	640646	641269	hypothetical protein, pyridoxamine 5'-phosphate oxidase	COG3576	R	273	303	-30
RCA23_c05970	642246	641284	acetyl-coenzyme A carboxylase carboxyl transferase alpha	COG0825	I	360	382	-22
RCA23_c05980	643307	642348	malyl-CoA ligase	COG2301	G	231	247	-16
RCA23_c05990	643594	644595	hypothetical protein DUF1611	COG3367	S	277	316	-39
RCA23_c06000	644712	645674	L-Ala-D/L-Glu epimerase YcjG	COG4948	M	259	446	-187
RCA23_c06010	645679	646539	D-alanine aminotransferase Dat	COG0115	E	337	480	-143
RCA23_c06020	646631	648778	hypothetical, OmpA-like			226	313	-87
RCA23_c06030	648803	649246	hypothetical protein	COG3743	S	411	402	9
RCA23_c06040	649545	649243	hypothetical protein DUF1244	COG3492	S	260	392	-132
RCA23_c06050	650297	649542	N-formylglutamate amidohydrolase	COG3931	E	349	368	-19

RCA23_c06060	650384	651832	pyruvate kinase PykF	COG0469	G	393	434	-41
RCA23_c06070	651832	652062	hypothetical protein			624	591	33
RCA23_c06080	652176	652376	50S ribosomal protein L35	COG0291	J	871	606	265
RCA23_c06090	652391	652753	50S ribosomal protein L20	COG0292	J	873	409	464
RCA23_c06100	652996	655311	putative subtilase family protein	COG1404	O	455	421	34
RCA23_c06110	655323	655715	hypothetical protein			809	570	239
RCA23_c06120	655813	656757	hypothetical protein, lipid A biosynthesis acyltransferase	COG1560	M	291	395	-104
RCA23_c06130	656869	657942	phenylalanyl-tRNA synthase alpha chain PheS	COG0016	J	304	451	-147
RCA23_c06140	657976	658701	glutamine amidotransferase class-I	COG0518	F	224	306	-82
RCA23_c06150	658701	661094	phenylalanyl-tRNA synthase beta chain PheT	COG0072	J	331	354	-23
RCA23_c06160	661282	661740	putative HTH-type transcriptional regulator	COG1522	K	798	641	157
RCA23_c06170	662019	661819	ribsomal protein S21	COG0828	J	1.005	915	90
RCA23_c06180	662122	662802	putative ubiquinone biosynthesis protein COQ9	COG5590	S	271	315	-44
RCA23_c06190	662830	663807	putative quinone oxidoreductase	COG0604	C	276	335	-59
RCA23_c06200	663977	664744	hypothetical protein DUF1013	COG3820	S	569	525	44
RCA23_c06210	665637	665044	recombination protein RecR	COG0353	L	391	437	-46
RCA23_c06220	665998	665654	hypothetical protein	COG0718	S	517	458	59
RCA23_c06230	667857	666067	DNA polymerase III subunit tau	COG2812	L	205	268	-63
RCA23_c06240	668143	669114	NADH pyrophosphatase Nudc	COG2816	L	277	428	-151
RCA23_c06250	669178	669648	hypothetical protein	COG3832	S	389	485	-96
RCA23_c06260	670475	669645	prephenate dehydratase PheA	COG0077	E	256	261	-5
RCA23_c06270	670580	671113	cytochrome c-552	COG3474	C	552	407	145
RCA23_c06280	671321	673186	ABC transporter extracellular solute-binding protein	COG4166	E	474	456	18
RCA23_c06290	673188	674267	ABC transporter permease protein	COG4174	R	595	473	122
RCA23_c06300	674264	675379	ABC transporter permease protein	COG4239	R	319	447	-128
RCA23_c06310	675376	676956	putative oligopeptide ABC transporter ATP-binding protein	COG4172	R	373	488	-115
RCA23_c06320	677633	676953	fumarylacetoacetate hydrolase family protein	COG0179	Q	215	253	-38
RCA23_c06330	679120	677702	D-alanyl-D-alanine carboxypeptidase DacF	COG1686	M	158	204	-46
RCA23_c06340	680015	679311	haloacid dehalogenase domain protein hydrolase	COG1011	R	406	341	65
RCA23_c06350	680085	680423	ATP-dependent Clp protease adapter protein ClpS	COG2127	S	541	439	102

RCA23_c06360	680427	681377	putative methyltransferase	COG2813	J		262	313	-51
RCA23_c06370	681417	682220	putative short chain dehydrogenase	COG1028	I		340	389	-49
RCA23_c06380	682229	683110	coproporphyrinogen 3 oxidase, aerobic	COG0408	H		253	334	-81
RCA23_c06390	683515	684843	oxidoreductase, FAD-binding protein	COG4097	P		282	382	-100
RCA23_c06400	686376	684955	D-lactate dehydrogenase	COG0277	C		399	414	-15
RCA23_c06420	686652	688244	integrase			GI 3	177	118	59
RCA23_c06430	690078	688801	mandelate racemase	COG4948	M	GI 3	36	16	20
RCA23_c06440	690933	690091	fumarylacetoacetate hydrolase family protein	COG0179	Q	GI 3	11	14	-3
RCA23_c06450	691708	690977	short-chain dehydrogenase/reductase	COG1028	I	GI 3	29	17	12
RCA23_c06460	692579	691701	amidohydrolase	COG3618	R	GI 3	20	19	1
RCA23_c06470	693475	692576	aldo/keto reductase	COG0667	C	GI 3	5	12	-7
RCA23_c06480	694269	693475	ABC transporter, permease protein	COG0600	P	GI 3	26	17	9
RCA23_c06490	695094	694273	ABC transporter, permease protein	COG0600	P	GI 3	32	16	16
RCA23_c06500	695378	695091	hypothetical protein			GI 3	18	11	7
RCA23_c06510	696208	695414	sulfonate/nitrate ABC transporter, ATPase	COG1116	P	GI 3	12	10	2
RCA23_c06520	697212	696214	ABC transporter, periplasmic substrate-binding protein	COG0715	P	GI 3	37	10	27
RCA23_c06530	698756	697293	arylsulfatase	COG3119	P	GI 3	18	10	8
RCA23_c06540	699421	698753	transcriptional regulator, GntR family	COG1802	K	GI 3	18	11	7
RCA23_c06550	700594	700028	hypothetical protein	COG4974	L	GI 3	18	32	-14
RCA23_c06560	701422	700781	hypothetical protein	COG1618	F	GI 3	59	54	5
RCA23_c06570	702310	701492	hypothetical protein			GI 3	393	273	120
RCA23_c06580	702492	703271	hypothetical protein			GI 3	142	92	50
RCA23_c06590	703826	703494	hypothetical protein			GI 3	885	731	154
RCA23_c06600	704277	703819	hypothetical protein			GI 3	675	432	243
RCA23_c06610	705323	704697	hypothetical protein			GI 3	515	372	143
RCA23_c06620	705475	705801	hypothetical protein			GI 3	472	348	124
RCA23_c06630	706774	707259	hypothetical protein	COG3000	I	GI 3	9	7	2
RCA23_c06640	707893	707468	transcriptional regulator, AsnC family	COG1522	K	GI 3	31	39	-8
RCA23_c06650	707989	709179	aspartate aminotransferase AspC	COG0436	E	GI 3	8	32	-24
RCA23_c06660	709204	709950	3-hydroxybutyrate dehydrogenase Bdh	COG1028	I	GI 3	5	26	-21

RCA23_c06670	709963	710817	putative fumarylacetoacetate hydrolase	COG0179	Q	GI 3	9	43	-34
RCA23_c06680	710820	711548	oxidoreductase, short chain dehydrogenase/reductase fami	COG1028	I	GI 3	11	45	-34
RCA23_c06690	711582	712247	dimethylmenaquinone methyltransferase	COG0684	H	GI 3	29	50	-21
RCA23_c06700	712286	713434	hypothetical protein	COG3970	R	GI 3	6	3	3
RCA23_c06710	713447	714952	NADP-dependent fatty aldehyde dehydrogenase AldH	COG1012	C	GI 3	3	2	1
RCA23_c06720	714981	715886	putative dihydrodipicolinate synthase	COG0329	E	GI 3	1	0	1
RCA23_c06730	716879	716205	transcriptional regulator protein, LacI family	COG1609	K	GI 3	33	53	-20
RCA23_c06740	717413	718414	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	GI 3	9	9	0
RCA23_c06750	718496	719083	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	GI 3	10	14	-4
RCA23_c06760	719077	720369	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	GI 3	9	11	-2
RCA23_c06770	720515	721783	D-amino acid dehydrogenase small subunit DadA	COG0665	E	GI 3	21	32	-11
RCA23_c06780	721815	722963	mandelate racemase/muconate lactonizing protein	COG4948	M	GI 3	30	54	-24
RCA23_c06790	722960	723874	dihydrodipicolinate synthase DapA	COG0329	E	GI 3	20	31	-11
RCA23_c06800	723995	725005	uncharacterized oxidoreductase, YjmC	COG2055	C	GI 3	22	34	-12
RCA23_c06810	725230	726555	sarcosine oxidase beta subunit SoxB	COG0665	E	GI 3	11	28	-17
RCA23_c06820	726567	726854	sarcosine oxidase delta subunit SoxD	COG4311	E	GI 3	13	24	-11
RCA23_c06830	726851	729808	sarcosine oxidase alpha subunit SoxA	COG0404	E	GI 3	11	27	-16
RCA23_c06840	729801	730364	sarcosine oxidase gamma subunit SoxG	COG4583	E	GI 3	61	59	2
RCA23_c06850	730372	731262	hypothetical protein, DUF6 transmembrane protein	COG0697	G	GI 3	44	63	-19
RCA23_c06860	732526	734097	putative phage integrase			GI 3	112	112	0
RCA23_c06870	734370	735797	putative amidase	COG0154	J	GI 3	49	26	23
RCA23_c06880	736558	735992	integrase	COG4974	L	GI 3	37	23	14
RCA23_c06890	737386	736745	putative helicase			GI 3	48	32	16
RCA23_c06900	737522	738220	transposase	COG3316	L	GI 3	491	601	-110
RCA23_c06910	739014	738397	HTH-type transcriptional regulator, LuxR family	COG2197	T	GI 3	34	36	-2
RCA23_c06920	739172	739651	putative oxidoreductase, molybdopterin binding	COG3915	S	GI 3	39	27	12
RCA23_c06930	739769	742216	signal transduction histidine kinase	COG0642	T	GI 3	52	35	17
RCA23_c06940	742420	743505	glucose-1-phosphate thymidylyltransferase RfbA	COG1209	M	GI 3	61	48	13
RCA23_c06950	743565	745631	hypothetical protein, chain length determinant protein	COG0489	D	GI 3	59	63	-4
RCA23_c06960	745621	747042	polysaccharide export protein	COG1596	M	GI 3	56	55	1

RCA23_c06970	747621	748013	hypothetical protein, VanZ-like	COG5652	S	GI 3	100	76	24
RCA23_c06980	749389	748190	type I secretion system membrane fusion protein, HlyD fam	COG1566	V	GI 3	76	75	1
RCA23_c06990	751598	749475	type I secretion system ATP-binding component, HlyB famil	COG2274	V	GI 3	96	91	5
RCA23_c07000	752950	751595	type I secretion outer membrane protein, TolC family	COG1538	M	GI 3	63	59	4
RCA23_c07010	754718	752955	putative serralysin-like metalloprotease	COG2931	Q	GI 3	104	122	-18
RCA23_c07020	755773	754919	dTDP-4-dehydrorhamnose reductase RfbD	COG1091	M	GI 3	20	33	-13
RCA23_c07030	756831	755770	dTDP-glucose 4,6-dehydratase RfbB	COG1088	M	GI 3	19	42	-23
RCA23_c07040	757401	756838	dTDP-4-dehydrorhamnose 3,5-epimerase RfbC	COG1898	M	GI 3	22	40	-18
RCA23_c07050	758505	757435	hypothetical protein			GI 3	148	223	-75
RCA23_c07060	759584	758505	glycosyltransferase	COG0438	M	GI 3	33	43	-10
RCA23_c07070	761170	759905	putative glycosyltransferase	COG0438	M	GI 3	38	25	13
RCA23_c07080	762479	761430	hypothetical protein, UDP-glycosyltransferase/glycogen pho	COG0438	M	GI 3	10	28	-18
RCA23_c07090	764009	762492	polysaccharide biosynthesis protein			GI 3	33	56	-23
RCA23_c07100	765109	764015	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotrar	COG0399	M	GI 3	24	51	-27
RCA23_c07110	765867	765106	hypothetical, WxcM-like	COG0110	R	GI 3	13	35	-22
RCA23_c07120	766099	765860	hypothetical protein, WxcM-like			GI 3	46	77	-31
RCA23_c07130	766651	766103	hypothetical protein, acyltransferase-like	COG1670	J	GI 3	42	46	-4
RCA23_c07140	767913	766648	UDP-glucose/GDP-mannose dehydrogenase family protein	COG0677	M	GI 3	73	172	-99
RCA23_c07150	768637	767906	S-adenosyl-L-methionine-dependent methyltransferase			GI 3	28	48	-20
RCA23_c07160	769668	768634	UDP-glucoronate 5'-epimerase LspL	COG0451	M	GI 3	105	167	-62
RCA23_c07170	770090	769713	transcriptional regulator, MarR family	COG1846	K	GI 3	42	63	-21
RCA23_c07180	770490	771002	transcription antitermination protein NusG	COG0250	K	GI 3	53	62	-9
RCA23_c07190	771959	771234	hypothetical protein	COG1434	S	GI 3	9	28	-19
RCA23_c07200	773130	771961	lipopolysaccharide core biosynthesis mannosyltransferase l	COG0438	M	GI 3	10	34	-24
RCA23_c07210	773759	773974	hypothetical protein			GI 3	21	47	-26
RCA23_c07220	774136	775224	undecaprenyl-phosphate alpha-N-acetylglucosaminyl 1-pho	COG0472	M	GI 3	54	68	-14
RCA23_c07230	775393	775839	hypothetical protein, transmembrane			GI 3	8	10	-2
RCA23_c07240	775844	776335	hypothetical protein, transmembrane			GI 3	5	12	-7
RCA23_c07250	777374	776907	transposase, IS4 family protein			GI 3	31	65	-34
RCA23_c07260	777803	777396	putative transposase	COG3039	L	GI 3	20	43	-23

RCA23_c07270	778001	778423	hypothetical protein, transmembrane			GI 3	15	10	5
RCA23_c07280	778428	778958	hypothetical protein, transmembrane			GI 3	20	11	9
RCA23_c07290	779627	780334	transposase	COG3316	L	GI 3	2	0	2
RCA23_c07300	780406	780888	glycogen synthase GlgA	COG0297	G	GI 3	30	39	-9
RCA23_c07310	780903	782972	glycogen debranching enzyme GlgX	COG1523	G	GI 3	71	40	31
RCA23_c07320	783008	784639	phosphoglucomutase Pgm	COG0033	G	GI 3	102	70	32
RCA23_c07330	786248	784641	putative alpha-glucosidase AglA	COG0366	G	GI 3	83	58	25
RCA23_c07340	786430	787128	transposase	COG3316	L	GI 3	425	503	-78
RCA23_c07350	787553	788764	uncharacterized hydrolase YtnL	COG1473	R	GI 3	507	377	130
RCA23_c07360	789242	790729	glutathione-binding protein GsiB	COG0747	E	GI 3	315	289	26
RCA23_c07370	790792	791688	glutathione transport system permease protein GsiC	COG0601	E	GI 3	208	284	-76
RCA23_c07380	791735	792535	dipeptide transport system permease protein	COG1173	E	GI 3	267	375	-108
RCA23_c07390	792532	794106	glutathione import ATP-binding protein GsiA	COG4172	R	GI 3	231	291	-60
RCA23_c07400	794103	795125	peptidase family S58	COG3191	E	GI 3	192	322	-130
RCA23_c07410	795122	796576	aldehyde dehydrogenase AldA	COG1012	C	GI 3	211	236	-25
RCA23_c07420	796585	797325	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	GI 3	252	304	-52
RCA23_c07430	797368	798825	amidase	COG0154	J	GI 3	305	284	21
RCA23_c07440	800908	799691	MFS-type transporter	COG2807	P	GI 3	377	182	195
RCA23_c07450	801403	802311	peptidase M20D amidohydrolase	COG1473	R	GI 3	8	4	4
RCA23_c07460	802343	802858	transposase	COG3316	L	GI 3	8	5	3
RCA23_c07470	802913	803992	integrase	COG2801	L	GI 3	0	0	0
RCA23_c07480	804013	804228	transposase			GI 3	0	3	-3
RCA23_c07490	804774	804520	hypothetical protein, DUF3764			GI 3	230	97	133
RCA23_c07500	805349	804861	hypothetical protein			GI 3	65	47	18
RCA23_c07510	805760	806716	hypothetical protein, beta-lactam-insensitive peptidoglycan t	COG1376	S	GI 3	158	81	77
RCA23_c07520	807253	807663	hypothetical protein			GI 3	286	208	78
RCA23_c07530	809049	807802	ABC transporter, ATP-binding protein, SbmA/BacA-like fam	COG1133	I	GI 3	53	47	6
RCA23_c07540	809116	809526	hypothetical protein			GI 3	67	64	3
RCA23_c07550	809659	809790	hypothetical protein			GI 3	141	42	99
RCA23_c07560	810944	809865	integrase	COG2801	L	GI 3	0	0	0

RCA23_c07570	811605	811252	hypothetical protein			GI 3	8	5	3
RCA23_c07580	812189	812058	hypothetical protein			GI 3	26	6	20
RCA23_c07590	814705	812846	flagellin FliC	COG1344	N	GI 3	26	16	10
RCA23_c07600	815060	816535	RNA polymerase sigma-54	COG1508	K	GI 3	135	86	49
RCA23_c07610	816917	816549	hypothetical protein, cyclic nucleotide-binding-like			GI 3	141	82	59
RCA23_c07620	817270	816917	hypothetical protein	COG0664	T	GI 3	133	74	59
RCA23_c07630	818209	817400	hypothetical protein			GI 3	213	115	98
RCA23_c07640	819031	818219	hypothetical protein			GI 3	368	158	210
RCA23_c07650	819467	819249	hypothetical protein			GI 3	386	127	259
RCA23_c07660	820626	819469	flagellar biosynthesis protein FlhB	COG1377	N	GI 3	257	130	127
RCA23_c07670	821440	820619	flagellar biosynthetic protein FliR	COG1684	N	GI 3	329	132	197
RCA23_c07680	821706	821437	flagellar biosynthetic protein FliQ	COG1987	N	GI 3	317	104	213
RCA23_c07690	822534	821713	flagellar biosynthetic protein FliP	COG1338	N	GI 3	239	115	124
RCA23_c07700	823188	822874	flagellar motor switch protein FliN	COG1886	N	GI 3	289	124	165
RCA23_c07710	824193	823201	flagellar motor switch proteins FliM	COG1868	N	GI 3	246	108	138
RCA23_c07720	824801	824208	protein FliL	COG1580	N	GI 3	391	159	232
RCA23_c07730	826771	824819	putative flagellar hook-length control protein FliK	COG3144	N	GI 3	140	72	68
RCA23_c07740	827248	826829	putative flagellar export protein FliJ			GI 3	144	84	60
RCA23_c07750	828581	827250	flagellar protein export ATPase FliI	COG1157	N	GI 3	114	77	37
RCA23_c07760	829597	828578	putative flagellar assembly protein FliH			GI 3	127	93	34
RCA23_c07770	830634	829594	flagellar motor switch protein G	COG1536	N	GI 3	199	111	88
RCA23_c07780	832414	830663	flagellar M-ring protein FliF	COG1766	N	GI 3	232	111	121
RCA23_c07790	832771	832457	flagellar hook-basal body complex subunit FliE	COG1677	N	GI 3	213	77	136
RCA23_c07800	833940	832786	transcriptional regulatory protein FlbD	COG2204	T	GI 3	241	104	137
RCA23_c07810	834232	834771	protein FliL	COG1580	N	GI 3	61	56	5
RCA23_c07820	834836	835009	hypothetical protein			GI 3	115	56	59
RCA23_c07830	835109	835480	flagellar protein FliS	COG1516	N	GI 3	167	93	74
RCA23_c07840	835599	836012	hypothetical protein			GI 3	232	133	99
RCA23_c07850	836532	836771	hypothetical protein			GI 3	93	80	13
RCA23_c07860	837607	836804	flagellar basal-body rod protein FlgG	COG4786	N	GI 3	112	62	50

RCA23_c07870	837665	839437	Cl- channel, voltage-gated family protein	COG0038	P	GI 3	104	68	36
RCA23_c07880	840278	839442	hypothetical protein, lysozym-like	COG0741	M	GI 3	195	92	103
RCA23_c07890	841546	840275	flagellar hook-associated protein FlgL	COG1344	N	GI 3	211	119	92
RCA23_c07900	845822	841581	flagellar hook-associated protein FlgK	COG1256	N	GI 3	153	98	55
RCA23_c07910	846175	845825	flagellar rod assembly protein/muramidase FlgJ	COG3951	M	GI 3	135	104	31
RCA23_c07920	847324	846188	flagellar P-ring protein FlgI	COG1706	N	GI 3	94	71	23
RCA23_c07930	847901	847338	flagellar L-ring protein FlgH	COG2063	N	GI 3	244	102	142
RCA23_c07940	848820	848032	flagellar basal-body rod protein FlgG	COG4786	N	GI 3	249	113	136
RCA23_c07950	849559	848882	putative flagellar basal body rod protein FlgF	COG4786	N	GI 3	295	125	170
RCA23_c07960	853429	849668	flagellar hook protein FlgE	COG1749	N	GI 3	209	100	109
RCA23_c07970	854213	853497	flagellar basal body rod modification protein FlgD	COG1843	N	GI 3	181	100	81
RCA23_c07980	854630	854214	flagellar basal-body rod protein FlgC	COG1558	N	GI 3	213	123	90
RCA23_c07990	855018	854632	flagellar basal-body rod protein FlgB	COG1815	N	GI 3	313	132	181
RCA23_c08000	855113	855655	hypothetical protein			GI 3	220	107	113
RCA23_c08010	855658	856014	hypothetical protein			GI 3	197	80	117
RCA23_c08020	856748	856023	RNA polymerase, sigma factor for flagellar operon FliA	COG1191	K	GI 3	268	92	176
RCA23_c08030	857551	856745	putative flagellar biosynthesis protein FlhG	COG0455	D	GI 3	198	108	90
RCA23_c08040	858864	857548	flagellar biosynthesis protein FlhF	COG1419	N	GI 3	195	100	95
RCA23_c08050	861023	858861	flagellar biosynthesis protein FlhA	COG1298	N	GI 3	263	115	148
RCA23_c08060	861272	861865	hypothetical membrane lipoprotein, DUF400			GI 3	172	125	47
RCA23_c08070	861871	863016	hypothetical protein			GI 3	101	55	46
RCA23_c08080	863013	863753	flagellar basal body P-ring biosynthesis protein FlgA	COG1261	N	GI 3	126	70	56
RCA23_c08090	863767	864087	putative negative regulator of flagellin synthesis FlgM			GI 3	206	93	113
RCA23_c08100	864117	864407	hypothetical protein			GI 3	258	96	162
RCA23_c08110	865057	864425	hypothetical protein, HCP-like	COG0790	R	GI 3	235	100	135
RCA23_c08120	865443	865054	hypothetical protein			GI 3	143	71	72
RCA23_c08130	865619	866770	sigma54 specific transcriptional regulator, Fis family	COG2204	T	GI 3	203	110	93
RCA23_c08140	868379	866742	flagellar hook-associated protein FliD	COG1345	N	GI 3	223	93	130
RCA23_c08150	869677	868454	chemotaxis protein MotB	COG1360	N	GI 3	231	134	97
RCA23_c08160	870477	869716	chemotaxis protein MotA	COG1291	N	GI 3	299	109	190

RCA23_c08170	870634	871101	hypothetical protein, DUF1566			GI 3	266	154	112
RCA23_c08180	872040	871126	putative transmembrane protein, DUF6			GI 3	397	243	154
RCA23_c08190	874242	872701	putative polyketide hydroxylase SchC	COG0654	H	GI 3	97	68	29
RCA23_c08200	875672	874248	aldehyde dehydrogenase	COG1012	C		60	52	8
RCA23_c08210	876685	875669	2-amino-3-carboxymuconate-6-semialdehyde decarboxylas	COG2159	R		118	61	57
RCA23_c08220	877470	876682	transcriptional regulator, IclR family	COG1414	K		66	69	-3
RCA23_c08230	879401	877467	TRAP transporter, 4TM/12TM fusion protein	COG4666	R		129	84	45
RCA23_c08240	880430	879408	TRAP transporter solute receptor, TAXI family	COG2358	R		86	70	16
RCA23_c08250	880496	882190	thiamine pyrophosphate enzyme-like TPP-binding	COG0028	E		76	72	4
RCA23_c08260	883788	882265	hypothetical protein	COG2268	S		86	60	26
RCA23_c08270	884588	883800	short chain dehydrogenase	COG1028	I		72	42	30
RCA23_c08280	886066	884591	putative aldehyde dehydrogenase yfmT	COG1012	C		119	66	53
RCA23_c08290	887201	886068	branched-chain amino acid ABC transporter, ATP-binding p	COG0410	E		113	70	43
RCA23_c08300	887913	887194	branched-chain amino acid ABC transporter, ATP-binding p	COG0411	E		146	96	50
RCA23_c08310	888863	887931	putative transporter, permease protein	COG4177	E		173	81	92
RCA23_c08320	889732	888860	putative transporter, permease protein	COG0559	E		110	69	41
RCA23_c08330	890937	889783	putative transporter, periplasmic binding protein	COG0683	E		122	74	48
RCA23_c08340	892154	891102	hypothetical protein, PrpF protein-like	COG2828	S		55	59	-4
RCA23_c08350	892833	892168	4-carboxy-4-hydroxy-2-oxoadipate aldolase/oxaloacetate de	COG0684	H		56	60	-4
RCA23_c08360	893753	893019	putative N-acetylglucosaminyl-phosphatidylinositol de-N-ac	COG2120	S		107	67	40
RCA23_c08370	894248	894994	hypothetical protein				161	125	36
RCA23_c08380	895147	895896	transcriptional regulator	COG1802	K		93	88	5
RCA23_c08390	895914	897314	aminomethyltransferase GcvT	COG0404	E		185	101	84
RCA23_c08400	897391	898368	TRAP dicarboxylate transporter, subunit DctP	COG1638	G		97	61	36
RCA23_c08410	898371	898883	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G		178	92	86
RCA23_c08420	898880	900184	TRAP dicarboxylate transporter, subunit DctM	COG1593	G		153	93	60
RCA23_c08430	900268	901209	metapyrocatechase XylE	COG2514	R		158	84	74
RCA23_c08440	901273	902451	MFS-type transporter	COG2814	G		111	95	16
RCA23_c08450	902749	903183	hypothetical protein, cytochrome c				143	79	64
RCA23_c08460	903218	903562	hypothetical protein, copper resistance protein C	COG2372	R		155	96	59

RCA23_c08470	903559	904410	hypothetical protein, copper resistance protein D	COG1276	P	165	111	54
RCA23_c08480	904753	905184	hypothetical protein	COG3613	F	132	101	31
RCA23_c08490	906165	905899	hypothetical protein			227	118	109
RCA23_c08500	906640	906996	hypothetical protein			170	169	1
RCA23_c08510	908415	907033	fumarate reductase flavoprotein subunit	COG1053	C	185	220	-35
RCA23_c08520	909253	908387	methylisocitrate lyase PrpB	COG2513	G	240	253	-13
RCA23_c08530	909870	909250	putative isochorismatase family protein	COG1335	Q	229	231	-2
RCA23_c08540	911929	909872	hydantoin utilization protein A	COG0145	E	178	191	-13
RCA23_c08550	913620	911926	hydantoin utilization protein B	COG0146	E	229	209	20
RCA23_c08560	914183	913617	hypothetical protein	COG1942	R	249	209	40
RCA23_c08570	914328	915128	HTH-type transcriptional regulator, GntR family	COG2188	K	138	206	-68
RCA23_c08580	915125	916531	3-isopropylmalate dehydratase large subunit LeuC	COG0065	E	226	214	12
RCA23_c08590	916558	917142	3-isopropylmalate dehydratase small subunit LeuD	COG0066	E	214	212	2
RCA23_c08600	917203	918216	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	530	304	226
RCA23_c08610	918309	918773	TRAP transporter, subunit DctQ	COG3090	G	514	328	186
RCA23_c08620	918773	920080	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	622	381	241
RCA23_c08630	920337	921674	putative glutamate synthase [NADPH] small chain	COG0493	E	375	303	72
RCA23_c08640	921674	922978	hypothetical protein, dihydrooronate dehydrogenase	COG0167	F	574	518	56
RCA23_c08650	923623	923021	putative HTH-type transcriptional regulator	COG1309	K	816	569	247
RCA23_c08660	923793	925058	N-carbamoyl-L-amino acid hydrolase AmaB	COG0624	E	600	475	125
RCA23_c08670	925095	926549	D-hydantoinase/dihydropyrimidinase Dht	COG0044	F	586	467	119
RCA23_c08680	926592	927374	ABC transporter ATP-binding protein	COG1116	P	627	416	211
RCA23_c08690	927402	928322	putative ABC transporter permease protein	COG0600	P	660	505	155
RCA23_c08700	928319	929161	putative ABC transporter permease protein	COG0600	P	913	619	294
RCA23_c08710	929210	930205	putative thiamine biosynthesis protein	COG0715	P	713	465	248
RCA23_c08720	931297	930413	putative integral membrane proein DUF6	COG0697	G	967	590	377
RCA23_c08730	932791	931913	hypothetical protein DUF6	COG0697	G	693	436	257
RCA23_c08740	933010	934104	alkaline phosphatase synthesis sensor protein PhoR	COG5002	T	428	407	21
RCA23_c08750	934233	935270	putative phosphate binding protein PstS	COG0226	P	488	443	45
RCA23_c08760	935370	936857	putative phosphate transport system permease protein PstC	COG0573	P	413	525	-112

RCA23_c08770	936857	938185	putative phosphate transport system permease protein PstA	COG0581	P	472	478	-6
RCA23_c08780	938190	938996	phosphate import ATP-binding protein PstB	COG1117	P	497	610	-113
RCA23_c08790	939007	939699	phosphate transport system regulatory protein PhoU	COG0704	P	608	544	64
RCA23_c08800	939724	940413	phosphate regulon transl protein PhoB	COG0745	T	1.061	623	438
RCA23_c08810	940589	941791	putative hippurate hydrolase protei	COG1473	R	206	127	79
RCA23_c08820	942193	942954	urease accessory protein UreD	COG0829	O	648	456	192
RCA23_c08830	942958	943260	urease gamma subunit UreA	COG0831	E	478	528	-50
RCA23_c08840	943270	943575	urease beta subunit UreB	COG0832	E	217	341	-124
RCA23_c08850	943575	945314	urease alpha subunit UreC	COG0804	E	415	369	46
RCA23_c08860	945314	945787	urease accessory protein UreE	COG2371	O	361	376	-15
RCA23_c08870	945957	946415	urease accessory protein UreF	COG0830	O	517	461	56
RCA23_c08880	946427	947074	urease accessory protein UreG	COG0378	O	698	524	174
RCA23_c08890	948686	947607	integrase	COG2801	L	0	0	0
RCA23_c08900	948917	949201	hypothetical protein	COG4585	T	780	501	279
RCA23_c08910	951070	949424	glucose-methanol-choline oxidoreductase AlkJ	COG2303	E	382	383	-1
RCA23_c08920	951279	951830	hypothetical protein, YaeQ family protein	COG4681	S	731	502	229
RCA23_c08930	953443	952526	hypothetical protein, DUF6 transmembrane protein	COG0697	G	592	558	34
RCA23_c08940	953803	954969	UDP-glucose 6-dehydrogenase Ugd	COG1004	M	456	389	67
RCA23_c08950	956602	955373	3-deoxy-D-manno-octulosonic-acid transferase WaaA	COG1519	M	251	308	-57
RCA23_c08960	957788	956628	UDP-N-acetylglucosamine 2-epimerase WecB	COG0381	M	421	408	13
RCA23_c08970	958846	957815	putative N-acetylneuramic acid synthase	COG2089	M	556	454	102
RCA23_c08980	959840	958926	hypothetical protein			595	416	179
RCA23_c08990	960088	960732	hypothetical protein	COG1651	O	387	414	-27
RCA23_c09000	960843	962327	choline-sulfatase BetC	COG3119	P	445	476	-31
RCA23_c09010	963713	962403	permease	COG2233	F	269	317	-48
RCA23_c09020	964049	964585	hypothetical protein			249	294	-45
RCA23_c09030	964737	965636	membrane dipeptidase	COG2355	E	283	328	-45
RCA23_c09040	965633	967285	hypothetical protein			232	375	-143
RCA23_c09050	967282	967791	hypothetical protein			138	237	-99
RCA23_c09060	968589	967816	tRNA pseudouridine synthase A	COG0101	J	255	224	31

RCA23_c09070	968867	970279	YcjX-like protein	COG3106	R		270	373	-103
RCA23_c09080	970276	970722	hypothetical protein, acetyltransferase-like	COG0456	R		298	384	-86
RCA23_c09090	970715	971731	hypothetical protein	COG3768	S		237	411	-174
RCA23_c09100	971937	974843	isoleucyl-tRNA synthase IleS	COG0060	J		370	394	-24
RCA23_c09110	974910	975746	putative integral membrane protein DUF6				282	479	-197
RCA23_c09120	976432	975743	phosphatidylcholine synthase Pcs	COG1183	I		316	382	-66
RCA23_c09130	977409	976468	putative tyrosine recombinase xerC	COG4974	L		184	268	-84
RCA23_c09140	978135	977416	hypothetical protein	COG3159	S		307	263	44
RCA23_c09150	978835	978182	transaldolase	COG0176	G		540	374	166
RCA23_c09160	978925	981114	primosomal protein N'	COG1198	L		227	311	-84
RCA23_c09170	983095	981122	methylmalonyl-CoA mutase McmA	COG1884	I		237	293	-56
RCA23_c09180	983235	984599	glycerol-3-phosphate acyltransferase PlsB	COG2937	I		324	355	-31
RCA23_c09190	984707	985993	crotonyl-CoA reductase Ccr	COG0604	C		367	438	-71
RCA23_c09200	986057	987634	putative ATPase	COG0507	L		241	328	-87
RCA23_c09210	989413	987845	trimethylamine methyltransferase MttB	COG5598	H		191	262	-71
RCA23_c09220	989431	989895	hypothetical protein DUF1052	COG5321	S		188	298	-110
RCA23_c09240	990785	991279	hypothetical protein, DUF411	COG3019	R	GI 4	13	25	-12
RCA23_c09250	991576	991869	hypothetical protein			GI 4	0	0	0
RCA23_c09260	992104	992391	hypothetical protein			GI 4	5	2	3
RCA23_c09270	993390	992584	MORN repeat	COG4642	S	GI 4	16	14	2
RCA23_c09280	993934	993512	hypothetical protein	COG3904	S	GI 4	0	7	-7
RCA23_c09290	994038	995099	putative DNA-binding protein, transposase-like	COG3415	L	GI 4	0	0	0
RCA23_c09300	996057	995158	HTH-type transcriptional regulator, LysR family	COG0583	K	GI 4	4	1	3
RCA23_c09310	996333	999947	protein of unknown function, DUF285/Bacterial Ig-like domain			GI 4	4	1	3
RCA23_c09320	1000260	1001186	Predicted esterase	COG0627	R	GI 4	0	0	0
RCA23_c09330	1002131	1001247	transposase	COG2801	L	GI 4	28	98	-70
RCA23_c09340	1002367	1002143	putative transposase			GI 4	35	50	-15
RCA23_c09350	1003498	1002938	hypothetical protein, ornithine cyclodeaminase-like	COG2423	E		0	1	-1
RCA23_c09360	1004444	1003728	LrgB-like protein	COG1346	M		289	250	39
RCA23_c09370	1004794	1004441	hypothetical protein	COG1380	R		347	346	1

RCA23_c09380	1005815	1004916	putative branched-chain-amino-acid aminotransferase IlvE	COG0115	E	518	431	87
RCA23_c09390	1006761	1005865	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I	343	293	50
RCA23_c09400	1008113	1006908	bicyclomycin resistance protein Bcr	COG2814	G	355	357	-2
RCA23_c09410	1008660	1008187	hypothetical protein			338	365	-27
RCA23_c09420	1010198	1008657	hypothetical protein, photolyase	COG3046	R	441	449	-8
RCA23_c09430	1010370	1011587	DNA photolyase, FAD-binding/cryptochrome	COG0415	L	245	277	-32
RCA23_c09440	1011944	1012426	carbon monoxide dehydrogenase small chain CoxS	COG2080	C	366	386	-20
RCA23_c09450	1012438	1014798	carbon monoxide dehydrogenase large chain CoxL	COG1529	C	282	324	-42
RCA23_c09460	1014811	1015602	carbon monoxide dehydrogenase medium chain CoxM	COG1319	C	294	374	-80
RCA23_c09470	1015645	1016556	MoxR-like ATPase	COG0714	R	367	401	-34
RCA23_c09480	1016556	1017776	CoxE-like protein	COG3552	R	237	272	-35
RCA23_c09490	1017882	1018688	5-deoxy-glucuronate isomerase IolB	COG3718	G	236	430	-194
RCA23_c09500	1019039	1018707	branched-chain amino acid transport protein, AzlD-like	COG4392	S	236	313	-77
RCA23_c09510	1019743	1019036	branched-chain amino acid transport protein, AzlC-like	COG1296	E	322	361	-39
RCA23_c09520	1021462	1019804	alpha-IPM synthase/homocitrate synthase	COG0119	E	268	314	-46
RCA23_c09530	1022656	1021499	cytochrome P450	COG2124	Q	179	259	-80
RCA23_c09540	1022810	1023811	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	437	429	8
RCA23_c09550	1023966	1024568	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	573	365	208
RCA23_c09560	1024573	1026420	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	283	338	-55
RCA23_c09570	1026486	1028369	putative citrate transporter	COG0471	P	356	376	-20
RCA23_c09580	1029097	1028366	triosephosphate isomerase TpiA	COG0149	G	264	327	-63
RCA23_c09590	1029454	1029137	putative regulator protein of competence-specific genes Tfo	COG3070	K	392	409	-17
RCA23_c09600	1029816	1029451	hypothetical protein, iron-sulfur cluster assembly protein	COG0316	S	520	454	66
RCA23_c09610	1030445	1029936	hypothetical protein DUF59	COG2151	R	242	320	-78
RCA23_c09620	1030469	1031599	queuine tRNA-ribosyltransferase Tgt	COG0343	J	237	292	-55
RCA23_c09630	1032099	1031611	hypothetical protein			216	256	-40
RCA23_c09640	1032374	1034782	ATP-dependent protease La	COG0466	O	795	491	304
RCA23_c09650	1034823	1035461	putative phosphoglycerate mutase family protein	COG0406	G	245	244	1
RCA23_c09670	1035777	1036430	protein-L-isoaspartate O-methyltransferase Pcm	COG2518	O	703	588	115
RCA23_c09680	1036466	1037830	outer membrane efflux protein	COG1538	M	327	323	4

RCA23_c09690	1037905	1038228	hypothetical protein			190	242	-52
RCA23_c09700	1039719	1038262	cobyric acid synthase CobQ	COG1492	H	274	372	-98
RCA23_c09710	1040228	1039917	hypothetical protein			1.161	530	631
RCA23_c09720	1040445	1041008	translation elongation factor P	COG0231	J	447	318	129
RCA23_c09730	1041736	1041005	hypothetical protein tRNA modifying protein YgfZ	COG0354	R	362	405	-43
RCA23_c09740	1041797	1042474	putative glycosyltransferase, family 2	COG0463	M	225	395	-170
RCA23_c09750	1044299	1042530	lipid A export ATP-binding/permease protein MsbA	COG1132	V	290	354	-64
RCA23_c09760	1044403	1045608	serine--glyoxylate aminotransferase SgaA	COG0075	E	201	278	-77
RCA23_c09770	1046719	1045610	histidinol-phosphate aminotransferase HisC	COG0079	E	210	281	-71
RCA23_c09780	1049682	1046755	valyl-tRNA synthase ValS	COG0525	J	332	382	-50
RCA23_c09790	1050340	1049744	hypothetical protein, metal-dependent phosphohydrolase	COG4341	R	276	359	-83
RCA23_c09800	1051209	1050337	putative phytanoyl-CoA dioxygenase family protein	COG5285	Q	412	427	-15
RCA23_c09810	1051351	1053636	molybdenum-containing hydroxylase	COG1529	C	264	374	-110
RCA23_c09820	1053643	1054722	hypothetical protein DUF2235	COG3673	S	304	419	-115
RCA23_c09830	1056046	1055225	5,10-methylenetetrahydrofolate reductase MetF	COG0685	E	229	270	-41
RCA23_c09840	1056143	1057048	HTH-type transcriptional regulator MetR	COG0583	K	226	312	-86
RCA23_c09850	1057134	1057898	inositol-1-monophosphatase SuhB	COG0483	G	292	283	9
RCA23_c09860	1057986	1058552	hypothetical protein	COG3063	N	278	328	-50
RCA23_c09880	1059277	1058798	transcriptional regulator, HxlR family	COG1733	K	534	437	97
RCA23_c09890	1059447	1059785	hypothetical protein, DUF3764			523	370	153
RCA23_c09900	1059839	1060168	hypothetical protein			293	227	66
RCA23_c09910	1060204	1060566	hypothetical protein, DsrE/F-like	COG1553	P	359	278	81
RCA23_c09920	1061259	1060684	glutathione peroxidase Gpo	COG0386	O	225	261	-36
RCA23_c09930	1062249	1061521	putative ion channel			421	440	-19
RCA23_c09950	1063201	1062944	hypothetical protein			347	292	55
RCA23_c09940	1063248	1064246	ribosomal large subunit pseudouridine synthase D	COG0564	J	255	339	-84
RCA23_c09960	1064374	1065270	RNA polymerase sigma-32 factor RpoH	COG0568	K	1.034	612	422
RCA23_c09970	1065351	1066337	putative oxidoreductase	COG0673	R	384	351	33
RCA23_c09980	1068332	1066503	oligoendopeptidase F	COG1164	E	339	303	36
RCA23_c09990	1068520	1069284	2-keto-3-deoxy-L-rhamnonate aldolase RhmA	COG3836	G	273	331	-58

RCA23_c10000	1070216	1069269	hypothetical protein, lysophospholipase L2	COG2267	I	278	289	-11
RCA23_c10010	1070506	1070216	putative sterol-binding protein	COG3255	I	284	291	-7
RCA23_c10020	1070580	1071140	hypothetical protein			335	350	-15
RCA23_c10030	1071160	1073967	ATP-dependent RNA helicase RhlB	COG4581	L	260	318	-58
RCA23_c10040	1074051	1074341	putative heat shock protein	COG1188	J	158	222	-64
RCA23_c10050	1074410	1074748	ferredoxin FdxA	COG1146	C	427	421	6
RCA23_c10060	1074901	1075419	putative transcriptional regulator, CarD family	COG1329	K	344	315	29
RCA23_c10070	1076242	1075490	cobalamin-5-phosphate synthase CobS	COG0368	H	135	170	-35
RCA23_c10080	1076326	1077339	nicotinate-nucleotide--dimethylbenzimidazole phosphoribos	COG2038	H	180	268	-88
RCA23_c10090	1079387	1077345	hypothetical protein			166	262	-96
RCA23_c10100	1079592	1080008	glyoxalase/bleomycin resistance protein	COG3565	R	433	479	-46
RCA23_c10110	1080998	1080018	hypothetical protein			392	375	17
RCA23_c10120	1081235	1082887	choline dehydrogenase BetA	COG2303	E	290	336	-46
RCA23_c10140	1083425	1084330	23S rRNA (guanosine-2'-O-)-methyltransferase RlmB	COG0566	J	274	297	-23
RCA23_c10150	1084647	1084342	phosphoribosyl-ATP pyrophosphatase HisE	COG0140	E	293	375	-82
RCA23_c10160	1085378	1084644	imidazole glycerol phosphate synthase subunit HisF	COG0107	E	188	248	-60
RCA23_c10170	1086150	1085434	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylider	COG0106	E	179	178	1
RCA23_c10180	1086818	1086180	imidazole glycerol phosphate synthase, glutamine amidotra	COG0118	E	345	348	-3
RCA23_c10190	1087410	1086823	imidazoleglycerol-phosphate dehydratase HisB	COG0131	E	327	348	-21
RCA23_c10210	1088751	1087954	hypothetical protein, calcineurin-like phosphoesterase-like	COG1409	R	418	358	60
RCA23_c10220	1092218	1088778	pyruvate carboxylase Pyc	COG1038	C	237	292	-55
RCA23_c10230	1093522	1092374	L-lactate dehydrogenase IldD	COG1304	C	281	304	-23
RCA23_c10240	1093655	1096048	putative DNA helicase II	COG0210	L	226	333	-107
RCA23_c10250	1096133	1097425	gamma-glutamylputrescine oxidoreductase PuuB	COG0665	E	199	255	-56
RCA23_c10260	1098046	1097444	putative cysteine/O-acetylserine efflux protein	COG1280	E	283	451	-168
RCA23_c10270	1098056	1098466	hypothetical protein			1.083	672	411
RCA23_c10280	1098622	1099686	putative protein Mrp	COG0489	D	232	370	-138
RCA23_c10290	1099713	1099967	hypothetical protein			599	383	216
RCA23_c10300	1100315	1100818	hypothetical protein, MraZ	COG2001	S	449	455	-6
RCA23_c10310	1100821	1101810	S-adenosyl-L-methionine-dependent methyltransferase Mra	COG0275	M	231	309	-78

RCA23_c10320	1101807	1102151	hypothetical protein	COG5462	S	369	361	8
RCA23_c10330	1102148	1103929	peptidoglycan synthase FtsI	COG0768	M	297	335	-38
RCA23_c10340	1103952	1105430	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminop	COG0769	M	234	338	-104
RCA23_c10350	1105430	1106848	putative UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alan	COG0770	M	223	280	-57
RCA23_c10360	1106866	1107948	phospho-N-acetylmuramoyl-pentapeptide-transferase MraY	COG0472	M	295	395	-100
RCA23_c10370	1108005	1109384	UDP-N-acetylmuramoylalanine--D-glutamate ligase MurD	COG0771	M	241	377	-136
RCA23_c10380	1109471	1110160	putative glycosyltransferase, sugar binding region	COG3774	M	549	438	111
RCA23_c10390	1110834	1110157	putative galactoside 2-alpha-L-fucosyltransferase 1			675	419	256
RCA23_c10400	1110875	1111054	hypothetical protein			439	354	85
RCA23_c10410	1111162	1111935	hypothetical protein	COG3306	M	476	358	118
RCA23_c10420	1113153	1111963	hypothetical protein HI0933	COG2081	R	190	222	-32
RCA23_c10430	1113229	1114434	cell division protein FtsW	COG0772	D	301	420	-119
RCA23_c10440	1114424	1115530	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide	COG0707	M	176	275	-99
RCA23_c10450	1115530	1116921	UDP-N-acetylmuramate--L-alanine ligase MurC	COG0773	M	264	342	-78
RCA23_c10460	1117803	1116925	permease of the drug/metabolite transporter superfamily	COG0697	G	226	272	-46
RCA23_c10470	1117895	1118863	UDP-N-acetylenolpyruvoylglucosamine reductase MurB	COG0812	M	208	372	-164
RCA23_c10480	1118953	1119843	D-alanine--D-alanine ligase Ddl	COG1181	M	235	430	-195
RCA23_c10490	1119879	1120751	putative cell division protein FtsQ			316	447	-131
RCA23_c10500	1120748	1122082	cell division protein FtsA	COG0849	D	275	404	-129
RCA23_c10510	1122236	1123780	cell division protein FtsZ	COG0206	D	341	361	-20
RCA23_c10520	1123992	1124912	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacet	COG0774	M	325	407	-82
RCA23_c10530	1125020	1125835	outer membrane assembly lipoprotein	COG4105	R	458	485	-27
RCA23_c10540	1125852	1127504	DNA repair protein RecN	COG0497	L	240	362	-122
RCA23_c10550	1127863	1127501	hypothetical protein DUF427	COG2343	S	394	461	-67
RCA23_c10560	1129669	1127879	putative Xaa-Pro aminopeptidase	COG0006	E	249	396	-147
RCA23_c10570	1131630	1129756	aerobic cobaltochelatase subunit CobT	COG4547	H	258	346	-88
RCA23_c10580	1132654	1131668	aerobic cobaltochelatase subunit CobS	COG0714	R	585	452	133
RCA23_c10590	1133471	1132833	hypothetical protein, DnaJ	COG0484	O	694	495	199
RCA23_c10600	1133514	1133801	putative stress-induced morphoprotein, BolA type	COG0271	T	335	420	-85
RCA23_c10610	1135336	1133825	aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit	COG0064	J	356	465	-109

RCA23_c10620	1137977	1135425	aminopeptidase N	COG0308	E		257	346	-89
RCA23_c10630	1138196	1140334	malate synthase GlcB	COG2225	C		371	390	-19
RCA23_c10640	1140396	1140824	hypothetical protein DUF336	COG3193	R		225	350	-125
RCA23_c10650	1140906	1141109	hypothetical protein				218	311	-93
RCA23_c10660	1141119	1141883	gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD	COG2071	R		209	324	-115
RCA23_c10670	1142635	1141880	putative D-beta-hydroxybutyrate dehydrogenase	COG1028	I		198	280	-82
RCA23_c10680	1142794	1144365	NAD(P) transhydrogenase alpha subunit PntA	COG3288	C		325	332	-7
RCA23_c10690	1144379	1145812	NAD(P) transhydrogenase beta subunit PntB	COG1282	C		435	457	-22
RCA23_c10700	1145880	1147163	hypothetical protein	COG4949	S		357	421	-64
RCA23_c10710	1147211	1148374	putative acetylornithine deacetylase ArgE	COG0624	E		296	398	-102
RCA23_c10720	1149375	1148371	molybdenum cofactor biosynthesis protein MoaA	COG2896	H		223	295	-72
RCA23_c10730	1149576	1151222	acetyl-coenzyme A synthase AcsA	COG0365	I		196	332	-136
RCA23_c10740	1151397	1151173	hypothetical protein	COG4321	R		261	309	-48
RCA23_c10750	1151573	1151397	hypothetical protein				286	384	-98
RCA23_c10760	1152961	1151570	fumarate hydratase class II	COG0114	C		301	362	-61
RCA23_c10770	1153500	1153036	hypothetical protein	COG3814	S		562	428	134
RCA23_c10780	1154840	1153572	putative chromate transport protein	COG2059	P		332	384	-52
RCA23_c10790	1155189	1157231	hydantoinase / oxoprolinase family protein	COG0145	E		170	273	-103
RCA23_c10800	1157331	1157729	hypothetical protein				647	502	145
RCA23_c10810	1158217	1159203	putative quinone oxidoreductase yhdH	COG0604	C		214	264	-50
RCA23_c10820	1159306	1160529	cysteine desulfurase SufS	COG0520	E		246	312	-66
RCA23_c10830	1160550	1161962	deoxyribodipyrimidine photo-lyase PhrB	COG0415	L		227	252	-25
RCA23_c10840	1162037	1163245	cyclopropane-fatty-acyl-phospholipid synthase Cfa	COG2230	M		546	408	138
RCA23_c10850	1163253	1164371	ADP-ribose pyrophosphatase NudF	COG0494	L		289	351	-62
RCA23_c10860	1164417	1165454	cysteine synthase CysK	COG0031	E		402	390	12
RCA23_c10870	1165441	1167759	mechanosensitive ion channel protein MscS	COG3264	M		353	377	-24
RCA23_c10890	1168102	1169268	putative phage integrase	COG4974	L	GI 5	13	11	2
RCA23_c10900	1173155	1169922	hypothetical protein			GI 5	18	15	3
RCA23_c10910	1173220	1175214	hypothetical protein, tetratricopetide domain TPR-1	COG0457	R	GI 5	22	11	11
RCA23_c10920	1177861	1176257	hypothetical protein, resolvase-like	COG1961	L	GI 5	799	557	242

RCA23_c10930	1178310	1177858	hypothetical protein, DUF2924			GI 5	237	234	3
RCA23_c10940	1178871	1179848	putative TRAP transporter solute receptor	COG2358	R	GI 5	239	214	25
RCA23_c10950	1179941	1182292	TRAP transporter, 4TM/12TM fusion protein	COG4666	R	GI 5	386	349	37
RCA23_c10960	1182304	1183887	putative sulfatase YidJ	COG3119	P	GI 5	325	398	-73
RCA23_c10970	1183893	1184789	arylsulfatase	COG1234	R	GI 5	306	236	70
RCA23_c10980	1184804	1185622	siderophore interactin protein, vibriobactin utilization protein	COG2375	P	GI 5	477	318	159
RCA23_c10990	1186917	1185640	hypothetical protein, metallo-beta-lactamase	COG2015	Q	GI 5	550	453	97
RCA23_c11000	1187177	1188193	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	GI 5	417	318	99
RCA23_c11010	1188369	1189034	glutathione S-transferase	COG0625	O	GI 5	357	294	63
RCA23_c11020	1189053	1189601	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	GI 5	459	306	153
RCA23_c11030	1189598	1190959	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	GI 5	284	244	40
RCA23_c11040	1190995	1191909	hypothetical protein, DUF6 transmembrane protein	COG0697	G	GI 5	228	198	30
RCA23_c11050	1192027	1192302	ferric reductase like transmembrane component family			GI 5	221	170	51
RCA23_c11060	1192990	1192559	hypothetical protein			GI 5	458	332	126
RCA23_c11070	1194176	1192992	hypothetical protein			GI 5	378	282	96
RCA23_c11080	1195854	1194217	sulfatase family protein	COG3119	P	GI 5	477	324	153
RCA23_c11090	1196618	1195860	short-chain dehydrogenase/reductase SDR	COG1028	I	GI 5	479	340	139
RCA23_c11100	1197395	1196637	maleylacetoacetate isomerase MaiA	COG0625	O	GI 5	545	391	154
RCA23_c11110	1198888	1197452	salicylaldehyde dehydrogenase	COG1012	C	GI 5	411	357	54
RCA23_c11120	1199107	1199895	transcriptional regulator, GntR family	COG1802	K	GI 5	419	355	64
RCA23_c11130	1200361	1200789	transposase	COG3316	L	GI 5	2	12	-10
RCA23_c11140	1200946	1202112	putative phage integrase	COG4974	L	GI 5	1	1	0
RCA23_c11150	1202483	1204498	hypothetical protein, alginate lyase-lyase			GI 5	2	3	-1
RCA23_c11160	1205812	1205189	HTH-type transcriptional regulator, GntR family	COG1802	K	GI 5	339	300	39
RCA23_c11170	1206137	1207324	branched-chain amino acid ABC transporter, periplasmic su	COG0683	E	GI 5	438	334	104
RCA23_c11180	1207409	1208311	branched-chain amino acid ABC transporter, permease prot	COG0559	E	GI 5	327	313	14
RCA23_c11190	1208316	1209296	branched-chain amino acid ABC transporter, permease prot	COG4177	E	GI 5	497	346	151
RCA23_c11200	1209320	1210075	branched-chain amino acid ABC transporter, ATP-binding p	COG0411	E	GI 5	414	325	89
RCA23_c11210	1210079	1210801	branched-chain amino acid ABC transporter, ATP-binding p	COG0410	E	GI 5	408	292	116
RCA23_c11220	1210808	1211491	hydantoin racemase HyuE	COG4126	E	GI 5	395	264	131

RCA23_c11230	1211851	1212558	transposase	COG3316	L	GI 5	66	52	14
RCA23_c11240	1213785	1212979	glutamine transport ATP-binding protein GlnQ	COG1126	E	GI 5	32	11	21
RCA23_c11250	1214500	1213802	putative inner membrane amino-acid ABC transporter perm	COG0765	E	GI 5	46	11	35
RCA23_c11260	1214805	1215512	transposase	COG3316	L	GI 5	10	6	4
RCA23_c11270	1216048	1217244	transcriptional regulator CoxC	COG3300	T	GI 5	224	150	74
RCA23_c11280	1217360	1218214	carbon monoxide dehydrogenase medium chain CoxM	COG1319	C	GI 5	83	69	14
RCA23_c11290	1218225	1218719	carbon monoxide dehydrogenase small chain CoxS	COG2080	C	GI 5	77	78	-1
RCA23_c11300	1218716	1221142	carbon monoxide dehydrogenase large chain CoxL	COG1529	C	GI 5	130	117	13
RCA23_c11310	1221188	1222063	AAA+ ATPase chaperone CoxD	COG0714	R	GI 5	129	129	0
RCA23_c11320	1222060	1223241	carbon monoxide dehydrogenase accessory protein CoxE	COG3552	R	GI 5	68	81	-13
RCA23_c11330	1223238	1224017	carbon monoxide dehydrogenase accessory protein CoxF	COG1975	O	GI 5	63	69	-6
RCA23_c11340	1224010	1224600	hypothetical protein	COG2068	R	GI 5	75	70	5
RCA23_c11350	1224727	1225182	carbon monoxide dehydrogenase protein CoxG	COG3427	S	GI 5	111	80	31
RCA23_c11360	1225187	1226101	carbon monoxide dehydrogenase accessory protein CoxI	COG1975	O	GI 5	120	110	10
RCA23_c11370	1226311	1226796	transposase	COG3316	L	GI 5	34	32	2
RCA23_c11380	1227061	1227807	dienelactone hydrolase	COG0412	Q	GI 5	337	347	-10
RCA23_c11390	1228850	1227804	aldo/keto reductase	COG0667	C	GI 5	243	227	16
RCA23_c11400	1230223	1229126	mandelate racemase MdlA	COG4948	M	GI 5	299	385	-86
RCA23_c11410	1231189	1230335	hypothetical protein DUF1498	COG3822	R	GI 5	363	506	-143
RCA23_c11420	1232598	1231231	two-component system, sensor histidine kinase protein	COG0642	T	GI 5	218	328	-110
RCA23_c11430	1233299	1232595	two-component system, response regulator protein	COG0745	T	GI 5	255	357	-102
RCA23_c11440	1233404	1234510	hypothetical protein	COG3181	S	GI 5	441	342	99
RCA23_c11450	1234593	1236608	putative tripartite tricarboxylate transporter (TTT) protein Tc	COG3333	S	GI 5	543	448	95
RCA23_c11460	1236605	1237696	hypothetical protein	COG2828	S	GI 5	288	325	-37
RCA23_c11470	1237724	1238692	D-isomer specific 2-hydroxyacid dehydrogenase	COG0111	H	GI 5	297	421	-124
RCA23_c11480	1239793	1238756	putative ammonia monooxygenase	COG3180	R	GI 5	263	351	-88
RCA23_c11490	1240440	1239790	transcriptional regulator	COG1802	K	GI 5	194	307	-113
RCA23_c11500	1240489	1241970	fumarate reductase flavoprotein subunit FccA	COG1053	C	GI 5	282	318	-36
RCA23_c11510	1242188	1243363	aspartate aminotransferase AspC	COG0436	E	GI 5	321	400	-79
RCA23_c11520	1244165	1243431	hypothetical protein			GI 5	269	385	-116

RCA23_c11530	1244807	1244235	aerobic glycerol-3-phosphate dehydrogenase GlpD	COG0578	C	GI 5	198	314	-116
RCA23_c11540	1245824	1244865	aerobic glycerol-3-phosphate dehydrogenase GlpD	COG0578	C	GI 5	239	288	-49
RCA23_c11550	1247291	1247935	hypothetical protein	COG1192	D	GI 5	29	30	-1
RCA23_c11560	1247928	1249136	chemotaxis protein CheW	COG0835	N	GI 5	27	26	1
RCA23_c11570	1249188	1251065	methyl-accepting chemotaxis protein II	COG0840	N	GI 5	62	37	25
RCA23_c11580	1251091	1254051	chemotaxis histidine kinase CheA	COG0643	N	GI 5	40	20	20
RCA23_c11590	1254048	1256066	response regulator CheY	COG2114	T	GI 5	36	14	22
RCA23_c11600	1256078	1256908	chemotaxis protein methyltransferase CheR	COG1352	N	GI 5	0	6	-6
RCA23_c11610	1256889	1257911	chemotaxis response regulator protein-glutamate methylest	COG2201	N	GI 5	13	11	2
RCA23_c11620	1258541	1259149	metal-dependent hydrolase	COG1235	R	GI 5	125	59	66
RCA23_c11630	1260667	1259588	integrase	COG2801	L	GI 5	0	0	0
RCA23_c11640	1261553	1260729	methyltransferase, FkbM family			GI 5	13	25	-12
RCA23_c11650	1262611	1261847	hypothetical protein, sugar transferase-like			GI 5	33	35	-2
RCA23_c11660	1263686	1262811	aldo/keto reductase	COG0667	C	GI 5	35	36	-1
RCA23_c11670	1264642	1264917	hypothetical protein			GI 5	44	45	-1
RCA23_c11680	1265571	1265173	hypothetical protein	COG0790	R	GI 5	13	13	0
RCA23_c11690	1266818	1266246	hypothetical protein	COG4430	S	GI 5	255	166	89
RCA23_c11700	1267451	1268101	hypothetical protein	COG2755	E	GI 5	0	0	0
RCA23_c11710	1268699	1269586	phytanoyl-CoA dioxygenase family protein	COG5285	Q	GI 5	343	209	134
RCA23_c11720	1270243	1270815	putative bacterial extracellular solute binding protein, family	COG1791	S	GI 5	67	79	-12
RCA23_c11730	1271031	1271468	hypothetical protein			GI 5	73	29	44
RCA23_c11740	1272798	1272199	SOUL heme-binding protein			GI 5	321	247	74
RCA23_c11750	1273478	1274995	arylsulfatase precursor	COG3119	P	GI 5	113	132	-19
RCA23_c11760	1275232	1275627	hypothetical protein DUF583			GI 5	7	10	-3
RCA23_c11770	1275804	1276883	integrase	COG2801	L	GI 5	0	0	0
RCA23_c11780	1277985	1276918	hypothetical protein	COG3616	E	GI 5	368	281	87
RCA23_c11790	1279810	1278509	glutamine synthase GlnA type I	COG0174	E	GI 5	3	7	-4
RCA23_c11800	1280652	1279807	putative HTH-type transcriptional regulator, RpiR family	COG1737	K	GI 5	11	5	6
RCA23_c11810	1280733	1281500	N-formylglutamate amidohydrolase	COG3931	E	GI 5	9	7	2
RCA23_c11820	1281497	1282138	putative isochorismatase family protein	COG1335	Q	GI 5	5	11	-6

RCA23_c11830	1282131	1282853	branched-chain amino acid ABC transporter, ATP-binding p	COG0411	E	GI 5	7	1	6
RCA23_c11840	1282882	1284033	branched-chain amino acid ABC transporter, periplasmic bi	COG0683	E	GI 5	10	5	5
RCA23_c11850	1284085	1284783	branched-chain amino acid ABC transporter, ATP-binding p	COG0410	E	GI 5	4	1	3
RCA23_c11860	1284783	1285679	branched-chain amino acid ABC transporter, permease pro	COG0559	E	GI 5	7	4	3
RCA23_c11870	1285754	1286698	branched-chain amino acid ABC transporter, permease pro	COG4177	E	GI 5	15	7	8
RCA23_c11880	1286708	1287646	acetamidase/formamidase family protein	COG2421	C	GI 5	1	6	-5
RCA23_c11890	1288192	1288902	hypothetical protein	COG1028	I	GI 5	468	321	147
RCA23_c11910	1289550	1290728	putative prophage integrase	COG0582	L	GI 5	13	8	5
RCA23_c11920	1290725	1291072	hypothetical protein			GI 5	16	9	7
RCA23_c11930	1291427	1292068	hypothetical protein			GI 5	79	49	30
RCA23_c11940	1292185	1293384	putative AAA ATPase	COG3598	L	GI 5	82	62	20
RCA23_c11950	1294040	1293531	hypothetical protein			GI 5	4	3	1
RCA23_c11960	1294894	1294163	hypothetical protein			GI 5	16	4	12
RCA23_c11970	1295139	1295579	hypothetical protein			GI 5	76	54	22
RCA23_c11980	1295579	1295947	hypothetical protein			GI 5	33	14	19
RCA23_c11990	1296908	1296210	branched-chain amino acid ABC transporter, ATP-binding p	COG0410	E	GI 5	45	35	10
RCA23_c12000	1297654	1296914	branched-chain amino acid ABC transporter, ATP-binding p	COG0411	E	GI 5	45	36	9
RCA23_c12010	1298634	1297651	branched-chain amino acid ABC transporter, permease pro	COG4177	E	GI 5	76	71	5
RCA23_c12020	1299559	1298666	branched-chain amino acid ABC transporter, permease pro	COG0559	E	GI 5	66	89	-23
RCA23_c12030	1300733	1299567	branched-chain amino acid ABC transporter, periplasmic bi	COG0683	E	GI 5	53	48	5
RCA23_c12040	1301568	1300768	enoyl-CoA hydratase/carnithine racemase	COG1024	I	GI 5	36	56	-20
RCA23_c12050	1303130	1301586	acetyl-CoA synthase-like protein	COG0318	I	GI 5	55	58	-3
RCA23_c12060	1303632	1303123	dehydrogenase iron-sulfur-binding subunit	COG2080	C	GI 5	49	59	-10
RCA23_c12070	1304444	1303620	dehydrogenase FAD-binding subunit	COG1319	C	GI 5	26	41	-15
RCA23_c12080	1307380	1304441	dehydrogenase molybdenum-binding subunit	COG1529	C	GI 5	41	48	-7
RCA23_c12090	1308419	1307487	putative HTH-type transcriptional regulator, AraC family	COG2207	K	GI 5	71	67	4
RCA23_c12100	1310095	1309085	periplasmic binding protein-like	COG1879	G	GI 5	18	16	2
RCA23_c12110	1311473	1310106	ROK family transcriptional repressor	COG1940	K	GI 5	9	11	-2
RCA23_c12120	1311642	1312454	xylose isomerase	COG1082	G	GI 5	8	13	-5
RCA23_c12130	1312504	1313622	sugar ABC transporter, periplasmic binding protein	COG1879	G	GI 5	11	9	2

RCA23_c12140	1313694	1314857	putative oxidoreductase	COG0673	R	GI 5	11	16	-5
RCA23_c12150	1314857	1316431	sugar ABC transporter, ATP-binding protein	COG1129	G	GI 5	11	11	0
RCA23_c12160	1316418	1317386	sugar ABC transporter, permease protein	COG1172	G	GI 5	12	24	-12
RCA23_c12170	1317389	1317691	hypothetical protein, monoxygenase-like	COG1359	S	GI 5	18	23	-5
RCA23_c12180	1317805	1318863	hypothetical protein	COG1082	G	GI 5	30	18	12
RCA23_c12190	1318867	1320018	putative oxidoreductase	COG0673	R	GI 5	29	25	4
RCA23_c12200	1320689	1321837	peptidase M20D, amidohydrolase	COG1473	R	GI 5	229	226	3
RCA23_c12210	1321840	1322985	X-Pro dipeptidase	COG0006	E	GI 5	289	228	61
RCA23_c12220	1324296	1323853	hypothetical protein			GI 5	18	9	9
RCA23_c12230	1325069	1324293	hypothetical protein, transmembrane			GI 5	12	12	0
RCA23_c12240	1325832	1327037	hypothetical protein, DNA breaking rejoining enzymes family protein-like			GI 5	37	26	11
RCA23_c12250	1327137	1328645	hypothetical protein			GI 5	1	0	1
RCA23_c12260	1328729	1329205	hypothetical protein			GI 5	10	5	5
RCA23_c12270	1329553	1329744	hypothetical protein			GI 5	0	13	-13
RCA23_c12280	1330088	1331644	type I restriction-modification system, M subunit	COG0286	V	GI 5	10	4	6
RCA23_c12290	1331637	1332842	type I restriction-modification system, S subunit	COG0732	V	GI 5	8	2	6
RCA23_c12300	1332842	1335937	type I restriction-modification system, R subunit	COG0610	V	GI 5	0	4	-4
RCA23_c12310	1336026	1337000	hypothetical protein	COG3012	S	GI 5	0	0	0
RCA23_c12320	1337142	1338152	hypothetical protein	COG2865	K	GI 5	0	0	0
RCA23_c12330	1338762	1338241	hypothetical protein			GI 5	0	0	0
RCA23_c12340	1340773	1338905	hypothetical protein, DNA helicase-like protein			GI 5	2	0	2
RCA23_c12350	1340970	1341416	hypothetical protein			GI 5	0	0	0
RCA23_c12360	1341557	1342156	integrase	COG0582	L	GI 5	0	1	-1
RCA23_c12370	1342597	1342181	hypothetical protein			GI 5	0	0	0
RCA23_c12380	1343211	1342702	hypothetical protein			GI 5	0	0	0
RCA23_c12390	1343382	1344026	hypothetical protein			GI 5	9	2	7
RCA23_c12400	1344489	1344037	hypothetical protein			GI 5	2	1	1
RCA23_c12410	1346190	1345639	hypothetical protein, DUF1994			GI 5	386	183	203
RCA23_c12420	1347575	1346172	exonuclease I	COG2925	L	GI 5	154	111	43
RCA23_c12430	1349548	1347575	transpeptidase, penicillin binding protein	COG4953	M	GI 5	127	64	63

RCA23_c12440	1354560	1349545	hypothetical protein	COG2373	R	GI 5	122	64	58
RCA23_c12450	1355067	1354687	hypothetical protein			GI 5	63	63	0
RCA23_c12460	1355322	1355774	hypothetical protein			GI 5	12	9	3
RCA23_c12470	1357492	1355828	type III restriction enzyme, res subunit	COG1061	K	GI 5	36	25	11
RCA23_c12480	1358034	1357498	putative type III restriction system protein, mod subunit			GI 5	25	14	11
RCA23_c12490	1358097	1359374	putative serine/threonine protein kinase	COG0515	R	GI 5	10	16	-6
RCA23_c12500	1359362	1360081	serine/threonine protein phosphatase PrpC	COG0631	T	GI 5	12	14	-2
RCA23_c12510	1362272	1360134	DNA helicase, UvrD/REP type	COG0210	L	GI 5	6	6	0
RCA23_c12520	1362274	1364124	hypothetical protein	COG1463	Q	GI 5	4	4	0
RCA23_c12530	1364121	1365416	hypothetical protein, OmpA/MotB-like	COG1360	N	GI 5	1	2	-1
RCA23_c12540	1365416	1367443	hypothetical protein			GI 5	2	1	1
RCA23_c12550	1367457	1370579	ATP-dependent helicase	COG0553	K	GI 5	5	5	0
RCA23_c12560	1371410	1370973	hypothetical protein			GI 5	7	3	4
RCA23_c12570	1372098	1371616	hypothetical protein			GI 5	11	7	4
RCA23_c12580	1372594	1372277	hypothetical protein			GI 5	4	4	0
RCA23_c12590	1372881	1373252	hypothetical protein			GI 5	14	2	12
RCA23_c12600	1373306	1374205	hypothetical protein			GI 5	9	5	4
RCA23_c12610	1374516	1374866	hypothetical protein			GI 5	0	0	0
RCA23_c12620	1374971	1375435	hypothetical protein			GI 5	0	0	0
RCA23_c12630	1375674	1376117	integrase	COG2801	L	GI 5	376	216	160
RCA23_c12640	1377304	1376135	hypothetical protein			GI 5	2	0	2
RCA23_c12650	1377893	1377396	hypothetical protein			GI 5	0	0	0
RCA23_c12660	1378200	1377955	DNA integration/recombination/invertion protein			GI 5	0	21	-21
RCA23_c12690	1379769	1378858	putative beta-lactamase-like protein	COG0491	R		250	305	-55
RCA23_c12680	1379768	1381060	osmolarity sensor protein EnvZ	COG0642	T		253	396	-143
RCA23_c12700	1381130	1381954	hypothetical protein				591	515	76
RCA23_c12710	1382060	1382731	50S ribosomal protein L21	COG0261	J		514	314	200
RCA23_c12720	1382738	1383007	50S ribosomal protein L27	COG0211	J		403	469	-66
RCA23_c12730	1383092	1383730	hypothetical protein, LysE type translocator	COG1280	E		245	351	-106
RCA23_c12740	1383727	1384263	putative acetyltransferase	COG1670	J		321	372	-51

RCA23_c12750	1384267	1384815	GCN5-like N-acetyltransferase	COG1670	J		202	299	-97
RCA23_c12760	1384812	1385840	GTP-binding protein Obg	COG0536	R		389	466	-77
RCA23_c12770	1385828	1386934	glutamate 5-kinase ProB	COG0263	E		322	395	-73
RCA23_c12780	1386950	1388206	gamma-glutamyl phosphate reductase ProA	COG0014	E		246	337	-91
RCA23_c12800	1388851	1388243	hypothetical protein	COG5385	S		346	369	-23
RCA23_c12790	1388850	1389026	hypothetical protein				186	377	-191
RCA23_c12810	1389099	1389872	hypothetical protein	COG3176	R		184	338	-154
RCA23_c12820	1389890	1390708	putative phosphate acyltransferase	COG0204	I		235	353	-118
RCA23_c12830	1391398	1390754	thiamine-phosphate pyrophosphorylase ThiE	COG0352	H		312	391	-79
RCA23_c12840	1391480	1392226	tRNA (cytidine/uridine-2'-O-)-methyltransferase TrmJ	COG0565	J		302	372	-70
RCA23_c12850	1392274	1393416	heme A synthase CtaA	COG1612	O		363	447	-84
RCA23_c12860	1393413	1394894	thermostable carboxypeptidase 1	COG2317	E		450	426	24
RCA23_c12870	1398094	1395344	DNA gyrase subunit A	COG0188	L		289	359	-70
RCA23_c12880	1399028	1398543	hypothetical protein	COG1495	O		189	308	-119
RCA23_c12890	1399603	1399025	DedA family protein	COG1238	S		295	288	7
RCA23_c12900	1400438	1399839	hypothetical protein			GI 6	458	361	97
RCA23_c12910	1401810	1401337	putative snoaL-like polyketide cyclase	COG3631	R	GI 6	222	195	27
RCA23_c12920	1402164	1402454	hypothetical protein, DUF1330	COG5470	S	GI 6	264	216	48
RCA23_c12930	1402883	1403188	hypothetical protein, DUF3303			GI 6	110	73	37
RCA23_c12940	1403986	1403495	hypothetical protein, transmembrane			GI 6	27	27	0
RCA23_c12950	1404278	1403991	hypothetical protein, transmembrane			GI 6	11	14	-3
RCA23_c12960	1404422	1405501	integrase	COG2801	L	GI 6	0	0	0
RCA23_c12970	1406068	1406265	hypothetical protein			GI 6	0	0	0
RCA23_c12980	1406863	1406288	site-specific recombinase, resolvase family protein	COG1961	L	GI 6	149	119	30
RCA23_c12990	1407072	1407455	hypothetical protein	COG4731	S	GI 6	437	256	181
RCA23_c13000	1407781	1407981	hypothetical protein			GI 6	17	6	11
RCA23_c13010	1409183	1408122	transposase	COG3415	L	GI 6	0	0	0
RCA23_c13020	1410237	1410437	ribsomal protein S21	COG0828	J	GI 6	301	250	51
RCA23_c13030	1410572	1410739	hypothetical protein			GI 6	484	243	241
RCA23_c13040	1410794	1411873	integrase	COG2801	L	GI 6	0	0	0

RCA23_c13050	1413278	1411980	hypothetical protein			GI 6	524	297	227
RCA23_c13060	1414337	1413294	putative phage integrase	COG0582	L	GI 6	476	307	169
RCA23_c13080	1414822	1415982	putative phage integrase	COG0582	L	GI 6	243	165	78
RCA23_c13090	1417647	1417237	hypothetical protein, GYD domain	COG4274	S	GI 6	33	42	-9
RCA23_c13100	1417942	1418304	hypothetical protein			GI 6	4	0	4
RCA23_c13110	1418856	1418572	hypothetical protein			GI 6	224	134	90
RCA23_c13120	1419178	1420584	hypothetical protein			GI 6	92	69	23
RCA23_c13130	1420645	1421724	integrase	COG2801	L	GI 6	0	0	0
RCA23_c13140	1422270	1422037	hypothetical protein	COG5586	S	GI 6	138	77	61
RCA23_c13150	1422868	1422398	hypothetical protein			GI 6	83	93	-10
RCA23_c13160	1423763	1423269	hypothetical protein			GI 6	217	204	13
RCA23_c13170	1424625	1424954	hypothetical protein	COG4274	S	GI 6	458	401	57
RCA23_c13180	1425131	1425541	hypothetical protein			GI 6	484	312	172
RCA23_c13190	1425553	1425708	hypothetical protein			GI 6	314	205	109
RCA23_c13210	1425849	1426889	phosphoribosylformylglycinamidine cyclo-ligase PurM	COG0150	F		228	287	-59
RCA23_c13220	1426886	1427464	phosphoribosylglycinamide formyltransferase PurN	COG0299	F		356	390	-34
RCA23_c13230	1427543	1428700	ribonuclease D	COG0349	J		533	523	10
RCA23_c13240	1428736	1429146	SufE-like protein	COG2166	R		396	545	-149
RCA23_c13250	1429745	1429143	hypothetical protein				328	397	-69
RCA23_c13260	1430578	1429871	putative methionine synthase (B12 dependent) subunit 2	COG5012	R		318	404	-86
RCA23_c13270	1431745	1430717	putative methionine synthase (B12 dependent) subunit 1	COG0646	E		255	293	-38
RCA23_c13280	1431913	1432788	phosphoribosylaminoimidazole-succinocarboxamide syntha	COG0152	F		544	497	47
RCA23_c13290	1432805	1433035	phosphoribosylformylglycinamidine (FGAM) synthase PurS	COG1828	F		429	477	-48
RCA23_c13300	1433035	1433703	phosphoribosylformylglycinamidine synthase PurQ	COG0047	F		256	362	-106
RCA23_c13310	1433749	1435512	C4-dicarboxylate transport sensor protein DctB	COG4191	T		424	436	-12
RCA23_c13320	1435515	1436849	C4-dicarboxylate transport transcriptional regulatory protein	COG2204	T		309	329	-20
RCA23_c13330	1437325	1440108	ribonuclease E	COG1530	J		255	309	-54
RCA23_c13340	1440411	1440178	putative sulfurtransferase tusA	COG0425	O		247	389	-142
RCA23_c13350	1440615	1441367	cytochrome C biogenesis protein transmembrane region	COG0785	O		246	383	-137
RCA23_c13360	1441476	1442669	hypothetical protein				167	250	-83

RCA23_c13370	1442666	1444039	putative cytochrome P450	COG2124	Q	149	243	-94
RCA23_c13380	1444734	1444045	DNA alkylation repair enzyme	COG4912	L	200	224	-24
RCA23_c13390	1445375	1444875	hypothetical protein	COG3153	R	535	551	-16
RCA23_c13400	1447484	1445661	glucosamine--fructose-6-phosphate aminotransferase GlmS	COG0449	M	266	262	4
RCA23_c13410	1448847	1447489	bifunctional protein GlmU	COG1207	M	241	239	2
RCA23_c13420	1448966	1449634	putative HAD-superfamily hydrolase	COG0546	R	158	222	-64
RCA23_c13440	1450935	1449619	MmgE/PrpD family protein	COG2079	R	213	307	-94
RCA23_c13430	1450926	1452245	putative pyridoxal-phosphate-dependent aminotransferase	COG0399	M	302	360	-58
RCA23_c13450	1452356	1453519	isovaleryl-CoA dehydrogenase	COG1960	I	393	399	-6
RCA23_c13460	1453516	1454001	hypothetical protein			340	392	-52
RCA23_c13470	1454014	1455594	methylcrotonoyl-CoA carboxylase beta subunit MccB	COG4799	I	176	239	-63
RCA23_c13480	1455563	1456069	hypothetical protein			283	455	-172
RCA23_c13490	1456066	1457997	methylcrotonoyl-CoA carboxylase alpha subunit MccA	COG4770	I	216	303	-87
RCA23_c13500	1457994	1458887	hydroxymethylglutaryl-CoA lyase MvaB	COG0119	E	169	249	-80
RCA23_c13510	1458877	1459662	putative methylglutaconyl-CoA hydratase	COG1024	I	202	268	-66
RCA23_c13520	1459788	1460153	NADH-quinone oxidoreductase subunit A	COG0838	C	661	586	75
RCA23_c13530	1460153	1460677	NADH-quinone oxidoreductase subunit NuoB	COG0377	C	486	555	-69
RCA23_c13540	1460682	1460888	hypothetical protein DUF2158	COG5475	S	652	610	42
RCA23_c13550	1460885	1461496	NADH-quinone oxidoreductase subunit C	COG0852	C	607	461	146
RCA23_c13560	1461557	1462396	hypothetical protein			533	476	57
RCA23_c13570	1462446	1463657	NADH-quinone oxidoreductase subunit D	COG0649	C	551	526	25
RCA23_c13580	1463723	1464529	NADH-quinone oxidoreductase subunit E	COG1905	C	409	436	-27
RCA23_c13590	1464577	1464804	hypothetical protein			413	373	40
RCA23_c13600	1464810	1466105	NADH-quinone oxidoreductase subunit F	COG1894	C	469	416	53
RCA23_c13610	1466139	1466543	hypothetical protein			322	392	-70
RCA23_c13620	1466546	1468564	NADH-quinone oxidoreductase subunit G	COG1034	C	305	351	-46
RCA23_c13630	1468569	1468826	hypothetical protein			466	384	82
RCA23_c13640	1468823	1469860	NADH-quinone oxidoreductase subunit H	COG1005	C	757	586	171
RCA23_c13650	1469865	1470359	NADH-quinone oxidoreductase subunit I	COG1143	C	602	413	189
RCA23_c13660	1470382	1470987	NADH-quinone oxidoreductase subunit J	COG0839	C	636	482	154

RCA23_c13670	1470984	1471289	NADH-quinone oxidoreductase subunit K	COG0713	C	821	637	184
RCA23_c13680	1471296	1473413	NADH-quinone oxidoreductase subunit L	COG1009	C	437	408	29
RCA23_c13690	1473413	1474963	NADH-quinone oxidoreductase subunit M	COG1008	C	532	545	-13
RCA23_c13700	1474980	1476422	NADH-quinone oxidoreductase subunit N	COG1007	C	401	459	-58
RCA23_c13710	1476424	1477155	biotin-[acetyl-CoA-carboxylase] ligase	COG0340	H	229	269	-40
RCA23_c13720	1477167	1477937	type III pantothenate kinase CoaX	COG1521	K	366	445	-79
RCA23_c13730	1477937	1479601	putative ribonuclease	COG0595	R	374	437	-63
RCA23_c13740	1481465	1479801	ATP-dependent RNA helicase RhlE	COG0513	L	354	399	-45
RCA23_c13750	1483232	1481634	peptide chain release factor 3	COG4108	J	313	275	38
RCA23_c13760	1484260	1483316	hypothetical protein			571	437	134
RCA23_c13770	1484470	1485249	putative short chain dehydrogenase	COG1028	I	262	331	-69
RCA23_c13780	1485369	1486409	arsenite methyltransferase	COG2226	H	400	459	-59
RCA23_c13790	1486465	1487541	Fe(3+) ions import ATP-binding protein FbpC	COG3842	E	250	357	-107
RCA23_c13800	1488367	1487555	putative helix-turn-helix protein	COG2378	K	292	291	1
RCA23_c13810	1488388	1488582	sec-independent protein translocase protein TatA	COG1826	U	505	457	48
RCA23_c13820	1488618	1489190	sec-independent protein translocase protein TatB	COG1826	U	330	343	-13
RCA23_c13830	1489187	1490077	sec-independent protein translocase protein TatC	COG0805	U	517	496	21
RCA23_c13840	1490074	1490913	hypothetical protein DUF815	COG2607	R	258	376	-118
RCA23_c13850	1492065	1490932	putative peptidoglycan-binding peptidase	COG0739	M	224	279	-55
RCA23_c13860	1492751	1492104	protein-L-isoaspartate O-methyltransferase Pcm	COG2518	O	271	340	-69
RCA23_c13870	1493504	1492755	5'-nucleotidase SurE	COG0496	R	249	361	-112
RCA23_c13880	1494293	1493628	putative short chain dehydrogenase	COG1028	I	171	181	-10
RCA23_c13890	1495797	1494328	amidophosphoribosyltransferase PurF	COG0034	F	308	326	-18
RCA23_c13900	1496469	1495963	putative colicin V production protein			695	461	234
RCA23_c13910	1497899	1496541	DNA repair protein RadA	COG1066	O	241	276	-35
RCA23_c13920	1498723	1497977	ABC transporter ATP-binding protein	COG1127	Q	240	329	-89
RCA23_c13930	1499508	1498720	hypothetical protein DUF140	COG0767	Q	233	367	-134
RCA23_c13940	1500548	1499505	alanine racemase, biosynthetic	COG0787	M	293	329	-36
RCA23_c13950	1502120	1500630	replicative DNA helicase DnaB	COG0305	L	362	394	-32
RCA23_c13960	1503031	1502354	orotate phosphoribosyltransferase PyrE	COG0461	F	442	359	83

RCA23_c13970	1504098	1503058	dihydroorotase PyrC	COG0418	F		202	251	-49
RCA23_c13980	1504207	1504548	hypothetical protein				487	337	150
RCA23_c13990	1504659	1504880	hypothetical protein				1.964	903	1.061
RCA23_c14000	1505732	1504914	hypothetical protein				387	351	36
RCA23_c14010	1507107	1506076	malate dehydrogenase Mdh	COG0039	C		211	210	1
RCA23_c14020	1507213	1508112	citrate lyase beta subunit CitE	COG2301	G		312	379	-67
RCA23_c14030	1508312	1509337	putative mesaconyl-CoA hydratase	COG2030	I		265	408	-143
RCA23_c14040	1509578	1509898	succinate dehydrogenase cytochrome b556 subunit SdhC	COG2009	C		557	392	165
RCA23_c14050	1509911	1510282	succinate dehydrogenase hydrophobic membrane anchor s	COG2142	C		433	353	80
RCA23_c14060	1510301	1512106	succinate dehydrogenase flavoprotein subunit SdhA	COG1053	C		278	316	-38
RCA23_c14070	1512106	1512408	hypothetical protein				327	285	42
RCA23_c14080	1512506	1513285	succinate dehydrogenase iron-sulfur subunit SdhB	COG0479	C		697	511	186
RCA23_c14090	1514072	1513368	purine nucleoside phosphorylase deoD-type	COG0813	F		306	402	-96
RCA23_c14100	1514517	1514077	acetyltransferase	COG1246	E		267	373	-106
RCA23_c14110	1515360	1514536	tryptophan synthase alpha chain TrpA	COG0159	E		426	401	25
RCA23_c14120	1515506	1516603	GTP-dependent nucleic acid-binding protein EngD	COG0012	J		325	304	21
RCA23_c14130	1516647	1517570	Non-specific ribonucleoside hydrolase rihC	COG1957	F		193	264	-71
RCA23_c14150	1518098	1518574	hypothetical protein DUF583	COG1664	M	GI 7	165	88	77
RCA23_c14160	1519456	1518554	protease HtpX	COG0501	O	GI 7	279	171	108
RCA23_c14170	1521395	1520490	histone deacetylase	COG0123	B	GI 7	75	50	25
RCA23_c14180	1521466	1521981	peripheral-type benzodiazepine receptor/signal transductior	COG3476	T	GI 7	102	83	19
RCA23_c14190	1522156	1522710	DNA-invertase Hin	COG1961	L	GI 7	122	64	58
RCA23_c14200	1523534	1524067	hypothetical protein			GI 7	24	29	-5
RCA23_c14210	1524259	1524588	hypothetical protein			GI 7	272	263	9
RCA23_c14220	1526125	1525748	hypothetical protein, RmlC-like cupin family	COG3450	R	GI 7	126	63	63
RCA23_c14230	1527629	1526337	integrase	COG4974	L	GI 7	158	126	32
RCA23_c14240	1528430	1529155	hypothetical membrane protein			GI 7	498	284	214
RCA23_c14260	1530033	1529710	hypothetical protein	COG4530	S		830	585	245
RCA23_c14270	1530131	1530841	hypothetical protein DUF45	COG1451	R		190	223	-33
RCA23_c14280	1530914	1531564	HTH-type transcriptional regulator, GntR family	COG1802	K		187	228	-41

RCA23_c14290	1533239	1531851	dihydrolipoyl dehydrogenase Lpd	COG1249	C	466	368	98
RCA23_c14300	1534890	1533394	dihydrolipoyllysine-residue succinyltransferase component (	COG0508	C	303	288	15
RCA23_c14310	1537852	1534895	2-oxoglutarate dehydrogenase E1 component SucA	COG0567	C	473	404	69
RCA23_c14320	1538893	1538009	succinyl-CoA ligase [ADP-forming] alpha subunit SucD	COG0074	C	358	285	73
RCA23_c14330	1540169	1538973	succinyl-CoA ligase [ADP-forming] beta subunit SucC	COG0045	C	690	360	330
RCA23_c14340	1541555	1540356	butyryl-CoA dehydrogenase	COG1960	I	341	300	41
RCA23_c14350	1542165	1541704	translation initiation factor IF-3	COG0290	J	630	362	268
RCA23_c14360	1543174	1542341	ferredoxin--NADP reductase Fpr	COG1018	C	336	418	-82
RCA23_c14370	1543641	1543234	hypothetical protein	COG3749	S	253	220	33
RCA23_c14380	1543877	1543641	cysH'	COG0175	E	232	182	50
RCA23_c14390	1545545	1543881	cysI/sir: sulfite reductase (ferredoxin)	COG0155	P	294	400	-106
RCA23_c14400	1545843	1545547	hypothetical protein			228	310	-82
RCA23_c14410	1547228	1545834	siroheme synthase CysG	COG0007	H	185	248	-63
RCA23_c14420	1547360	1547827	HTH-type transcriptional regulator, AsnC family	COG1522	K	378	390	-12
RCA23_c14450	1549602	1548220	peptidase	COG0624	E	247	316	-69
RCA23_c14460	1549872	1551119	hypothetical protein	COG1426	S	263	283	-20
RCA23_c14470	1551156	1552289	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase IspG	COG0821	I	231	303	-72
RCA23_c14480	1553599	1552286	TPR-repeat containing protein	COG4783	R	216	239	-23
RCA23_c14490	1553648	1554796	aspartate aminotransferase	COG0436	E	243	353	-110
RCA23_c14500	1554989	1557526	penicillin-binding protein 1A	COG5009	M	262	302	-40
RCA23_c14510	1557611	1558735	peptide chain release factor 2	COG1186	J	431	474	-43
RCA23_c14520	1559286	1558798	hypothetical protein DUF583	COG1664	M	284	384	-100
RCA23_c14530	1560550	1559276	putative peptidase, M23 family	COG0739	M	336	341	-5
RCA23_c14540	1561570	1560740	hypothetical protein DUF455	COG2833	S	240	368	-128
RCA23_c14550	1562037	1561567	peroxiredoxin Bcp	COG1225	O	290	273	17
RCA23_c14560	1562118	1565399	hypothetical protein			277	331	-54
RCA23_c14570	1565470	1566519	S-adenosylmethionine:tRNA ribosyltransferase-isomerase (	COG0809	J	222	302	-80
RCA23_c14580	1566591	1567847	MFS-type transporter			516	456	60
RCA23_c14590	1567922	1568500	hypothetical protein DUF924	COG3803	S	230	272	-42
RCA23_c14600	1568608	1570005	dihydrolipoyl dehydrogenase Lpd	COG1249	C	446	363	83

RCA23_c14610	1570379	1573255	uvrABC system protein A	COG0178	L	350	381	-31
RCA23_c14620	1574577	1573252	MmgE/PrpD family protein	COG2079	R	259	359	-100
RCA23_c14630	1575459	1574587	3-hydroxyisobutyrate dehydrogenase MmsB	COG2084	I	182	158	24
RCA23_c14640	1576498	1575461	3-hydroxyisobutyryl-CoA hydrolase	COG1024	I	249	295	-46
RCA23_c14650	1577670	1576495	isobutyryl-CoA dehydrogenase	COG1960	I	231	285	-54
RCA23_c14660	1579264	1577765	methylmalonate-semialdehyde dehydrogenase MmsA	COG1012	C	289	334	-45
RCA23_c14670	1579359	1580273	putative HTH-type transcriptional regulator, LysR family	COG0583	K	195	319	-124
RCA23_c14680	1580359	1580856	phosphopantetheine adenylyltransferase CoaD	COG0669	H	397	393	4
RCA23_c14690	1581884	1580877	glyceraldehyde-3-phosphate dehydrogenase Gap	COG0057	G	443	402	41
RCA23_c14700	1583056	1582103	glyceraldehyde-3-phosphate dehydrogenase Gap	COG0057	G	317	344	-27
RCA23_c14710	1585311	1583296	transketolase TktA	COG0021	G	257	307	-50
RCA23_c14720	1585511	1585894	hypothetical protein			236	322	-86
RCA23_c14730	1585887	1586291	cell division protein ZapA	COG3027	S	366	260	106
RCA23_c14740	1586719	1586354	putative glutaredoxin	COG0278	O	643	378	265
RCA23_c14750	1586969	1586709	hypothetical protein			556	382	174
RCA23_c14760	1587213	1586971	hypothetical protein, BolA-like	COG0271	T	488	401	87
RCA23_c14770	1589513	1587267	phosphoribosylformylglycinamidine synthase PurL	COG0046	F	254	340	-86
RCA23_c14780	1590634	1589726	HTH-type transcriptional regulator, LysR family	COG0583	K	322	383	-61
RCA23_c14790	1590763	1594146	hypothetical protein, pyruvate ferredoxin/flavodoxin oxidore	COG4231	C	194	266	-72
RCA23_c14800	1594236	1595048	glutamate racemase MurI	COG0796	M	356	334	22
RCA23_c14810	1595120	1596148	N-acetyl-gamma-glutamyl-phosphate reductase ArgC	COG0002	E	564	441	123
RCA23_c14820	1596148	1596600	cytochrome c-type biogenesis protein CcmE	COG2332	O	291	279	12
RCA23_c14830	1596707	1598668	cytochrome c-type biogenesis protein CcmF	COG1138	O	404	410	-6
RCA23_c14840	1598722	1599117	cytochrome c-type biogenesis protein CcmH	COG3088	O	315	462	-147
RCA23_c14850	1599159	1599935	putative enoyl-CoA hydratase FadB	COG1024	I	285	353	-68
RCA23_c14860	1600258	1600728	hypothetical protein			544	317	227
RCA23_c14870	1602134	1600809	citrate synthase GltA	COG0372	C	642	370	272
RCA23_c14880	1603561	1602152	glutamyl-tRNA synthase 2	COG0008	J	390	320	70
RCA23_c14890	1603619	1605670	hypothetical protein competence protein E	COG0658	R	263	389	-126
RCA23_c14900	1606374	1605667	LexA repressor	COG1974	K	337	368	-31

RCA23_c14910	1607624	1606449	molybdopterin biosynthesis protein MoeA	COG0303	H	152	180	-28
RCA23_c14920	1608103	1607621	molybdenum cofactor biosynthesis protein MoaC	COG0315	H	261	210	51
RCA23_c14930	1608916	1608104	indole-3-glycerol phosphate synthase TrpC	COG0134	E	180	257	-77
RCA23_c14940	1609952	1608933	anthranilate phosphoribosyltransferase TrpD	COG0547	E	184	199	-15
RCA23_c14950	1610530	1609949	anthranilate synthase component TrpG	COG0512	E	229	330	-101
RCA23_c14960	1610679	1611788	hypothetical protein, divergent polysaccharide deacetylase			278	342	-64
RCA23_c14970	1612782	1611811	hypothetical protein			318	371	-53
RCA23_c14990	1613796	1613200	hypothetical protein	COG3108	S	294	342	-48
RCA23_c15000	1614095	1615693	putative L,D-transpeptidase YcbB	COG2989	S	337	409	-72
RCA23_c15010	1615723	1616820	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransfera	COG1044	M	307	313	-6
RCA23_c15020	1616840	1617097	acyl carrier protein	COG0236	I	752	635	117
RCA23_c15030	1617101	1618309	3-oxoacyl-[acyl-carrier-protein] synthase FabF	COG0304	I	360	368	-8
RCA23_c15040	1618913	1618323	hypothetical protein, invasion protein B	COG5342	R	470	365	105
RCA23_c15050	1620585	1619287	inner membrane protein	COG4536	P	535	395	140
RCA23_c15060	1621505	1620585	tyrosine recombinase XerD	COG4974	L	356	425	-69
RCA23_c15070	1622956	1621502	hypothetical protein			267	381	-114
RCA23_c15080	1623167	1622958	hypothetical protein			401	504	-103
RCA23_c15090	1623212	1623763	shikimate kinase AroK	COG0703	E	556	623	-67
RCA23_c15100	1623760	1624878	3-dehydroquinate synthase AroB	COG0337	E	227	392	-165
RCA23_c15110	1625462	1624944	single-stranded DNA-binding protein	COG0629	L	285	427	-142
RCA23_c15120	1627772	1625688	glutathione import ATP-binding protein GsiA	COG0444	E	341	366	-25
RCA23_c15130	1629378	1627777	putative ABC transporter inner membrane component	COG1173	E	271	317	-46
RCA23_c15140	1630432	1629392	putative ABC transporter permease protein	COG0601	E	407	407	0
RCA23_c15150	1632178	1630532	ABC transporter extracellular solute-binding protein	COG0747	E	250	204	46
RCA23_c15160	1633423	1632533	tRNA delta(2)-isopentenylpyrophosphate transferase MiaA	COG0324	J	194	325	-131
RCA23_c15170	1633522	1634253	uridylate kinase PyrH	COG0528	F	352	335	17
RCA23_c15180	1634275	1634841	ribosome recycling factor Frr	COG0233	J	393	363	30
RCA23_c15190	1634856	1635563	undecaprenyl pyrophosphate synthase UppS	COG0020	I	318	350	-32
RCA23_c15200	1635560	1636345	putative cytidylyltransferase	COG0575	I	183	309	-126
RCA23_c15210	1636345	1637505	1-deoxy-D-xylulose 5-phosphate reductoisomerase Dxr	COG0743	I	155	311	-156

RCA23_c15220	1637498	1638829	RIP metalloprotease RseP	COG0750	M	293	380	-87
RCA23_c15230	1638949	1641297	putative outer membrane assembly factor	COG4775	M	444	390	54
RCA23_c15240	1641399	1641863	putative outer tein			412	324	88
RCA23_c15250	1641972	1642433	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase Fal	COG0764	I	326	374	-48
RCA23_c15260	1642430	1643227	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acy	COG1043	M	283	369	-86
RCA23_c15270	1643232	1644035	hypothetical protein DUF1009	COG3494	S	199	315	-116
RCA23_c15280	1644032	1645186	lipid-A-disaccharide synthase LpxB	COG0763	M	225	289	-64
RCA23_c15290	1646325	1645189	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransfera	COG0482	J	270	301	-31
RCA23_c15300	1646437	1646721	hypothetical protein			191	235	-44
RCA23_c15310	1646802	1647515	cell cycle transcriptional regulator	COG0745	T	662	573	89
RCA23_c15320	1647766	1649886	DNA ligase LigA	COG0272	L	331	351	-20
RCA23_c15330	1649883	1651973	ATP-dependent DNA helicase RecG	COG1200	L	233	326	-93
RCA23_c15340	1653286	1652261	hypothetical protein			341	418	-77
RCA23_c15350	1653842	1653390	hypothetical protein	COG0822	C	206	275	-69
RCA23_c15360	1653910	1654266	phosphoribosyl-AMP cyclohydrolase HisI	COG0139	E	288	223	65
RCA23_c15370	1655153	1654284	glutamyl-Q tRNA(Asp) synthase GluQ	COG0008	J	175	221	-46
RCA23_c15380	1655731	1655150	hypothetical protein, methyltransferase	COG4976	R	245	288	-43
RCA23_c15390	1657149	1655812	tRNA uridine 5-carboxymethylaminomethyl modification enz	COG1206	J	217	276	-59
RCA23_c15400	1657978	1657190	putative crotonase	COG1024	I	321	314	7
RCA23_c15410	1658033	1658455	putative thioesterase	COG2050	Q	219	259	-40
RCA23_c15420	1658600	1659058	50S ribosomal protein L13	COG0102	J	681	369	312
RCA23_c15430	1659062	1659550	30S ribosomal protein S9	COG0103	J	324	197	127
RCA23_c15440	1659763	1661193	integrase			1	0	1
RCA23_c15460	1663623	1661959	choline dehydrogenase BetA	COG2303	E	197	307	-110
RCA23_c15470	1663885	1664619	ABC transporter, permease protein	COG4662	H	250	357	-107
RCA23_c15480	1664613	1665329	ABC transporter, ATP-binding protein	COG1131	V	208	301	-93
RCA23_c15490	1665326	1666123	ABC transporter, periplasmic substrate-binding protein	COG2998	H	273	269	4
RCA23_c15500	1666130	1666561	hypothetical membrane protein	COG1238	S	257	388	-131
RCA23_c15510	1666576	1667487	hypothetical protein DUF6 family, transmembrane	COG2962	R	263	325	-62
RCA23_c15520	1668078	1667497	hypothetical protein	COG2068	R	284	389	-105

RCA23_c15530	1670501	1668111	dimethylglycine dehydrogenase	COG0404	E	282	264	18
RCA23_c15540	1670705	1673146	dimethylglycine dehydrogenase	COG0404	E	272	341	-69
RCA23_c15550	1674027	1673149	short-chain dehydrogenase/reductase SDR	COG1028	I	334	424	-90
RCA23_c15560	1674074	1674460	hypothetical transmembrane protein	COG3788	R	458	440	18
RCA23_c15570	1675087	1674470	hypothetical protein, tetR regulator	COG1309	K	206	240	-34
RCA23_c15580	1676054	1675137	homoserine O-succinyltransferase MetA	COG1897	E	361	398	-37
RCA23_c15590	1676950	1676060	hypothetical protein			409	301	108
RCA23_c15600	1677863	1677015	putative integral membrane protein	COG2510	S	444	337	107
RCA23_c15610	1678994	1678410	Gcn5-like N-acetyltransferase	COG1670	J	276	171	105
RCA23_c15620	1681717	1680161	GMP synthase GuaA	COG0519	F	282	260	22
RCA23_c15630	1681941	1683413	trimethylamine methyltransferase MttB	COG5598	H	243	240	3
RCA23_c15640	1683982	1684989	hypothetical protein			313	290	23
RCA23_c15650	1685077	1686024	lipoyl synthase LipA	COG0320	H	597	400	197
RCA23_c15660	1686218	1686039	hypothetical protein			488	401	87
RCA23_c15670	1686751	1686215	hypoxanthine phosphoribosyltransferase Hpt	COG0634	F	754	569	185
RCA23_c15680	1686817	1687272	hypothetical protein	COG2867	I	446	375	71
RCA23_c15690	1688506	1687319	ammonium transporter AmtB	COG0004	P	357	340	17
RCA23_c15700	1689126	1688647	putative competence-damaged protein	COG1546	R	144	266	-122
RCA23_c15710	1689584	1689123	phosphatidylglycerophosphatase A	COG1267	I	217	316	-99
RCA23_c15720	1690723	1689581	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase IspF	COG0245	I	197	251	-54
RCA23_c15730	1690934	1691875	tRNA-dihydrouridine synthase B	COG0042	J	294	356	-62
RCA23_c15740	1691872	1692936	histidine kinase, nitrogen regulation protein NtrB	COG3852	T	361	518	-157
RCA23_c15750	1693055	1694299	nitrogen regulation protein NtrC	COG2204	T	515	426	89
RCA23_c15760	1694375	1696633	histidine kinase, nitrogen regulation protein NtrY	COG5000	T	356	352	4
RCA23_c15770	1696630	1698030	nitrogen assimilation regulatory protein NtrX	COG2204	T	504	383	121
RCA23_c15780	1698108	1699484	Trk system potassium uptake protein TrkA	COG0569	P	427	389	38
RCA23_c15790	1699487	1700971	putative Trk system potassium uptake protein trkH	COG0168	P	482	534	-52
RCA23_c15800	1701080	1701313	RNA-binding protein Hfq	COG1923	R	693	491	202
RCA23_c15810	1701319	1702623	GTP-binding protein HflX	COG2262	R	414	423	-9
RCA23_c15820	1705060	1702583	acyl-homoserine lactone acylase QuiP	COG2366	R	244	335	-91

RCA23_c15830	1705687	1705130	hypothetical protein	COG2930	S	247	295	-48
RCA23_c15840	1706785	1705775	delta-aminolevulinic acid dehydratase HemB	COG0113	H	238	353	-115
RCA23_c15850	1706883	1707368	hypothetical protein			400	448	-48
RCA23_c15860	1707395	1710835	transcription-repair-coupling factor Mfd	COG1197	L	287	350	-63
RCA23_c15870	1711473	1710832	putative DSBA-like thioredoxin family protein	COG2761	Q	297	350	-53
RCA23_c15880	1712966	1711470	long-chain-fatty-acid-CoA ligase	COG0318	I	183	288	-105
RCA23_c15890	1713007	1713654	HTH-type transcriptional regulator	COG1396	K	344	355	-11
RCA23_c15900	1714310	1713651	aquaporin AqpZ	COG0580	G	476	472	4
RCA23_c15910	1716201	1714342	putative extracellular solute-binding protein	COG4166	E	453	381	72
RCA23_c15920	1716981	1716376	lysine exporter protein LysE	COG1279	R	534	383	151
RCA23_c15930	1717676	1716978	hypothetical protein DUF502	COG2928	S	372	301	71
RCA23_c15940	1718674	1717751	pseudouridine-5'-phosphate glycosidase PsuG	COG2313	Q	184	280	-96
RCA23_c15950	1719592	1718678	hypothetical protein, pfkB family carbohydrate kinase	COG0524	G	194	242	-48
RCA23_c15960	1719832	1720038	cold shock protein CspA	COG1278	K	361	291	70
RCA23_c15970	1720408	1721112	30S ribosomal protein S2	COG0052	J	423	286	137
RCA23_c15980	1721186	1722061	elongation factor Ts	COG0264	J	443	397	46
RCA23_c15990	1723202	1722168	hypothetical protein			215	203	12
RCA23_c16000	1723712	1723203	hypothetical protein, acetyltransferase-like	COG1670	J	194	134	60
RCA23_c16010	1724127	1723705	hypothetical protein, transcription factor NusA like			194	156	38
RCA23_c16020	1724961	1724428	putative phenylacetic acid degradation protein	COG0663	R	265	270	-5
RCA23_c16030	1725592	1724951	guanylate kinase Gmk	COG0194	F	475	457	18
RCA23_c16040	1726466	1725597	hypothetical protein, YicC-like	COG1561	S	323	347	-24
RCA23_c16050	1726645	1727271	hypothetical protein, DUF1457	COG5388	S	144	183	-39
RCA23_c16060	1728661	1727291	phospho-2-dehydro-3-deoxyheptonate aldolase	COG3200	E	330	309	21
RCA23_c16070	1728824	1729837	putative HTH-type transcriptional regulator, AraC family	COG4977	K	326	340	-14
RCA23_c16080	1729969	1731150	putative amino-acid binding protein	COG0683	E	196	180	16
RCA23_c16090	1731226	1732005	putative branched-chain amino acid transport ATP-binding p	COG0411	E	241	251	-10
RCA23_c16100	1732002	1732715	putative branched-chain amino acid transport ATP-binding p	COG0410	E	163	218	-55
RCA23_c16110	1732719	1733732	putative branched-chain amino acid transport system perme	COG0559	E	255	366	-111
RCA23_c16120	1733729	1735063	putative branched-chain amino acid transport system perme	COG4177	E	234	308	-74

RCA23_c16130	1735123	1736940	GTP-binding protein TypA	COG1217	T	275	334	-59
RCA23_c16140	1737304	1737014	hypothetical protein DUF1330	COG5470	S	210	320	-110
RCA23_c16150	1740028	1737311	alanyl-tRNA synthase AlaS	COG0013	J	300	335	-35
RCA23_c16160	1741231	1740119	protein RecA	COG0468	L	701	482	219
RCA23_c16170	1741383	1741979	hypothetical protein			495	354	141
RCA23_c16180	1743391	1741988	sensor transduction histidine kianse	COG0642	T	488	383	105
RCA23_c16190	1744278	1743409	hypothetical protein			338	362	-24
RCA23_c16200	1745526	1744369	hypothetical protein, NOL1/NOP2/sun family	COG0144	J	284	351	-67
RCA23_c16210	1746815	1745526	inosine-5'-monophosphate dehydrogenase GuaB	COG0516	F	165	267	-102
RCA23_c16220	1747235	1749274	hypothetical protein	COG3894	R	252	383	-131
RCA23_c16240	1750674	1749451	cysteine desulfurase SufS	COG0520	E	271	289	-18
RCA23_c16250	1751249	1750671	hypothetical protein			453	462	-9
RCA23_c16260	1751722	1751246	hypothetical protein			225	326	-101
RCA23_c16270	1752993	1751728	FeS assembly protein SufD	COG0719	O	389	321	68
RCA23_c16280	1753748	1752993	FeS assembly ATPase SufC	COG0396	O	763	557	206
RCA23_c16290	1755321	1753810	FeS assembly protein SufB	COG0719	O	654	401	253
RCA23_c16300	1756376	1755327	putative cysteine desulfurase	COG1104	E	226	294	-68
RCA23_c16310	1756801	1756373	HTH-type transcriptional regulator Rrf2	COG1959	K	288	297	-9
RCA23_c16320	1757012	1757665	hypothetical protein	COG2945	R	375	324	51
RCA23_c16330	1757780	1758907	putative oxidoreductase	COG0673	R	247	281	-34
RCA23_c16340	1758960	1759811	inosose dehydratase IolE	COG1082	G	288	364	-76
RCA23_c16350	1759823	1760977	putative oxidoreductase	COG0673	R	306	350	-44
RCA23_c16360	1762000	1760999	HTH-type transcriptional repressor, LacI family	COG1609	K	336	279	57
RCA23_c16370	1762217	1762014	hypothetical protein			574	457	117
RCA23_c16380	1762996	1762214	sugar (ribose) ABC-transport system ATP binding protein	COG1129	G	573	414	159
RCA23_c16390	1764085	1763012	sugar (ribose) ABC transporter permease protein	COG1172	G	736	471	265
RCA23_c16400	1765194	1764175	sugar (ribose) ABC transporter periplasmic binding protein	COG1879	G	541	324	217
RCA23_c16410	1766724	1765405	beta-glucosidase BglA	COG2723	G	315	289	26
RCA23_c16420	1767779	1766721	maltose/maltodextrin import ATP-binding protein MalK	COG3839	G	335	415	-80
RCA23_c16430	1769330	1767780	putative alpha-glucosidase AglA	COG0366	G	330	387	-57

RCA23_c16440	1770025	1769333	alpha-glucoside transport system permease protein AglG	COG0395	G	485	419	66
RCA23_c16450	1770558	1770022	ABC transporter, membrane spanning protein	COG0395	G	344	327	17
RCA23_c16460	1771823	1770570	alpha-glucoside transport system permease protein AglF	COG1175	G	696	565	131
RCA23_c16470	1773262	1771922	alpha-glucosides-binding periplasmic protein AglE	COG1653	G	753	462	291
RCA23_c16480	1773635	1774675	HTH-type transcriptional regulator	COG1609	K	265	360	-95
RCA23_c16490	1775566	1774682	oxidoreductase	COG4989	R	197	228	-31
RCA23_c16500	1776723	1775566	alkanesulfonate monooxygenase SsuD	COG2141	C	310	287	23
RCA23_c16510	1776787	1777521	putative regulatory DNA binding protein	COG2188	K	311	294	17
RCA23_c16520	1777531	1778616	putative oxidoreductase	COG0673	R	256	312	-56
RCA23_c16530	1780202	1778613	glucose-6-phosphate isomerase Pgi	COG0166	G	219	289	-70
RCA23_c16540	1780889	1780212	6-phosphogluconolactonase Pgl	COG0363	G	228	320	-92
RCA23_c16550	1782340	1780889	glucose-6-phosphate 1-dehydrogenase Zwf	COG0364	G	280	364	-84
RCA23_c16560	1782518	1783450	hypothetical protein, radical SAM	COG0535	R	211	260	-49
RCA23_c16570	1784031	1783447	putative 5-methylcytosine-specific restriction enzyme McrA	COG1403	V	336	360	-24
RCA23_c16580	1784786	1784112	putative phospholipase/carboxylesterase	COG0400	R	320	433	-113
RCA23_c16590	1785406	1784783	HhH-GPD superfamily base excision DNA repair protein	COG0122	L	201	319	-118
RCA23_c16600	1786128	1785397	precorrin-6A reductase CobK	COG2099	H	207	313	-106
RCA23_c16610	1787144	1786125	cobalamin (vitamin B12) biosynthesis CbiDprotein CbiD	COG1903	H	234	296	-62
RCA23_c16620	1787265	1788116	uroporphyrinogen-III C-methyltransferase CobA	COG0007	H	228	319	-91
RCA23_c16630	1788113	1789378	cobyrinic acid A,C-diamide synthase CobB	COG1797	H	219	233	-14
RCA23_c16640	1789397	1790620	putative major facilitator superfamily transporter			279	340	-61
RCA23_c16650	1791369	1790629	hypothetical protein	COG1562	I	196	257	-61
RCA23_c16660	1792730	1791366	cysteinyl-tRNA synthase CysS	COG0215	J	268	283	-15
RCA23_c16670	1792939	1794108	aspartate aminotransferase AspC	COG0436	E	174	265	-91
RCA23_c16680	1794310	1795845	trimethylamine methyltransferase MttB	COG5598	H	206	357	-151
RCA23_c16690	1797683	1795869	ABC transporter ATP-binding protein Uup	COG0488	R	291	374	-83
RCA23_c16700	1798221	1797715	hypothetical protein	COG3176	R	334	381	-47
RCA23_c16710	1798563	1798961	peptide methionine sulfoxide reductase MsrB	COG0229	O	307	366	-59
RCA23_c16720	1799290	1798958	hypothetical protein, lipoprotein	COG2913	J	179	299	-120
RCA23_c16730	1799541	1800038	hypothetical protein	COG1399	R	327	476	-149

RCA23_c16740	1800483	1801568	fatty acid/phospholipid synthesis protein PlsX	COG0416	I	370	356	14
RCA23_c16750	1801571	1802533	3-oxoacyl-[acyl-carrier-protein] synthase FabH	COG0332	I	321	335	-14
RCA23_c16760	1802638	1802916	integration host factor alpha subunit IhfA	COG0776	L	908	657	251
RCA23_c16770	1802922	1803605	hypothetical protein, MerR family regulatory protein	COG0789	K	363	366	-3
RCA23_c16790	1803802	1804875	deoxyguanosinetriphosphate triphosphohydrolase-like prote	COG0717	F	284	273	11
RCA23_c16800	1805746	1805078	putative segregation and condesation protein B	COG1386	K	292	333	-41
RCA23_c16810	1806524	1805739	hypothetical protein, segregation and condensation protein	COG1354	S	264	347	-83
RCA23_c16820	1807519	1806524	beta-hexosaminidase NagZ	COG1472	G	233	290	-57
RCA23_c16830	1808420	1807533	hypothetical protein			203	222	-19
RCA23_c16840	1810236	1808491	arginyl-tRNA synthase	COG0018	J	357	365	-8
RCA23_c16850	1811427	1810285	deoxyguanosinetriphosphate triphosphohydrolase-like prote	COG0232	F	269	313	-44
RCA23_c16860	1811504	1811830	putative iron-sulfur insertion protein erpA	COG0316	S	581	371	210
RCA23_c16870	1811876	1812664	exodeoxyribonuclease III	COG0708	L	439	421	18
RCA23_c16880	1813149	1812685	putative cytochrome B561	COG3038	C	276	347	-71
RCA23_c16890	1813755	1813333	putative dnak suppressor protein dksA	COG1734	T	459	450	9
RCA23_c16900	1813911	1814750	hypothetical protein	COG0714	R	623	506	117
RCA23_c16910	1814970	1816151	hypothetical protein	COG3825	S	326	385	-59
RCA23_c16920	1816264	1816953	putative peptidase, M48 family	COG0501	O	364	324	40
RCA23_c16930	1817481	1816948	hypothetical protein			233	300	-67
RCA23_c16940	1818677	1817481	putative ribosomal RNA large subunit methyltransferase	COG1092	R	266	302	-36
RCA23_c16950	1818800	1820605	phosphogluconate dehydratase Edd	COG0129	E	268	330	-62
RCA23_c16960	1820631	1821248	KHG/KDPG aldolase Eda	COG0800	G	181	211	-30
RCA23_c16970	1821507	1821208	hypothetical protein			267	405	-138
RCA23_c16980	1824217	1821497	glutamate-ammonia-ligase adenylyltransferase GlnE	COG1391	O	227	331	-104
RCA23_c16990	1824296	1824769	hypothetical protein	COG2606	S	270	333	-63
RCA23_c17000	1824858	1825280	hypothetical protein			1.898	707	1.191
RCA23_c17010	1825558	1825782	hypothetical protein			304	243	61
RCA23_c17020	1826061	1825840	hypothetical protein			247	200	47
RCA23_c17030	1827217	1826054	putative aromatic-ring-hydroxylating dioxygenase	COG4638	P	320	342	-22
RCA23_c17040	1827932	1827375	putative acetolactate synthase small subunit IlvH	COG0440	E	419	452	-33

RCA23_c17050	1829695	1827944	acetolactate synthase isozyme large subunit IlvI	COG0028	E	380	318	62
RCA23_c17060	1830055	1831023	hypothetical membrane protein, porin-like			724	550	174
RCA23_c17070	1831298	1832551	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	407	376	31
RCA23_c17080	1832674	1833531	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	711	490	221
RCA23_c17090	1833543	1835897	TRAP dicarboxylate transporter, subunit DctM	COG4664	Q	337	395	-58
RCA23_c17100	1836779	1835958	putative arginyl-tRNA--protein transferase Ate	COG2935	O	448	457	-9
RCA23_c17110	1840606	1836959	vitamin B12-dependent ribonucleotide reductase NrdJ	COG0209	F	477	377	100
RCA23_c17120	1840806	1841066	hypothetical protein			175	144	31
RCA23_c17140	1841706	1841278	hypothetical protein DUF192	COG1430	S	474	371	103
RCA23_c17150	1842275	1841745	cold shock protein	COG1278	K	334	425	-91
RCA23_c17160	1843096	1842491	pyridoxine/pyridoxamine 5'-phosphate oxidase PdxH	COG0259	H	276	337	-61
RCA23_c17170	1843179	1843967	enoyl-[acyl-carrier-protein] reductase FabI	COG0623	I	385	443	-58
RCA23_c17180	1844001	1844504	xanthine phosphoribosyltransferase Gpt	COG0503	F	270	379	-109
RCA23_c17190	1844553	1845488	pyrimidine-specific ribonucleoside hydrolase RihA	COG1957	F	181	247	-66
RCA23_c17200	1845485	1846876	hypothetical protein UPF0061	COG0397	S	227	261	-34
RCA23_c17210	1846923	1848170	putative sodium/hydrogen exchanger	COG0025	P	328	487	-159
RCA23_c17220	1848192	1848626	ferric uptake regulator protein Fur	COG0735	P	556	482	74
RCA23_c17230	1848619	1849494	putative S-adenosylmethionine uptake transporter	COG0697	G	301	356	-55
RCA23_c17240	1849532	1850851	enolase Eno	COG0148	G	302	407	-105
RCA23_c17250	1852135	1851050	anhydro-N-acetylmuramic acid kinase AnmK	COG2377	O	202	334	-132
RCA23_c17260	1852202	1853452	tyrosyl-tRNA synthase TyrS	COG0162	J	477	463	14
RCA23_c17270	1853449	1854615	aspartate aminotransferase AspC	COG0436	E	258	354	-96
RCA23_c17280	1854680	1856539	hypothetical protein	COG0760	O	297	306	-9
RCA23_c17290	1856545	1858056	anthranilate synthase component TrpE	COG0147	E	224	249	-25
RCA23_c17300	1858187	1859110	protein soxG	COG0491	R	323	315	8
RCA23_c17310	1859114	1860058	protein soxH	COG0491	R	362	449	-87
RCA23_c17320	1861293	1860073	5-aminolevulinate synthase HemA	COG0156	H	337	393	-56
RCA23_c17330	1861747	1861295	sulfide dehydrogenase flavoprotein chain SoxF	COG3439	S	404	468	-64
RCA23_c17340	1862211	1861777	domain of unknown function DUF1791	COG1416	S	351	361	-10
RCA23_c17350	1863325	1862297	sulfite oxidase cytochrome subunit SoxD	COG3474	C	410	350	60

RCA23_c17360	1864577	1863309	sulfite oxidase molybdopterin subunit SoxC	COG2041	R	409	366	43
RCA23_c17370	1866311	1864608	sulfur oxidation protein SoxB	COG0737	F	518	409	109
RCA23_c17380	1867248	1866412	diheme cytochrome c	COG3258	C	536	298	238
RCA23_c17390	1867605	1867276	protein SoxZ			570	470	100
RCA23_c17400	1868053	1867631	protein SoxY	COG5501	S	301	243	58
RCA23_c17410	1868532	1868083	cytochrome c	COG2010	C	404	363	41
RCA23_c17420	1869206	1868649	thioredoxin SoxW	COG2143	O	639	479	160
RCA23_c17430	1869959	1869219	cytochrome c-type biogenesis protein SoxV	COG0785	O	588	396	192
RCA23_c17440	1870030	1870389	protein SoxS			435	382	53
RCA23_c17450	1870466	1870801	HTH-type transcriptional regulator, ArsR family	COG0640	K	321	345	-24
RCA23_c17460	1870811	1871866	putative soxT, transmembrane protein DUF395	COG2391	R	242	311	-69
RCA23_c17470	1872074	1873135	hypothetical protein DUF395	COG2391	R	383	379	4
RCA23_c17480	1874065	1873184	NAD-binding protein	COG2084	I	184	255	-71
RCA23_c17490	1875045	1874071	glyoxylate reductase GyaR	COG1052	C	161	206	-45
RCA23_c17500	1876573	1875236	glutamyl-tRNA(Gln) amidotransferase subunit A	COG0154	J	174	225	-51
RCA23_c17510	1877949	1876570	TRAP dicarboxylate transporter, subunit DctM	COG4664	Q	358	314	44
RCA23_c17520	1878581	1877964	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	460	318	142
RCA23_c17530	1879745	1878651	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	421	267	154
RCA23_c17540	1880759	1880388	hypothetical protein			351	196	155
RCA23_c17550	1883014	1881539	glycerol kinase GlpK	COG0554	C	190	274	-84
RCA23_c17560	1883797	1883087	ribosomal RNA large subunit methyltransferase J	COG0293	J	286	438	-152
RCA23_c17570	1884919	1883804	putative Ppx/GppA phosphatase family protein	COG0248	F	321	343	-22
RCA23_c17580	1885166	1885453	hypothetical protein			298	453	-155
RCA23_c17590	1885491	1886417	hypothetical protein	COG0685	E	250	310	-60
RCA23_c17600	1886430	1887494	pterin domain containing enzyme	COG1410	E	257	263	-6
RCA23_c17610	1888307	1887537	HpcH/HpaI aldolase family protein	COG3836	G	262	427	-165
RCA23_c17620	1888466	1889464	glucokinase Glk	COG0837	G	320	380	-60
RCA23_c17630	1889451	1890644	hypothetical protein DUF1006	COG3214	S	188	320	-132
RCA23_c17640	1893859	1890659	reductive dehalogenase	COG1018	C	279	315	-36
RCA23_c17650	1894859	1893852	hypothetical protein, XdhC and CoxI	COG1975	O	205	328	-123

RCA23_c17660	1895830	1894940	2-hydroxy-3-oxopropionate reductase GlxR	COG2084	I	232	253	-21
RCA23_c17670	1896292	1897032	3-hydroxybutyrate dehydrogenase	COG1028	I	356	301	55
RCA23_c17680	1897074	1898861	dihydroxy-acid dehydratase IlvD	COG0129	E	259	298	-39
RCA23_c17690	1898872	1899720	fumarylacetoacetate hydrolase	COG0179	Q	276	301	-25
RCA23_c17700	1899793	1900119	hypothetical protein			128	158	-30
RCA23_c17710	1900591	1900169	nucleoside diphosphate kinase Ndk	COG0105	F	455	256	199
RCA23_c17720	1900744	1902597	hypothetical protein			97	96	1
RCA23_c17730	1902629	1904470	ABC transporter ATP-binding protein	COG0488	R	222	265	-43
RCA23_c17740	1904463	1905104	MarC family integral membrane protein	COG2095	U	403	394	9
RCA23_c17750	1905138	1905722	hypothetical protein	COG3577	R	305	350	-45
RCA23_c17760	1906191	1905760	putative DNA polymerase III chi subunit, HolC	COG2927	L	209	241	-32
RCA23_c17770	1907713	1906223	cytosol aminopeptidase PepA	COG0260	E	268	371	-103
RCA23_c17780	1907835	1908953	puative permease, YjgP/YjgQ family	COG0795	R	524	334	190
RCA23_c17790	1908950	1910047	puative permease, YjgP/YjgQ family	COG0795	R	326	382	-56
RCA23_c17800	1910078	1912183	putative organic solvent tolerance protein	COG1452	M	375	389	-14
RCA23_c17810	1912196	1913395	hypothetical protein, SurA	COG0760	O	287	313	-26
RCA23_c17820	1913462	1914475	4-hydroxythreonine-4-phosphate dehydrogenase PdxA	COG1995	H	207	256	-49
RCA23_c17830	1914465	1915301	dimethyladenosine transferase KsgA	COG0030	J	179	238	-59
RCA23_c17840	1916086	1915556	hypothetical protein			300	334	-34
RCA23_c17850	1917167	1916316	modification methylase, hemK family	COG2890	J	353	442	-89
RCA23_c17860	1918213	1917164	peptide chain release factor 1	COG0216	J	222	312	-90
RCA23_c17870	1918660	1918274	hypothetical protein	COG4446	S	205	218	-13
RCA23_c17880	1919709	1918672	agmatinase SpeB	COG0010	E	232	318	-86
RCA23_c17890	1919932	1920441	alpha/beta hydrolase			162	268	-106
RCA23_c17900	1921451	1920480	agmatinase SpeB	COG0010	E	226	300	-74
RCA23_c17910	1922618	1921452	hippurate hydrolase HipO	COG1473	R	195	320	-125
RCA23_c17920	1922668	1923480	protein MazG	COG3956	R	366	371	-5
RCA23_c17930	1923546	1924556	ABC transporter periplasmic iron-binding protein FutA	COG1840	P	270	407	-137
RCA23_c17940	1925179	1924631	putative peptidyl-prolyl cis-trans isomerase Ppi	COG0652	O	284	433	-149
RCA23_c17950	1925684	1925172	putative peptidyl-prolyl cis-trans isomerase Ppi	COG0652	O	415	489	-74

RCA23_c17960	1925805	1926983	phosphoglycerate kinase Pgk	COG0126	G	223	274	-51
RCA23_c17970	1927091	1927984	fructose-bisphosphate aldolase class 1	COG3588	G	219	411	-192
RCA23_c17980	1928074	1928373	hypothetical protein, septum formation initiator	COG2919	D	299	324	-25
RCA23_c17990	1928491	1929489	pyruvate dehydrogenase E1 component alpha subunit PdhA	COG1071	C	443	418	25
RCA23_c18000	1929493	1930863	pyruvate dehydrogenase E1 component beta subunit PdhB	COG0022	C	414	346	68
RCA23_c18010	1930874	1932154	dihydrolipoyllysine-residue acetyltransferase PdhC	COG0508	C	253	284	-31
RCA23_c18020	1933034	1932237	serine acetyltransferase CysE	COG1045	E	711	414	297
RCA23_c18030	1934250	1933132	putative gene transfer agent protein			109	156	-47
RCA23_c18040	1935148	1934639	putative gene transfer agent large terminase part 1	COG5323	S	8	55	-47
RCA23_c18050	1936044	1935133	putative gene transfer agent large terminase part 2	COG5323	S	121	191	-70
RCA23_c18060	1937543	1936392	aminodeoxychorismate lyase	COG1559	R	246	282	-36
RCA23_c18070	1938774	1937545	3-oxoacyl-[acyl-carrier-protein] synthase FabF	COG0304	I	379	520	-141
RCA23_c18080	1939180	1938947	acyl carrier protein AcpP	COG0236	I	739	630	109
RCA23_c18090	1940136	1939399	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	252	290	-38
RCA23_c18100	1941123	1940194	malonyl CoA-acyl carrier protein transacylase	COG0331	I	204	295	-91
RCA23_c18110	1941409	1941774	30S ribosomal protein S6	COG0360	J	443	354	89
RCA23_c18120	1941788	1942015	30S ribosomal protein S18	COG0238	J	560	410	150
RCA23_c18130	1942052	1942645	50S ribosomal protein L9	COG0359	J	546	404	142
RCA23_c18140	1944185	1942854	trigger factor (TF)	COG0544	O	758	494	264
RCA23_c18160	1944635	1944973	nitrogen regulatory protein P-II 1	COG0347	E	931	485	446
RCA23_c18170	1945000	1946403	glutamine synthase GlnA type I	COG0174	E	786	431	355
RCA23_c18190	1947274	1946483	hypothetical protein			262	401	-139
RCA23_c18180	1947253	1948665	dimethylpropiothetin dethiomethylase DddP	COG0006	E	254	366	-112
RCA23_c18200	1949338	1948739	Biotin transporter BioY	COG1268	R	244	323	-79
RCA23_c18210	1949882	1949370	hypothetical protein			250	326	-76
RCA23_c18220	1949991	1951268	adenylosuccinate lyase PurB	COG0015	F	256	334	-78
RCA23_c18230	1951352	1951684	putative nitrile hydratase, beta subunit			324	231	93
RCA23_c18240	1951681	1951938	hypothetical protein			218	264	-46
RCA23_c18250	1951938	1952600	nitrile hydratase alpha subunit NthA			166	214	-48
RCA23_c18260	1952597	1953469	hypothetical protein DUF6 transmembrane	COG0697	G	213	283	-70

RCA23_c18270	1953462	1954361	putative lipid A biosynthesis lauroyl acyltransferase	COG1560	M	347	371	-24
RCA23_c18280	1955107	1954409	hypothetical protein	COG5429	S	489	407	82
RCA23_c18290	1955215	1957980	aconitate hydratase AcnA	COG1048	C	388	396	-8
RCA23_c18300	1958697	1958161	cytochrome c biogenesis protein CcmG	COG0526	O	243	314	-71
RCA23_c18310	1959609	1958839	cytochrome c-type biogenesis protein CcmC	COG0755	O	314	515	-201
RCA23_c18320	1960296	1959640	cytochrome c-type biogenesis protein CcmB	COG2386	O	273	384	-111
RCA23_c18330	1960907	1960293	cytochrome c biogenesis ATP-binding export protein CcmA	COG4133	O	227	402	-175
RCA23_c18340	1961264	1960911	hypothetical protein	COG3737	S	379	432	-53
RCA23_c18350	1962232	1961264	protein export membrane protein SecF	COG0341	U	436	416	20
RCA23_c18360	1963867	1962236	protein export membrane protein SecD	COG0342	U	300	327	-27
RCA23_c18370	1964193	1963912	putative immunogenic membrane protein YajC	COG1862	U	415	411	4
RCA23_c18380	1964441	1965733	seryl-tRNA synthase SerS	COG0172	J	267	313	-46
RCA23_c18390	1966135	1967112	hypothetical protein, alpha/beta hydrolase-like	COG2267	I	57	33	24
RCA23_c18400	1968740	1967283	GTP-binding protein EngA	COG1160	R	288	375	-87
RCA23_c18410	1970103	1968790	putative quinoprotein	COG1520	S	242	318	-76
RCA23_c18420	1970807	1970169	hypothetical protein DUF2133	COG4649	S	207	259	-52
RCA23_c18430	1971036	1972271	RND efflux transporter, MFP subunit	COG0845	M	278	245	33
RCA23_c18440	1972307	1975867	RND efflux transporter, permease protein	COG0841	V	413	402	11
RCA23_c18450	1977696	1975885	ABC transporter ATP-binding protein	COG5265	O	384	385	-1
RCA23_c18460	1979340	1977742	hypothetical protein, peptidoglycan-binding protein domain I	COG1652	S	241	344	-103
RCA23_c18470	1980799	1979921	protein RarD	COG2962	R	792	473	319
RCA23_c18480	1981500	1980901	superoxide dismutase SodB	COG0605	P	814	669	145
RCA23_c18490	1982152	1981604	sarcosine oxidase subunit SoxG	COG4583	E	434	445	-11
RCA23_c18500	1985156	1982145	sarcosine oxidase alpha subunit SoxA	COG0404	E	308	339	-31
RCA23_c18510	1985553	1985224	sarcosine oxidase subunit SoxD	COG4311	E	215	300	-85
RCA23_c18520	1986862	1985618	sarcosine oxidase beta subunit SoxB	COG0665	E	258	354	-96
RCA23_c18530	1987085	1988227	cytochrome c-type biogenesis protein CycH	COG4235	O	177	223	-46
RCA23_c18540	1988224	1988700	hypothetical protein	COG0816	L	177	338	-161
RCA23_c18550	1988697	1988945	hypothetical protein DUF1289	COG3313	R	234	344	-110
RCA23_c18560	1989819	1989013	hypothetical protein			204	305	-101

RCA23_c18570	1990628	1989810	hypothetical protein DUF81	COG0730	R	187	292	-105
RCA23_c18580	1991653	1990628	tRNA-dihydrouridine synthase Dus	COG0042	J	192	241	-49
RCA23_c18590	1992384	1991857	putative peroxiredoxin (thioredoxin reductase)	COG0678	O	16	6	10
RCA23_c18600	1993473	1992418	benzaldehyde dehydrogenase	COG1012	C	10	6	4
RCA23_c18610	1994144	1993836	hypothetical protein			22	7	15
RCA23_c18620	1995454	1994423	hypothetical protein	COG2130	R	336	356	-20
RCA23_c18630	1995676	1996863	MFS-type transporter	COG2814	G	221	315	-94
RCA23_c18640	1996989	1997762	sorbitol dehydrogenase	COG1028	I	202	295	-93
RCA23_c18650	1997773	1999044	6-hydroxynicotinate 3-monooxygenase	COG0654	H	179	183	-4
RCA23_c18660	1999655	1999212	hypothetical protein			176	334	-158
RCA23_c18670	2001036	1999702	glycerate kinase	COG2379	G	170	266	-96
RCA23_c18680	2001093	2001482	putative carboxymuconolactone decarboxylase	COG0599	S	406	540	-134
RCA23_c18690	2001563	2003575	2,4-dienoyl-CoA reductase [NADPH]	COG1902	C	246	338	-92
RCA23_c18700	2005239	2003572	thiamine pyrophosphate protein	COG0028	E	182	231	-49
RCA23_c18710	2006314	2005394	rffG/rfbB: dTDP-glucose 4,6-dehydratase	COG1088	M	332	308	24
RCA23_c18720	2007489	2006419	hypothetical protein			322	352	-30
RCA23_c18730	2007506	2009269	choline dehydrogenase BetA	COG2303	E	198	304	-106
RCA23_c18740	2010131	2009487	succinyl-CoA:3-ketoacid-coenzyme A transferase subunit S	COG2057	I	338	422	-84
RCA23_c18750	2010826	2010131	succinyl-CoA:3-ketoacid-coenzyme A transferase subunit S	COG1788	I	799	597	202
RCA23_c18760	2011295	2011669	hypothetical membrane protein			419	471	-52
RCA23_c18770	2012857	2011751	putative dimethyl sulfoniopropionate demethylase DmdA	COG0404	E	288	455	-167
RCA23_c18780	2013130	2015229	hydantoin utilization protein A	COG0145	E	214	300	-86
RCA23_c18790	2015226	2016950	hydantoin utilization protein B	COG0146	E	190	304	-114
RCA23_c18800	2016951	2018159	FAD dependent oxidoreductase	COG0665	E	250	294	-44
RCA23_c18810	2018159	2018995	hypothetical protein, 3-beta hydroxysteroid dehydrogenase	COG0451	M	280	321	-41
RCA23_c18820	2019084	2019956	hypothetical protein, serine/threonine-protein kinase	COG1262	S	160	199	-39
RCA23_c18830	2021104	2019980	hypothetical protein	COG0457	R	168	236	-68
RCA23_c18840	2022102	2021257	hypothetical protein DUF1989	COG3665	S	128	236	-108
RCA23_c18850	2022203	2023267	putative diaminopropionate ammonia-lyase	COG1171	E	156	175	-19
RCA23_c18860	2023267	2024406	hypothetical protein, metallopeptidase M24	COG0006	E	270	264	6

RCA23_c18870	2024408	2025154	hypothetical protein, Asp/Glu/hydantoin racemase	COG3473	Q	224	239	-15
RCA23_c18880	2025960	2025172	putative 3-oxoadipate enol-lactonase 2, alpha/beta hydrolas	COG0596	R	268	310	-42
RCA23_c18890	2026549	2025953	hypothetical protein DUF1185			387	557	-170
RCA23_c18900	2027634	2026561	limonene 1,2-monooxygenase LimB	COG2141	C	291	316	-25
RCA23_c18910	2029075	2027600	aldehyde dehydrogenase, cytosolic	COG1012	C	197	267	-70
RCA23_c18920	2030336	2029185	ABC transporter, spermidine/putrescine import, permease p	COG1177	E	398	447	-49
RCA23_c18930	2031682	2030432	ABC transporter, spermidine/putrescine import, permease p	COG1176	E	339	358	-19
RCA23_c18940	2032842	2031745	ABC transporter, spermidine/putrescine import, substrate bi	COG0687	E	397	254	143
RCA23_c18950	2034010	2032913	ABC transporter, spermidine/putrescine import, ATP-binding	COG3842	E	529	415	114
RCA23_c18960	2034377	2034865	putative nitrilotriacetate monooxygenase component B	COG1853	R	273	303	-30
RCA23_c18970	2034990	2035700	hypothetical protein, DUF268			334	290	44
RCA23_c18980	2036599	2035859	hypothetical protein DUF28	COG0217	S	633	385	248
RCA23_c18990	2038600	2036810	sodium/sulphate symporter	COG0471	P	338	413	-75
RCA23_c19000	2039486	2038578	hypothetical protein	COG0697	G	338	439	-101
RCA23_c19010	2040307	2039483	hypothetical protein	COG1692	S	237	351	-114
RCA23_c19020	2040825	2040358	osmotically inducible OsmC-like protein	COG1764	O	222	392	-170
RCA23_c19030	2043111	2041045	3-hydroxyacyl-CoA dehydrogenase, NAD-binding	COG1250	I	224	311	-87
RCA23_c19040	2044071	2043151	putative transcriptional regulator	COG0703	E	205	334	-129
RCA23_c19050	2044508	2044068	thioesterase-like protein	COG0824	R	238	311	-73
RCA23_c19060	2045700	2044510	benzoyl-CoA oxygenase component A	COG0369	P	188	277	-89
RCA23_c19070	2047239	2045734	benzoyl-CoA oxygenase component B	COG3396	S	320	346	-26
RCA23_c19080	2048905	2047250	benzoyl-CoA-dihydrodiol lyase BoxC	COG1024	I	315	360	-45
RCA23_c19090	2049492	2049049	hypothetical protein DUF309			204	271	-67
RCA23_c19100	2050211	2049489	alpha/beta hydrolase	COG0596	R	179	279	-100
RCA23_c19110	2051839	2050304	benzoate-coenzyme A ligase	COG0318	I	187	281	-94
RCA23_c19120	2052630	2052070	putative 5-formyltetrahydrofolate cyclo-ligase family protein	COG0212	H	187	344	-157
RCA23_c19130	2054000	2052627	magnesium transporter MgtE	COG2239	P	255	384	-129
RCA23_c19140	2054072	2055358	guanine deaminase GuaD	COG0402	F	196	307	-111
RCA23_c19150	2056765	2055422	putative hydroxydechloroatrazine ethylaminohydrolase	COG0402	F	189	260	-71
RCA23_c19160	2057582	2056758	putative inositol monophosphatase family protein	COG0483	G	194	271	-77

RCA23_c19170	2057984	2057688	putative helix-turn-helix protein	COG1396	K	318	258	60
RCA23_c19180	2058365	2059330	putative alcohol dehydrogenase	COG0604	C	268	331	-63
RCA23_c19190	2059441	2060259	putative inner membrane protein	COG0670	R	1.541	607	934
RCA23_c19200	2060615	2061193	hypothetical protein			1.346	656	690
RCA23_c19210	2061960	2061328	putative N-acetylmuramoyl-L-alanine amidase amiD	COG3023	V	151	239	-88
RCA23_c19220	2062550	2062017	hypothetical protein			243	299	-56
RCA23_c19230	2064075	2062591	glutamyl-tRNA(Gln) amidotransferase subunit A	COG0154	J	294	350	-56
RCA23_c19240	2064362	2064075	aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit	COG0721	J	232	432	-200
RCA23_c19250	2064883	2064449	putative deaminase	COG0590	F	205	227	-22
RCA23_c19260	2064947	2065729	putative ribosomal large subunit pseudouridine synthase B	COG1187	J	199	273	-74
RCA23_c19270	2066811	2065744	molybdate ABC transporter, ATP-binding protein ModC	COG4148	P	209	359	-150
RCA23_c19280	2067494	2066808	molybdate ABC transporter, pemease protein ModB	COG4149	P	241	332	-91
RCA23_c19290	2068318	2067491	molybdate ABC transporter, substrate binding protein ModA	COG0725	P	231	310	-79
RCA23_c19300	2068328	2068789	hypothetical protein, NUDIX hydrolase	COG0494	L	206	453	-247
RCA23_c19310	2068866	2069288	hypothetical protein DUF1178	COG5319	S	451	394	57
RCA23_c19320	2069388	2070611	aspartokinase LysC	COG0527	E	283	299	-16
RCA23_c19330	2070626	2072872	phosphoenolpyruvate-protein phosphotransferase PstI	COG3605	T	212	271	-59
RCA23_c19340	2073508	2073008	hypothetical protein, acetyltransferase-like	COG3153	R	164	260	-96
RCA23_c19350	2073581	2074186	hypothetical protein	COG1853	R	202	219	-17
RCA23_c19360	2075848	2074217	putative sulfate transporter	COG0659	P	264	369	-105
RCA23_c19370	2076031	2076906	methylthioadenosine phosphorylase MtnP	COG0005	F	204	304	-100
RCA23_c19380	2076963	2077523	adenine phosphoribosyltransferase Apt	COG0503	F	276	337	-61
RCA23_c19390	2078104	2077520	ribosomal-protein-alanine acetyltransferase RimJ	COG1670	J	240	287	-47
RCA23_c19400	2079370	2078108	uncharacterized zinc protease YmxG	COG0612	R	351	418	-67
RCA23_c19410	2080752	2079367	threonine synthase ThrC	COG0498	E	301	345	-44
RCA23_c19420	2081473	2080763	hypothetical protein, SURF1	COG3346	S	421	448	-27
RCA23_c19430	2082328	2081525	cytochrome c oxidase subunit 3	COG1845	C	694	471	223
RCA23_c19440	2082916	2082347	cytochrome c oxidase assembly protein CtaG	COG3175	O	694	524	170
RCA23_c19450	2084033	2083098	protoheme IX farnesyltransferase CtaB	COG0109	O	436	515	-79
RCA23_c19460	2084904	2084056	cytochrome c oxidase subunit 2 precursor	COG1622	C	794	480	314

RCA23_c19470	2085093	2086514	protein TldD	COG0312	R	190	276	-86
RCA23_c19480	2086648	2087799	SMF family protein	COG0758	L	160	264	-104
RCA23_c19490	2087889	2090483	DNA topoisomerase TopA	COG0550	L	240	343	-103
RCA23_c19500	2091332	2090493	fructose-bisphosphate aldolase Fba	COG0191	G	226	349	-123
RCA23_c19510	2092171	2091329	5-deoxy-glucuronate isomerase IolB	COG3718	G	378	459	-81
RCA23_c19520	2093164	2092175	5-dehydro-2-deoxygluconokinase IolC	COG0524	G	203	302	-99
RCA23_c19530	2095038	2093167	3D-(3,5/4)-trihydroxycyclohexane-1,2-dione hydrolase IolD	COG3962	E	218	303	-85
RCA23_c19540	2096366	2095143	hypothetical protein DUF989	COG3748	S	398	405	-7
RCA23_c19550	2096731	2096378	5-hydroxyisourate hydrolase UraH	COG2351	R	330	328	2
RCA23_c19560	2096794	2098206	uric acid degradation bifunctional protein PucL	COG3195	S	210	248	-38
RCA23_c19570	2099161	2098337	putative allantoin catabolism protein YlbA	COG3257	R	252	300	-48
RCA23_c19580	2099379	2100608	sarcosine oxidase beta subunit SoxB	COG0665	E	248	344	-96
RCA23_c19590	2100622	2100909	sarcosine oxidase subunit SoxD	COG4311	E	235	330	-95
RCA23_c19600	2100906	2103860	sarcosine oxidase alpha subunit SoxA	COG0404	E	185	228	-43
RCA23_c19610	2103853	2104413	sarcosine oxidase subunit SoxG	COG4583	E	217	262	-45
RCA23_c19620	2105450	2104446	putative HTH-type transcriptional regulator, AraC family	COG4977	K	207	326	-119
RCA23_c19630	2105519	2106172	putative DNA-binding protein	COG1396	K	275	407	-132
RCA23_c19640	2106599	2106183	putative membrane protein			218	337	-119
RCA23_c19650	2107064	2107477	hypothetical protein			576	510	66
RCA23_c19660	2108227	2107835	hypothetical protein, thioesterase	COG2050	Q	412	379	33
RCA23_c19670	2108892	2108230	hypothetical protein, calcineurin-like phosphoesterase-like	COG1407	R	271	407	-136
RCA23_c19680	2111360	2108931	putative DEAD/DEAH box helicase	COG1201	R	207	359	-152
RCA23_c19690	2112302	2111400	bidunctional enzyme FolD	COG0190	H	301	440	-139
RCA23_c19700	2114064	2112388	formate--tetrahydrofolate ligase Fhs	COG2759	F	356	344	12
RCA23_c19710	2114959	2114357	hypothetical protein			275	383	-108
RCA23_c19720	2116937	2115018	cell division protease FtsH	COG0465	O	464	415	49
RCA23_c19730	2118254	2117031	tRNA(Ile)-lysidine synthase TilS	COG0037	D	188	268	-80
RCA23_c19740	2119039	2118251	hypothetical protein	COG1729	S	214	219	-5
RCA23_c19750	2119539	2119039	peptidoglycan-associated lipoprotein Pal	COG2885	M	189	256	-67
RCA23_c19760	2120916	2119594	Tol-Pal system beta propeller repeat protein TolB	COG0823	U	269	322	-53

RCA23_c19770	2121908	2120913	hypothetical protein, TolA-like			194	283	-89
RCA23_c19780	2122520	2122056	biopolymer transport protein TolR	COG0848	U	204	389	-185
RCA23_c19790	2123204	2122530	biopolymer transport protein TolQ	COG0811	U	243	362	-119
RCA23_c19800	2123640	2123296	acyl-CoA thioester hydrolase	COG0824	R	206	269	-63
RCA23_c19810	2124311	2123682	hypothetical protein			225	325	-100
RCA23_c19820	2125447	2124398	holliday junction ATP-dependent DNA helicase RuvB	COG2255	L	212	330	-118
RCA23_c19830	2126118	2125444	holliday junction ATP-dependent DNA helicase RuvA	COG0632	L	241	443	-202
RCA23_c19840	2126621	2126115	crossover junction endodeoxyribonuclease RuvC	COG0817	L	211	426	-215
RCA23_c19850	2128058	2127186	putative ribosomal protein L11 methyltransferase	COG2264	J	199	331	-132
RCA23_c19860	2128842	2128189	peptide methionine sulfoxide reductase MsrA	COG0225	O	470	535	-65
RCA23_c19870	2130139	2128931	hypothetical protein, major facilitator superfamily transporter			568	531	37
RCA23_c19880	2130289	2131452	L-lactate dehydrogenase lldD	COG1304	C	384	429	-45
RCA23_c19890	2133377	2131542	BCCT family transporter involved in DMSP uptake	COG1292	M	415	398	17
RCA23_c19900	2134841	2133696	putative MFS-type transporter			335	395	-60
RCA23_c19910	2134970	2135605	50S ribosomal protein L25	COG1825	J	907	573	334
RCA23_c19920	2135732	2136418	peptidyl-tRNA hydrolase Pth	COG0193	J	315	395	-80
RCA23_c19930	2136415	2137437	hypothetical protein	COG4427	S	189	250	-61
RCA23_c19940	2137439	2137810	hypothetical protein	COG3651	S	233	303	-70
RCA23_c19950	2139068	2137836	tryptophan synthase beta chain TrpB	COG0133	E	369	326	43
RCA23_c19960	2139789	2139151	N-(5'-phosphoribosyl)anthranilate isomerase TrpF	COG0135	E	185	217	-32
RCA23_c19970	2140121	2139786	hypothetical protein			393	402	-9
RCA23_c19980	2140422	2140141	integration host factor beta subunit IhfB	COG0776	L	601	452	149
RCA23_c19990	2142345	2140684	30S ribosomal protein S1	COG0539	J	695	460	235
RCA23_c20000	2142605	2142934	hypothetical protein			405	426	-21
RCA23_c20010	2143968	2142946	hypothetical protein, lacI family HTH-type regulatory protein	COG1609	K	159	288	-129
RCA23_c20020	2144050	2145192	putative phytanoyl-CoA dioxygenase	COG5285	Q	353	567	-214
RCA23_c20030	2146137	2145199	aldo/keto reductase	COG0667	C	253	316	-63
RCA23_c20040	2146855	2146184	cytidylate kinase Cmk	COG0283	F	198	213	-15
RCA23_c20050	2148201	2146852	3-phosphoshikimate 1-carboxyvinyltransferase AroA	COG0128	E	215	383	-168
RCA23_c20060	2148997	2148275	tRNA (guanine-N(7)-)-methyltransferase TrmB	COG0220	R	237	310	-73

RCA23_c20070	2150223	2149042	S-adenosylmethionine synthase MetK	COG0192	H	360	341	19
RCA23_c20080	2151767	2150271	apolipoprotein N-acyltransferase Lnt	COG0815	M	237	319	-82
RCA23_c20090	2152641	2151742	putative magnesium and cobalt efflux protein	COG1253	R	383	427	-44
RCA23_c20100	2153141	2152638	hypothetical protein UPF0054	COG0319	R	368	393	-25
RCA23_c20110	2154136	2153138	PhoH-like protein	COG1702	T	275	401	-126
RCA23_c20120	2155564	2154233	RNA modification enzyme, MiaB family	COG0621	J	275	292	-17
RCA23_c20130	2155629	2156543	hypothetical protein			513	459	54
RCA23_c20140	2156986	2156552	putative ferric uptake regulator family protein	COG0735	P	537	555	-18
RCA23_c20150	2157143	2157652	3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase FabA	COG0764	I	376	331	45
RCA23_c20160	2157667	2158896	3-oxoacyl-[acyl-carrier-protein] synthase FabB	COG0304	I	361	339	22
RCA23_c20170	2158901	2159689	enoyl-[acyl-carrier-protein] reductase FabI	COG0623	I	292	375	-83
RCA23_c20180	2160715	2159738	threo-3-hydroxyaspartate ammonia-lyase SRY	COG1171	E	167	199	-32
RCA23_c20190	2161402	2160719	2-haloalkanoic acid dehalogenase	COG1011	R	258	323	-65
RCA23_c20200	2161460	2162248	3-oxoadipate enol-lactonase CatD	COG0596	R	289	372	-83
RCA23_c20210	2162268	2163494	putative tetracycline resistance protein, class C	COG2814	G	193	322	-129
RCA23_c20220	2163510	2164649	NAD/mycothiol-dependent formaldehyde dehydrogenase	COG1062	C	189	249	-60
RCA23_c20230	2164704	2165945	D-galactonate dehydratase DgoD	COG4948	M	320	382	-62
RCA23_c20240	2168433	2165986	dimethylglycine dehydrogenase	COG0404	E	269	340	-71
RCA23_c20250	2170044	2168488	trimethylamine methyltransferase MttB	COG5598	H	270	344	-74
RCA23_c20260	2170193	2170597	hypothetical protein	COG0607	P	248	216	32
RCA23_c20270	2170629	2171999	uncharacterized aminotransferase	COG0161	H	325	341	-16
RCA23_c20280	2172154	2173251	aminomethyltransferase, mitochondrial	COG0404	E	364	312	52
RCA23_c20290	2173265	2173630	glycine cleavage system protein GcvH	COG0509	E	591	456	135
RCA23_c20300	2173679	2176525	glycine dehydrogenase GcvP	COG1003	E	291	292	-1
RCA23_c20310	2177847	2176510	D-amino acid dehydrogenase small subunit DadA	COG0665	E	184	266	-82
RCA23_c20320	2177927	2178970	putrescine ABC transport system putrescine-binding peripla	COG0687	E	343	375	-32
RCA23_c20330	2180246	2179164	putrescine ABC transport system ATP-binding protein PotG	COG3842	E	257	331	-74
RCA23_c20340	2181058	2180243	putrescine ABC transport system permease protein PotI	COG1177	E	470	610	-140
RCA23_c20350	2181888	2181055	putrescine ABC transport system permease protein PotB	COG1176	E	607	570	37
RCA23_c20360	2183133	2181934	rieske 2Fe-2S domain protein	COG4638	P	465	415	50

RCA23_c20370	2184561	2183245	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	441	547	-106
RCA23_c20380	2185139	2184558	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	521	546	-25
RCA23_c20390	2186242	2185214	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	328	358	-30
RCA23_c20400	2186550	2187362	putative phytanoyl-CoA dioxygenase	COG5285	Q	167	221	-54
RCA23_c20410	2188449	2187382	putative phenylacetic acid degradation NADH oxidoreducta	COG1018	C	203	273	-70
RCA23_c20420	2188902	2188462	phenylacetate-CoA oxygenase, subunit PaaD	COG2151	R	161	250	-89
RCA23_c20430	2189696	2188944	phenylacetate-CoA oxygenase, subunit PaaC	COG3396	S	230	299	-69
RCA23_c20440	2189986	2189696	phenylacetate-CoA oxygenase, subunit PaaB	COG3460	Q	272	347	-75
RCA23_c20450	2190999	2190007	phenylacetate-CoA oxygenase, subunit PaaA	COG3396	S	292	384	-92
RCA23_c20460	2191871	2191050	phenylacetic acid degradation operon negative regulatory p	COG3327	K	194	338	-144
RCA23_c20470	2193916	2191868	phenylacetic acid degradation protein PaaZ	COG1012	C	200	257	-57
RCA23_c20480	2194006	2194791	putative enoyl-CoA hydratase PaaG	COG1024	I	157	198	-41
RCA23_c20490	2194793	2195233	acyl-coenzyme A thioesterase PaaI	COG2050	Q	207	193	14
RCA23_c20500	2195230	2196432	beta-ketoadipyl CoA thiolase PaaJ	COG0183	I	159	218	-59
RCA23_c20510	2196443	2197726	phenylacetate-CoA ligase PaaK	COG1541	H	258	377	-119
RCA23_c20520	2197730	2198344	putative HTH-type transcriptional regulator, TetR family	COG1309	K	255	313	-58
RCA23_c20530	2198826	2198341	3-dehydroquinate dehydratase AroQ	COG0757	E	480	484	-4
RCA23_c20540	2200260	2198869	tyrosine decarboxylase	COG0076	E	256	302	-46
RCA23_c20550	2200544	2201383	putative short chain dehydrogenase	COG0300	R	180	216	-36
RCA23_c20560	2202389	2201403	hypothetical protein, acetyltransferase-like	COG1247	M	188	325	-137
RCA23_c20570	2202793	2202455	protein CsaA	COG0073	R	280	393	-113
RCA23_c20580	2203601	2202786	pyrroline-5-carboxylate reductase ProC	COG0345	E	179	341	-162
RCA23_c20590	2204138	2203635	hypothetical protein DUF1790	COG5465	S	484	467	17
RCA23_c20600	2204711	2204427	hypothetical protein DUF526	COG2960	S	356	344	12
RCA23_c20610	2204859	2205740	prolipoprotein diacylglyceryl transferase Lgt	COG0682	M	217	274	-57
RCA23_c20620	2205707	2206798	hypothetical protein DUF185	COG1565	S	183	256	-73
RCA23_c20630	2206795	2207550	multi-copper polyphenol oxidoreductase laccase	COG1496	S	267	296	-29
RCA23_c20640	2208728	2207769	hypothetical protein			176	115	61
RCA23_c20650	2209482	2208988	leucine-responsive regulatory protein Lrp	COG1522	K	252	363	-111
RCA23_c20660	2209668	2210609	thioredoxin reductase TrxB	COG0492	O	277	273	4

RCA23_c20670	2210930	2212633	bifunctional sulfate adenylyltransferase / adenylyl-sulfate kir	COG2046	P	237	293	-56
RCA23_c20680	2213026	2212808	hypothetical protein DUF1150			755	497	258
RCA23_c20690	2213444	2213028	small heat shock protein IbpA	COG0071	O	739	516	223
RCA23_c20700	2213600	2213815	hypothetical protein DUF465	COG5481	S	416	481	-65
RCA23_c20710	2213872	2214354	N5-carboxyaminoimidazole ribonucleotide mutase PurE	COG0041	F	324	250	74
RCA23_c20720	2214347	2215408	N5-carboxyaminoimidazole ribonucleotide synthase PurK	COG0026	F	192	252	-60
RCA23_c20730	2216205	2215438	hypothetical protein	COG1075	R	302	371	-69
RCA23_c20740	2217963	2216314	chaperonin GroEL	COG0459	O	1.092	481	611
RCA23_c20750	2218302	2218015	chaperonin GroS	COG0234	O	1.580	549	1.031
RCA23_c20760	2218533	2219483	hypothetical protein			272	228	44
RCA23_c20770	2219687	2220529	hypothetical protein			311	282	29
RCA23_c20780	2220548	2221510	hypothetical protein			256	251	5
RCA23_c20790	2221605	2222753	creatinase	COG0006	E	207	242	-35
RCA23_c20800	2222847	2223821	putative membrane lipoprotein			262	439	-177
RCA23_c20810	2224547	2223837	hypothetical protein DUF261	COG5482	S	192	370	-178
RCA23_c20820	2224600	2225517	putative manganese-dependent inorganic pyrophosphatase	COG1227	C	323	325	-2
RCA23_c20830	2225535	2226407	hypothetical protein, HAD family hydrolase	COG0647	G	261	348	-87
RCA23_c20840	2226492	2226935	putative MaoC-like dehydratase	COG2030	I	300	344	-44
RCA23_c20850	2226997	2227944	riboflavin biosynthesis protein RibF	COG0196	H	281	423	-142
RCA23_c20860	2227941	2228405	hypothetical protein UPF0260	COG2983	S	370	349	21
RCA23_c20870	2228409	2229458	low specificity L-threonine aldolase ItaE	COG2008	E	279	242	37
RCA23_c20880	2230157	2229510	hypothetical protein, 2-hydroxychromene-2-carboxylate isor	COG3917	Q	254	239	15
RCA23_c20890	2230217	2231236	ribose-phosphate pyrophosphokinase Prs	COG0462	F	351	337	14
RCA23_c20900	2231567	2231247	hypothetical protein			339	407	-68
RCA23_c20910	2232095	2231688	ATP synthase epsilon chain AtpC	COG0355	C	217	213	4
RCA23_c20920	2233519	2232095	ATP synthase beta subunit AtpD	COG0055	C	765	423	342
RCA23_c20930	2234405	2233536	ATP synthase gamma chain AtpG	COG0224	C	785	455	330
RCA23_c20940	2235953	2234418	ATP synthase alpha subunit AtpA	COG0056	C	653	400	253
RCA23_c20950	2236520	2235954	ATP synthase delta chain AtpH	COG0712	C	742	430	312
RCA23_c20960	2237506	2236745	hypothetical protein			238	284	-46

RCA23_c20970	2237568	2238365	hydroxyacylglutathione hydrolase GloB	COG0491	R	203	292	-89
RCA23_c20980	2238504	2240834	ATP-dependent Clp protease ATP-binding subunit ClpA	COG0542	O	394	338	56
RCA23_c20990	2243121	2240878	hypothetical protein, OmpA	COG1360	N	305	321	-16
RCA23_c21000	2244303	2243131	hypothetical protein			257	346	-89
RCA23_c21010	2244901	2244371	putative cation transport protein ChaC	COG3703	P	210	316	-106
RCA23_c21020	2245169	2246512	hypothetical protein	COG1357	S	297	248	49
RCA23_c21030	2247465	2246560	protein TyrC	COG0287	E	207	274	-67
RCA23_c21040	2248547	2247462	histidinol-phosphate aminotransferase HisC	COG0079	E	204	333	-129
RCA23_c21050	2249402	2248746	30S ribosomal protein S4	COG0522	J	558	359	199
RCA23_c21060	2251712	2249745	cold-shock DEAD box protein A	COG0513	L	271	328	-57
RCA23_c21070	2252217	2251846	hypothetical protein DUF393	COG3011	S	295	410	-115
RCA23_c21080	2254594	2252333	vitamin B12-dependent ribonucleoside-diphosphate reducta	COG0209	F	354	420	-66
RCA23_c21090	2254754	2255152	hypothetical protein	COG5458	S	170	273	-103
RCA23_c21100	2255830	2255156	ATP-phosphoribosyltransferase HisG	COG0040	E	238	409	-171
RCA23_c21110	2256873	2255827	ATP phosphoribosyltransferase regulatory subunit HisZ	COG3705	E	159	258	-99
RCA23_c21120	2258375	2256870	histidyl-tRNA synthase HisS	COG0124	J	253	369	-116
RCA23_c21130	2258424	2258699	hypothetical protein	COG2900	S	209	308	-99
RCA23_c21140	2262238	2258738	DNA polymerase III 1 alpha subunit DnaE	COG0587	L	327	386	-59
RCA23_c21150	2262629	2263969	xanthine dehydrogenase XdhA	COG4630	F	236	242	-6
RCA23_c21160	2263966	2266269	xanthine dehydrogenase XdhB	COG4631	F	224	287	-63
RCA23_c21170	2266266	2267186	xanthine dehydrogenase accessory protein XdhC	COG1975	O	208	245	-37
RCA23_c21180	2267183	2268688	sugar ABC transporter, ATP-binding protein	COG3845	R	202	224	-22
RCA23_c21190	2268675	2269760	sugar ABC transporter, permease protein	COG4603	R	330	307	23
RCA23_c21200	2269741	2270742	ABC transporter permease protein	COG1079	R	283	408	-125
RCA23_c21210	2270770	2271849	putative basic membrane protein	COG1744	R	455	356	99
RCA23_c21220	2271933	2273012	putative FAD dependent oxidoreductase	COG0665	E	164	201	-37
RCA23_c21230	2275563	2273065	dimethylglycine dehydrogenase	COG0404	E	264	326	-62
RCA23_c21240	2275772	2276626	pantothenate synthase PanC	COG0414	H	240	384	-144
RCA23_c21250	2276623	2277435	3-methyl-2-oxobutanoate hydroxymethyltransferase PanB	COG0413	H	307	438	-131
RCA23_c21260	2277469	2277984	hypothetical protein	COG3807	S	226	318	-92

RCA23_c21270	2280010	2277986	peptidyl-dipeptidase Dcp	COG0339	E	227	267	-40
RCA23_c21280	2281046	2279997	molybdopterin biosynthesis protein MoeB	COG0476	H	211	292	-81
RCA23_c21290	2281504	2281043	deoxyuridine 5'-triphosphate nucleotidohydrolase Dut	COG0756	F	194	283	-89
RCA23_c21300	2282691	2281501	coenzyme A biosynthesis bifunctional protein CoaBC	COG0452	H	292	263	29
RCA23_c21310	2283626	2282748	RNA polymerase sigma-32 factor RpoH	COG0568	K	992	490	502
RCA23_c21320	2283716	2284249	bifunctional adenosylcobalamin biosynthesis protein CobP	COG2087	H	179	216	-37
RCA23_c21330	2284246	2284818	putative phosphoglycerate mutase family protein	COG0406	G	125	238	-113
RCA23_c21340	2284882	2285307	hypothetical protein, glutathione S-transferase	COG0625	O	225	320	-95
RCA23_c21350	2285323	2285757	glutathione S-transferase family protein			215	332	-117
RCA23_c21360	2285829	2286233	magnesium chelatase-like protein	COG0606	O	225	250	-25
RCA23_c21370	2286230	2287420	magnesium chelatase	COG0606	O	163	269	-106
RCA23_c21380	2288109	2287342	putative transmembrane protein			338	385	-47
RCA23_c21390	2289026	2288109	putative sulfite exporter TauE/SafE	COG0730	R	370	455	-85
RCA23_c21400	2290274	2289135	branched-chain amino acid transporter, permease compone	COG1079	R	267	471	-204
RCA23_c21410	2291340	2290279	branched-chain amino acid transporter, permease compone	COG4603	R	290	336	-46
RCA23_c21420	2292916	2291381	ABC transporter ATP-binding protein	COG3845	R	317	388	-71
RCA23_c21430	2293990	2292992	putative ABC transporter, periplasmic substrate-binding pro	COG1744	R	354	382	-28
RCA23_c21440	2294587	2294177	hypothetical protein, acetyltransferase-like	COG0456	R	209	279	-70
RCA23_c21450	2295161	2294598	NifU-like protein	COG0694	O	409	435	-26
RCA23_c21460	2295762	2295250	putative universal stress protein	COG0589	T	518	514	4
RCA23_c21470	2296878	2295862	tryptophanyl-tRNA synthase TrpS	COG0180	J	372	400	-28
RCA23_c21480	2296952	2297626	hypothetical protein, rhomboid protease	COG0705	R	167	212	-45
RCA23_c21490	2299173	2297638	virulence factor MviN homolg	COG0728	R	237	404	-167
RCA23_c21500	2301938	2299170	uridylyltransferase GlnD	COG2844	O	288	413	-125
RCA23_c21510	2303131	2301941	hypothetical protein	COG0683	E	173	274	-101
RCA23_c21520	2303164	2304075	putative tetrapyrrole methylase	COG0313	R	201	203	-2
RCA23_c21530	2304080	2304448	hypothetical protein UPF0102	COG0792	L	245	322	-77
RCA23_c21540	2304547	2305485	glutathione synthase GshB	COG0189	H	319	394	-75
RCA23_c21550	2306330	2305503	putative esterase/lipase	COG0657	I	327	376	-49
RCA23_c21560	2306806	2306327	NusB-like protein NusB	COG0781	K	272	421	-149

RCA23_c21570	2307324	2306803	6,7-dimethyl-8-ribityllumazine synthase RibH	COG0054	H	240	301	-61
RCA23_c21580	2308460	2307324	riboflavin biosynthesis protein RibB	COG0108	H	374	499	-125
RCA23_c21590	2309051	2308464	riboflavin synthase alpha chain RibE	COG0307	H	371	609	-238
RCA23_c21600	2309210	2310511	putative capsule polysaccharide export protein	COG3562	M	318	279	39
RCA23_c21610	2310600	2311748	putative polysaccharide export protein	COG1596	M	203	186	17
RCA23_c21620	2311750	2313762	putative capsule polysaccharide export protein	COG3563	M	150	241	-91
RCA23_c21630	2314853	2313759	riboflavin biosynthesis protein RibD	COG0117	H	190	222	-32
RCA23_c21640	2315317	2314850	transcriptional repressor NrdR	COG1327	K	348	414	-66
RCA23_c21650	2317485	2315515	RNA polymerase sigma factor RpoD	COG0568	K	608	437	171
RCA23_c21660	2319548	2317602	DNA primase DnaG	COG0358	L	355	423	-68
RCA23_c21670	2320108	2319557	sarcosine oxidase subunit SoxG	COG4583	E	231	384	-153
RCA23_c21680	2323031	2320101	sarcosine oxidase alpha subunit SoxA	COG0404	E	196	307	-111
RCA23_c21690	2323270	2323034	sarcosine oxidase subunit SoxD	COG4311	E	308	398	-90
RCA23_c21700	2324611	2323358	sarcosine oxidase beta subunit SoxB	COG0665	E	245	350	-105
RCA23_c21710	2325416	2324724	phosphatidylserine decarboxylase proenzyme Psd	COG0688	I	550	425	125
RCA23_c21720	2325810	2325421	diacylglycerol kinase	COG0818	M	881	641	240
RCA23_c21730	2327399	2325807	putative sulfatase	COG2194	R	919	638	281
RCA23_c21740	2328122	2327421	hypothetical protein, CDP-alcohol phosphatidyltransferase-l	COG1183	I	809	632	177
RCA23_c21750	2328718	2328254	hypothetical protein	COG0671	I	526	430	96
RCA23_c21760	2329421	2328897	hypothetical protein	COG2194	R	558	446	112
RCA23_c21770	2330356	2329541	hypothetical protein, DUF1705	COG2194	R	624	467	157
RCA23_c21780	2331919	2330483	glutamate dehydrogenase GluD	COG0334	E	314	332	-18
RCA23_c21790	2332495	2331998	hypothetical protein			258	365	-107
RCA23_c21800	2333529	2332585	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I	329	359	-30
RCA23_c21810	2333544	2334782	reductive dehalogenase	COG1600	C	264	315	-51
RCA23_c21820	2334839	2335750	pirin	COG1741	R	430	440	-10
RCA23_c21830	2336333	2335734	transcriptional regulator	COG1309	K	274	419	-145
RCA23_c21840	2336413	2336667	hypothetical protein			261	359	-98
RCA23_c21850	2336787	2338076	homoserine dehydrogenase Hom	COG0460	E	298	326	-28
RCA23_c21860	2338209	2339174	fructose-1,6-bisphosphatase class II	COG1494	G	221	267	-46

RCA23_c21870	2339176	2340912	putative single-stranded-DNA-specific exonuclease	COG0608	L	211	311	-100
RCA23_c21890	2341884	2341321	sulfopyruvate decarboxylase beta subunit ComE	COG0028	E	560	522	38
RCA23_c21900	2342362	2341886	sulfopyruvate decarboxylase alpha subunit ComD	COG4032	R	797	611	186
RCA23_c21910	2343377	2342394	zinc-type alcohol dehydrogenase	COG1063	E	316	480	-164
RCA23_c21920	2344144	2343386	gluconate 5-dehydrogenase Gno	COG1028	I	200	301	-101
RCA23_c21930	2345464	2344148	histidinol dehydrogenase HisD	COG0141	E	306	379	-73
RCA23_c21940	2345518	2346576	HTH-type transcriptional regulator, LacI family	COG1609	K	291	268	23
RCA23_c21950	2347369	2346614	HTH-type transcriptional regulator, IclR family	COG1414	K	263	354	-91
RCA23_c21960	2347514	2349115	alcohol dehydrogenase AlkJ	COG2303	E	278	286	-8
RCA23_c21970	2349149	2349817	integral membrane protein TerC	COG0861	P	322	278	44
RCA23_c21980	2349936	2351048	putative TRAP transporter, DctP subunit	COG4663	Q	464	312	152
RCA23_c21990	2351168	2353642	TRAP transporter, DctM/DctQ subunit	COG4664	Q	296	253	43
RCA23_c22000	2355428	2353818	trimethylamine methyltransferase	COG5598	H	411	403	8
RCA23_c22010	2356156	2355458	putative HAD-family hydrolase	COG0546	R	278	354	-76
RCA23_c22020	2356437	2356156	hypothetical protein			306	534	-228
RCA23_c22030	2356766	2356512	hypothetical protein			187	301	-114
RCA23_c22040	2357079	2356789	hypothetical protein			143	183	-40
RCA23_c22050	2357261	2357656	hypothetical protein	COG5349	S	172	246	-74
RCA23_c22060	2357741	2358352	hypothetical protein, NUDIX hydrolase	COG0494	L	185	264	-79
RCA23_c22070	2359231	2358332	hypothetical protein			240	315	-75
RCA23_c22080	2360146	2359235	putative fatty acid desaturase	COG3239	I	238	271	-33
RCA23_c22090	2360233	2361102	putative helix-turn-helix protein	COG1396	K	271	380	-109
RCA23_c22100	2361854	2361099	short chain dehydrogenase	COG1028	I	199	337	-138
RCA23_c22110	2363320	2361851	putative 3-hydroxyacyl-CoA dehydrogenase	COG1250	I	222	316	-94
RCA23_c22130	2364364	2363357	hypothetical protein, DUF849	COG3246	S	198	375	-177
RCA23_c22120	2364345	2365316	transcriptional regulator, AraC family	COG4977	K	181	300	-119
RCA23_c22140	2366365	2365322	protein Tas	COG0667	C	249	289	-40
RCA23_c22150	2366465	2367088	hypothetical protein	COG4544	S	139	266	-127
RCA23_c22160	2367182	2368417	putative MFS-type transporter	COG2211	G	250	323	-73
RCA23_c22170	2368475	2370187	methionyl-tRNA synthase MetG	COG0143	J	246	334	-88

RCA23_c22190	2370689	2372218	integrase			GI 8	337	236	101
RCA23_c22200	2372685	2372215	integrase	COG4974	L	GI 8	417	345	72
RCA23_c22210	2373139	2373477	putative nucleotidyl transferase	COG1210	M	GI 8	343	199	144
RCA23_c22220	2375142	2374411	hypothetical protein			GI 8	371	226	145
RCA23_c22230	2375873	2375142	hypothetical protein			GI 8	424	250	174
RCA23_c22240	2376433	2375873	putative replication factor C, small subunit	COG2256	L	GI 8	578	315	263
RCA23_c22250	2377432	2376614	hypothetical protein			GI 8	378	220	158
RCA23_c22260	2377643	2378410	hypothetical protein			GI 8	363	260	103
RCA23_c22270	2380540	2379620	ParB-like nuclease	COG1475	K	GI 8	484	357	127
RCA23_c22280	2380679	2383024	cadmium-transporting ATPase CadA	COG2217	P	GI 8	160	160	0
RCA23_c22290	2383190	2383600	hypothetical protein	COG3034	S	GI 8	193	223	-30
RCA23_c22300	2384025	2383591	putative lipoprotein signal peptidase	COG0597	M	GI 8	324	287	37
RCA23_c22310	2384826	2384065	putative ZIP zinc transporter	COG0428	P	GI 8	206	270	-64
RCA23_c22320	2385290	2384874	hypothetical protein DUF411	COG3019	R	GI 8	365	300	65
RCA23_c22330	2385943	2385320	SCO-like protein	COG1999	R	GI 8	251	222	29
RCA23_c22340	2386320	2385943	hypothetical protein	COG2847	S	GI 8	208	146	62
RCA23_c22350	2387089	2386430	hypothetical protein	COG1651	O	GI 8	162	177	-15
RCA23_c22360	2387511	2387086	hypothetical protein, disulfide bond formation protein	COG1495	O	GI 8	213	286	-73
RCA23_c22370	2388080	2387508	SCO-like protein	COG1999	R	GI 8	225	229	-4
RCA23_c22380	2388277	2388705	putative HTH-type transcriptional regulator	COG0789	K	GI 8	204	271	-67
RCA23_c22390	2388708	2389034	hypothetical protein UPF0060	COG1742	S	GI 8	176	169	7
RCA23_c22400	2389105	2389329	hypothetical protein			GI 8	272	193	79
RCA23_c22410	2390264	2389434	hypothetical protein			GI 8	84	88	-4
RCA23_c22420	2392406	2391699	transposase	COG3316	L	GI 8	2	3	-1
RCA23_c22430	2393515	2392478	MORN motif precursor	COG4642	S	GI 8	283	186	97
RCA23_c22440	2394592	2393885	transposase	COG3316	L	GI 8	8	0	8
RCA23_c22450	2394953	2394711	hypothetical protein			GI 8	13	0	13
RCA23_c22460	2395528	2395211	hypothetical protein			GI 8	2.261	1.424	837
RCA23_c22470	2396464	2395883	hypothetical protein	COG3295	S	GI 8	42	18	24
RCA23_c22480	2397201	2396467	hypothetical protein			GI 8	64	18	46

RCA23_c22490	2397865	2397518	hypothetical protein			GI 8	0	0	0
RCA23_c22500	2398105	2398617	transcriptional regulator, LuxR family	COG2197	T	GI 8	24	14	10
RCA23_c22510	2399503	2398796	hypothetical protein	COG3316	L	GI 8	5	1	4
RCA23_c22520	2401019	2399598	hypothetical protein, HTH-type transcriptional regulator, Lux	COG5616	S	GI 8	40	75	-35
RCA23_c22530	2401309	2402388	cysteine synthase CysK	COG0031	E	GI 8	64	75	-11
RCA23_c22540	2402668	2402847	hypothetical protein			GI 8	137	182	-45
RCA23_c22550	2402904	2403158	hypothetical protein			GI 8	135	123	12
RCA23_c22560	2403266	2403550	hypothetical protein			GI 8	150	154	-4
RCA23_c22570	2403731	2404039	hypothetical protein			GI 8	67	62	5
RCA23_c22580	2404165	2404464	hypothetical protein	COG5470	S	GI 8	91	36	55
RCA23_c22590	2404900	2406312	coniferyl aldehyde dehydrogenase CalB	COG1012	C	GI 8	349	196	153
RCA23_c22600	2407339	2406443	regulatory protein NocR	COG0583	K	GI 8	256	176	80
RCA23_c22610	2407478	2408791	sn-glycerol-3-phosphate-binding periplasmic protein UgpB	COG1653	G	GI 8	431	258	173
RCA23_c22620	2408867	2409748	sn-glycerol-3-phosphate transport system permease protein	COG1175	G	GI 8	568	361	207
RCA23_c22630	2409790	2410626	sn-glycerol-3-phosphate transport system permease protein	COG0395	G	GI 8	590	318	272
RCA23_c22640	2410634	2411680	sn-glycerol-3-phosphate import ATP-bindingprotein UgpC	COG3839	G	GI 8	527	263	264
RCA23_c22650	2411680	2412570	putative glycerophosphoryl diester phosphodiesterase	COG0584	C	GI 8	368	259	109
RCA23_c22660	2414704	2413001	aerobic glycerol-3-phosphate dehydrogenase GlpD	COG0578	C	GI 8	212	146	66
RCA23_c22670	2416422	2414794	hypothetical protein, Na+/Pi-cotransporter	COG1283	P	GI 8	144	136	8
RCA23_c22680	2417269	2416478	hypothetical protein, calcineurin-like phosphoesterase-like	COG1409	R	GI 8	166	115	51
RCA23_c22690	2418075	2417266	ABC transporter permease protein	COG0395	G	GI 8	173	119	54
RCA23_c22700	2418965	2418072	ABC transporter permease protein	COG1175	G	GI 8	166	109	57
RCA23_c22710	2420318	2419050	ABC transporter extracellular solute-binding protein	COG1653	G	GI 8	305	157	148
RCA23_c22720	2421380	2420382	ABC transporter ATP-binding protein	COG3839	G	GI 8	233	135	98
RCA23_c22730	2422258	2421398	HTH-type transcriptional regulator, DeoR family	COG1349	K	GI 8	194	176	18
RCA23_c22740	2422599	2423399	HAD-superfamily hydrolase, subfamily IIB	COG0561	R	GI 8	192	150	42
RCA23_c22750	2423921	2424865	putative inner membrane transporter	COG0697	G	GI 8	285	181	104
RCA23_c22760	2424987	2426117	soluble aldose sugar dehydrogenase YliI	COG2133	G	GI 8	223	155	68
RCA23_c22770	2426301	2427080	putative FKBP-type peptidyl-prolyl cis-trans isomerase	COG0545	O	GI 8	37	48	-11
RCA23_c22780	2427117	2427665	putative NnrU family protein	COG4094	S	GI 8	65	34	31

RCA23_c22790	2428870	2429508	hypothetical protein			GI 8	23	19	4
RCA23_c22800	2429684	2430973	putative adenine methyltransferase	COG0863	L	GI 8	71	54	17
RCA23_c22810	2430970	2431335	hypothetical protein			GI 8	33	23	10
RCA23_c22820	2431343	2432776	phage uncharacterised protein	COG5410	S	GI 8	38	24	14
RCA23_c22830	2433238	2434899	hypothetical protein, resolvase-like	COG1961	L	GI 8	34	19	15
RCA23_c22840	2435361	2436626	putative prophage integrase	COG0582	L	GI 8	0	3	-3
RCA23_c22850	2436623	2437081	hypothetical protein, putative phage-like protein			GI 8	0	0	0
RCA23_c22860	2437153	2437380	hypothetical protein	COG3311	K	GI 8	15	3	12
RCA23_c22870	2437482	2437922	hypothetical protein			GI 8	3	3	0
RCA23_c22880	2438000	2439085	hypothetical protein			GI 8	85	61	24
RCA23_c22890	2439085	2440413	DNA polymerase III	COG5545	R	GI 8	168	114	54
RCA23_c22900	2440954	2441754	hypothetical protein, periplasmic binding protein-like	COG2998	H	GI 8	14	7	7
RCA23_c22910	2441772	2441930	hypothetical protein			GI 8	126	134	-8
RCA23_c22920	2442714	2442007	transposase	COG3316	L	GI 8	1	0	1
RCA23_c22930	2444420	2443158	molybdopterin biosynthesis protein MoeA	COG0303	H	GI 8	53	66	-13
RCA23_c22940	2444914	2444417	molybdopterin-guanine dinucleotide biosynthesis protein Mo	COG1763	H	GI 8	82	88	-6
RCA23_c22950	2445231	2444911	hypothetical protein, MobA-like	COG0746	H	GI 8	207	166	41
RCA23_c22960	2446280	2445228	formate dehydrogenase family accessory protein FdhD	COG1526	C	GI 8	171	128	43
RCA23_c22970	2447181	2446738	molybdopterin-converting factor subunit MoaE	COG0314	H	GI 8	65	59	6
RCA23_c22980	2447430	2447185	molybdopterin-converting factor subunit MoaD	COG1977	H	GI 8	163	66	97
RCA23_c22990	2448117	2447674	transposase	COG3328	L	GI 8	158	124	34
RCA23_c23000	2449773	2448766	molybdenum cofactor biosynthesis protein MoaA	COG2896	H	GI 8	105	125	-20
RCA23_c23010	2450390	2450662	transposase			GI 8	163	189	-26
RCA23_c23020	2451451	2450738	ABC transporter, ATP-binding cassette protein	COG3839	G	GI 8	113	81	32
RCA23_c23030	2452152	2451448	ABC transporter, permease protein	COG4662	H	GI 8	107	61	46
RCA23_c23040	2453771	2452395	formate dehydrogenase, gamma subunit FdhI	COG2864	C	GI 8	213	101	112
RCA23_c23050	2454447	2453854	formate dehydrogenase, iron-sulfur subunit FdhB	COG0437	C	GI 8	155	92	63
RCA23_c23060	2457355	2454461	formate dehydrogenase alpha subunit FdhA	COG0243	C	GI 8	116	75	41
RCA23_c23070	2457663	2457472	putative twin-arginine translocation pathway signal sequence	domain		GI 8	62	49	13
RCA23_c23080	2458323	2457718	hypothetical protein, cytoplasmic chaperon TorD	COG3381	R	GI 8	126	84	42

RCA23_c23090	2458949	2458320	hypothetical protein			GI 8	196	135	61
RCA23_c23100	2459476	2458946	putative molybdopterin-guanine dinucleotide biosynthesis protein A			GI 8	117	86	31
RCA23_c23110	2459807	2461798	hypothetical protein, 4Fe-4S ferredoxin-like	COG1148	C	GI 8	51	69	-18
RCA23_c23120	2462517	2461792	hypothetical protein			GI 8	81	65	16
RCA23_c23130	2462613	2462804	hypothetical protein			GI 8	89	131	-42
RCA23_c23140	2462805	2463860	hypothetical protein, Mrp/NBP35 family protein	COG0489	D	GI 8	86	65	21
RCA23_c23150	2463857	2464570	hypothetical protein	COG0340	H	GI 8	67	78	-11
RCA23_c23160	2464572	2465081	hypothetical protein			GI 8	92	92	0
RCA23_c23170	2466057	2465644	putative phage terminase			GI 8	15	22	-7
RCA23_c23180	2467057	2466350	transposase	COG3316	L	GI 8	1	1	0
RCA23_c23190	2467131	2467988	putative phage helicase	COG3378	R	GI 8	70	48	22
RCA23_c23200	2468581	2468237	hypothetical protein			GI 8	73	97	-24
RCA23_c23210	2468997	2470031	putative pyridoxal 4-dehydrogenase	COG0667	C	GI 8	57	17	40
RCA23_c23220	2470845	2470078	2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase KduD	COG1028	I	GI 8	75	17	58
RCA23_c23230	2471994	2470855	putative L-rhamnonate dehydratase rhamD	COG4948	M	GI 8	64	17	47
RCA23_c23240	2472826	2471987	ureidoglycolate lyase	COG0179	Q	GI 8	90	30	60
RCA23_c23250	2473812	2472823	D-3-phosphoglycerate dehydrogenase	COG1052	C	GI 8	68	18	50
RCA23_c23260	2474558	2473809	uncharacterized oxidoreductase	COG1028	I	GI 8	35	19	16
RCA23_c23270	2474872	2474555	putative rhamnose mutarotase RhaM	COG3254	S	GI 8	51	30	21
RCA23_c23280	2475636	2474872	putative acetoacetate decarboxylase Adc	COG4689	Q	GI 8	88	30	58
RCA23_c23290	2476659	2475637	uncharacterized oxidoreductase	COG0673	R	GI 8	77	19	58
RCA23_c23300	2477416	2476673	transcriptional regulator, GntR family	COG2186	K	GI 8	81	23	58
RCA23_c23310	2478408	2477413	putative ribose ABC transport system, permease protein Rb	COG1172	G	GI 8	111	41	70
RCA23_c23320	2479472	2478405	putative ribose ABC transport system, permease protein Rb	COG1172	G	GI 8	87	30	57
RCA23_c23330	2480980	2479472	ribose import ABC transporter, ATP-binding protein RbsA	COG1129	G	GI 8	68	26	42
RCA23_c23340	2482093	2481056	putative rhamnose ABC transport system, substrate-binding	COG1879	G	GI 8	47	36	11
RCA23_c23350	2482310	2483446	galactonate dehydratase	COG4948	M	GI 8	46	19	27
RCA23_c23360	2483461	2484276	xylose isomerase-like	COG1082	G	GI 8	66	19	47
RCA23_c23370	2484419	2484691	putative aldo/keto reductase	COG0667	C	GI 8	56	23	33
RCA23_c23380	2484851	2485741	amidohydrolase	COG3618	R	GI 8	75	36	39

RCA23_c23390	2486071	2486280	hypothetical protein			GI 8	58	21	37
RCA23_c23400	2487635	2486928	transposase	COG3316	L	GI 8	31	25	6
RCA23_c23410	2488060	2487743	L-rhamnose mutarotase RhaM	COG3254	S	GI 8	80	22	58
RCA23_c23420	2489347	2488439	ureidoglycolate lyase	COG0179	Q	GI 8	23	20	3
RCA23_c23430	2490360	2489458	dihydrodipicolinate synthase DapA	COG0329	E	GI 8	43	31	12
RCA23_c23440	2491404	2490508	amidohydrolase	COG3618	R	GI 8	67	20	47
RCA23_c23450	2492228	2491446	3-ketoacyl-(acyl-carrier-protein) reductase FabG	COG1028	I	GI 8	85	52	33
RCA23_c23460	2493559	2492393	L-rhamnonate dehydratase RhmD	COG4948	M	GI 8	28	27	1
RCA23_c23470	2494711	2493587	polyamine ABC transporter, ATP-binding protein PotA	COG3842	E	GI 8	14	8	6
RCA23_c23480	2495571	2494786	polyamine ABC transporter, permease protein	COG1177	E	GI 8	22	19	3
RCA23_c23490	2496561	2495575	polyamine ABC transporter, permease protein	COG1176	E	GI 8	21	7	14
RCA23_c23500	2497743	2496676	polyamine ABC transporter, substrate binding protein	COG0687	E	GI 8	13	9	4
RCA23_c23510	2498385	2498131	hypothetical protein			GI 8	21	11	10
RCA23_c23520	2498724	2498401	hypothetical protein			GI 8	12	4	8
RCA23_c23530	2499506	2498760	uncharacterized oxidoreductase	COG1028	I	GI 8	31	16	15
RCA23_c23540	2500818	2499526	L-rhamnonate dehydratase	COG4948	M	GI 8	25	20	5
RCA23_c23550	2501838	2500861	ribose ABC transport system, permease protein RbsC	COG1172	G	GI 8	5	21	-16
RCA23_c23560	2502656	2501835	ribose import ABC transporter, ATP-binding protein RbsA	COG1129	G	GI 8	17	16	1
RCA23_c23570	2503752	2502640	ribose import ABC transporter, substrate binding protein Rb	COG1879	G	GI 8	14	13	1
RCA23_c23580	2503924	2504961	HTH-type transcriptional regulator, GntR family	COG1609	K	GI 8	18	19	-1
RCA23_c23590	2505104	2505304	hypothetical protein	COG3311	K	GI 8	39	9	30
RCA23_c23600	2506163	2506816	putative DNA-binding protein	COG1396	K	GI 8	29	12	17
RCA23_c23610	2507367	2508263	glutamine amidotransferase-like protein GlxB	COG0067	E	GI 8	41	12	29
RCA23_c23620	2508265	2508942	glutamate synthase alpha subunit GlxC	COG0070	E	GI 8	20	3	17
RCA23_c23630	2508954	2510294	glutamate synthase large subunit GlxD	COG0069	E	GI 8	18	6	12
RCA23_c23640	2510364	2511698	glutamine synthase GlnA type III	COG0174	E	GI 8	17	8	9
RCA23_c23650	2512409	2513053	hypothetical protein			GI 8	48	23	25
RCA23_c23660	2513221	2514510	methyltrasnferase	COG0863	L	GI 8	66	37	29
RCA23_c23670	2514507	2514890	hypothetical protein			GI 8	38	17	21
RCA23_c23680	2514880	2516313	putative terminase, large subunit	COG5410	S	GI 8	31	15	16

RCA23_c23690	2516395	2516823	hypothetical protein, DUF2924			GI 8	42	22	20
RCA23_c23700	2516820	2518436	hypothetical protein, resolvase-like	COG1961	L	GI 8	19	21	-2
RCA23_c23720	2520249	2518873	signal transduction histidine kinase	COG0642	T		324	351	-27
RCA23_c23730	2521684	2520269	mercuric reductase MerA	COG1249	C		253	349	-96
RCA23_c23740	2522418	2521687	hypothetical protein	COG0398	S		532	524	8
RCA23_c23750	2522803	2523192	putative ribonuclease P	COG0594	J		226	280	-54
RCA23_c23760	2523189	2523428	hypothetical protein DUF37	COG0759	S		269	308	-39
RCA23_c23770	2523749	2523964	hypothetical protein				216	397	-181
RCA23_c23780	2524020	2524874	tRNA 2-thiocytidine biosynthesis protein TtcA	COG0037	D		246	264	-18
RCA23_c23790	2524928	2525959	putative diguanylate phosphodiesterase	COG2200	T		373	385	-12
RCA23_c23800	2526067	2527851	inner membrane protein OxaA	COG0706	U		371	367	4
RCA23_c23810	2527853	2528602	hypothetical protein, molybdenum cofactor sulfurase	COG3217	R		301	323	-22
RCA23_c23820	2528599	2529249	putative GTP-binding protein EngB	COG0218	R		261	317	-56
RCA23_c23830	2529343	2530209	acetylglutamate kinase ArgB	COG0548	E		288	298	-10
RCA23_c23840	2530224	2531015	putative fatty acid hydroxylase	COG3000	I		407	521	-114
RCA23_c23850	2531012	2531632	hypothetical protein	COG1148	C		236	288	-52
RCA23_c23860	2531644	2532126	hypothetical protein, phosphogylcerate mutase	COG2062	T		239	338	-99
RCA23_c23870	2532914	2532150	glutamate/glutamine/aspartate/asparagine transport ATP-bi	COG1126	E		530	458	72
RCA23_c23880	2534220	2532928	glutamate/glutamine/aspartate/asparagine transport system	COG0765	E		785	600	185
RCA23_c23890	2535391	2534222	glutamate/glutamine/aspartate/asparagine transport system	COG4597	E		312	250	62
RCA23_c23900	2536594	2535548	glutamate/glutamine/aspartate/asparagine-binding protein E	COG0834	E		743	471	272
RCA23_c23910	2537482	2536778	ATP chaperone protein	COG5387	O		326	440	-114
RCA23_c23920	2538147	2537479	putative phosphoglycerate phosphatase	COG0546	R		309	416	-107
RCA23_c23930	2539190	2538144	ribosomal large subunit pseudouridine synthase C	COG0564	J		228	289	-61
RCA23_c23940	2539558	2539181	putative ccrb-like protein	COG0239	D		321	383	-62
RCA23_c23950	2540903	2539590	replication-associated recombination protein A	COG2256	L		230	384	-154
RCA23_c23960	2541581	2540970	HTH-type transcriptional regulator, LuxR family	COG2197	T		227	315	-88
RCA23_c23970	2542163	2541810	50S ribosomal protein L17	COG0203	J		383	328	55
RCA23_c23980	2543490	2542336	DNA-directed RNA polymerase alpha subunit RpoA	COG0202	K		941	507	434
RCA23_c23990	2543849	2543460	30S ribosomal protein S11	COG0100	J		1.196	675	521

RCA23_c24000	2544228	2543860	30S ribosomal protein S13	COG0099	J	1.607	755	852
RCA23_c24010	2545048	2544407	adenylate kinase Adk	COG0563	F	470	419	51
RCA23_c24020	2546403	2545045	preprotein translocase subunit SecY	COG0201	U	515	462	53
RCA23_c24030	2546975	2546508	50S ribosomal protein L15	COG0200	J	648	387	261
RCA23_c24040	2547397	2547209	50S ribosomal protein L30	COG1841	J	495	458	37
RCA23_c24050	2548008	2547409	30S ribosomal protein S5	COG0098	J	562	385	177
RCA23_c24060	2548445	2548086	50S ribosomal protein L18	COG0256	J	613	407	206
RCA23_c24070	2548990	2548457	50S ribosomal protein L6	COG0097	J	960	459	501
RCA23_c24080	2549392	2549000	30S ribosomal protein S8	COG0096	J	851	363	488
RCA23_c24090	2549710	2549405	30S ribosomal protein S14	COG0199	J	1.103	430	673
RCA23_c24100	2550282	2549722	50S ribosomal protein L5	COG0094	J	1.295	531	764
RCA23_c24110	2550587	2550282	50S ribosomal protein L24	COG0198	J	1.045	441	604
RCA23_c24120	2550957	2550589	50S ribosomal protein L14	COG0093	J	1.214	626	588
RCA23_c24130	2551255	2551025	30S ribosomal protein S17	COG0186	J	1.128	542	586
RCA23_c24140	2551467	2551261	50S ribosomal protein L29	COG0255	J	919	575	344
RCA23_c24150	2551692	2552312	hypothetical protein			491	514	-23
RCA23_c24160	2552703	2552407	50S ribosomal protein L23	COG0089	J	913	531	382
RCA23_c24170	2553317	2552700	50S ribosomal protein L4	COG0088	J	843	572	271
RCA23_c24180	2554045	2553314	50S ribosomal protein L3	COG0087	J	919	565	354
RCA23_c24190	2554368	2554060	30S ribosomal protein S10	COG0051	J	561	460	101
RCA23_c24200	2555649	2554474	elongation factor Tu (EF-Tu)	COG0050	J	41	15	26
RCA23_c24210	2557851	2555734	elongation factor FusA	COG0480	J	997	541	456
RCA23_c24220	2558343	2557873	30S ribosomal protein S7	COG0049	J	558	374	184
RCA23_c24230	2558583	2558356	30S ribosomal protein S12	COG0048	J	966	883	83
RCA23_c24240	2559706	2559092	hypothetical protein	COG1072	H	197	222	-25
RCA23_c24250	2560443	2559703	ABC transporter ATP-binding protein	COG1129	G	377	332	45
RCA23_c24260	2561552	2560479	ABC tansporter permease protein	COG1172	G	316	328	-12
RCA23_c24270	2562676	2561672	putative ABC transporter periplasmic binding protein	COG1879	G	376	304	72
RCA23_c24280	2562858	2564048	putative transcriptional repressor	COG1940	K	257	274	-17
RCA23_c24290	2564800	2564060	hypothetical protein, acyl-CoA thioesterase-like	COG2755	E	330	432	-102

RCA23_c24300	2564921	2565451	ABC transporter ATP-binding protein	COG1136	V		271	378	-107
RCA23_c24310	2565448	2567970	putative ABC transporter permease protein	COG3127	Q		212	276	-64
RCA23_c24320	2568328	2569251	hypothetical protein, gylcosyl transferase family 2	COG0463	M		218	237	-19
RCA23_c24330	2570639	2569266	aminotransferase class-III	COG0161	H		358	316	42
RCA23_c24340	2571879	2570683	putative ectoine utilization protein EutD	COG0006	E		406	426	-20
RCA23_c24350	2572018	2572968	putative integral membrane proein DUF6				260	386	-126
RCA23_c24360	2573067	2574635	4-coumarate--CoA ligase	COG0318	I		229	302	-73
RCA23_c24370	2575672	2574737	hypothetical protein	COG0679	R		315	415	-100
RCA23_c24380	2576572	2575760	hypothetical protein DUF125	COG1814	S		191	301	-110
RCA23_c24390	2577015	2576587	transcriptional regulator, AsnC family	COG1522	K		269	276	-7
RCA23_c24400	2577133	2578062	arginase ArcA	COG0010	E		294	304	-10
RCA23_c24410	2578064	2579122	ornithine cyclodeaminase ArcB	COG2423	E		294	233	61
RCA23_c24420	2579119	2582481	bifunctional protein PutA	COG4230	C		254	302	-48
RCA23_c24430	2583144	2582581	hypothetical protein				459	415	44
RCA23_c24440	2583406	2583128	hypothetical protein				349	222	127
RCA23_c24460	2583863	2583615	hypothetical protein				249	201	48
RCA23_c24450	2583834	2585051	creatinase	COG0006	E		209	197	12
RCA23_c24470	2585998	2585276	putative hydrolase	COG0596	R		44	43	1
RCA23_c24480	2587548	2585995	glycine betaine/L-proline transport system permease protein	COG4176	E		108	62	46
RCA23_c24490	2588600	2587545	glycine betaine/L-proline transport ATP-binding protein ProV	COG4175	E		112	59	53
RCA23_c24500	2589635	2588670	glycine betaine-binding periplasmic protein ProX	COG2113	E		102	46	56
RCA23_c24510	2589879	2590763	putative HTH-type transcriptional regluator, LysR family	COG0583	K		244	166	78
RCA23_c24520	2591855	2590770	D-cysteine desulfhydrase DcyD	COG2515	E		285	248	37
RCA23_c24530	2591995	2592351	hypothetical protein				794	435	359
RCA23_c24540	2593298	2592375	high-affinity zinc uptake system protein ZnuA	COG4531	P		270	294	-24
RCA23_c24550	2593355	2593870	zinc uptake regulator	COG0735	P		213	353	-140
RCA23_c24560	2593867	2594616	zinc import ATP-binding protein ZnuC	COG1121	P		361	352	9
RCA23_c24570	2594616	2595410	high-affinity zinc uptake system membrane protein ZnuB	COG1108	P		234	351	-117
RCA23_c24580	2595534	2597435	4-hydroxyphenylpyruvate dioxygenase Hpd	COG3185	E		358	304	54
RCA23_c24590	2598436	2598104	hypothetical protein			GI 9	626	538	88

RCA23_c24600	2598748	2598503	hypothetical protein			GI 9	0	3	-3
RCA23_c24610	2599041	2599844	hypothetical protein	COG1028	I	GI 9	169	138	31
RCA23_c24620	2599863	2600606	hypothetical protein	COG1402	R	GI 9	102	95	7
RCA23_c24630	2601568	2600642	putative gluconolactonase	COG3386	G	GI 9	198	166	32
RCA23_c24640	2602967	2602047	ribokinase RbsK	COG0524	G	GI 9	266	185	81
RCA23_c24650	2603416	2602964	ribose ABC transporter protein RbsD	COG4154	G	GI 9	233	166	67
RCA23_c24660	2604556	2603489	putative ribose ABC transporter, ATP-binding protein RbsA	COG3839	G	GI 9	298	217	81
RCA23_c24670	2604859	2606439	putative ribose ABC transporter, substrate binding protein R	COG1653	G	GI 9	236	184	52
RCA23_c24680	2606509	2607615	putative ribose ABC transporter, permease protein RbsC	COG1175	G	GI 9	314	141	173
RCA23_c24690	2607617	2608534	putative ribose ABC transporter, permease protein RbsC	COG0395	G	GI 9	231	152	79
RCA23_c24700	2610024	2608537	hypothetical protein, L-fucose isomerase-like			GI 9	155	137	18
RCA23_c24710	2610626	2610162	sugar isomerase	COG0794	M	GI 9	162	121	41
RCA23_c24720	2612206	2610716	xylulose kinase XylB	COG1070	G	GI 9	199	157	42
RCA23_c24730	2613273	2612203	aminopeptidase SgcX	COG1363	G	GI 9	223	193	30
RCA23_c24740	2614073	2613270	photosystem I biogenesis protein BtpA	COG0434	R	GI 9	253	182	71
RCA23_c24750	2614166	2615188	pyridoxal 4-dehydrogenase Pld	COG0667	C	GI 9	294	224	70
RCA23_c24760	2615185	2616021	putative amidohydrolase	COG3618	R	GI 9	205	195	10
RCA23_c24770	2616791	2616009	putative carbohydrate kinase, pfkB family	COG0524	G	GI 9	131	162	-31
RCA23_c24780	2617753	2616854	xylose isomerase family protein	COG4952	M	GI 9	185	163	22
RCA23_c24790	2618529	2617822	putative HTH-type transcriptional regulator, GntR family	COG1802	K	GI 9	254	163	91
RCA23_c24800	2620064	2618739	mandelate racemase/muconate lactonizing enzyme	COG4948	M	GI 9	222	159	63
RCA23_c24810	2620722	2621792	fatty acid desaturase	COG3239	I	GI 9	209	154	55
RCA23_c24820	2622254	2621892	hypothetical protein			GI 9	16	5	11
RCA23_c24830	2623076	2622369	transposase	COG3316	L	GI 9	43	27	16
RCA23_c24840	2624076	2623558	transporter, LysE family	COG1280	E	GI 9	0	1	-1
RCA23_c24850	2624531	2624851	hypothetical protein			GI 9	26	35	-9
RCA23_c24860	2625998	2624925	putative transporter, periplasmic binding protein	COG2358	R	GI 9	8	1	7
RCA23_c24870	2626948	2627676	ribitol 2-dehydrogenase RbtD	COG4221	R	GI 9	50	26	24
RCA23_c24880	2631855	2628631	hypothetical protein			GI 9	26	16	10
RCA23_c24890	2631920	2634466	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltra	COG0457	R	GI 9	52	46	6

RCA23_c24900	2636319	2635342	putative phage integrase	COG4974	L	GI 9	100	70	30
RCA23_c24920	2636842	2638131	D-alanyl-D-alanine carboxypeptidase DacC	COG1686	M		251	324	-73
RCA23_c24930	2638132	2638746	thymidylate kinase Tmk	COG0125	F		197	301	-104
RCA23_c24940	2638743	2639867	hypothetical protein, DNA polymerase III delta subunit	COG2812	L		192	277	-85
RCA23_c24950	2639864	2640652	TatD family deoxyribonuclease	COG0084	L		240	435	-195
RCA23_c24960	2640649	2641437	hypothetical protein, metallo-beta-lactamase	COG1235	R		292	403	-111
RCA23_c24970	2641439	2642368	membrane transport protein	COG0679	R		307	401	-94
RCA23_c24980	2643457	2642372	FAD dependent oxidoreductase	COG0665	E		183	305	-122
RCA23_c24990	2643645	2645834	UvrABC system protein B	COG0556	L		246	358	-112
RCA23_c25000	2647627	2645909	hypothetical protein				553	485	68
RCA23_c25010	2648247	2647642	hypothetical protein				444	381	63
RCA23_c25020	2648732	2648331	hypothetical protein				640	595	45
RCA23_c25030	2650370	2648796	lysyl-tRNA synthase LysS	COG1384	J		506	461	45
RCA23_c25040	2650510	2650884	hypothetical protein				359	434	-75
RCA23_c25050	2652333	2650876	putative D-Ala-D-Ala carboxypeptidase 3 (S13) family prote	COG2027	M		272	431	-159
RCA23_c25060	2652562	2653368	hypothetical protein				240	271	-31
RCA23_c25070	2653967	2653365	nicotinic acid mononucleotide adenyltransferase NadD	COG1057	H		207	270	-63
RCA23_c25080	2654128	2655756	ABC transporter, ATP-binding cassette protein, ChvD family	COG0488	R		488	465	23
RCA23_c25090	2655999	2656205	cold shock protein CspA	COG1278	K		304	373	-69
RCA23_c25100	2657646	2656327	glutamyl-tRNA synthase 1	COG0008	J		276	321	-45
RCA23_c25110	2659351	2657714	glutamine-dependent NAD(+) synthase NadE	COG0171	H		286	411	-125
RCA23_c25120	2659555	2660997	putative phosphatidylinositol-4-phosphate 5-kinase	COG4642	S		504	467	37
RCA23_c25130	2661772	2661011	hypothetical protein	COG1028	I		216	333	-117
RCA23_c25140	2661935	2663500	2-isopropylmalate synthase LeuA	COG0119	E		392	420	-28
RCA23_c25150	2663641	2664678	rod shape-determining protein MreB	COG1077	D		324	359	-35
RCA23_c25160	2664709	2665593	rod shape-determining protein MreC	COG1792	M		222	292	-70
RCA23_c25170	2665590	2666129	hypothetical protein				310	423	-113
RCA23_c25180	2666126	2668078	penicillin-binding protein 2	COG0768	M		309	352	-43
RCA23_c25190	2668075	2669226	rod shape-determining protein RodA	COG0772	D		498	483	15
RCA23_c25200	2669223	2670152	glyoxylate/hydroxypyruvate reductase GhrA	COG0111	H		293	472	-179

RCA23_c25210	2670938	2670162	hypothetical protein			226	269	-43
RCA23_c25220	2671545	2670925	hypothetical protein	COG1573	L	248	272	-24
RCA23_c25230	2671592	2672890	Na+/H+ antiporter NhaA	COG3004	P	338	391	-53
RCA23_c25240	2673062	2673721	capsule polysaccharide export protein KpsT	COG1134	G	568	561	7
RCA23_c25250	2673699	2675342	capsule polysaccharide export protein KpsE	COG3524	M	361	359	2
RCA23_c25270	2675994	2675545	hypothetical protein	COG1832	R	412	456	-44
RCA23_c25280	2676353	2675997	ferredoxin PetF	COG0633	C	421	371	50
RCA23_c25290	2677702	2676350	selenium binding protein			596	556	40
RCA23_c25300	2679228	2677789	aldehyde dehydrogenase	COG1012	C	264	321	-57
RCA23_c25310	2680435	2679296	CoA-transferase family III protein involved in DMSP degrad	COG1804	C	268	406	-138
RCA23_c25320	2680698	2681036	nitrogen regulatory protein P-II 2	COG0347	E	654	505	149
RCA23_c25330	2681036	2682373	ammonium transporter	COG0004	P	507	419	88
RCA23_c25340	2682887	2682426	homoprotocatechuate degradative operon repressor	COG1846	K	311	282	29
RCA23_c25350	2683007	2684515	5-carboxymethyl-2-hydroxymuconic semialdehyde dehydro	COG1012	C	402	427	-25
RCA23_c25360	2684539	2685522	3,4-dihydroxyphenylacetate 2,3-dioxygenase HpaD	COG0346	E	355	344	11
RCA23_c25370	2686538	2685534	3-carboxy-cis,cis-muconate cycloisomerase PcaB	COG0015	F	222	260	-38
RCA23_c25380	2687377	2686535	hypothetical protein DUF849	COG3246	S	286	319	-33
RCA23_c25390	2687930	2687370	protocatechuate 3,4-dioxygenase alpha chain PcaG	COG3485	Q	216	313	-97
RCA23_c25400	2688658	2687930	protocatechuate 3,4-dioxygenase beta chain PcaH	COG3485	Q	172	217	-45
RCA23_c25410	2689041	2688655	4-carboxymuconolactone decarboxylase PcaC	COG0599	S	193	228	-35
RCA23_c25420	2689451	2689260	hypothetical protein			343	373	-30
RCA23_c25430	2689498	2690184	transcriptional regulatory protein	COG0745	T	365	418	-53
RCA23_c25440	2690190	2691731	two component signal transduction histidine kinase ChvG	COG0642	T	479	455	24
RCA23_c25450	2691980	2691744	hypothetical protein			545	451	94
RCA23_c25460	2694029	2692536	Na(+)-phosphate symporter Pit	COG0306	P	375	421	-46
RCA23_c25470	2694422	2694246	hypothetical protein			625	496	129
RCA23_c25480	2695336	2694491	2-dehydro-3-deoxyphosphooctonate aldolase KdsA	COG2877	M	308	347	-39
RCA23_c25490	2695480	2696481	arabinose 5-phosphate isomerase KdsD	COG0794	M	281	328	-47
RCA23_c25500	2697297	2696491	3-deoxy-manno-octulosonate cytidylyltransferase KdsB	COG1212	M	476	394	82
RCA23_c25510	2698445	2697384	hypothetical protein			122	146	-24

RCA23_c25520	2703372	2698459	bacterial surface protein			86	44	42
RCA23_c25530	2703948	2706137	catalase-peroxidase KatG	COG0376	P	267	267	0
RCA23_c25540	2706353	2707432	hypothetical protein	COG2067	I	134	138	-4
RCA23_c25550	2707908	2707480	thioesterase/thiol ester dehydrase-isomerase	COG0824	R	50	135	-85
RCA23_c25560	2709059	2707905	iron-containing alcohol dehydrogenase	COG1454	C	136	190	-54
RCA23_c25570	2710310	2709075	hypothetical protein	COG1960	I	224	359	-135
RCA23_c25580	2710403	2710906	nodulation protein N-like protein	COG2030	I	274	303	-29
RCA23_c25590	2710903	2711943	putative phosphotransferase, eukaryotic acyl-CoA dehydrog	COG3173	R	286	369	-83
RCA23_c25600	2711940	2712548	putative HTH-type transcriptional regulator, TetR family	COG1309	K	280	365	-85
RCA23_c25610	2712623	2714713	fatty acid oxidation complex alpha subunit FadJ	COG1250	I	230	283	-53
RCA23_c25620	2714718	2715914	acyl-CoA dehydrogenase	COG1960	I	529	366	163
RCA23_c25630	2715918	2717012	acyl-CoA dehydrogenase	COG1960	I	266	245	21
RCA23_c25640	2717231	2717998	hypothetical protein, MarR family			285	270	15
RCA23_c25650	2718086	2719246	acyl-CoA dehydrogenase	COG1960	I	495	432	63
RCA23_c25660	2720282	2719305	quinone oxidoreductase (1.6.5.5.)	COG0604	C	467	472	-5
RCA23_c25670	2720400	2720990	hypothetical protein			218	348	-130
RCA23_c25680	2721719	2720952	putative DNA repair protein RecO	COG1381	L	297	529	-232
RCA23_c25690	2722155	2721823	hypothetical protein	COG5447	S	328	366	-38
RCA23_c25700	2723060	2722152	GTP-binding protein Era	COG1159	R	435	524	-89
RCA23_c25710	2723740	2723057	ribonuclease 3	COG0571	K	271	398	-127
RCA23_c25720	2724636	2723767	signal peptidase I	COG0681	U	348	422	-74
RCA23_c25730	2725101	2724688	holo-[acyl-carrier-protein] synthase AcpS	COG0736	I	188	392	-204
RCA23_c25740	2725841	2725098	pyridoxine 5'-phosphate synthase PdxJ	COG0854	H	284	320	-36
RCA23_c25750	2726515	2725865	hypothetical protein DUF2062	COG3216	S	343	410	-67
RCA23_c25760	2728673	2726508	guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase	COG0317	T	439	454	-15
RCA23_c25770	2729039	2728686	DNA-directed RNA polymerase subunit omega	COG1758	K	1.123	669	454
RCA23_c25780	2729719	2729135	putative 2-amino-4-hydroxy-6-hydroxymethyldihydropteridin	COG0801	H	399	370	29
RCA23_c25790	2729822	2730388	hypothetical protein DUF88	COG1432	S	663	386	277
RCA23_c25800	2730453	2731418	4-hydroxy-3-methylbut-2-enyl diphosphate reductase IspH	COG0761	I	269	305	-36
RCA23_c25810	2731411	2731998	hypothetical protein, methyltransferase	COG2227	H	203	289	-86

RCA23_c25820	2731991	2732443	ribonuclease H	COG0328	L		255	386	-131
RCA23_c25830	2732467	2732703	hypothetical protein				262	428	-166
RCA23_c25840	2733584	2732700	methionyl-tRNA formyltransferase Fmt	COG0223	J		177	313	-136
RCA23_c25850	2734084	2733590	peptide deformylase Def	COG0242	J		298	457	-159
RCA23_c25860	2734710	2734081	peptide deformylase Def	COG0242	J		400	459	-59
RCA23_c25870	2734743	2735909	aminotransferase	COG1168	E		342	372	-30
RCA23_c25880	2736272	2735922	hypothetical protein				315	306	9
RCA23_c25890	2737133	2736360	precorrin-4 C(11)-methyltransferase CobM	COG2875	H		269	297	-28
RCA23_c25900	2738971	2737130	precorrin-3B C(17)-methyltransferase CobJ	COG1010	H		190	309	-119
RCA23_c25910	2739663	2738962	precorrin-2 C(20)-methyltransferase CobI	COG2243	H		215	381	-166
RCA23_c25920	2740859	2739660	precorrin-6Y C(5,15)-methyltransferase CobL	COG2242	H		190	368	-178
RCA23_c25930	2741494	2740856	precorrin-8X methylmutase CobH	COG2082	H		243	357	-114
RCA23_c25940	2742653	2741541	sirohydrochlorin cobaltochelatase CbiX	COG2138	S		330	408	-78
RCA23_c25950	2743257	2744771	methyltransferase, FkbM family	COG3774	M		151	145	6
RCA23_c25960	2744880	2745842	hypothetical protein, glutathione S-transferase	COG0435	O		555	486	69
RCA23_c25970	2745920	2746291	hypothetical protein DUF1636	COG5469	S		250	284	-34
RCA23_c25980	2746291	2747238	threonine-phosphate decarboxylase CobC	COG0079	E		212	334	-122
RCA23_c25990	2747235	2748140	cobalamin biosynthesis protein CobD	COG1270	H		288	321	-33
RCA23_c26000	2748210	2749349	putative peptidoglycan-binding lytic murein transglycosylase	COG2951	M		269	370	-101
RCA23_c26010	2752960	2749505	putative chromosome partition protein smc	COG1196	D		238	324	-86
RCA23_c26020	2753354	2753145	hypothetical protein				402	371	31
RCA23_c26030	2753524	2754378	hypothetical protein	COG1028	I		304	334	-30
RCA23_c26040	2754434	2755204	hypothetical protein DUF81	COG0730	R		615	430	185
RCA23_c26050	2755268	2756272	50S ribosomal protein L2	COG0090	J		670	458	212
RCA23_c26060	2756318	2756554	30S ribosomal protein S19	COG0185	J		875	505	370
RCA23_c26070	2756558	2756938	50S ribosomal protein L22	COG0091	J		1.115	598	517
RCA23_c26080	2756938	2757657	30S ribosomal protein S3	COG0092	J		748	364	384
RCA23_c26090	2757702	2758082	50S ribosomal protein L16	COG0197	J		391	269	122
RCA23_c26100	2758763	2758521	hypothetical protein			GI 10	525	482	43
RCA23_c26120	2760596	2759328	hypothetcial protein	COG0666	R	GI 10	10	9	1

RCA23_c26130	2761026	2760724	transcriptional regulator			GI 10	245	94	151
RCA23_c26140	2761304	2761564	hypothetical protein			GI 10	158	107	51
RCA23_c26150	2762158	2761697	hypothetical protein, UPF0311			GI 10	298	163	135
RCA23_c26160	2762238	2762687	transcriptional regulator, MarR family	COG1846	K	GI 10	324	211	113
RCA23_c26170	2762756	2764528	long-chain-fatty-acid-CoA-ligase	COG0318	I	GI 10	171	159	12
RCA23_c26180	2765853	2764525	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	GI 10	163	218	-55
RCA23_c26190	2766366	2765854	TRAP dicarboxylate transporter, subunit DctQ			GI 10	225	194	31
RCA23_c26200	2767464	2766427	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	GI 10	171	138	33
RCA23_c26210	2768296	2767511	3-hydroxybutyryl-CoA dehydratase Crt	COG1024	I	GI 10	128	207	-79
RCA23_c26220	2768372	2769280	metal-dependent hydrolase	COG2159	R	GI 10	381	230	151
RCA23_c26230	2769438	2770715	FAD dependent monooxygenase	COG0654	H	GI 10	204	180	24
RCA23_c26240	2772437	2770791	arylsulfatase	COG3119	P	GI 10	137	165	-28
RCA23_c26250	2772940	2772434	transcriptional regulator, MarR family	COG1846	K	GI 10	149	189	-40
RCA23_c26260	2773067	2774122	putative transporter, periplasmic binding protein	COG2358	R	GI 10	113	141	-28
RCA23_c26270	2774186	2776480	TRAP transporter, 4TM/12TM fusion protein	COG4666	R	GI 10	256	221	35
RCA23_c26280	2778081	2776852	hypothetical protein, UPF0261	COG5441	S	GI 10	500	428	72
RCA23_c26290	2779513	2778236	gamma-glutamylputrescine synthase PuuA	COG0174	E	GI 10	493	386	107
RCA23_c26300	2780750	2779515	cytochrome P450	COG2124	Q	GI 10	546	377	169
RCA23_c26310	2781471	2780740	glutamine amidotransferase class-I	COG0518	F	GI 10	630	421	209
RCA23_c26320	2782504	2781536	tripartite tricarboxylate transporter (TTT) protein TctC	COG3181	S	GI 10	491	360	131
RCA23_c26330	2783997	2782504	tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S	GI 10	471	384	87
RCA23_c26340	2784742	2784008	hypothetical protein, transmembrane			GI 10	794	831	-37
RCA23_c26350	2786274	2784739	gamma-glutamyl-gamma-aminobutyraldehyde dehydrogena	COG1012	C	GI 10	526	374	152
RCA23_c26360	2786992	2787321	hypothetical protein			GI 10	20	18	2
RCA23_c26370	2787569	2787901	hypothetical protein			GI 10	52	13	39
RCA23_c26380	2787903	2788388	hypothetical protein			GI 10	18	11	7
RCA23_c26390	2789043	2790062	putative fucosyltransferase			GI 10	54	41	13
RCA23_c26400	2790398	2790138	integrase	COG2801	L	GI 10	156	153	3
RCA23_c26410	2791528	2790497	putative extracellular solute-binding protein	COG1840	P	GI 10	271	192	79
RCA23_c26420	2792889	2791498	two-component system, sensor histidine kinase protein	COG0642	T	GI 10	137	118	19

RCA23_c26430	2793554	2792886	two-component system, response regulator protein	COG0745	T	GI 10	214	156	58
RCA23_c26440	2793656	2794624	tripartite tricarboxylate transporter (TTT) protein TctC	COG3181	S	GI 10	713	468	245
RCA23_c26450	2794624	2795112	tripartite tricarboxylate transporter (TTT) protein TctB			GI 10	634	402	232
RCA23_c26460	2795114	2796619	tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S	GI 10	592	369	223
RCA23_c26470	2796621	2797979	hypothetical protein			GI 10	474	353	121
RCA23_c26480	2797972	2798277	hypothetical protein			GI 10	592	419	173
RCA23_c26490	2798419	2798775	transposase A			GI 10	341	254	87
RCA23_c26510	2799713	2799357	hypothetical protein			GI 10	1	0	1
RCA23_c26520	2800934	2801293	hypothetical protein	COG1709	K	GI 10	0	4	-4
RCA23_c26530	2802543	2802343	hypothetical protein			GI 10	0	0	0
RCA23_c26540	2806319	2802786	putative RTX toxin and hemolysin-type calcium binding protein			GI 10	1	2	-1
RCA23_c26550	2809639	2806427	putative RTX toxin and hemolysin-type calcium binding protein			GI 10	89	22	67
RCA23_c26560	2809943	2811319	type I secretion outer membrane protein, TolC family	COG1538	M	GI 10	267	163	104
RCA23_c26570	2811316	2813484	type I secretion system ATP-binding component	COG2274	V	GI 10	378	289	89
RCA23_c26580	2813500	2814810	type I RTX secretion system membrane fusion protein, HlyD	COG1566	V	GI 10	493	289	204
RCA23_c26590	2814853	2815215	hypothetical protein			GI 10	680	516	164
RCA23_c26600	2815394	2815645	hypothetical protein			GI 10	1.043	506	537
RCA23_c26610	2815655	2815879	hypothetical protein			GI 10	1.033	554	479
RCA23_c26620	2815898	2816089	hypothetical protein			GI 10	1.041	460	581
RCA23_c26630	2816211	2817194	hypothetical protein DUF2125			GI 10	666	361	305
RCA23_c26640	2817683	2818480	Asp/Glu racemase	COG3473	Q	GI 10	500	349	151
RCA23_c26650	2818577	2819731	cystathionine beta-lyase	COG0626	E	GI 10	549	536	13
RCA23_c26660	2820022	2821635	deoxyribodipyrimidine photo-lyase PhrB	COG0415	L	GI 10	921	514	407
RCA23_c26670	2821821	2822804	hypothetical protein	COG3380	R	GI 10	564	376	188
RCA23_c26680	2823798	2822806	O-sialoglycoprotein endopeptidase Gcp	COG0533	O	GI 10	284	330	-46
RCA23_c26690	2824420	2823791	Sua5/YciO/YrdC/YwlC family protein	COG0009	J	GI 10	299	317	-18
RCA23_c26700	2825070	2824417	putative glycoprotease family protein	COG1214	O	GI 10	294	484	-190
RCA23_c26710	2825579	2825061	hypothetical protein UPF0079	COG0802	R	GI 10	267	388	-121
RCA23_c26720	2825795	2826562	xylose isomerase family protein	COG3622	G	GI 10	264	388	-124
RCA23_c26730	2826823	2828073	putative phage integrase	COG0582	L	GI 10	171	127	44

RCA23_c26740	2828066	2828713	hypothetical protein			GI 10	240	116	124
RCA23_c26750	2828793	2828993	putative prophage regulatory protein	COG3311	K	GI 10	75	92	-17
RCA23_c26760	2828990	2829565	hypothetical protein			GI 10	99	70	29
RCA23_c26770	2829691	2831313	DNA polymerase	COG0749	L	GI 10	144	96	48
RCA23_c26780	2832293	2831898	hypothetical protein			GI 10	65	51	14
RCA23_c26790	2833890	2832826	hypothetical protein, snoaL-like polyketide cyclase	COG5485	R	GI 10	73	81	-8
RCA23_c26800	2834472	2834092	hypothetical protein			GI 10	105	69	36
RCA23_c26810	2834841	2834566	glyoxalase/bleomycin resistance protein/dioxygenase superfamily protein			GI 10	58	114	-56
RCA23_c26820	2835971	2834964	hypothetical protein			GI 10	55	74	-19
RCA23_c26830	2836859	2835975	putative ABC transporter permease protein	COG0395	G	GI 10	92	82	10
RCA23_c26840	2837837	2836863	putative ABC transporter permease protein	COG1175	G	GI 10	79	120	-41
RCA23_c26850	2839229	2837985	putative ABC transporter extracellular solute binding protein	COG1653	G	GI 10	74	71	3
RCA23_c26860	2839509	2840555	ABC transporter ATP-binding protein	COG3839	G	GI 10	64	75	-11
RCA23_c26870	2840606	2841937	histidinol dehydrogenase HisD	COG0141	E	GI 10	55	65	-10
RCA23_c26880	2841934	2842959	hypothetical protein, snoaL-like polyketide cyclase	COG5485	R	GI 10	39	67	-28
RCA23_c26890	2842980	2843720	short chain dehydrogenase	COG1028	I	GI 10	32	52	-20
RCA23_c26900	2843720	2844733	putative HTH-type transcriptional regulator LacI family	COG1609	K	GI 10	46	72	-26
RCA23_c26910	2844730	2845731	hypothetical protein, snoaL-like polyketide cyclase			GI 10	44	75	-31
RCA23_c26920	2845728	2846693	hypothetical protein			GI 10	58	93	-35
RCA23_c26930	2846826	2847896	hypothetical protein, 2-hydroxypropyl-CoM lyase-like	COG0620	E	GI 10	53	79	-26
RCA23_c26940	2847893	2848561	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	GI 10	80	90	-10
RCA23_c26950	2848567	2849259	hypothetical protein	COG0684	H	GI 10	49	85	-36
RCA23_c26960	2850269	2851174	3-hydroxyisobutyrate dehydrogenase MmsB	COG2084	I	GI 10	171	236	-65
RCA23_c26970	2851171	2852874	dihydroxy-acid dehydratase IlvD	COG0129	E	GI 10	91	124	-33
RCA23_c26980	2852884	2854224	hypothetical protein	COG4091	E	GI 10	80	132	-52
RCA23_c26990	2854264	2855901	putative choline (or alcohol) dehydrogenase	COG2303	E	GI 10	82	154	-72
RCA23_c27000	2855926	2857404	aldehyde dehydrogenase	COG1012	C	GI 10	132	164	-32
RCA23_c27010	2858020	2857418	transporter, LysE family	COG1280	E	GI 10	78	82	-4
RCA23_c27020	2859546	2858080	TRAP dicarboxylate transporter, subunit DctM	COG4664	Q	GI 10	186	208	-22
RCA23_c27030	2860109	2859546	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	GI 10	193	194	-1

RCA23_c27040	2861235	2860195	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	GI 10	219	154	65
RCA23_c27050	2862106	2861444	transcriptional regulator, LysR family	COG0583	K	GI 10	265	258	7
RCA23_c27060	2864132	2863029	mandelate racemase/muconate lactonizing enzyme	COG4948	M	GI 10	75	142	-67
RCA23_c27070	2865241	2864129	Zn-dependant oxidoreductase	COG0604	C	GI 10	72	182	-110
RCA23_c27090	2866870	2865992	HTH-type transcriptional regulator, LysR family	COG0583	K	GI 10	109	112	-3
RCA23_c27100	2866984	2867739	class II aldolase	COG0235	G	GI 10	127	145	-18
RCA23_c27110	2867742	2868770	fatty acid desaturase	COG3239	I	GI 10	59	122	-63
RCA23_c27120	2868772	2869620	3-methyl-2-oxobutanoate hydroxymethyltransferase PanB	COG0413	H	GI 10	113	150	-37
RCA23_c27130	2869651	2869959	putative rieske [2Fe-2S] protein	COG2146	P	GI 10	181	173	8
RCA23_c27150	2871527	2870697	hypothetical protein	COG1396	K		597	445	152
RCA23_c27160	2872169	2873122	arginase family protein	COG0010	E		371	394	-23
RCA23_c27170	2873972	2873130	hypothetical protein	COG1082	G		609	538	71
RCA23_c27180	2874311	2875399	putative HTH-type transcriptional regulator, AraC family	COG2169	F		277	286	-9
RCA23_c27190	2876347	2875367	hypothetical protein. fatty acid desaturase	COG3239	I		497	509	-12
RCA23_c27200	2878323	2876440	AMP-dependent synthase / ligase	COG0318	I		392	482	-90
RCA23_c27210	2880212	2878422	ABC transporter ATP-binding protein	COG1132	V		279	330	-51
RCA23_c27220	2880650	2881150	spore coat family protein				310	326	-16
RCA23_c27230	2881370	2882092	PapD-like chaperone involved in fimbrial biogenesis	COG3121	N		313	228	85
RCA23_c27240	2882279	2884033	fimbrial biogenesis outer membrane usher protein	COG3188	N		235	333	-98
RCA23_c27250	2884043	2884906	fimbrial biogenesis outer membrane usher protein	COG3188	N		320	412	-92
RCA23_c27260	2885061	2885357	hypothetical protein				275	288	-13
RCA23_c27280	2886473	2885598	hypothetical protein	COG0697	G		437	532	-95
RCA23_c27290	2890851	2886601	DNA-directed RNA polymerase beta subunit RpoC	COG0086	K		605	404	201
RCA23_c27300	2895044	2890911	DNA-directed RNA polymerase beta subunit RpoB	COG0085	K		596	421	175
RCA23_c27310	2895693	2895316	50S ribosomal protein L7/L12	COG0222	J		1.149	666	483
RCA23_c27320	2896250	2895762	50S ribosomal protein L10	COG0244	J		1.251	786	465
RCA23_c27330	2897346	2896648	50S ribosomal protein L1	COG0081	J		541	451	90
RCA23_c27340	2897821	2897348	50S ribosomal protein L11	COG0080	J		723	502	221
RCA23_c27350	2898476	2897928	transcription antitermination protein NusG	COG0250	K		711	380	331
RCA23_c27360	2898893	2898627	preprotein translocase, subunit SecE	COG0690	U		366	293	73

RCA23_c27370	2899800	2898895	hypothetical protein			315	334	-19
RCA23_c27430	2905830	2905985	hypothetical protein			433	427	6
RCA23_c27440	2906449	2906021	hypothetical protein, transmembrane			477	559	-82
RCA23_c27450	2907137	2906511	hypothetical protein			166	344	-178
RCA23_c27460	2908720	2907134	putative acetolactate synthase isozyme 1 large subunit	COG0028	E	255	334	-79
RCA23_c27470	2910168	2908720	putative cation transporter	COG0168	P	535	514	21
RCA23_c27480	2911228	2910209	GTP cyclohydrolase FolE	COG1469	S	454	382	72
RCA23_c27490	2912575	2911334	O-succinylhomoserine sulfhydrylase MetZ	COG0626	E	484	552	-68
RCA23_c27500	2912751	2913443	hypothetical protein	COG0625	O	355	420	-65
RCA23_c27510	2914130	2913531	putative intracellular septation protein	COG2917	D	725	845	-120
RCA23_c27520	2915048	2914146	hypothetical protein	COG0697	G	373	538	-165
RCA23_c27530	2916226	2915045	putative cell division protein	COG0552	U	244	359	-115
RCA23_c27540	2918008	2916323	NADH-quinone oxidoreductase subunits E/F (fused)	COG1894	C	257	373	-116
RCA23_c27550	2919686	2918133	exodeoxyribonuclease 7 large subunit XseA	COG1570	L	243	310	-67
RCA23_c27560	2919757	2921019	phosphoribosylamine--glycine ligase PurD	COG0151	F	245	400	-155
RCA23_c27570	2921398	2921075	hypothetical protein, ferredoxin	COG0633	C	495	375	120
RCA23_c27580	2921502	2921894	hypothetical protein	COG0587	L	158	249	-91
RCA23_c27590	2923443	2921956	periplasmic serine protease DO-like	COG0265	O	449	643	-194
RCA23_c27600	2924668	2923781	putative protein hflC	COG0330	O	553	459	94
RCA23_c27610	2925849	2924668	protein HflK	COG0330	O	588	454	134
RCA23_c27620	2927271	2925913	glutathione reductase Gor	COG1249	C	302	406	-104
RCA23_c27630	2928055	2927336	ribose-5-phosphate isomerase A	COG0120	G	357	434	-77
RCA23_c27640	2928355	2928125	hypothetical protein			300	310	-10
RCA23_c27650	2928419	2929792	L-serine dehydratase SdaB	COG1760	E	239	300	-61
RCA23_c27660	2929823	2930707	putative integral membrane protein	COG0697	G	437	492	-55
RCA23_c27670	2931346	2930672	hypothetical protein, thiamine pyrophosphokinase	COG1564	H	326	453	-127
RCA23_c27680	2931480	2931917	hypothetical protein			176	341	-165
RCA23_c27690	2932339	2933484	hypothetical protein			512	344	168
RCA23_c27700	2933779	2933552	hypothetical protein			335	362	-27
RCA23_c27710	2935071	2933779	adenylosuccinate synthase PurA	COG0104	F	436	466	-30

RCA23_c27720	2935313	2935666	preprotein translocase, SecG subunit SecG			324	415	-91
RCA23_c27730	2935819	2937462	CTP synthase	COG0504	F	449	408	41
RCA23_c27740	2937481	2937774	hypothetical protein, DUF1330	COG5470	S	276	315	-39
RCA23_c27750	2937874	2938227	hypothetical protein DUF1332	COG4103	S	274	238	36
RCA23_c27760	2938232	2938948	hypothetical protein	COG3221	P	184	292	-108
RCA23_c27770	2939027	2939800	amino acid transport ATP-binding protein	COG4598	E	412	570	-158
RCA23_c27780	2939834	2940547	putative amino acid transport extracellular solute binding pr	COG0834	E	476	436	40
RCA23_c27790	2940762	2941547	putative ABC transporter inner membrane component	COG4215	E	351	411	-60
RCA23_c27800	2941544	2942347	putative amino acid transport permease protein	COG4160	E	271	371	-100
RCA23_c27810	2942344	2943621	gamma-glutamylputrescine synthase PuuA	COG0174	E	366	351	15
RCA23_c27820	2943682	2944383	hypothetical protein, glutamine amidotransferase class I	COG0518	F	410	506	-96
RCA23_c27830	2944380	2945732	gamma-glutamylputrescine synthase PuuA	COG0174	E	319	376	-57
RCA23_c27840	2945725	2947029	gamma-glutamylputrescine oxidoreductase PuuB	COG0665	E	194	296	-102
RCA23_c27850	2947227	2948594	aminotransferase class III	COG0161	H	481	435	46
RCA23_c27860	2949102	2948632	putative PHB synthesis repressor protein	COG5394	S	616	440	176
RCA23_c27870	2949760	2949338	hypothetical protein			962	483	479
RCA23_c27880	2951592	2949850	poly(R)-hydroxyalkanoic acid synthase PhaC	COG3243	I	772	600	172
RCA23_c27890	2951693	2952952	polyhydroxyalkanoate depolymerase PhaZ	COG4553	I	424	429	-5
RCA23_c27900	2952949	2953992	putative ribosome biogenesis GTPase RsgA	COG1162	R	273	335	-62
RCA23_c27910	2954694	2953999	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	295	397	-102
RCA23_c27920	2955122	2954814	hypothetical protein			640	594	46
RCA23_c27930	2955968	2955144	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	489	563	-74
RCA23_c27940	2956110	2958059	threonyl-tRNA synthase ThrS	COG0441	J	302	309	-7
RCA23_c27950	2958494	2958111	hypothetical protein, ArsC	COG1393	P	275	345	-70
RCA23_c27960	2959802	2958891	thymidylate synthase	COG1351	F	515	437	78
RCA23_c27970	2959890	2960318	putative glyoxalase/bleomycin resistance protein/dioxygena	COG0346	E	379	413	-34
RCA23_c27980	2960315	2961049	hypothetical protein			489	471	18
RCA23_c27990	2961528	2961043	putative transcriptional regulatorm marR family	COG1846	K	717	600	117
RCA23_c28000	2961866	2961600	hypothetical protein DUF339	COG2938	S	807	701	106
RCA23_c28010	2962255	2961863	hypothetical protein, DNA binding	COG1813	K	611	550	61

RCA23_c28020	2962344	2963546	aspartate aminotransferase AatA	COG0436	E	538	517	21
RCA23_c28030	2963696	2963965	hypothetical protein			520	538	-18
RCA23_c28040	2965311	2963962	multidrug resistance protein NorM	COG0534	V	289	371	-82
RCA23_c28050	2965347	2967407	DNA topoisomerase 4 subunit B	COG0187	L	401	416	-15
RCA23_c28060	2967432	2967866	hypothetical protein	COG3238	S	246	456	-210
RCA23_c28070	2968865	2967876	putative malate/L-lactate dehydrogenase	COG2055	C	232	307	-75
RCA23_c28080	2969623	2968946	lipoprotein-releasing system ATP-binding protein LolD	COG1136	V	324	503	-179
RCA23_c28090	2970938	2969616	lipoprotein-releasing system transmembrane protein, lolC/E	COG4591	M	391	420	-29
RCA23_c28100	2972322	2970991	prolyl-tRNA synthase ProS	COG0442	J	552	522	30
RCA23_c28110	2972479	2973552	hypothetical protein DUF20	COG0628	R	376	457	-81
RCA23_c28120	2973581	2974198	hypothetical protein	COG0593	L	404	491	-87
RCA23_c28130	2974261	2976429	polyphosphate kinase Ppk	COG0855	P	432	474	-42
RCA23_c28140	2976502	2978046	putative phosphatase	COG0248	F	329	415	-86
RCA23_c28150	2978758	2978078	hypothetical protein, DnaJ	COG1076	O	357	406	-49
RCA23_c28160	2979420	2978806	hypothetical protein			360	363	-3
RCA23_c28170	2981568	2979442	methylmalonyl-CoA mutase McmA	COG1884	I	438	453	-15
RCA23_c28180	2981692	2982117	hypothetical protein			443	452	-9
RCA23_c28190	2984179	2982182	biotin carboxylase AccC	COG4770	I	463	438	25
RCA23_c28200	2984788	2984429	hypothetical protein			688	418	270
RCA23_c28210	2986426	2984807	propionyl-CoA carboxylase beta chain, mitochondrial precu	COG4799	I	431	395	36
RCA23_c28220	2986614	2987720	putative major facilitator superfamily transporter	COG2814	G	324	367	-43
RCA23_c28230	2987752	2989146	hypothetical protein	COG3800	R	268	299	-31
RCA23_c28240	2990629	2989175	betaine aldehyde dehydrogenase BetB	COG1012	C	271	382	-111
RCA23_c28250	2991519	2990626	hypothetical protein	COG0697	G	358	670	-312
RCA23_c28260	2992971	2991571	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	456	452	4
RCA23_c28270	2993600	2992971	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	526	453	73
RCA23_c28280	2994698	2993682	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	773	495	278
RCA23_c28290	2995125	2994823	transcriptional regulator, LacI family	COG1609	K	194	295	-101
RCA23_c28300	2996093	2995341	hypothetical protein, DUF81 family	COG0730	R	642	487	155
RCA23_c28310	2997240	2996074	serine--glyoxylate aminotransferase SgaA	COG0075	E	671	457	214

RCA23_c28320	2997311	3000124	D-lactate dehydrogenase Dld	COG0277	C	433	316	117
RCA23_c28330	3000108	3000872	hypothetical protein			591	406	185
RCA23_c28340	3002359	3000869	AMP-binding enzyme	COG0318	I	627	378	249
RCA23_c28350	3003741	3002356	succinate-semialdehyde dehydrogenase SucD	COG1012	C	714	387	327
RCA23_c28360	3005263	3003773	hypothetical protein	COG3333	S	551	345	206
RCA23_c28370	3005815	3005270	hypothetical protein			566	428	138
RCA23_c28380	3006899	3005886	putative tripartite tricarboxylate transporter family receptor	COG3181	S	478	291	187
RCA23_c28390	3008174	3007812	putative ETC complex I subunit			483	380	103
RCA23_c28450	3014946	3014311	sulfoxide reductase heme-binding subunit YedZ	COG2717	S	606	414	192
RCA23_c28460	3015970	3015032	sulfoxide reductase catalytic subunit YedY	COG2041	R	699	498	201
RCA23_c28470	3016084	3016839	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	217	388	-171
RCA23_c28480	3017484	3016816	putative haloacid dehalogenase-like hydrolase	COG0637	R	241	355	-114
RCA23_c28490	3020158	3017549	chaperone protein ClpB	COG0542	O	571	466	105
RCA23_c28500	3020382	3021098	orotidine 5'-phosphate decarboxylase PyrF	COG0284	F	231	360	-129
RCA23_c28510	3021376	3021095	hypothetical protein			303	257	46
RCA23_c28520	3022266	3021376	SPFH domain/band 7 family protein	COG0330	O	332	320	12
RCA23_c28530	3022399	3023655	DNA polymerase IV	COG0389	L	244	376	-132
RCA23_c28540	3024583	3023660	putative N-formylglutamate amidohydrolase	COG3741	E	220	384	-164
RCA23_c28560	3026236	3025031	hypothetical protein, phenylacetate-coenzyme A ligase	COG1541	H	354	468	-114
RCA23_c28570	3027095	3026262	high-affinity branched-chain amino acid transporter ATP-bin	COG0410	E	388	363	25
RCA23_c28580	3028426	3027146	hypothetical protein	COG0683	E	356	266	90
RCA23_c28590	3029567	3028491	high-affinity branched-chain amino acid transporter permeas	COG4177	E	514	348	166
RCA23_c28600	3030631	3029630	high-affinity branched-chain amino acid transporter permeas	COG0559	E	662	420	242
RCA23_c28610	3031517	3030699	high-affinity branched-chain amino acid transporter ATP-bin	COG0411	E	616	480	136
RCA23_c28620	3033530	3031569	putative long-chain-fatty-acid-CoA ligase	COG1022	I	540	469	71
RCA23_c28630	3034611	3033706	tRNA pseudouridine synthase B	COG0130	J	444	471	-27
RCA23_c28650	3035184	3034663	ribosome-binding factor A	COG0858	J	390	589	-199
RCA23_c28640	3035183	3035992	dihydrodipicolinate synthase DapA	COG0289	E	232	307	-75
RCA23_c28660	3036242	3036039	hypothetical protein DUF1674	COG5508	S	308	405	-97
RCA23_c28670	3036304	3037596	putative ribosomal RNA small subunit methyltransferase B	COG0144	J	319	342	-23

RCA23_c28680	3037762	3039378	hypothetical protein, heparinase II/III	COG5360	S	307	415	-108
RCA23_c28690	3039420	3041006	bifunctional purine biosynthesis protein PurH	COG0138	F	334	417	-83
RCA23_c28700	3041008	3041490	signal peptidase II	COG0597	M	444	466	-22
RCA23_c28710	3041564	3042898	uncharacterized zinc protease y4wA	COG0612	R	259	271	-12
RCA23_c28720	3042895	3044199	uncharacterized zinc protease y4wB	COG0612	R	343	387	-44
RCA23_c28730	3044234	3046045	DNA mismatch repair protein MutL	COG0323	L	277	394	-117
RCA23_c28740	3046042	3047181	putative RmuC family protein	COG1322	S	453	493	-40
RCA23_c28750	3047826	3047176	transcriptional activator ChrR	COG3806	T	328	336	-8
RCA23_c28760	3048473	3047823	RNA polymerase sigma factor SigK	COG1595	K	815	442	373
RCA23_c28770	3048595	3049896	hypothetical protein	COG2907	R	645	423	222
RCA23_c28780	3049893	3050648	hypothetical protein DUF1365	COG3496	S	321	368	-47
RCA23_c28790	3050645	3051886	hypothetical protein	COG2211	G	246	374	-128
RCA23_c28800	3051883	3052434	hypothetical protein			338	406	-68
RCA23_c28810	3052441	3053175	putative short chain dehydrogenase	COG4221	R	338	412	-74
RCA23_c28820	3054242	3053181	saccharopine dehydrogenase (NAD+,L-lysine-forming)	COG3288	C	244	380	-136
RCA23_c28830	3054838	3054239	putative glutathione S-transferase	COG0625	O	410	462	-52
RCA23_c28840	3055518	3054859	putative phosphoglycerate mutase family protein	COG0406	G	408	433	-25
RCA23_c28850	3055678	3058473	formate dehydrogenase Fdh	COG3383	R	256	355	-99
RCA23_c28860	3059684	3058476	hypothetical protein	COG1357	S	149	173	-24
RCA23_c28870	3059845	3060351	hypothetical protein DUF1643	COG4333	S	165	251	-86
RCA23_c28880	3060348	3061232	haloacetate dehalogenase DehH	COG0596	R	216	302	-86
RCA23_c28890	3061378	3061647	30S ribosomal protein S15	COG0184	J	532	436	96
RCA23_c28900	3061763	3062671	hypothetical protein, DUF6 transmembrane protein	COG0697	G	291	363	-72
RCA23_c28910	3062837	3064972	polyribonucleotide nucleotidyltransferase Pnp	COG1185	J	566	452	114
RCA23_c28920	3065297	3066739	aldehyde dehydrogenase	COG1012	C	367	317	50
RCA23_c28930	3067514	3066870	ribosomal large subunit pseudouridine synthase A	COG0564	J	330	420	-90
RCA23_c28940	3067641	3068903	protein OtnG			239	284	-45
RCA23_c28950	3068952	3069767	hypothetical protein, DUF940 putative lipoprotein			460	502	-42
RCA23_c28960	3069770	3070432	hypothetical protein			502	397	105
RCA23_c28970	3070451	3071461	UDP-glucose 4-epimerase GalE	COG1087	M	409	393	16

RCA23_c28980	3072918	3071458	putative FAD linked oxidase	COG0277	C	266	432	-166
RCA23_c28990	3073207	3074703	signal recognition particle protein Ffh	COG0541	U	323	360	-37
RCA23_c29000	3074735	3075031	putative chorismate mutase type II	COG1605	E	471	593	-122
RCA23_c29010	3075077	3075463	30S ribosomal protein S16	COG0228	J	454	521	-67
RCA23_c29020	3075599	3076204	ribosome maturation factor RimM	COG0806	J	531	440	91
RCA23_c29030	3076201	3076983	tRNA (guanine-N(1)-)-methyltransferase TrmD	COG0336	J	262	287	-25
RCA23_c29040	3077274	3077645	50S ribosomal protein L19	COG0335	J	1.180	564	616
RCA23_c29050	3077657	3077878	50S ribosomal protein L31	COG0254	J	843	515	328
RCA23_c29060	3078055	3078864	ATPase MipZ	COG1192	D	467	427	40
RCA23_c29070	3079793	3078861	fructokinase	COG0524	G	353	434	-81
RCA23_c29080	3080009	3081187	P-hydroxybenzoate hydroxylase PobA	COG0654	H	481	429	52
RCA23_c29090	3081251	3082447	kynureninase KynU	COG3844	E	389	318	71
RCA23_c29100	3083041	3083694	HTH-type transcriptional regulator, GntR family	COG1802	K	449	308	141
RCA23_c29110	3083761	3084600	hypothetical protein, glutathione S-transferase	COG0625	O	529	404	125
RCA23_c29120	3084701	3085639	L-threonine 3-dehydrogenase	COG0451	M	348	353	-5
RCA23_c29130	3085650	3086444	enoyl-CoA hydratase	COG1024	I	185	255	-70
RCA23_c29140	3087386	3086463	enoyl-CoA hydratase	COG1250	I	205	333	-128
RCA23_c29150	3088163	3087396	3-hydroxyacyl-CoA dehydrogenase FadN	COG1028	I	242	312	-70
RCA23_c29160	3088298	3088942	hypothetical protein			381	386	-5
RCA23_c29170	3089925	3088969	integrase	COG4974	L	458	427	31
RCA23_c29180	3090052	3091107	hypothetical protein, porin-like			520	390	130
RCA23_c29190	3091254	3093845	leucyl-tRNA synthase LeuS	COG0495	J	433	380	53
RCA23_c29200	3093904	3094302	hypothetical protein DUF2159			571	462	109
RCA23_c29210	3094299	3095324	hypothetical protein			274	392	-118
RCA23_c29220	3095325	3095912	hypothetical protein, glutathione S-transferase	COG0625	O	500	442	58
RCA23_c29230	3097108	3095885	hypothetical protein	COG2081	R	237	270	-33
RCA23_c29240	3098243	3097311	hypothetical protein, porin			1.707	409	1.298
RCA23_c29250	3098481	3099143	hypothetical protein, alanine racemase-like	COG0325	R	273	282	-9
RCA23_c29260	3099644	3099150	hypothetical protein, YkuD	COG3786	S	218	327	-109
RCA23_c29270	3100696	3099641	GTP cyclohydrolase II RibA	COG0807	H	321	376	-55

RCA23_c29280	3100840	3101526	response regulator receiver protein	COG0745	T	801	648	153
RCA23_c29290	3102576	3101779	2-oxo-hepta-3-ene-1,7-dioic acid hydratase HpcG	COG3971	Q	378	339	39
RCA23_c29300	3103430	3102573	fumarylacetoacetate hydrolase family protein	COG0179	Q	416	428	-12
RCA23_c29310	3103834	3103427	5-carboxymethyl-2-hydroxymuconate isomerase	COG3232	E	367	337	30
RCA23_c29320	3104009	3104446	homoprotocatechuate degradation operon regulator HpaR	COG1846	K	513	682	-169
RCA23_c29330	3104443	3105903	4-hydroxyphenylacetate 3-monooxygenase oxygenase com	COG2368	Q	396	413	-17
RCA23_c29340	3106978	3105926	ABC sugar transporter, ATPase subunit	COG3839	G	417	427	-10
RCA23_c29350	3108378	3106966	mannitol 2-dehydrogenase MtlK	COG0246	G	541	470	71
RCA23_c29360	3109277	3108384	transcriptional regulator, AraC family	COG2207	K	845	568	277
RCA23_c29370	3110179	3109343	ABC sugar transporter, permease protein	COG0395	G	673	486	187
RCA23_c29380	3111128	3110190	ABC sugar transporter, permease protein	COG1175	G	627	368	259
RCA23_c29390	3112520	3111240	ABC transporter, periplasmic substrate-binding protein	COG1653	G	482	349	133
RCA23_c29400	3113045	3113410	putative flavohemoglobin / bacterial hemoglobin	COG1017	C	172	151	21
RCA23_c29410	3113570	3113848	hypothetical protein			357	239	118
RCA23_c29420	3115053	3113854	Na+/H+ antiporter NhaA	COG3004	P	869	651	218
RCA23_c29430	3115555	3115160	hypothetical protein			343	377	-34
RCA23_c29440	3116046	3115633	hypothetical protein DUF55	COG2947	S	575	446	129
RCA23_c29450	3116318	3116046	protein YciI	COG2350	S	444	407	37
RCA23_c29460	3117283	3116318	glycerol-3-phosphate dehydrogenase GpsA	COG0240	C	295	365	-70
RCA23_c29470	3117383	3118087	hypothetical protein, uroporophyrinogen-III synthase HemD	COG1587	H	261	373	-112
RCA23_c29480	3118112	3119260	hypothetical protein	COG3264	M	297	444	-147
RCA23_c29490	3119271	3120746	hypothetical protein, HemY	COG3898	S	255	322	-67
RCA23_c29500	3121065	3120736	transcriptional regulatory protein, Ars family	COG0640	K	351	449	-98
RCA23_c29510	3122055	3121183	magnesium-chelatase subunit BchO	COG0596	R	368	539	-171
RCA23_c29520	3123707	3122055	magnesium-chelatase subunit BchD	COG1240	H	228	347	-119
RCA23_c29530	3124716	3123709	magnesium-chelatase subunit BchI	COG1239	H	259	357	-98
RCA23_c29540	3125459	3124713	spheroidene monooxygenase CrtA			327	471	-144
RCA23_c29550	3125514	3127076	phytoene dehydrogenase CrtI	COG1233	Q	248	310	-62
RCA23_c29560	3127073	3128113	phytoene synthase CrtB	COG1562	I	282	412	-130
RCA23_c29570	3129022	3128114	hydroxyneurosporene dehydrogenase CrtC			315	427	-112

RCA23_c29580	3130449	3129025	methoxyneurosporene dehydrogenase CrtD	COG1233	Q	258	382	-124
RCA23_c29590	3130597	3131511	geranylgeranyl pyrophosphate synthase CrtE	COG0142	H	317	337	-20
RCA23_c29600	3131511	3132656	hydroxyneurosporene methyltransferase CrtF			351	342	9
RCA23_c29610	3132720	3133685	2-desacetyl-2-hydroxyethyl bacteriochlorophyllide A dehydr	COG1063	E	357	361	-4
RCA23_c29620	3133682	3134686	chlorophyllide reductase BchX	COG1348	P	519	502	17
RCA23_c29630	3134679	3136250	chlorophyllide reductase BchY			341	319	22
RCA23_c29640	3136250	3137710	chlorophyllide reductase subunit BchZ	COG2710	C	536	488	48
RCA23_c29650	3137707	3137940	protein PufQ			537	427	110
RCA23_c29660	3138087	3138242	light-harvesting protein B-870 beta chain PufB			6.978	500	6.478
RCA23_c29670	3138259	3138411	light-harvesting protein B-870 alpha chain PufA			6.270	527	5.743
RCA23_c29680	3138531	3139361	reaction center protein L chain PufL			874	581	293
RCA23_c29690	3139375	3140298	reaction center protein M chain PufM			677	578	99
RCA23_c29700	3140323	3140562	protein PufX			968	617	351
RCA23_c29710	3140748	3142640	1-deoxy-D-xylulose-5-phosphate synthase Dxs	COG1154	H	354	396	-42
RCA23_c29720	3143177	3142659	isopentenyl-diphosphate delta-isomerase Idi	COG1443	I	410	573	-163
RCA23_c29730	3144355	3143177	geranylgeranyl reductase BchP	COG0644	C	323	416	-93
RCA23_c29740	3145650	3144358	bacteriochlorophyll synthase 44.5 kDa chain			331	411	-80
RCA23_c29750	3146546	3145647	bacteriochlorophyll synthase BchG	COG0382	H	279	365	-86
RCA23_c29760	3146699	3147133	cytochrome c-551	COG3474	C	2.610	694	1.916
RCA23_c29770	3148204	3147191	uroporphyrinogen decarboxylase HemE	COG0407	H	678	490	188
RCA23_c29780	3148383	3149273	porphobilinogen deaminase HemC	COG0181	H	296	376	-80
RCA23_c29790	3149254	3150213	hypothetical protein, NmrA-like	COG0702	M	262	383	-121
RCA23_c29800	3151424	3150210	5-aminolevulinate synthase HemA	COG0156	H	552	474	78
RCA23_c29810	3152220	3151414	hypothetical protein			439	465	-26
RCA23_c29820	3153347	3152220	aerobic Mg-protoporphyrin IX monomethyl ester oxidative cyclase AcsF			646	487	159
RCA23_c29830	3153640	3153344	hypothetical protein			666	516	150
RCA23_c29840	3154124	3153657	hypothetical protein			392	356	36
RCA23_c29850	3154768	3154130	hypothetical protein			466	594	-128
RCA23_c29860	3155600	3154827	reaction center protein PuhA			424	380	44
RCA23_c29870	3157046	3155619	protein PucC			416	549	-133

RCA23_c29880	3157708	3157043	magnesium-protoporphyrin O-methyltransferase BchM	COG2227	H	353	470	-117
RCA23_c29890	3158604	3157708	light-independent protochlorophyllide reductase iron-sulfur A	COG1348	P	615	601	14
RCA23_c29900	3162147	3158638	magnesium-chelatase subunit BchH	COG1429	H	346	408	-62
RCA23_c29910	3163717	3162173	light-independent protochlorophyllide reductase subunit Bch	COG2710	C	379	421	-42
RCA23_c29920	3164946	3163714	light-independent protochlorophyllide reductase subunit Bch	COG2710	C	243	340	-97
RCA23_c29930	3165482	3164994	2-vinyl bacteriochlorophyllide hydratase BchF			606	518	88
RCA23_c29940	3165800	3166510	putative transcriptional regulator PpaA			654	578	76
RCA23_c29950	3166571	3167995	transcriptional regulator PpsR	COG3829	K	507	360	147
RCA23_c29960	3168781	3168005	hypothetical protein	COG5012	R	383	415	-32
RCA23_c29970	3169806	3169207	transcriptional regulator protein FixJ	COG4566	T	571	377	194
RCA23_c29980	3170351	3169872	peripheral-type benzodiazepine receptor/signal transductior	COG3476	T	659	589	70
RCA23_c29990	3170538	3170972	hypothetical protein, integral membrane proteins YeeE/Yedl	COG2391	R	475	462	13
RCA23_c30000	3171010	3171402	hypothetical protein, integral membrane proteins YeeE/Yedl	COG2391	R	439	527	-88
RCA23_c30010	3171440	3172324	putative beta-lactamase hydrolase-like protein	COG0491	R	519	451	68
RCA23_c30020	3172386	3172811	hypothetical protein	COG3453	S	400	482	-82
RCA23_c30030	3172854	3174578	sulphate transporter	COG0659	P	577	541	36
RCA23_c30040	3174623	3176065	amidase	COG0154	J	392	412	-20
RCA23_c30050	3176200	3176790	hypothetical protein			555	390	165
RCA23_c30060	3176778	3177194	aminoglycoside phosphotransferase Aph			655	451	204
RCA23_c30070	3177342	3178808	hypothetical protein	COG2885	M	447	429	18
RCA23_c30080	3179766	3178849	chaperone protein DnaJ	COG0484	O	337	327	10
RCA23_c30090	3181980	3180073	chaperone protein DnaK	COG0443	O	2.596	868	1.728
RCA23_c30100	3182136	3182774	putative alpha-ketoglutarate-dependent dioxygenase AlkB	COG3145	L	444	449	-5
RCA23_c30110	3183607	3182780	putative ABC-2 type transporter	COG1682	G	559	435	124
RCA23_c30120	3183719	3184516	3'(2'),5'-bisphosphate nucleotidase CysQ	COG1218	P	487	424	63
RCA23_c30130	3184712	3185623	UTP-glucose-1-phosphate uridylyltransferase GalU	COG1210	M	100	76	24
RCA23_c30140	3185671	3186672	hypothetical protein	COG0463	M	29	29	0
RCA23_c30150	3186755	3187789	hypothetical protein			61	22	39
RCA23_c30160	3187804	3189483	hypothetical protein beta-1,6-N-acetylglucosaminyltransferases			64	40	24
RCA23_c30170	3189485	3190912	hypothetical protein			83	93	-10

RCA23_c30180	3191373	3190909	nitrogen regulatory protein PtsN	COG1762	G	462	459	3
RCA23_c30190	3191964	3191398	putative sigma 54 modulation protein	COG1544	J	716	558	158
RCA23_c30200	3192915	3192157	lipopolysaccharide export system ATP-binding protein LptB	COG1137	R	444	438	6
RCA23_c30210	3193403	3192915	putative lipopolysaccharide export system protein LptA	COG1934	S	280	352	-72
RCA23_c30220	3194020	3193412	hypothetical protein			689	490	199
RCA23_c30230	3194662	3194048	putative 3'-5'-exonuclease	COG0349	J	741	509	232
RCA23_c30240	3194850	3195968	glycine amidinotransferase	COG1834	E	521	526	-5
RCA23_c30250	3196817	3196053	hypothetical protein	COG4123	R	631	493	138
RCA23_c30260	3197481	3196918	hypothetical protein, metal-dependent phosphohydrolase	COG1896	R	457	580	-123
RCA23_c30270	3197634	3199025	S-adenosyl-L-homocysteine hydrolase AhcY	COG0499	H	603	490	113
RCA23_c30280	3199890	3199339	photosynthetic apparatus regulatory protein RegA	COG4567	T	407	550	-143
RCA23_c30290	3200570	3199947	protein SenC	COG1999	R	543	578	-35
RCA23_c30300	3200649	3202016	sensor histidine kinase RegB	COG0642	T	358	402	-44
RCA23_c30310	3202073	3203299	hypothetical protein			342	444	-102
RCA23_c30320	3203450	3204445	hypothetical protein, aminoglycoside phosphotransferase	COG3178	R	494	540	-46
RCA23_c30330	3204445	3205107	hypothetical protein, nucleotidyl transferase	COG1208	M	458	486	-28
RCA23_c30340	3205100	3208018	double-strand break repair protein AddB	COG3893	L	333	483	-150
RCA23_c30350	3208015	3211371	double-strand break repair helicase AddA	COG1074	L	483	495	-12
RCA23_c30360	3211430	3211750	thioredoxin TrxA	COG3118	O	926	781	145
RCA23_c30370	3211858	3212415	ATP-dependent protease HslV	COG5405	O	570	616	-46
RCA23_c30380	3212412	3213719	ATP-dependent hsl protease ATP-binding subunit HslU	COG1220	O	505	486	19
RCA23_c30390	3213757	3214968	MFS-type transporter	COG2814	G	450	609	-159
RCA23_c30400	3215543	3214965	hypothetical protein, Smr protein/MutS2	COG2840	S	359	384	-25
RCA23_c30410	3216574	3215546	putative lytic murein transglycosylase	COG2821	M	516	462	54
RCA23_c30420	3217245	3216571	hypothetical protein, TIM44	COG4395	S	603	469	134
RCA23_c30430	3217333	3217791	putative cytoplasmic membrane protein FxsA	COG3030	R	1.292	879	413
RCA23_c30440	3217847	3218353	protein-export protein SecB	COG1952	U	980	584	396
RCA23_c30450	3219045	3218350	DNA polymerase III subunit epsilon	COG0847	L	740	456	284
RCA23_c30460	3219631	3219038	dephospho-CoA kinase CoaE	COG0237	H	607	420	187
RCA23_c30470	3221043	3219628	shikimate 5-dehydrogenase AroE	COG0169	E	384	325	59

RCA23_c30480	3222030	3223286	transcription termination factor Rho	COG1158	K	799	587	212
RCA23_c30490	3223296	3224582	tRNA modification GTPase MnmE	COG0486	R	677	610	67
RCA23_c30500	3224597	3226456	tRNA uridine 5-carboxymethylaminomethyl modification enz	COG0445	D	754	479	275
RCA23_c30510	3226453	3227076	ribosomal RNA small subunit methyltransferase G	COG0357	M	923	624	299
RCA23_c30520	3227063	3227839	chromosome-partitioning protein ParA	COG1192	D	715	536	179
RCA23_c30530	3227913	3228749	chromosome-partitioning protein ParB	COG1475	K	590	401	189
RCA23_c30540	3229894	3228746	oxygen-independent coproporphyrinogen III oxidase HemN	COG0635	H	502	365	137
RCA23_c30550	3230502	3229894	nucleoside-triphosphatase RdgB	COG0127	F	443	404	39
RCA23_c30560	3231212	3230502	ribonuclease PH	COG0689	J	433	583	-150
RCA23_c30570	3231309	3232340	heat-inducible transcription repressor HrcA	COG1420	K	677	565	112
RCA23_c30580	3232347	3232901	protein GrpE	COG0576	O	715	541	174
RCA23_c30590	3235734	3232948	DNA mismatch repair protein MutS	COG0249	L	382	470	-88
RCA23_c30600	3235847	3238102	NADP-dependent malic enzyme MaeB	COG0281	C	751	604	147
RCA23_c30610	3238099	3238974	pfkB family carbohydrate kinase	COG0524	G	385	437	-52
RCA23_c30620	3239064	3239972	putative aminotransferase class IV	COG0115	E	420	511	-91
RCA23_c30630	3239969	3240679	hypothetical protein, branched-chain-amino-acid aminotransferase-like			281	358	-77
RCA23_c30640	3241914	3240697	argininosuccinate synthase ArgG	COG0137	E	456	465	-9
RCA23_c30650	3242076	3243311	threonine dehydratase, biosynthetic	COG1171	E	565	492	73
RCA23_c30660	3243760	3243320	putative NUDIX hydrolase	COG0494	L	859	585	274
RCA23_c30670	3243841	3244851	heat shock protein 33	COG1281	O	563	552	11
RCA23_c30680	3244832	3245428	putative NUDIX hydrolase	COG0494	L	312	397	-85
RCA23_c30690	3245425	3246585	hypothetical protein, poly A polymerase	COG0617	J	231	347	-116
RCA23_c30700	3246667	3248505	ABC transporter ATP-binding/permease protein	COG1132	V	582	671	-89
RCA23_c30710	3248505	3249719	hypothetical protein	COG2265	J	472	532	-60
RCA23_c30720	3249726	3250487	ion transport protein			613	593	20
RCA23_c30730	3250547	3251041	hypothetical protein	COG3034	S	812	587	225
RCA23_c30740	3251589	3251068	hypothetical protein, SCP-like extracellular protein	COG2340	S	495	433	62
RCA23_c30750	3251740	3252339	hypothetical protein	COG1376	S	713	514	199
RCA23_c30760	3253788	3252730	ferrochelatase HemH	COG0276	H	912	584	328
RCA23_c30770	3254654	3253839	hypothetical protein			409	512	-103

RCA23_c30780	3254684	3255250	hypothetical protein, ComF/GntX family	COG1040	R	415	446	-31
RCA23_c30790	3255612	3255869	glutaredoxin GrxC	COG0695	O	580	614	-34
RCA23_c30800	3255866	3256708	putative carbon-nitrogen hydrolase	COG0388	R	467	593	-126
RCA23_c30810	3256692	3257135	putative HTH-type transcriptional regulator, MarR family	COG1846	K	401	502	-101
RCA23_c30820	3257890	3257132	3-demethylubiquinone-9 3-O-methyltransferase UbiG	COG2227	H	394	398	-4
RCA23_c30830	3257891	3258916	proline iminopeptidase Pip	COG0596	R	482	453	29
RCA23_c30840	3259000	3259989	putative peptide transport system permease protein	COG0601	E	927	665	262
RCA23_c30850	3259998	3261185	putative peptide transport system permease protein	COG1173	E	830	594	236
RCA23_c30860	3261190	3262836	peptide transport system ATP-binding protein	COG1123	R	405	486	-81
RCA23_c30870	3262914	3264458	putative periplasmic peptide-binding protein	COG0747	E	501	416	85
RCA23_c30880	3264952	3265524	hypothetical protein UPF0090	COG0779	S	735	510	225
RCA23_c30890	3265524	3267137	transcription elongation protein NusA	COG0195	K	693	541	152
RCA23_c30900	3267138	3267740	hypothetical protein DUF448	COG2740	K	467	523	-56
RCA23_c30910	3267775	3270240	translation initiation factor IF-2	COG0532	J	419	390	29
RCA23_c30920	3270924	3270526	putative mutator MutT protein	COG1051	F	388	382	6
RCA23_c30930	3272138	3270921	arginine biosynthesis bifunctional protein ArgJ	COG1364	E	357	419	-62
RCA23_c30940	3272976	3272143	putative peptidylprolyl isomerase	COG0760	O	441	438	3
RCA23_c30950	3273165	3275849	protein translocase subunit SecA	COG0653	U	662	490	172
RCA23_c30960	3276115	3277296	putative O-acetyltransferase OatA	COG1835	I	719	508	211
RCA23_c30970	3278383	3277550	hypothetical protein	COG4922	S	341	243	98
RCA23_c30980	3278502	3278876	hypothetical protein, transcriptional regulator-like	COG1733	K	257	195	62
RCA23_c30990	3278997	3280049	UDP-glucoronate 5'-epimerase LspL	COG0451	M	224	174	50
RCA23_c31000	3281732	3280023	hypothetical protein			214	200	14
RCA23_c31010	3283201	3281858	phosphoglucosamine mutase GlmM	COG1109	G	394	488	-94
RCA23_c31020	3284844	3283315	putative ubiquinone biosynthesis protein UbiB	COG0661	R	400	432	-32
RCA23_c31030	3285621	3284845	ubiquinone/menaquinone biosynthesis methyltransferase U	COG2226	H	488	408	80
RCA23_c31040	3285680	3286531	formamidopyrimidine-DNA glycosylase MutM	COG0266	L	369	371	-2
RCA23_c31050	3286594	3287370	enoyl-CoA hydratase/isomerase	COG1024	I	637	567	70
RCA23_c31060	3287519	3287782	30S ribosomal protein S20	COG0268	J	521	411	110

Table S7: COG categories, genomic islands and coverage of *P. temperata* RC23 genes from combined stations GS02-13 of the GOS data set and Norwegian fjord metagenome (Gilbert et al. 2008).

Coverages of <80% are marked pink and genomic islands gray

locus_tag	annotation	COG id	COG category	genomic islands	GOS stations GS02-GS13 gene coverage [%]	Fjord gDNA gene coverage [%]
RCA23_c00010	chromosomal replication initiatior protein DnaA	COG0593	L		100.00	100.00
RCA23_c00020	DNA polymerase III beta subunit DnaN	COG0592	L		100.00	99.55
RCA23_c00030	DNA replication and repair protein RecF	COG1195	L		100.00	100.00
RCA23_c00040	hypothetical protein, LysE type translocator	COG1280	E		100.00	100.00
RCA23_c00050	DNA gyrase subunit B	COG0187	L		100.00	99.92
RCA23_c00060	FAD dependent oxidoreductase	COG0665	E		100.00	100.00
RCA23_c00070	hypothetical protein, rhodanese-like sulphurtransferase	COG1054	R		100.00	100.00
RCA23_c00080	pyrazinamidase/nicotinamidase PncA	COG1335	Q		100.00	100.00
RCA23_c00090	nicotinate phosphoribosyltransferase PncB	COG1488	H		100.00	100.00
RCA23_c00100	aminoglycoside phosphotransferase	COG2334	R		100.00	100.00
RCA23_c00110	putative aminotransferase class III	COG0160	E		100.00	100.00
RCA23_c00120	putative peptide chain release factor	COG1186	J		100.00	100.00
RCA23_c00130	hypothetical integral membrane protein	COG1738	S		100.00	100.00
RCA23_c00140	penicillin-insensitive murein endopeptidase MepA	COG3770	M		100.00	99.57
RCA23_c00150	putative MFS-type transporter	COG2270	R		100.00	100.00
RCA23_c00160	hypothetical protein	COG0824	R		100.00	100.00
RCA23_c00170	hypothetical protein, YGGT family	COG0762	S		100.00	100.00
RCA23_c00180	ATP-dependent DNA helicase RecQ	COG0514	L		100.00	100.00
RCA23_c00190	hypothetical protein DUF328	COG3022	S		100.00	100.00
RCA23_c00200	isopentenyl-diphosphate delta-isomerase Idi	COG1443	I		100.00	100.00
RCA23_c00210	5-aminolevulinate synthase HemA	COG0156	H		100.00	100.00
RCA23_c00220	cytochrome c2	COG3474	C		100.00	100.00
RCA23_c00230	acetyl-CoA acetyltransferase PhaA	COG0183	I		100.00	100.00
RCA23_c00240	DNA polymerase III 2 alpha subunit DnaE	COG0587	L		100.00	100.00

RCA23_c00250	transcription antitermination protein NusG	COG0250	K	100.00	100.00
RCA23_c00260	putative lytic murein transglycosylase	COG2951	M	100.00	100.00
RCA23_c00270	putative ribonuclease R	COG0557	K	100.00	99.24
RCA23_c00280	hypothetical protein DUF461	COG2847	S	100.00	100.00
RCA23_c00290	chorismate synthase AroC	COG0082	E	100.00	100.00
RCA23_c00300	hypothetical protein			100.00	100.00
RCA23_c00310	thiamine-binding periplasmic protein ThiB	COG4143	H	100.00	100.00
RCA23_c00320	thiamine transport system permease protein ThiP	COG1178	P	100.00	100.00
RCA23_c00330	thiamine import ATP-binding protein ThiQ	COG3840	H	100.00	100.00
RCA23_c00340	cytochrome c1	COG2857	C	100.00	100.00
RCA23_c00350	cytochrome b	COG1290	C	100.00	97.53
RCA23_c00360	ubiquinol-cytochrome c reductase iron-sulfur subunit PetA	COG0723	C	100.00	100.00
RCA23_c00370	hypothetical protein, glutathione S-transferase	COG0625	O	100.00	100.00
RCA23_c00380	inositol 2-dehydrogenase IdhA	COG0673	R	100.00	100.00
RCA23_c00390	protein PmbA	COG0312	R	100.00	100.00
RCA23_c00400	putative inositol monophosphatase	COG0483	G	100.00	100.00
RCA23_c00410	putative 3-deoxy-D-manno-octulosonic-acid transferase	COG1519	M	100.00	100.00
RCA23_c00420	tetraacyldisaccharide 4'-kinase LpxK	COG1663	M	100.00	100.00
RCA23_c00430	hypothetical protein, thioredoxin	COG1651	O	100.00	100.00
RCA23_c00440	hypothetical protein DUF721	COG5389	S	100.00	100.00
RCA23_c00450	A/G-specific adenine glycosylase YfhQ	COG1194	L	100.00	100.00
RCA23_c00460	alkane 1-monooxygenase AlkB			100.00	100.00
RCA23_c00470	modification methylase CcrM	COG0863	L	100.00	100.00
RCA23_c00480	ribonuclease HII	COG0164	L	100.00	100.00
RCA23_c00490	putative exodeoxyribonuclease III	COG0708	L	100.00	100.00
RCA23_c00500	hypothetical protein, thioredoxin	COG3118	O	100.00	100.00
RCA23_c00510	putative ATP-dependent protease La (LON)	COG2802	R	100.00	100.00
RCA23_c00520	hypothetical protein, Trm112p-like	COG2835	S	100.00	100.00
RCA23_c00530	putative 2-octaprenyl-6-methoxyphenol hydroxylase	COG0654	H	100.00	100.00
RCA23_c00540	glutamyl-tRNA(Gln) amidotransferase subunit A	COG0154	J	100.00	98.02
RCA23_c00550	LL-diaminopimelate aminotransferase DapL	COG0436	E	100.00	100.00
RCA23_c00560	putative DNA transclocase FtsK	COG1674	D	100.00	100.00
RCA23_c00570	putative outer membrane lipoprotein carrier protein LolA	COG2834	M	100.00	93.55

RCA23_c00580	hypothetical protein	COG4764	S		100.00	100.00
RCA23_c00590	hemimethylated DNA-binding protein, YccV like	COG3785	S		100.00	100.00
RCA23_c00600	putative gamma-glutamyltransferase ywrD	COG0405	E		100.00	100.00
RCA23_c00610	acetyl-CoA acetyltransferase ThlA	COG0183	I		100.00	98.13
RCA23_c00620	hypothetical protein	COG1956	T		100.00	100.00
RCA23_c00630	putative HTH-type transcriptional regulator	COG1396	K		100.00	100.00
RCA23_c00640	hypothetical protein, homoserine/homoserine lactone efflux protein	COG1280	E		100.00	100.00
RCA23_c00650	LysE-type translocator	COG1280	E		100.00	87.72
RCA23_c00660	dimethylglycine dehydrogenase	COG0404	E		100.00	95.83
RCA23_c00670	putative homocysteine S-methyltransferase	COG2040	E		100.00	100.00
RCA23_c00680	pyridoxamine 5'-phosphate oxidase-like protein, FMN-binding	COG3576	R		100.00	99.34
RCA23_c00690	sarcosine dehydrogenase	COG0404	E		100.00	100.00
RCA23_c00700	hypothetical protein	COG0790	R		100.00	100.00
RCA23_c00710	putative agmatine deiminase AguA	COG2957	E		100.00	100.00
RCA23_c00720	Cl- channel, voltage-gated family protein	COG0038	P		100.00	85.75
RCA23_c00730	Cl- channel, voltage-gated family protein	COG0038	P		100.00	89.68
RCA23_c00750	diaminopimelate epimerase DapF	COG0253	E	GI 1	100.00	100.00
RCA23_c00760	(dimethylallyl)adenosine tRNA methylthiotransferase MiaB	COG0621	J	GI 1	100.00	100.00
RCA23_c00770	hypothetical protein, glutathione S-transferotein	COG0625	O	GI 1	100.00	100.00
RCA23_c00780	HTH-type transcriptional regulator, LacI family	COG1609	K	GI 1	71.79	100.00
RCA23_c00790	hypothetical protein			GI 1	100.00	100.00
RCA23_c00800	hypothetical protein	COG4274	S	GI 1	100.00	100.00
RCA23_c00810	hypothetical protein, phosphoglycerate mutase-like	COG0406	G	GI 1	95.68	100.00
RCA23_c00820	hypothetical protein	COG0790	R	GI 1	100.00	100.00
RCA23_c00830	hypothetical protein, glyoxalase/bleomycin resistance protein/dihydroxybiphe	COG0346	E	GI 1	100.00	100.00
RCA23_c00840	putative phage integrase	COG4974	L	GI 1	100.00	100.00
RCA23_c00850	hypothetical protein			GI 1	0.00	67.44
RCA23_c00860	hypothetical protein			GI 1	0.00	32.85
RCA23_c00870	putative endonuclease			GI 1	0.00	0.00
RCA23_c00880	DNA integration/recombination/invertion protein	COG1961	L	GI 1	0.00	100.00
RCA23_c00890	hypothetical protein			GI 1	0.00	100.00
RCA23_c00900	hypothetical protein, peptidoglycan binding-like			GI 1	0.00	0.00
RCA23_c00910	hypothetical protein			GI 1	0.00	0.00

RCA23_c00920	putative phage integrase	COG4974	L	GI 1	0.00	50.33
RCA23_c00940	putative HTH-type transcriptional repressor, ArsR family	COG0640	K		100.00	100.00
RCA23_c00950	ATP synthase protein I				100.00	100.00
RCA23_c00960	ATP synthase subunit AtpB	COG0356	C		100.00	100.00
RCA23_c00970	ATP synthase subunit c				100.00	100.00
RCA23_c01000	pyruvate dehydrogenase complex repressor	COG2186	K		100.00	100.00
RCA23_c01010	inner membrane lipoprotein YiaD	COG2885	M		100.00	100.00
RCA23_c01020	endonuclease III	COG0177	L		100.00	100.00
RCA23_c01030	putative pfkB family carbohydrate kinase	COG0524	G		100.00	100.00
RCA23_c01040	hypothetical protein				100.00	100.00
RCA23_c01050	glycosyl transferase family 14				100.00	100.00
RCA23_c01060	hypothetical protein				100.00	100.00
RCA23_c01070	HIT-like protein	COG0537	F		100.00	100.00
RCA23_c01080	ABC transporter ATP binding protein	COG1131	V		100.00	96.42
RCA23_c01090	hypothetical protein	COG4391	S		100.00	100.00
RCA23_c01100	hypothetical protein				100.00	100.00
RCA23_c01110	DNA polymerase I	COG0749	L		100.00	100.00
RCA23_c01120	cystathionine gamma-synthase MetB	COG0626	E		100.00	100.00
RCA23_c01130	ribosomal large subunit pseudouridine synthase RluE	COG1187	J		100.00	100.00
RCA23_c01140	hypothetical protein DUF1285	COG3816	S		100.00	100.00
RCA23_c01150	ATPase, MoxR type	COG0714	R		100.00	100.00
RCA23_c01160	hypothetical protein	COG1721	R		100.00	94.56
RCA23_c01170	hypothetical protein				100.00	100.00
RCA23_c01180	hypothetical protein DUF1355	COG5426	S		100.00	100.00
RCA23_c01190	glycolate oxidase subunit GlcD	COG0277	C		100.00	80.83
RCA23_c01200	glycolate oxidase subunit GlcE	COG0277	C		100.00	100.00
RCA23_c01210	glycolate oxidase iron-sulfur subunit GlcF	COG0247	C		100.00	100.00
RCA23_c01220	hypothetical protein, trypsin	COG3591	E		100.00	100.00
RCA23_c01230	putative chitinase	COG3325	G		15.64	33.69
RCA23_c01240	small heat shock protein IbpA	COG0071	O		100.00	100.00
RCA23_c01250	succinate-semialdehyde dehydrogenase GabD	COG1012	C		100.00	100.00
RCA23_c01260	hypothetical protein	COG0657	I		100.00	100.00
RCA23_c01270	hypothetical protein	COG5617	S		100.00	100.00

RCA23_c01280	putative cyclopentanol dehydrogenase CpnA	COG1028	I	100.00	100.00
RCA23_c01290	hypothetical protein			100.00	100.00
RCA23_c01300	hypothetical protein DUF1523			100.00	100.00
RCA23_c01310	aldehyde dehydrogenase	COG1012	C	100.00	100.00
RCA23_c01320	deoxyribose-phosphate aldolase DeoC	COG0274	F	100.00	100.00
RCA23_c01330	ribulose-phosphate 3-epimerase, chromosomal	COG0036	G	100.00	100.00
RCA23_c01340	hemolysin-type calcium-binding region			100.00	100.00
RCA23_c01350	5-oxoprolinase (ATP-hydrolyzing)	COG0145	E	95.58	99.00
RCA23_c01360	aminotransferase class-III	COG0161	H	100.00	100.00
RCA23_c01370	ureidoglycolate hydrolase AllA	COG3194	F	100.00	100.00
RCA23_c01380	aldose 1-epimerase GalM	COG2017	G	100.00	100.00
RCA23_c01390	beta-galactosidase BgaB	COG1874	G	100.00	96.33
RCA23_c01400	putative gluconolactonase	COG3386	G	100.00	100.00
RCA23_c01410	2-dehydro-3-deoxy-6-phosphogalactonate aldolase DgoA	COG0800	G	100.00	100.00
RCA23_c01420	2-dehydro-3-deoxygalactonokinase DgoK	COG3734	G	100.00	100.00
RCA23_c01430	short chain dehydrogenase	COG1028	I	100.00	100.00
RCA23_c01440	alpha-galactosidase RafA	COG3345	G	100.00	100.00
RCA23_c01450	putative ABC transporter inner membrane component	COG0395	G	100.00	98.89
RCA23_c01460	putative ABC transporter inner membrane component	COG1175	G	100.00	100.00
RCA23_c01470	putative extracellular solute-binding protein	COG1653	G	42.04	70.25
RCA23_c01480	HTH-type transcriptional regulator, IclR family	COG1414	K	51.78	84.73
RCA23_c01490	sugar ABC transporter ATP-binding protein	COG3839	G	100.00	87.72
RCA23_c01500	Sulfite exporter TauE/SafE			100.00	100.00
RCA23_c01510	putative transcriptional regulator, gntR family	COG1167	K	100.00	100.00
RCA23_c01520	taurine--pyruvate aminotransferase Tpa	COG0161	H	95.52	95.92
RCA23_c01530	taurine ABC transporter, periplasmic binding protein TauA	COG4521	P	100.00	100.00
RCA23_c01540	taurine ABC transport system ATP-binding protein TauB	COG1116	P	94.55	100.00
RCA23_c01550	taurine ABC transport system permease protein TauC	COG0600	P	100.00	100.00
RCA23_c01560	dimethylglycine dehydrogenase	COG0404	E	100.00	99.21
RCA23_c01570	trimethylamine methyltransferase MttB	COG5598	H	100.00	100.00
RCA23_c01580	3-hydroxyacyl-CoA dehydrogenase FadN	COG1250	I	100.00	100.00
RCA23_c01590	hypothetical protein	COG0596	R	100.00	100.00
RCA23_c01600	O-acetylhomoserine (thiol)-lyase CysD	COG2873	E	100.00	97.02

RCA23_c01610	putative signaling protein	COG2200	T	100.00	100.00
RCA23_c01620	sulfoacetaldehyde acetyltransferase Xsc	COG0028	E	100.00	100.00
RCA23_c01630	phosphate acetyltransferase Pta	COG0280	C	100.00	100.00
RCA23_c01640	DMSO reductase chain A	COG0243	C	100.00	98.54
RCA23_c01650	alpha/beta hydrolase	COG0596	R	100.00	93.42
RCA23_c01660	DMSO reductase chain B	COG0437	C	100.00	100.00
RCA23_c01670	DMSO reductase chain C	COG3302	R	100.00	100.00
RCA23_c01680	sodium:alanine symporter	COG1115	E	100.00	100.00
RCA23_c01690	thymidine kinase Tdk	COG1435	F	100.00	100.00
RCA23_c01700	NADPH-dependent FMN reductase	COG0431	R	100.00	100.00
RCA23_c01710	carbamoyl-phosphate synthase large chain CarB	COG0458	E	73.66	100.00
RCA23_c01720	HTH-type transcriptional regulator, AsnC family	COG1522	K	44.51	100.00
RCA23_c01730	hypothetical protein			100.00	100.00
RCA23_c01740	aspartyl-tRNA synthase AspS	COG0173	J	100.00	100.00
RCA23_c01750	hypothetical protein			100.00	100.00
RCA23_c01760	putative signal transduction response regulator receiver protein	COG2197	T	100.00	64.91
RCA23_c01770	methylmalonyl-CoA epimerase	COG0346	E	100.00	100.00
RCA23_c01780	hypothetical protein DUF1467	COG5454	S	100.00	100.00
RCA23_c01790	hypothetical protein DUF540	COG2981	E	39.77	100.00
RCA23_c01800	putative nitroreductase	COG0778	C	100.00	100.00
RCA23_c01810	histone deacetylase-like amidohydrolase HdaH	COG0123	B	100.00	100.00
RCA23_c01820	hypothetical protein			100.00	100.00
RCA23_c01830	hypothetical protein, peptidase family M48	COG0501	O	100.00	100.00
RCA23_c01840	hypothetical protein			100.00	100.00
RCA23_c01850	ribosomal protein S12 methylthiotransferase RimO	COG0621	J	100.00	100.00
RCA23_c01860	hypothetical protein, transmembrane			100.00	100.00
RCA23_c01870	GcrA cell cycle regulator	COG5352	S	100.00	96.34
RCA23_c01880	ABC-2 type transport system membrane protein	COG0842	V	100.00	100.00
RCA23_c01890	acetylornithine aminotransferase ArgD	COG4992	E	100.00	100.00
RCA23_c01900	ornithine carbamoyltransferase ArgF	COG0078	E	100.00	100.00
RCA23_c01910	hypothetical protein			100.00	100.00
RCA23_c01920	6-phosphogluconate dehydrogenase GntZ	COG0362	G	100.00	100.00
RCA23_c01930	ATP-dependent RNA helicase HrpB	COG1643	L	100.00	98.78

RCA23_c01940	LAO/AO transport system ATPase	COG1703	E	100.00	100.00
RCA23_c01950	50S ribosomal protein L28	COG0227	J	100.00	100.00
RCA23_c01960	hypothetical protein			100.00	100.00
RCA23_c01970	GTP-binding protein LepA	COG0481	M	100.00	93.71
RCA23_c01980	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	100.00	100.00
RCA23_c01990	putative ring-cleaving dioxygenase	COG0346	E	100.00	100.00
RCA23_c02000	serine hydroxymethyltransferase GlyA	COG0112	E	100.00	100.00
RCA23_c02010	putative inorganic polyphosphate/ATP-NAD kinase PpnK	COG0061	G	100.00	100.00
RCA23_c02020	propionate--CoA ligase PrpE	COG0365	I	100.00	100.00
RCA23_c02030	hypothetical protein	COG1917	S	86.94	100.00
RCA23_c02040	cytidine deaminase Cdd	COG0295	F	100.00	100.00
RCA23_c02050	thymidine phosphorylase DeoA	COG0213	F	100.00	93.50
RCA23_c02060	phosphopentomutase DeoB	COG1015	G	100.00	91.69
RCA23_c02070	adenosine deaminase Add	COG1816	F	100.00	100.00
RCA23_c02080	uracil phosphoribosyltransferase Upp	COG0035	F	100.00	100.00
RCA23_c02090	hypothetical protein			100.00	100.00
RCA23_c02100	3-deoxy-D-manno-octulosonic-acid transferase WaaA	COG1519	M	100.00	100.00
RCA23_c02110	L-sorbose 1-dehydrogenase	COG2303	E	100.00	100.00
RCA23_c02120	putative N-acylneuraminate cytidylyltransferase	COG1083	M	100.00	100.00
RCA23_c02130	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase ArnB	COG0399	M	100.00	100.00
RCA23_c02140	hypothetical protein	COG3119	P	0.00	77.21
RCA23_c02150	hypothetical protein, nucleoside triphosphate hydrolases-like	COG1122	P	0.00	69.70
RCA23_c02160	S-adenosyl-L-methionine-dependent methyltransferase	COG2226	H	0.00	100.00
RCA23_c02170	hypothetical protein			100.00	99.89
RCA23_c02180	medium-chain-fatty-acid--CoA ligase AlkK	COG0318	I	100.00	100.00
RCA23_c02190	hypothetical protein DUF6 transmembrane	COG0697	G	100.00	100.00
RCA23_c02200	fatty acid oxidation complex alpha subunit FadJ	COG1250	I	100.00	100.00
RCA23_c02210	hypothetical protein			100.00	96.38
RCA23_c02220	short chain dehydrogenase	COG0183	I	100.00	100.00
RCA23_c02230	putative glutathione S-transferase	COG0625	O	100.00	100.00
RCA23_c02240	acyl-CoA dehydrogenase MmgC	COG1960	I	100.00	100.00
RCA23_c02250	putative HTH-type transcriptional regulator, MerR family	COG0789	K	100.00	100.00
RCA23_c02260	putative HTH-type transcriptional regulator, MerR family	COG0789	K	100.00	100.00

RCA23_c02270	hypothetical protein, transmembrane protein DUF2899			100.00	100.00
RCA23_c02280	hypothetical protein, thioesterase	COG2050	Q	100.00	100.00
RCA23_c02290	hypothetical protein, thioesterase	COG2050	Q	100.00	100.00
RCA23_c02300	DNA-damage-inducible protein F	COG0534	V	100.00	100.00
RCA23_c02310	dihydroorotate dehydrogenase PyrD	COG0167	F	100.00	100.00
RCA23_c02320	hypothetical protein DUF952	COG3502	S	100.00	100.00
RCA23_c02330	5'-nucleotidase SurE	COG0737	F	100.00	97.26
RCA23_c02340	hypothetical protein	COG3409	M	100.00	100.00
RCA23_c02350	glycyl-tRNA synthase alpha subunit GlyQ	COG0752	J	100.00	100.00
RCA23_c02360	hypothetical protein			100.00	100.00
RCA23_c02370	glycyl-tRNA synthase beta subunit GlyS	COG0751	J	100.00	100.00
RCA23_c02380	pyruvate, phosphate dikinase PpdK	COG0574	G	100.00	98.35
RCA23_c02390	putative cell wall hydrolase	COG3773	M	100.00	100.00
RCA23_c02400	hypothetical protein, dihydroneopterin aldolase	COG1539	H	100.00	100.00
RCA23_c02410	dihydropteroate synthase FolP	COG0294	H	100.00	100.00
RCA23_c02420	putative integral membrane protein DUF6	COG0697	G	100.00	100.00
RCA23_c02430	ketol-acid reductoisomerase IlvC	COG0059	E	100.00	100.00
RCA23_c02440	putative transcriptional regulator, asnC family	COG1522	K	100.00	100.00
RCA23_c02450	putative transcriptional regulator, asnC family	COG1522	K	100.00	100.00
RCA23_c02460	hypothetical protein	COG0075	E	94.44	100.00
RCA23_c02470	2-octaprenyl-6-methoxyphenol hydroxylase UbiH	COG0654	H	100.00	100.00
RCA23_c02480	putative pyrimidine 5-nucleotidase	COG1011	R	100.00	100.00
RCA23_c02490	putative HTH-type transcriptional regulator, GntR family	COG1802	K	100.00	100.00
RCA23_c02500	uncharacterized glycosyltransferase YdaM	COG1215	M	100.00	96.52
RCA23_c02510	carbamoyl-phosphate synthase small chain CarA	COG0505	E	100.00	100.00
RCA23_c02520	GatB/YqeY family protein	COG1610	S	100.00	100.00
RCA23_c02530	hypothetical protein	COG5488	S	100.00	100.00
RCA23_c02540	cytochrome c oxidase subunit 1	COG0843	C	100.00	100.00
RCA23_c02550	octanoyltransferase LipB	COG0321	H	100.00	100.00
RCA23_c02560	hypothetical protein, LytTr transcriptional regulator	COG3279	K	100.00	99.61
RCA23_c02570	membrane protein-like	COG5395	S	100.00	100.00
RCA23_c02590	arylsulfatase	COG3119	P	100.00	100.00
RCA23_c02600	NADPH dehydrogenas	COG1902	C	100.00	100.00

RCA23_c02610	L-idonate 5-dehydrogenase IdnD	COG1063	E	100.00	100.00
RCA23_c02620	gluconate 5-dehydrogenase Gno	COG1028	I	100.00	100.00
RCA23_c02630	uncharacterized oxidoreductase YgbJ	COG2084	I	100.00	100.00
RCA23_c02640	hypothetical protein, xylose isomerase-like	COG1082	G	100.00	100.00
RCA23_c02650	glyoxylate reductase GyaR	COG1052	C	100.00	100.00
RCA23_c02660	hypothetical protein			100.00	100.00
RCA23_c02670	hypothetical protein, AzlC-like	COG1296	E	100.00	27.02
RCA23_c02680	cbbT/tktB: transketolase	COG0021	G	43.25	14.48
RCA23_c02690	hypothetical protein	COG4091	E	0.00	0.00
RCA23_c02700	transketolase, alpha subunit	COG3959	G	43.65	0.00
RCA23_c02710	transketolase, beta-subunit	COG3958	G	100.00	22.22
RCA23_c02720	putative 3-hydroxyisobutyrate dehydrogenase	COG2084	I	49.38	0.00
RCA23_c02730	putative 3-hydroxyisobutyrate dehydrogenase	COG2084	I	100.00	22.31
RCA23_c02740	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	100.00	0.00
RCA23_c02750	HTH-type transcriptional regulator, LysR family	COG0583	K	100.00	0.10
RCA23_c02760	putative HTH-type transcriptional regulator, GntR family	COG1802	K	100.00	15.72
RCA23_c02770	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	14.10	0.00
RCA23_c02780	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	100.00	44.33
RCA23_c02800	hypothetical protein	COG0240	C	100.00	0.00
RCA23_c02810	hypothetical protein DUF1537	COG3395	S	46.88	0.00
RCA23_c02820	putative ribulose bisphosphate carboxylase large chain	COG1850	G	72.70	0.00
RCA23_c02830	hypothetical protein, NAD dependent epimerase / dehydratase family	COG0451	M	100.00	0.00
RCA23_c02840	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I	100.00	26.84
RCA23_c02850	hypothetical protein			7.71	0.00
RCA23_c02860	altronate hydrolase UxaA	COG2721	G	31.20	94.15
RCA23_c02870	putative oxidoreductase	COG0673	R	100.00	100.00
RCA23_c02880	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I	100.00	98.20
RCA23_c02890	S-adenosylmethionine uptake transporter Sam	COG0697	G	100.00	100.00
RCA23_c02900	2-dehydro-3-deoxygluconokinase KdgK	COG0524	G	100.00	100.00
RCA23_c02910	D-mannonate oxidoreductase UxuB	COG0246	G	100.00	100.00
RCA23_c02920	uronate isomerase UxaC	COG1904	G	100.00	100.00
RCA23_c02930	putative oxidoreductase	COG0673	R	100.00	100.00
RCA23_c02940	putative oxidoreductase	COG0673	R	100.00	100.00

RCA23_c02950	putative oxidoreductase	COG0673	R	100.00	100.00
RCA23_c02960	mannonate dehydratase UxuA	COG1312	G	100.00	100.00
RCA23_c02970	long-chain-fatty-acid--CoA ligase	COG0318	I	100.00	100.00
RCA23_c02980	D-beta-hydroxybutyrate dehydrogenase BdhA	COG1028	I	100.00	100.00
RCA23_c02990	copper-transporting P-type ATPase ActP	COG2217	P	100.00	100.00
RCA23_c03000	HTH-type transcriptional regulator (copper efflux regulator)	COG0789	K	100.00	100.00
RCA23_c03010	hypothetical protein	COG3937	S	100.00	100.00
RCA23_c03020	hypothetical protein			100.00	100.00
RCA23_c03030	isocitrate dehydrogenase Icd	COG2838	C	91.78	100.00
RCA23_c03040	integral membrane protein DUF6	COG0697	G	100.00	100.00
RCA23_c03050	flavocytochrome c cytochrome subunit SoxE	COG3474	C	100.00	100.00
RCA23_c03060	sulfide dehydrogenase [flavocytochrome c] flavoprotein chain SoxF	COG0446	R	100.00	100.00
RCA23_c03070	hypothetical protein, OsmC-like	COG1764	O	100.00	100.00
RCA23_c03080	putative dimethyl sulfoniopropionate demethylase DmdA	COG0404	E	100.00	99.83
RCA23_c03090	transcriptional regulator, AsnC family	COG1522	K	100.00	100.00
RCA23_c03100	glutathione S-transferase	COG0625	O	0.00	0.00
RCA23_c03120	hypothetical protein, ribonuclease, E/G family	COG1530	J	87.39	100.00
RCA23_c03130	Maf-like protein	COG0424	D	100.00	100.00
RCA23_c03140	translation initiation factor IF-1	COG0361	J	100.00	100.00
RCA23_c03150	hypothetical protein, low molecular weight phosphotyrosine protein phospha	COG0394	T	76.44	100.00
RCA23_c03160	hypothetical protein UPF0262	COG5328	S	32.08	100.00
RCA23_c03170	histidinol dehydrogenase HisD	COG0141	E	100.00	100.00
RCA23_c03180	hypothetical protein			100.00	100.00
RCA23_c03190	UDP-N-acetylglucosamine 1-carboxyvinyltransferase MurA	COG0766	M	100.00	100.00
RCA23_c03210	PAS/PAC sensor hybrid histidine kinase	COG0642	T	100.00	95.59
RCA23_c03220	2OG-Fe(II) oxygenase	COG3491	R	99.62	85.93
RCA23_c03230	putative lysine exporter protein	COG1280	E	77.52	67.91
RCA23_c03240	putative carbon monoxide dehydrogenase subunit G	COG3427	S	100.00	100.00
RCA23_c03250	hypothetical protein, XdhC and CoxI family	COG1975	O	100.00	95.16
RCA23_c03260	putative MFS-type transporter	COG2814	G	93.00	100.00
RCA23_c03270	AFG1-like ATPase	COG1485	R	73.77	89.30
RCA23_c03280	bifunctional protein FolC	COG0285	H	100.00	100.00
RCA23_c03290	acetyl-coenzyme A carboxylase carboxyl transferase beta subunit AccD	COG0777	I	100.00	100.00

RCA23_c03300	hypothetical protein, CAAX amino terminal protease-like	COG1266	R	100.00	100.00
RCA23_c03310	dihydroxy-acid dehydratase IlvD	COG0129	E	100.00	96.08
RCA23_c03320	hypothetical protein			100.00	100.00
RCA23_c03330	hypothetical protein			100.00	100.00
RCA23_c03340	hypothetical protein			100.00	100.00
RCA23_c03350	hypothetical protein, OmpA family	COG2885	M	100.00	100.00
RCA23_c03360	hypothetical protein, peroxidase-like protein	COG2128	S	100.00	100.00
RCA23_c03370	hypothetical protein, acetyltransferase-like	COG0456	R	100.00	100.00
RCA23_c03380	hypothetical protein, probably molybdopterin binding	COG1058	R	100.00	100.00
RCA23_c03390	sugar fermentation stimulation protein SfsA	COG1489	R	100.00	100.00
RCA23_c03400	methionine aminopeptidase Map	COG0024	J	100.00	100.00
RCA23_c03410	ATP-dependent RNA helicase RhlE	COG0513	L	100.00	100.00
RCA23_c03420	putative ribosomal RNA small subunit methyltransferase D	COG0742	L	100.00	100.00
RCA23_c03430	rhodocoxin reductase ThcD	COG0446	R	100.00	100.00
RCA23_c03440	peroxiredoxin	COG0678	O	100.00	100.00
RCA23_c03450	pterin-4-alpha-carbinolamine dehydratase	COG2154	H	100.00	100.00
RCA23_c03460	hypothetical protein DUF482	COG3146	S	100.00	97.64
RCA23_c03470	putative glycerophosphoryl diester phosphodiesterase	COG0584	C	100.00	100.00
RCA23_c03480	putative endoribonuclease L-PSP	COG0251	J	100.00	100.00
RCA23_c03490	type I secretion system protein, HlyD family	COG1566	V	100.00	87.35
RCA23_c03500	type I secretion system ATP-binding component	COG4618	R	100.00	100.00
RCA23_c03510	putative type I secretion system protein, transmembrane domain	COG4618	R	100.00	100.00
RCA23_c03520	VacJ like lipoprotein	COG2853	M	100.00	100.00
RCA23_c03530	putative toluene tolerance protein	COG2854	Q	100.00	98.18
RCA23_c03540	penicillin-binding protein 1B	COG0744	M	100.00	97.57
RCA23_c03550	aromatic-amino-acid aminotransferase TyrB	COG1448	E	100.00	100.00
RCA23_c03560	3-mercaptopyruvate sulfurtransferase SseA	COG2897	P	100.00	100.00
RCA23_c03570	SsrA-binding protein SmpB	COG0691	O	100.00	100.00
RCA23_c03580	dihydrodipicolinate synthase DapA	COG0329	E	100.00	100.00
RCA23_c03590	soluble lytic murein transglycosylase Slt	COG0741	M	100.00	100.00
RCA23_c03600	hypothetical protein DUF6 transmembrane	COG0697	G	100.00	100.00
RCA23_c03610	hypothetical protein DUF752	COG4121	S	100.00	97.44
RCA23_c03620	putative FAD-dependent oxidoreductase	COG0665	E	100.00	100.00

RCA23_c03630	hypothetical protein, glyoxalase/dioxygenase superfamily	COG0346	E	100.00	100.00
RCA23_c03640	hypothetical protein			100.00	100.00
RCA23_c03650	hypothetical protein			100.00	100.00
RCA23_c03660	cob(I)yrinic acid a,c-diamide adenosyltransferase CobO	COG2109	H	100.00	100.00
RCA23_c03670	hypothetical protein			100.00	100.00
RCA23_c03680	oligopeptide/dipeptide ABC transporter, ATP-binding protein	COG0444	E	100.00	100.00
RCA23_c03690	oligopeptide/dipeptide ABC transporter, ATP-binding protein	COG0444	E	100.00	87.27
RCA23_c03700	oligopeptide/dipeptide ABC transporter, permease protein	COG1173	E	100.00	100.00
RCA23_c03710	oligopeptide/dipeptide ABC transporter, permease protein	COG0601	E	100.00	100.00
RCA23_c03720	oligopeptide/dipeptide ABC transporter, periplasmic substrate-binding protei	COG0747	E	100.00	90.14
RCA23_c03730	HTH-type transcriptional regulator, LysR family	COG0583	K	100.00	99.78
RCA23_c03740	putative amidohydrolase 3	COG1574	R	100.00	84.16
RCA23_c03750	cobalamin biosynthesis CobW-like	COG0523	R	100.00	100.00
RCA23_c03760	hypothetical protein, restriction endonuclease type IV-like	COG4127	S	100.00	93.58
RCA23_c03770	glycine betaine transport ATP-binding protein OpuAA	COG4175	E	100.00	97.85
RCA23_c03780	glycine betaine transport system permease protein OpuAB	COG4176	E	95.35	100.00
RCA23_c03790	glycine betaine transporter substrate-binding protein OpuAC	COG2113	E	100.00	100.00
RCA23_c03800	fatty acid desaturase	COG3239	I	100.00	100.00
RCA23_c03810	CoA-transferase family III protein involved in DMSP degradation	COG1804	C	100.00	100.00
RCA23_c03820	glyceraldehyde-3-phosphate dehydrogenase Gap	COG0057	G	100.00	100.00
RCA23_c03830	ATP-dependent Clp protease proteolytic subunit ClpP	COG0740	O	91.95	100.00
RCA23_c03840	ATP-dependent Clp protease ATP-binding subunit ClpX	COG1219	O	100.00	99.53
RCA23_c03850	putative NADH ubiquinone oxidoreductase subunit NDUFA12	COG3761	C	100.00	100.00
RCA23_c03860	ABC-type transport system involved in resistance to organic solvents, periplas	COG1463	Q	100.00	100.00
RCA23_c03870	Uncharacterized protein conserved in bacteria (DUF2155)	COG4765	S	100.00	100.00
RCA23_c03880	leucyl/phenylalanyl-tRNA--protein transferase Aat	COG2360	O	100.00	100.00
RCA23_c03890	biotin carboxylase AccC	COG4770	I	100.00	100.00
RCA23_c03900	acetyl-coenzyme A synthase AcsA	COG0365	I	100.00	100.00
RCA23_c03910	high-affinity branched-chain amino acid transport ATP-binding protein LivF	COG0410	E	100.00	100.00
RCA23_c03920	high-affinity branched-chain amino acid transport ATP-binding protein BraF	COG0411	E	14.62	100.00
RCA23_c03930	putative branched-chain amino acid transport system permease protein	COG4177	E	100.00	92.60
RCA23_c03940	putative branched-chain amino acid transport system permease protein	COG0559	E	100.00	100.00
RCA23_c03950	hypothetical protein	COG0683	E	100.00	100.00

RCA23_c03960	hypothetical protein, DNA binding helix-turn helix proteins	COG1396	K		100.00	100.00
RCA23_c03970	putative response regulator receiver protein, CheY like	COG0745	T		100.00	100.00
RCA23_c03980	hypothetical protein				100.00	100.00
RCA23_c03990	sensor protein kinase WalK	COG0591	E		100.00	100.00
RCA23_c04000	hypothetical protein DUF442	COG3453	S		100.00	100.00
RCA23_c04010	hypothetical protein				100.00	100.00
RCA23_c04020	hypothetical protein	COG3750	S		100.00	100.00
RCA23_c04030	hypothetical protein				100.00	100.00
RCA23_c04040	putative Fe(3+)-transport system protein SfuB	COG1178	P		100.00	100.00
RCA23_c04060	integrase	COG0582	L	GI 2	33.16	20.26
RCA23_c04070	hypothetical protein	COG3945	S	GI 2	0.00	0.00
RCA23_c04080	hypothetical protein			GI 2	0.00	22.80
RCA23_c04090	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase subunit	COG0457	R	GI 2	0.00	15.98
RCA23_c04100	esterase, SGNH hydrolase-type	COG2755	E	GI 2	0.00	0.00
RCA23_c04110	hypothetical protein			GI 2	0.00	97.49
RCA23_c04120	hypothetical protein, tetratricopeptide repeat	COG0790	R	GI 2	0.00	47.47
RCA23_c04130	hypothetical protein			GI 2	100.00	89.46
RCA23_c04140	hypothetical protein			GI 2	100.00	100.00
RCA23_c04150	hypothetical protein			GI 2	100.00	100.00
RCA23_c04160	hypothetical protein, RmlC-like cupin family	COG3450	R	GI 2	0.00	39.42
RCA23_c04170	hypothetical protein			GI 2	0.00	71.47
RCA23_c04180	hypothetical protein, DUF3127			GI 2	0.00	100.00
RCA23_c04190	putative nucleoside triphosphate hydrolase, ATPase domain	COG1066	O	GI 2	0.00	29.78
RCA23_c04200	zinc-dependent metalloprotease	COG2931	Q	GI 2	65.16	41.30
RCA23_c04210	HTH-type transcriptional regulator, AsnC family	COG1522	K	GI 2	100.00	100.00
RCA23_c04220	integrase	COG2801	L	GI 2	0.00	0.00
RCA23_c04230	putative ion channel			GI 2	96.98	100.00
RCA23_c04240	integrase	COG4974	L	GI 2	100.00	72.51
RCA23_c04250	Gcn5-like N-acetyltransferase	COG1670	J	GI 2	0.00	0.00
RCA23_c04260	integrase	COG0582	L	GI 2	36.10	94.11
RCA23_c04280	hypothetical protein, NAD dependent epimerase / dehydratase family	COG0702	M		100.00	100.00
RCA23_c04290	undecaprenyl-diphosphatase UppP	COG1968	V		100.00	100.00
RCA23_c04300	glutamate synthase [NADPH] small chain GltD	COG0493	E		100.00	100.00

RCA23_c04310	glutamate synthase [NADPH] large chain GltB	COG0069	E	100.00	99.98
RCA23_c04320	monofunctional biosynthetic peptidoglycan transglycosylase MtgA	COG0744	M	100.00	97.45
RCA23_c04330	putative glutathione S-transferase	COG0625	O	100.00	100.00
RCA23_c04340	putative electron transport protein yjeS	COG1600	C	100.00	100.00
RCA23_c04360	hypothetical protein	COG5266	P	100.00	100.00
RCA23_c04350	hypothetical protein			100.00	100.00
RCA23_c04370	branched-chain-amino-acid aminotransferase IlvE	COG0115	E	100.00	100.00
RCA23_c04380	HTH-type transcriptional regulator PetP	COG1846	K	100.00	100.00
RCA23_c04390	protein PetR	COG0745	T	100.00	100.00
RCA23_c04400	putative deacytelase, histone deacetylase superfamily protein-like	COG0123	B	100.00	100.00
RCA23_c04410	exodeoxyribonuclease 7 small subunit XseB	COG1722	L	100.00	100.00
RCA23_c04420	geranyltranstransferase IspA	COG0142	H	91.64	100.00
RCA23_c04430	1-deoxy-D-xylulose-5-phosphate synthase Dxs	COG1154	H	100.00	97.90
RCA23_c04440	hypothetical protein, NAD dependent epimerase/dehydratase	COG0451	M	100.00	100.00
RCA23_c04450	hypothetical protein			100.00	100.00
RCA23_c04460	hypothetical protein			100.00	100.00
RCA23_c04470	carnitinyl-CoA dehydratase CaiD	COG1024	I	100.00	100.00
RCA23_c04480	hypothetical protein, CoA-binding	COG1042	C	100.00	98.47
RCA23_c04490	putative acyl-CoA dehydrogenase YngJ	COG1960	I	100.00	100.00
RCA23_c04500	transcriptional regulator, AraC family	COG4977	K	100.00	100.00
RCA23_c04510	class II aldolase	COG0235	G	100.00	100.00
RCA23_c04520	arsenate reductase ArsC	COG1393	P	100.00	100.00
RCA23_c04530	adenine deaminase Ade	COG1001	F	100.00	100.00
RCA23_c04540	AMP nucleosidase Amn	COG0775	F	100.00	92.89
RCA23_c04550	DNA-binding protein HU	COG0776	L	100.00	100.00
RCA23_c04560	inner membrane protein DUF6	COG0697	G	100.00	100.00
RCA23_c04570	hypothetical protein, cytochrome C biogenesis protein	COG0785	O	100.00	100.00
RCA23_c04580	hypothetical protein, thioredoxin	COG0526	O	100.00	100.00
RCA23_c04590	hypothetical protein	COG3222	S	100.00	100.00
RCA23_c04600	succinyl-diaminopimelate desuccinylase DapE	COG0624	E	100.00	98.32
RCA23_c04610	hypothetical protein			100.00	100.00
RCA23_c04620	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase DapD	COG2171	E	100.00	100.00
RCA23_c04630	L-threonine ammonia-lyase	COG1171	E	100.00	100.00

RCA23_c04640	ribosomal RNA large subunit methyltransferase N	COG0820	R	100.00	100.00
RCA23_c04650	hypothetical protein			100.00	100.00
RCA23_c04660	L-asparaginase II	COG4448	E	100.00	100.00
RCA23_c04670	phosphoserine phosphatase SerB	COG0560	E	100.00	100.00
RCA23_c04680	hypothetical protein			100.00	100.00
RCA23_c04690	phosphoserine aminotransferase SerC	COG1932	H	100.00	100.00
RCA23_c04700	D-3-phosphoglycerate dehydrogenase SerA	COG0111	H	100.00	100.00
RCA23_c04710	metallophosphoesterase			100.00	100.00
RCA23_c04720	RNA pyrophosphohydrolase RppH	COG0494	L	100.00	100.00
RCA23_c04730	carboxy-terminal-processing protease CtpA	COG0793	M	100.00	100.00
RCA23_c04740	2,3-bisphosphoglycerate-independent phosphoglycerate mutase GpmI	COG0696	G	100.00	100.00
RCA23_c04750	putative protein ImuB			100.00	100.00
RCA23_c04760	hypothetical protein			100.00	100.00
RCA23_c04770	hypothetical protein, NlpC/P60	COG0791	M	100.00	100.00
RCA23_c04780	putative cytosol aminopeptidase PepA	COG0260	E	100.00	100.00
RCA23_c04790	hypothetical protein			100.00	100.00
RCA23_c04800	carbonic anhydrase CynT	COG0288	P	100.00	100.00
RCA23_c04810	aspartate-semialdehyde dehydrogenase Asd	COG0136	E	100.00	100.00
RCA23_c04820	MFS-type transporter	COG2223	P	100.00	93.28
RCA23_c04830	short chain dehydrogenase	COG4221	R	100.00	100.00
RCA23_c04840	FAD dependent oxidoreductase	COG0665	E	100.00	100.00
RCA23_c04850	trimethylamine methyltransferase MttB	COG5598	H	100.00	100.00
RCA23_c04860	aldehyde dehydrogenase	COG1012	C	100.00	100.00
RCA23_c04870	dihydrodipicolinate synthase DapA	COG0329	E	100.00	100.00
RCA23_c04880	transcriptional regulator, GntR family	COG1802	K	100.00	100.00
RCA23_c04890	hypothetical protein			100.00	100.00
RCA23_c04900	isopropylmalate dehydrogenase LeuB	COG0473	C	100.00	100.00
RCA23_c04910	hypothetical protein			100.00	100.00
RCA23_c04920	3-isopropylmalate dehydratase small subunit LeuD	COG0066	E	100.00	100.00
RCA23_c04930	3-isopropylmalate dehydratase large subunit LeuC	COG0065	E	100.00	99.50
RCA23_c04940	hypothetical protein, DUF2975			100.00	92.13
RCA23_c04950	hypothetical protein DUF143	COG0799	S	100.00	100.00
RCA23_c04960	ribosomal RNA large subunit methyltransferase H	COG1576	S	100.00	100.00

RCA23_c04970	alcohol dehydrogenase class-3 AdhI	COG1062	C	100.00	100.00
RCA23_c04980	putative membrane transport protein	COG0679	R	100.00	100.00
RCA23_c04990	S-formylglutathione hydrolase YeiG	COG0627	R	100.00	100.00
RCA23_c05000	soluble pyridine nucleotide transhydrogenase	COG1249	C	100.00	100.00
RCA23_c05010	hypothetical protein DUF188	COG1671	S	100.00	100.00
RCA23_c05020	putative integral membrane protein	COG0697	G	100.00	100.00
RCA23_c05030	putative HAD-superfamily hydrolase	COG1011	R	100.00	100.00
RCA23_c05040	putative ornithine cyclodeaminase	COG2423	E	100.00	97.47
RCA23_c05050	aerobic cobaltochelatase subunit CobN	COG1429	H	100.00	100.00
RCA23_c05060	hypothetical protein, acetyltransferase-like	COG0456	R	100.00	100.00
RCA23_c05070	protein CobW	COG0523	R	100.00	91.48
RCA23_c05080	hypothetical protein DUF1636	COG5469	S	100.00	100.00
RCA23_c05090	possible cobalt transporter, subunit CbtA	COG5446	S	100.00	100.00
RCA23_c05100	Cytochrome c peroxidase	COG1858	P	100.00	100.00
RCA23_c05110	hypothetical protein	COG4188	R	100.00	87.50
RCA23_c05120	hypothetical protein			42.94	100.00
RCA23_c05130	xylose repressor XylR	COG1940	K	90.90	100.00
RCA23_c05140	D-xylose-binding periplasmic protein XylF	COG4213	G	100.00	100.00
RCA23_c05150	xylose transport system permease protein XylH	COG4214	G	100.00	100.00
RCA23_c05160	xylose ABC transporter, ATP-binding protein XylG	COG1129	G	100.00	100.00
RCA23_c05170	xylulose kinase XylB	COG1070	G	100.00	100.00
RCA23_c05180	xylose isomerase XylA	COG2115	G	100.00	95.85
RCA23_c05190	long-chain-fatty-acid--CoA ligase LcfB	COG0318	I	100.00	98.16
RCA23_c05200	malonyl-CoA decarboxylase			100.00	98.72
RCA23_c05210	acyl-CoA dehydrogenase	COG1960	I	100.00	98.98
RCA23_c05220	hypothetical protein DUF81			100.00	99.63
RCA23_c05230	phosphoenolpyruvate carboxykinase PckA	COG1866	C	100.00	98.71
RCA23_c05240	two component signal transduction response regulator receiver protein ChvI	COG0745	T	100.00	96.15
RCA23_c05250	two component signal transduction histidine kinase ChvG	COG0642	T	100.00	100.00
RCA23_c05260	hypothetical protein, HPr serine kinase	COG1493	T	100.00	76.41
RCA23_c05270	putative P-loop containing ATPase	COG1660	R	100.00	100.00
RCA23_c05280	hypothetical protein, PTS system mannose-specific EIIA component	COG2893	G	100.00	100.00
RCA23_c05290	phosphocarrier protein NPr	COG1925	G	100.00	100.00

RCA23_c05300	electron transfer flavoprotein alpha subunit EtfA	COG2025	C	100.00	93.66
RCA23_c05310	electron transfer flavoprotein beta subunit EtfB	COG2086	C	100.00	100.00
RCA23_c05320	cob(I)yrinic acid a,c-diamide adenosyltransferase CobO	COG2096	S	100.00	100.00
RCA23_c05330	putative short chain dehydrogenase	COG4221	R	100.00	100.00
RCA23_c05340	DNA topoisomerase 4 subunit A	COG0188	L	100.00	100.00
RCA23_c05350	hypothetical protein DUF898	COG4269	S	54.73	98.47
RCA23_c05360	peptidase M48 Ste24p	COG4783	R	96.95	93.73
RCA23_c05370	hypothetical protein			100.00	100.00
RCA23_c05380	elongation factor Tu (EF-Tu)	COG0050	J	100.00	34.35
RCA23_c05400	hypothetical protein, DUF560			100.00	100.00
RCA23_c05410	hippurate hydrolase HipO	COG1473	R	100.00	98.89
RCA23_c05420	glutathione import ATP-binding protein GsiA	COG1123	R	100.00	90.41
RCA23_c05430	oligopeptide-binding protein AppA	COG0747	E	100.00	100.00
RCA23_c05440	oligopeptide transport system permease protein AppB	COG0601	E	100.00	100.00
RCA23_c05450	oligopeptide transport system permease protein AppC	COG1173	E	100.00	100.00
RCA23_c05460	ABC transporter ATP-binding protein	COG1131	V	100.00	100.00
RCA23_c05470	inner membrane transport permease	COG0842	V	100.00	95.54
RCA23_c05480	hypothetical protein, peptidase family S49	COG0616	O	100.00	99.75
RCA23_c05490	putative sodium/calcium exchanger protein	COG0530	P	100.00	100.00
RCA23_c05500	putative short cprotein	COG1028	I	100.00	100.00
RCA23_c05510	UvrABC system protein C	COG0322	L	100.00	100.00
RCA23_c05520	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	COG0558	I	100.00	100.00
RCA23_c05530	molybdopterin-converting factor subunit MoaD	COG1977	H	100.00	100.00
RCA23_c05540	molybdopterin synthase catalytic subunit MoaE	COG0314	H	100.00	100.00
RCA23_c05550	hypothetical protein			100.00	89.09
RCA23_c05560	hypothetical protein, OmpA	COG2885	M	100.00	100.00
RCA23_c05570	UbiA prenyltransferase	COG0382	H	100.00	100.00
RCA23_c05580	RNA methyltransferase	COG1385	S	100.00	100.00
RCA23_c05590	hypothetical protein			100.00	100.00
RCA23_c05600	glutamate--cysteine ligase	COG3572	H	100.00	100.00
RCA23_c05610	glycerol-3-phosphate acyltransferase PlsY	COG0344	S	100.00	100.00
RCA23_c05620	dihydroorotase PyrC	COG0044	F	100.00	95.87
RCA23_c05630	aspartate carbamoyltransferase PyrB	COG0540	F	100.00	100.00

RCA23_c05640	uracil DNA glycosylase family protein	COG1573	L	100.00	100.00
RCA23_c05650	molybdenum cofactor biosynthesis protein MoaB	COG0521	H	100.00	100.00
RCA23_c05660	putative efflux transporter, RND family, membrane fusion protein	COG1566	V	100.00	100.00
RCA23_c05670	transporter, AcrB/AcrD/AcrF family	COG0841	V	100.00	98.97
RCA23_c05680	hypothetical protein	COG4233	O	100.00	100.00
RCA23_c05690	hypothetical protein DUF179	COG1678	K	100.00	100.00
RCA23_c05700	acyl-CoA dehydrogenase	COG1960	I	100.00	98.18
RCA23_c05710	putative metallo-beta-lactamase family protein	COG0491	R	100.00	100.00
RCA23_c05720	hypothetical protein			100.00	100.00
RCA23_c05730	putative branched-chain amino acid transport protein	COG4392	S	100.00	100.00
RCA23_c05740	protein AzlC	COG1296	E	100.00	85.33
RCA23_c05750	formate dehydrogenase accessory protein FdhD	COG1526	C	100.00	100.00
RCA23_c05760	molybdopterin-guanine dinucleotide biosynthesis protein MobA	COG0746	H	100.00	100.00
RCA23_c05770	molybdopterin-guanine dinucleotide biosynthesis protein MobB	COG1763	H	100.00	97.79
RCA23_c05780	molybdopterin biosynthesis protein MoeA	COG0303	H	54.18	100.00
RCA23_c05790	transcription elongation factor GreA	COG0782	K	100.00	100.00
RCA23_c05800	electron transfer flavoprotein-ubiquinone oxidoreductase	COG0644	C	100.00	100.00
RCA23_c05810	tetratricopeptide repeat-containing protein	COG0457	R	100.00	100.00
RCA23_c05820	4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate synthase IspE	COG1947	I	100.00	100.00
RCA23_c05830	octaprenyl-diphosphate synthase IspB	COG0142	H	100.00	99.80
RCA23_c05840	hypothetical protein, methyltransferase	COG4123	R	100.00	100.00
RCA23_c05850	acetoacetyl-CoA reductase PhaB	COG1028	I	100.00	100.00
RCA23_c05860	acetyl-CoA acetyltransferase PhaA	COG0183	I	100.00	100.00
RCA23_c05870	signaling protein	COG2200	T	100.00	100.00
RCA23_c05880	DNA-3-methyladenine glycosylase 1	COG2818	L	100.00	100.00
RCA23_c05890	thiol:disulfide interchange protein TlpA	COG0526	O	100.00	100.00
RCA23_c05900	argininosuccinate lyase ArgH	COG0165	E	100.00	100.00
RCA23_c05910	hypothetical protein			100.00	100.00
RCA23_c05920	diaminopimelate decarboxylase LysA	COG0019	E	100.00	100.00
RCA23_c05930	hypothetical protein	COG1196	D	86.19	99.92
RCA23_c05940	putative acyltransferase	COG0204	I	100.00	100.00
RCA23_c05950	glyoxalase/bleomycin resistance protein	COG2764	S	100.00	100.00
RCA23_c05960	hypothetical protein, pyridoxamine 5'-phosphate oxidase	COG3576	R	100.00	100.00

RCA23_c05970	acetyl-coenzyme A carboxylase carboxyl transferase alpha subunit AccA	COG0825	I	100.00	100.00
RCA23_c05980	malyl-CoA ligase	COG2301	G	100.00	92.08
RCA23_c05990	hypothetical protein DUF1611	COG3367	S	100.00	100.00
RCA23_c06000	L-Ala-D/L-Glu epimerase YcjG	COG4948	M	100.00	100.00
RCA23_c06010	D-alanine aminotransferase Dat	COG0115	E	100.00	100.00
RCA23_c06020	hypothetical, OmpA-like			93.16	100.00
RCA23_c06030	hypothetical protein	COG3743	S	100.00	100.00
RCA23_c06040	hypothetical protein DUF1244	COG3492	S	100.00	98.35
RCA23_c06050	N-formylglutamate amidohydrolase	COG3931	E	100.00	100.00
RCA23_c06060	pyruvate kinase PykF	COG0469	G	100.00	100.00
RCA23_c06070	hypothetical protein			100.00	100.00
RCA23_c06080	50S ribosomal protein L35	COG0291	J	100.00	100.00
RCA23_c06090	50S ribosomal protein L20	COG0292	J	100.00	100.00
RCA23_c06100	putative subtilase family protein	COG1404	O	74.74	100.00
RCA23_c06110	hypothetical protein			100.00	100.00
RCA23_c06120	hypothetical protein, lipid A biosynthesis acyltransferase	COG1560	M	100.00	94.50
RCA23_c06130	phenylalanyl-tRNA synthase alpha chain PheS	COG0016	J	100.00	91.34
RCA23_c06140	glutamine amidotransferase class-I	COG0518	F	100.00	100.00
RCA23_c06150	phenylalanyl-tRNA synthase beta chain PheT	COG0072	J	100.00	100.00
RCA23_c06160	putative HTH-type transcriptional regulator	COG1522	K	100.00	100.00
RCA23_c06170	ribsomal protein S21	COG0828	J	100.00	100.00
RCA23_c06180	putative ubiquinone biosynthesis protein COQ9	COG5590	S	100.00	100.00
RCA23_c06190	putative quinone oxidoreductase	COG0604	C	100.00	100.00
RCA23_c06200	hypothetical protein DUF1013	COG3820	S	100.00	93.10
RCA23_c06210	recombination protein RecR	COG0353	L	100.00	100.00
RCA23_c06220	hypothetical protein	COG0718	S	100.00	100.00
RCA23_c06230	DNA polymerase III subunit tau	COG2812	L	100.00	98.44
RCA23_c06240	NADH pyrophosphatase Nudc	COG2816	L	100.00	100.00
RCA23_c06250	hypothetical protein	COG3832	S	100.00	100.00
RCA23_c06260	prephenate dehydratase PheA	COG0077	E	100.00	100.00
RCA23_c06270	cytochrome c-552	COG3474	C	100.00	100.00
RCA23_c06280	ABC transporter extracellular solute-binding protein	COG4166	E	100.00	100.00
RCA23_c06290	ABC transporter permease protein	COG4174	R	100.00	100.00

RCA23_c06300	ABC transporter permease protein	COG4239	R		100.00	100.00
RCA23_c06310	putative oligopeptide ABC transporter ATP-binding protein	COG4172	R		100.00	98.99
RCA23_c06320	fumarylacetoacetate hydrolase family protein	COG0179	Q		100.00	98.68
RCA23_c06330	D-alanyl-D-alanine carboxypeptidase DacF	COG1686	M		100.00	100.00
RCA23_c06340	haloacid dehalogenase domain protein hydrolase	COG1011	R		100.00	100.00
RCA23_c06350	ATP-dependent Clp protease adapter protein ClpS	COG2127	S		100.00	100.00
RCA23_c06360	putative methyltransferase	COG2813	J		100.00	100.00
RCA23_c06370	putative short chain dehydrogenase	COG1028	I		100.00	99.50
RCA23_c06380	coproporphyrinogen 3 oxidase, aerobic	COG0408	H		100.00	100.00
RCA23_c06390	oxidoreductase, FAD-binding protein	COG4097	P		100.00	93.60
RCA23_c06400	D-lactate dehydrogenase	COG0277	C		100.00	100.00
RCA23_c06420	integrase			GI 3	98.18	62.52
RCA23_c06430	mandelate racemase	COG4948	M	GI 3	0.00	0.00
RCA23_c06440	fumarylacetoacetate hydrolase family protein	COG0179	Q	GI 3	0.00	0.00
RCA23_c06450	short-chain dehydrogenase/reductase	COG1028	I	GI 3	0.00	0.00
RCA23_c06460	amidohydrolase	COG3618	R	GI 3	0.00	0.00
RCA23_c06470	aldo/keto reductase	COG0667	C	GI 3	0.00	0.00
RCA23_c06480	ABC transporter, permease protein	COG0600	P	GI 3	0.00	0.00
RCA23_c06490	ABC transporter, permease protein	COG0600	P	GI 3	0.00	0.00
RCA23_c06500	hypothetical protein			GI 3	0.00	0.00
RCA23_c06510	sulfonate/nitrate ABC transporter, ATPase	COG1116	P	GI 3	0.00	0.00
RCA23_c06520	ABC transporter, periplasmic substrate-binding protein	COG0715	P	GI 3	0.00	0.00
RCA23_c06530	arylsulfatase	COG3119	P	GI 3	51.23	0.00
RCA23_c06540	transcriptional regulator, GntR family	COG1802	K	GI 3	100.00	0.00
RCA23_c06550	hypothetical protein	COG4974	L	GI 3	1.94	0.00
RCA23_c06560	hypothetical protein	COG1618	F	GI 3	100.00	0.00
RCA23_c06570	hypothetical protein			GI 3	100.00	99.63
RCA23_c06580	hypothetical protein			GI 3	100.00	56.03
RCA23_c06590	hypothetical protein			GI 3	100.00	89.19
RCA23_c06600	hypothetical protein			GI 3	94.99	100.00
RCA23_c06610	hypothetical protein			GI 3	100.00	100.00
RCA23_c06620	hypothetical protein			GI 3	100.00	100.00
RCA23_c06630	hypothetical protein	COG3000	I	GI 3	0.00	72.84

RCA23_c06640	transcriptional regulator, AsnC family	COG1522	K	GI 3	100.00	100.00
RCA23_c06650	aspartate aminotransferase AspC	COG0436	E	GI 3	100.00	94.71
RCA23_c06660	3-hydroxybutyrate dehydrogenase Bdh	COG1028	I	GI 3	99.20	100.00
RCA23_c06670	putative fumarylacetoacetate hydrolase	COG0179	Q	GI 3	0.00	67.72
RCA23_c06680	oxidoreductase, short chain dehydrogenase/reductase family protein	COG1028	I	GI 3	0.00	75.17
RCA23_c06690	dimethylmenaquinone methyltransferase	COG0684	H	GI 3	0.00	98.95
RCA23_c06700	hypothetical protein	COG3970	R	GI 3	0.00	94.60
RCA23_c06710	NADP-dependent fatty aldehyde dehydrogenase AldH	COG1012	C	GI 3	0.00	75.96
RCA23_c06720	putative dihydrodipicolinate synthase	COG0329	E	GI 3	0.00	70.75
RCA23_c06730	transcriptional regulator protein, LacI family	COG1609	K	GI 3	0.00	73.33
RCA23_c06740	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	GI 3	0.00	80.34
RCA23_c06750	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	GI 3	0.00	62.24
RCA23_c06760	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	GI 3	0.00	78.96
RCA23_c06770	D-amino acid dehydrogenase small subunit DadA	COG0665	E	GI 3	0.00	83.92
RCA23_c06780	mandelate racemase/muconate lactonizing protein	COG4948	M	GI 3	53.61	78.24
RCA23_c06790	dihydrodipicolinate synthase DapA	COG0329	E	GI 3	100.00	64.59
RCA23_c06800	uncharacterized oxidoreductase, YjmC	COG2055	C	GI 3	57.07	88.13
RCA23_c06810	sarcosine oxidase beta subunit SoxB	COG0665	E	GI 3	82.58	81.90
RCA23_c06820	sarcosine oxidase delta subunit SoxD	COG4311	E	GI 3	0.00	100.00
RCA23_c06830	sarcosine oxidase alpha subunit SoxA	COG0404	E	GI 3	41.21	75.69
RCA23_c06840	sarcosine oxidase gamma subunit SoxG	COG4583	E	GI 3	100.00	77.66
RCA23_c06850	hypothetical protein, DUF6 transmembrane protein	COG0697	G	GI 3	100.00	78.68
RCA23_c06860	putative phage integrase			GI 3	46.25	27.23
RCA23_c06870	putative amidase	COG0154	J	GI 3	0.00	0.00
RCA23_c06880	integrase	COG4974	L	GI 3	100.00	55.20
RCA23_c06890	putative helicase			GI 3	48.29	0.00
RCA23_c06900	transposase	COG3316	L	GI 3	0.00	99.00
RCA23_c06910	HTH-type transcriptional regulator, LuxR family	COG2197	T	GI 3	0.00	100.00
RCA23_c06920	putative oxidoreductase, molybdopterin binding	COG3915	S	GI 3	0.00	100.00
RCA23_c06930	signal transduction histidine kinase	COG0642	T	GI 3	0.00	73.12
RCA23_c06940	glucose-1-phosphate thymidylyltransferase RfbA	COG1209	M	GI 3	0.00	61.33
RCA23_c06950	hypothetical protein, chain length determinant protein	COG0489	D	GI 3	0.00	69.96
RCA23_c06960	polysaccharide export protein	COG1596	M	GI 3	0.00	79.68

RCA23_c06970	hypothetical protein, VanZ-like	COG5652	S	GI 3	0.00	100.00
RCA23_c06980	type I secretion system membrane fusion protein, HlyD family	COG1566	V	GI 3	0.00	22.42
RCA23_c06990	type I secretion system ATP-binding component, HlyB family	COG2274	V	GI 3	0.00	67.28
RCA23_c07000	type I secretion outer membrane protein, TolC family	COG1538	M	GI 3	0.00	52.58
RCA23_c07010	putative serralysin-like metalloprotease	COG2931	Q	GI 3	0.00	26.87
RCA23_c07020	dTDP-4-dehydrorhamnose reductase RfbD	COG1091	M	GI 3	0.00	79.65
RCA23_c07030	dTDP-glucose 4,6-dehydratase RfbB	COG1088	M	GI 3	0.00	23.63
RCA23_c07040	dTDP-4-dehydrorhamnose 3,5-epimerase RfbC	COG1898	M	GI 3	0.00	88.30
RCA23_c07050	hypothetical protein			GI 3	0.00	80.02
RCA23_c07060	glycosyltransferase	COG0438	M	GI 3	0.00	21.48
RCA23_c07070	putative glycosyltransferase	COG0438	M	GI 3	0.00	46.21
RCA23_c07080	hypothetical protein, UDP-glycosyltransferase/glycogen phosphorylase-like	COG0438	M	GI 3	0.00	23.52
RCA23_c07090	polysaccharide biosynthesis protein			GI 3	0.00	31.69
RCA23_c07100	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase ArnB	COG0399	M	GI 3	0.00	0.00
RCA23_c07110	hypothetical, WxcM-like	COG0110	R	GI 3	0.00	53.54
RCA23_c07120	hypothetical protein, WxcM-like			GI 3	0.00	0.00
RCA23_c07130	hypothetical protein, acyltransferase-like	COG1670	J	GI 3	0.00	0.00
RCA23_c07140	UDP-glucose/GDP-mannose dehydrogenase family protein	COG0677	M	GI 3	0.00	68.17
RCA23_c07150	S-adenosyl-L-methionine-dependent methyltransferase			GI 3	0.00	25.41
RCA23_c07160	UDP-glucoronate 5'-epimerase LspL	COG0451	M	GI 3	0.00	53.53
RCA23_c07170	transcriptional regulator, MarR family	COG1846	K	GI 3	0.00	62.43
RCA23_c07180	transcription antitermination protein NusG	COG0250	K	GI 3	0.00	87.33
RCA23_c07190	hypothetical protein	COG1434	S	GI 3	0.00	81.82
RCA23_c07200	lipopolysaccharide core biosynthesis mannosyltransferase LpcC	COG0438	M	GI 3	0.00	69.74
RCA23_c07210	hypothetical protein			GI 3	0.00	100.00
RCA23_c07220	undecaprenyl-phosphate alpha-N-acetylglucosaminyl 1-phosphate transferase	COG0472	M	GI 3	0.00	80.44
RCA23_c07230	hypothetical protein, transmembrane			GI 3	0.00	84.12
RCA23_c07240	hypothetical protein, transmembrane			GI 3	0.00	68.70
RCA23_c07250	transposase, IS4 family protein			GI 3	0.00	100.00
RCA23_c07260	putative transposase	COG3039	L	GI 3	0.00	56.86
RCA23_c07270	hypothetical protein, transmembrane			GI 3	0.00	74.94
RCA23_c07280	hypothetical protein, transmembrane			GI 3	0.00	49.34
RCA23_c07290	transposase	COG3316	L	GI 3	0.00	0.00

RCA23_c07300	glycogen synthase GlgA	COG0297	G	GI 3	0.00	49.69
RCA23_c07310	glycogen debranching enzyme GlgX	COG1523	G	GI 3	0.00	35.65
RCA23_c07320	phosphoglucomutase Pgm	COG0033	G	GI 3	0.00	32.35
RCA23_c07330	putative alpha-glucosidase AglA	COG0366	G	GI 3	0.00	45.27
RCA23_c07340	transposase	COG3316	L	GI 3	0.00	100.00
RCA23_c07350	uncharacterized hydrolase YtnL	COG1473	R	GI 3	100.00	62.38
RCA23_c07360	glutathione-binding protein GsiB	COG0747	E	GI 3	89.85	100.00
RCA23_c07370	glutathione transport system permease protein GsiC	COG0601	E	GI 3	100.00	94.20
RCA23_c07380	dipeptide transport system permease protein	COG1173	E	GI 3	100.00	100.00
RCA23_c07390	glutathione import ATP-binding protein GsiA	COG4172	R	GI 3	100.00	97.59
RCA23_c07400	peptidase family S58	COG3191	E	GI 3	100.00	100.00
RCA23_c07410	aldehyde dehydrogenase AldA	COG1012	C	GI 3	100.00	100.00
RCA23_c07420	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	GI 3	100.00	97.71
RCA23_c07430	amidase	COG0154	J	GI 3	100.00	98.01
RCA23_c07440	MFS-type transporter	COG2807	P	GI 3	100.00	82.51
RCA23_c07450	peptidase M20D amidohydrolase	COG1473	R	GI 3	0.00	0.00
RCA23_c07460	transposase	COG3316	L	GI 3	0.00	0.00
RCA23_c07470	integrase	COG2801	L	GI 3	0.00	0.00
RCA23_c07480	transposase			GI 3	0.00	0.00
RCA23_c07490	hypothetical protein, DUF3764			GI 3	67.84	28.63
RCA23_c07500	hypothetical protein			GI 3	100.00	0.00
RCA23_c07510	hypothetical protein, beta-lactam-insensitive peptidoglycan transpeptidase Yk	COG1376	S	GI 3	100.00	56.43
RCA23_c07520	hypothetical protein			GI 3	100.00	100.00
RCA23_c07530	ABC transporter, ATP-binding protein, SbmA/BacA-like family	COG1133	I	GI 3	100.00	62.26
RCA23_c07540	hypothetical protein			GI 3	100.00	50.36
RCA23_c07550	hypothetical protein			GI 3	0.00	0.00
RCA23_c07560	integrase	COG2801	L	GI 3	0.00	0.00
RCA23_c07570	hypothetical protein			GI 3	0.00	0.00
RCA23_c07580	hypothetical protein			GI 3	0.00	0.00
RCA23_c07590	flagellin FliC	COG1344	N	GI 3	0.00	0.00
RCA23_c07600	RNA polymerase sigma-54	COG1508	K	GI 3	84.01	49.32
RCA23_c07610	hypothetical protein, cyclic nucleotide-binding-like			GI 3	100.00	0.00
RCA23_c07620	hypothetical protein	COG0664	T	GI 3	100.00	51.13

RCA23_c07630	hypothetical protein			GI 3	0.00	21.60
RCA23_c07640	hypothetical protein			GI 3	0.00	33.83
RCA23_c07650	hypothetical protein			GI 3	0.00	59.36
RCA23_c07660	flagellar biosynthesis protein FlhB	COG1377	N	GI 3	69.34	29.88
RCA23_c07670	flagellar biosynthetic protein FliR	COG1684	N	GI 3	100.00	67.52
RCA23_c07680	flagellar biosynthetic protein FliQ	COG1987	N	GI 3	100.00	56.30
RCA23_c07690	flagellar biosynthetic protein FliP	COG1338	N	GI 3	100.00	0.00
RCA23_c07700	flagellar motor switch protein FliN	COG1886	N	GI 3	100.00	34.92
RCA23_c07710	flagellar motor switch proteins FliM	COG1868	N	GI 3	100.00	42.20
RCA23_c07720	protein FliL	COG1580	N	GI 3	100.00	59.26
RCA23_c07730	putative flagellar hook-length control protein FliK	COG3144	N	GI 3	45.98	38.66
RCA23_c07740	putative flagellar export protein FliJ			GI 3	0.00	45.00
RCA23_c07750	flagellar protein export ATPase FliI	COG1157	N	GI 3	0.00	78.38
RCA23_c07760	putative flagellar assembly protein FliH			GI 3	45.29	41.18
RCA23_c07770	flagellar motor switch protein G	COG1536	N	GI 3	100.00	48.70
RCA23_c07780	flagellar M-ring protein FliF	COG1766	N	GI 3	26.08	71.29
RCA23_c07790	flagellar hook-basal body complex subunit FliE	COG1677	N	GI 3	0.00	59.68
RCA23_c07800	transcriptional regulatory protein FlbD	COG2204	T	GI 3	40.95	29.18
RCA23_c07810	protein FliL	COG1580	N	GI 3	100.00	66.48
RCA23_c07820	hypothetical protein			GI 3	100.00	71.26
RCA23_c07830	flagellar protein FliS	COG1516	N	GI 3	36.83	43.01
RCA23_c07840	hypothetical protein			GI 3	33.33	0.00
RCA23_c07850	hypothetical protein			GI 3	100.00	0.00
RCA23_c07860	flagellar basal-body rod protein FlgG	COG4786	N	GI 3	100.00	14.18
RCA23_c07870	Cl- channel, voltage-gated family protein	COG0038	P	GI 3	89.79	12.80
RCA23_c07880	hypothetical protein, lysozym-like	COG0741	M	GI 3	100.00	17.44
RCA23_c07890	flagellar hook-associated protein FlgL	COG1344	N	GI 3	96.70	40.02
RCA23_c07900	flagellar hook-associated protein FlgK	COG1256	N	GI 3	95.40	20.46
RCA23_c07910	flagellar rod assembly protein/muramidase FlgJ	COG3951	M	GI 3	100.00	71.51
RCA23_c07920	flagellar P-ring protein FlgI	COG1706	N	GI 3	100.00	54.35
RCA23_c07930	flagellar L-ring protein FlgH	COG2063	N	GI 3	100.00	45.74
RCA23_c07940	flagellar basal-body rod protein FlgG	COG4786	N	GI 3	45.75	38.28
RCA23_c07950	putative flagellar basal body rod protein FlgF	COG4786	N	GI 3	0.00	40.41

RCA23_c07960	flagellar hook protein FlgE	COG1749	N	GI 3	52.92	46.57
RCA23_c07970	flagellar basal body rod modification protein FlgD	COG1843	N	GI 3	100.00	71.27
RCA23_c07980	flagellar basal-body rod protein FlgC	COG1558	N	GI 3	83.45	49.16
RCA23_c07990	flagellar basal-body rod protein FlgB	COG1815	N	GI 3	0.00	63.31
RCA23_c08000	hypothetical protein			GI 3	0.00	68.88
RCA23_c08010	hypothetical protein			GI 3	0.00	68.35
RCA23_c08020	RNA polymerase, sigma factor for flagellar operon FliA	COG1191	K	GI 3	13.77	44.08
RCA23_c08030	putative flagellar biosynthesis protein FlhG	COG0455	D	GI 3	100.00	64.44
RCA23_c08040	flagellar biosynthesis protein FlhF	COG1419	N	GI 3	91.12	25.89
RCA23_c08050	flagellar biosynthesis protein FlhA	COG1298	N	GI 3	100.00	47.99
RCA23_c08060	hypothetical membrane lipoprotein, DUF400			GI 3	46.46	9.26
RCA23_c08070	hypothetical protein			GI 3	0.00	58.90
RCA23_c08080	flagellar basal body P-ring biosynthesis protein FlgA	COG1261	N	GI 3	95.68	41.84
RCA23_c08090	putative negative regulator of flagellin synthesis FlgM			GI 3	100.00	94.39
RCA23_c08100	hypothetical protein			GI 3	100.00	40.89
RCA23_c08110	hypothetical protein, HCP-like	COG0790	R	GI 3	100.00	60.82
RCA23_c08120	hypothetical protein			GI 3	100.00	0.00
RCA23_c08130	sigma54 specific transcriptional regulator, Fis family	COG2204	T	GI 3	100.00	70.49
RCA23_c08140	flagellar hook-associated protein FliD	COG1345	N	GI 3	15.63	57.75
RCA23_c08150	chemotaxis protein MotB	COG1360	N	GI 3	68.55	24.59
RCA23_c08160	chemotaxis protein MotA	COG1291	N	GI 3	42.13	35.70
RCA23_c08170	hypothetical protein, DUF1566			GI 3	7.69	25.85
RCA23_c08180	putative transmembrane protein, DUF6			GI 3	100.00	100.00
RCA23_c08190	putative polyketide hydroxylase SchC	COG0654	H	GI 3	67.51	100.00
RCA23_c08200	aldehyde dehydrogenase	COG1012	C		100.00	64.07
RCA23_c08210	2-amino-3-carboxymuconate-6-semialdehyde decarboxylase	COG2159	R		100.00	89.09
RCA23_c08220	transcriptional regulator, IclR family	COG1414	K		100.00	100.00
RCA23_c08230	TRAP transporter, 4TM/12TM fusion protein	COG4666	R		100.00	92.04
RCA23_c08240	TRAP transporter solute receptor, TAXI family	COG2358	R		100.00	81.43
RCA23_c08250	thiamine pyrophosphate enzyme-like TPP-binding	COG0028	E		95.75	89.62
RCA23_c08260	hypothetical protein	COG2268	S		100.00	67.72
RCA23_c08270	short chain dehydrogenase	COG1028	I		100.00	62.23
RCA23_c08280	putative aldehyde dehydrogenase yfmT	COG1012	C		4.67	92.95

RCA23_c08290	branched-chain amino acid ABC transporter, ATP-binding protein	COG0410	E	0.00	100.00
RCA23_c08300	branched-chain amino acid ABC transporter, ATP-binding protein	COG0411	E	0.00	80.97
RCA23_c08310	putative transporter, permease protein	COG4177	E	28.51	89.92
RCA23_c08320	putative transporter, permease protein	COG0559	E	100.00	80.18
RCA23_c08330	putative transporter, periplasmic binding protein	COG0683	E	0.00	75.58
RCA23_c08340	hypothetical protein, PrpF protein-like	COG2828	S	63.44	99.81
RCA23_c08350	4-carboxy-4-hydroxy-2-oxoadipate aldolase/oxaloacetate decarboxylase	COG0684	H	67.87	63.96
RCA23_c08360	putative N-acetylglucosaminyl-phosphatidylinositol de-N-acetylase,	COG2120	S	26.67	73.47
RCA23_c08370	hypothetical protein			44.71	100.00
RCA23_c08380	transcriptional regulator	COG1802	K	0.00	100.00
RCA23_c08390	aminomethyltransferase GcvT	COG0404	E	92.15	95.36
RCA23_c08400	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	100.00	85.38
RCA23_c08410	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	100.00	59.65
RCA23_c08420	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	71.26	100.00
RCA23_c08430	metapyrocatechase XylE	COG2514	R	78.34	86.31
RCA23_c08440	MFS-type transporter	COG2814	G	100.00	87.79
RCA23_c08450	hypothetical protein, cytochrome c			100.00	100.00
RCA23_c08460	hypothetical protein, copper resistance protein C	COG2372	R	100.00	100.00
RCA23_c08470	hypothetical protein, copper resistance protein D	COG1276	P	24.18	100.00
RCA23_c08480	hypothetical protein	COG3613	F	100.00	100.00
RCA23_c08490	hypothetical protein			100.00	100.00
RCA23_c08500	hypothetical protein			100.00	100.00
RCA23_c08510	fumarate reductase flavoprotein subunit	COG1053	C	100.00	100.00
RCA23_c08520	methylisocitrate lyase PrpB	COG2513	G	100.00	99.54
RCA23_c08530	putative isochorismatase family protein	COG1335	Q	100.00	93.56
RCA23_c08540	hydantoin utilization protein A	COG0145	E	84.74	100.00
RCA23_c08550	hydantoin utilization protein B	COG0146	E	79.94	99.76
RCA23_c08560	hypothetical protein	COG1942	R	100.00	100.00
RCA23_c08570	HTH-type transcriptional regulator, GntR family	COG2188	K	100.00	91.76
RCA23_c08580	3-isopropylmalate dehydratase large subunit LeuC	COG0065	E	100.00	100.00
RCA23_c08590	3-isopropylmalate dehydratase small subunit LeuD	COG0066	E	100.00	100.00
RCA23_c08600	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	100.00	100.00
RCA23_c08610	TRAP transporter, subunit DctQ	COG3090	G	100.00	100.00

RCA23_c08620	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	100.00	100.00
RCA23_c08630	putative glutamate synthase [NADPH] small chain	COG0493	E	100.00	100.00
RCA23_c08640	hypothetical protein, dihydrooronate dehydrogenase	COG0167	F	100.00	98.93
RCA23_c08650	putative HTH-type transcriptional regulator	COG1309	K	100.00	100.00
RCA23_c08660	N-carbamoyl-L-amino acid hydrolase AmaB	COG0624	E	100.00	100.00
RCA23_c08670	D-hydantoinase/dihydropyrimidinase Dht	COG0044	F	100.00	99.59
RCA23_c08680	ABC transporter ATP-binding protein	COG1116	P	100.00	100.00
RCA23_c08690	putative ABC transporter permease protein	COG0600	P	100.00	100.00
RCA23_c08700	putative ABC transporter permease protein	COG0600	P	100.00	100.00
RCA23_c08710	putative thiamine biosynthesis protein	COG0715	P	100.00	100.00
RCA23_c08720	putative integral membrane proein DUF6	COG0697	G	100.00	100.00
RCA23_c08730	hypothetical protein DUF6	COG0697	G	100.00	100.00
RCA23_c08740	alkaline phosphatase synthesis sensor protein PhoR	COG5002	T	100.00	100.00
RCA23_c08750	putative phosphate binding protein PstS	COG0226	P	100.00	97.50
RCA23_c08760	putative phosphate transport system permease protein PstC	COG0573	P	93.28	100.00
RCA23_c08770	putative phosphate transport system permease protein PstA	COG0581	P	100.00	100.00
RCA23_c08780	phosphate import ATP-binding protein PstB	COG1117	P	100.00	100.00
RCA23_c08790	phosphate transport system regulatory protein PhoU	COG0704	P	100.00	96.83
RCA23_c08800	phosphate regulon transl protein PhoB	COG0745	T	100.00	100.00
RCA23_c08810	putative hippurate hydrolase protei	COG1473	R	100.00	49.13
RCA23_c08820	urease accessory protein UreD	COG0829	O	100.00	100.00
RCA23_c08830	urease gamma subunit UreA	COG0831	E	100.00	100.00
RCA23_c08840	urease beta subunit UreB	COG0832	E	100.00	100.00
RCA23_c08850	urease alpha subunit UreC	COG0804	E	100.00	98.62
RCA23_c08860	urease accessory protein UreE	COG2371	O	100.00	100.00
RCA23_c08870	urease accessory protein UreF	COG0830	O	2.83	100.00
RCA23_c08880	urease accessory protein UreG	COG0378	O	0.00	100.00
RCA23_c08890	integrase	COG2801	L	0.00	0.00
RCA23_c08900	hypothetical protein	COG4585	T	53.33	100.00
RCA23_c08910	glucose-methanol-choline oxidoreductase AlkJ	COG2303	E	92.11	100.00
RCA23_c08920	hypothetical protein, YaeQ family protein	COG4681	S	100.00	100.00
RCA23_c08930	hypothetical protein, DUF6 transmembrane protein	COG0697	G	100.00	100.00
RCA23_c08940	UDP-glucose 6-dehydrogenase Ugd	COG1004	M	100.00	100.00

RCA23_c08950	3-deoxy-D-manno-octulosonic-acid transferase WaaA	COG1519	M		100.00	100.00
RCA23_c08960	UDP-N-acetylglucosamine 2-epimerase WecB	COG0381	M		100.00	100.00
RCA23_c08970	putative N-acetylneuramic acid synthase	COG2089	M		42.44	100.00
RCA23_c08980	hypothetical protein				67.65	100.00
RCA23_c08990	hypothetical protein	COG1651	O		100.00	100.00
RCA23_c09000	choline-sulfatase BetC	COG3119	P		100.00	94.68
RCA23_c09010	permease	COG2233	F		100.00	100.00
RCA23_c09020	hypothetical protein				100.00	100.00
RCA23_c09030	membrane dipeptidase	COG2355	E		100.00	100.00
RCA23_c09040	hypothetical protein				100.00	100.00
RCA23_c09050	hypothetical protein				100.00	100.00
RCA23_c09060	tRNA pseudouridine synthase A	COG0101	J		100.00	96.38
RCA23_c09070	YcjX-like protein	COG3106	R		100.00	93.14
RCA23_c09080	hypothetical protein, acetyltransferase-like	COG0456	R		100.00	100.00
RCA23_c09090	hypothetical protein	COG3768	S		100.00	92.33
RCA23_c09100	isoleucyl-tRNA synthase IleS	COG0060	J		100.00	100.00
RCA23_c09110	putative integral membrane protein DUF6				100.00	100.00
RCA23_c09120	phosphatidylcholine synthase Pcs	COG1183	I		100.00	100.00
RCA23_c09130	putative tyrosine recombinase xerC	COG4974	L		100.00	100.00
RCA23_c09140	hypothetical protein	COG3159	S		100.00	100.00
RCA23_c09150	transaldolase	COG0176	G		100.00	100.00
RCA23_c09160	primosomal protein N'	COG1198	L		100.00	100.00
RCA23_c09170	methylmalonyl-CoA mutase McmA	COG1884	I		100.00	93.52
RCA23_c09180	glycerol-3-phosphate acyltransferase PlsB	COG2937	I		100.00	98.68
RCA23_c09190	crotonyl-CoA reductase Ccr	COG0604	C		100.00	100.00
RCA23_c09200	putative ATPase	COG0507	L		100.00	98.61
RCA23_c09210	trimethylamine methyltransferase MttB	COG5598	H		100.00	100.00
RCA23_c09220	hypothetical protein DUF1052	COG5321	S		100.00	100.00
RCA23_c09240	hypothetical protein, DUF411	COG3019	R	GI 4	0.00	0.00
RCA23_c09250	hypothetical protein			GI 4	0.00	0.00
RCA23_c09260	hypothetical protein			GI 4	0.00	0.00
RCA23_c09270	hypothetical protein	COG4642	S	GI 4	0.00	0.00
RCA23_c09280	hypothetical protein	COG3904	S	GI 4	0.00	0.00

RCA23_c09290	putative DNA-binding protein, transposase-like	COG3415	L	GI 4	0.00	0.00
RCA23_c09300	HTH-type transcriptional regulator, LysR family	COG0583	K	GI 4	0.00	0.00
RCA23_c09310	hypothetical protein			GI 4	0.00	0.00
RCA23_c09320	hypothetical protein, hydrolase-like	COG0627	R	GI 4	0.00	0.00
RCA23_c09330	putative integrase	COG2801	L	GI 4	0.00	44.52
RCA23_c09340	putative transposase			GI 4	0.00	100.00
RCA23_c09350	hypothetical protein, ornithine cyclodeaminase-like	COG2423	E		0.00	0.00
RCA23_c09360	LrgB-like protein	COG1346	M		100.00	100.00
RCA23_c09370	hypothetical protein	COG1380	R		100.00	100.00
RCA23_c09380	putative branched-chain-amino-acid aminotransferase IlvE	COG0115	E		100.00	100.00
RCA23_c09390	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I		100.00	100.00
RCA23_c09400	bicyclomycin resistance protein Bcr	COG2814	G		100.00	100.00
RCA23_c09410	hypothetical protein				100.00	71.10
RCA23_c09420	hypothetical protein, photolyase	COG3046	R		100.00	97.08
RCA23_c09430	DNA photolyase, FAD-binding/cryptochrome	COG0415	L		100.00	91.30
RCA23_c09440	carbon monoxide dehydrogenase small chain CoxS	COG2080	C		100.00	100.00
RCA23_c09450	carbon monoxide dehydrogenase large chain CoxL	COG1529	C		100.00	100.00
RCA23_c09460	carbon monoxide dehydrogenase medium chain CoxM	COG1319	C		100.00	100.00
RCA23_c09470	MoxR-like ATPase	COG0714	R		100.00	100.00
RCA23_c09480	CoxE-like protein	COG3552	R		100.00	89.68
RCA23_c09490	5-deoxy-glucuronate isomerase IolB	COG3718	G		100.00	100.00
RCA23_c09500	branched-chain amino acid transport protein, AzlD-like	COG4392	S		100.00	100.00
RCA23_c09510	branched-chain amino acid transport protein, AzlC-like	COG1296	E		100.00	100.00
RCA23_c09520	alpha-IPM synthase/homocitrate synthase	COG0119	E		100.00	95.84
RCA23_c09530	cytochrome P450	COG2124	Q		93.18	94.13
RCA23_c09540	TRAP dicarboxylate transporter, subunit DctP	COG1638	G		100.00	100.00
RCA23_c09550	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G		100.00	100.00
RCA23_c09560	TRAP dicarboxylate transporter, subunit DctM	COG1593	G		100.00	99.40
RCA23_c09570	putative citrate transporter	COG0471	P		100.00	100.00
RCA23_c09580	triosephosphate isomerase TpiA	COG0149	G		100.00	100.00
RCA23_c09590	putative regulator protein of competence-specific genes TfoX	COG3070	K		100.00	98.11
RCA23_c09600	hypothetical protein, iron-sulfur cluster assembly protein	COG0316	S		100.00	100.00
RCA23_c09610	hypothetical protein DUF59	COG2151	R		100.00	100.00

RCA23_c09620	queuine tRNA-ribosyltransferase Tgt	COG0343	J	100.00	100.00
RCA23_c09630	hypothetical protein			100.00	100.00
RCA23_c09640	ATP-dependent protease La	COG0466	O	100.00	99.83
RCA23_c09650	putative phosphoglycerate mutase family protein	COG0406	G	100.00	100.00
RCA23_c09670	protein-L-isoaspartate O-methyltransferase Pcm	COG2518	O	100.00	100.00
RCA23_c09680	outer membrane efflux protein	COG1538	M	100.00	100.00
RCA23_c09690	hypothetical protein			100.00	100.00
RCA23_c09700	cobyric acid synthase CobQ	COG1492	H	100.00	100.00
RCA23_c09710	hypothetical protein			100.00	100.00
RCA23_c09720	translation elongation factor P	COG0231	J	100.00	100.00
RCA23_c09730	hypothetical protein tRNA modifying protein YgfZ	COG0354	R	100.00	100.00
RCA23_c09740	putative glycosyltransferase, family 2	COG0463	M	100.00	100.00
RCA23_c09750	lipid A export ATP-binding/permease protein MsbA	COG1132	V	100.00	100.00
RCA23_c09760	serine--glyoxylate aminotransferase SgaA	COG0075	E	100.00	93.37
RCA23_c09770	histidinol-phosphate aminotransferase HisC	COG0079	E	100.00	100.00
RCA23_c09780	valyl-tRNA synthase ValS	COG0525	J	100.00	98.26
RCA23_c09790	hypothetical protein, metal-dependent phosphohydrolase	COG4341	R	100.00	100.00
RCA23_c09800	putative phytanoyl-CoA dioxygenase family protein	COG5285	Q	100.00	96.68
RCA23_c09810	molybdenum-containing hydroxylase	COG1529	C	100.00	99.43
RCA23_c09820	hypothetical protein DUF2235	COG3673	S	100.00	100.00
RCA23_c09830	5,10-methylenetetrahydrofolate reductase MetF	COG0685	E	100.00	100.00
RCA23_c09840	HTH-type transcriptional regulator MetR	COG0583	K	100.00	100.00
RCA23_c09850	inositol-1-monophosphatase SuhB	COG0483	G	100.00	100.00
RCA23_c09860	hypothetical protein	COG3063	N	100.00	100.00
RCA23_c09880	transcriptional regulator, HxlR family	COG1733	K	100.00	100.00
RCA23_c09890	hypothetical protein, DUF3764			100.00	100.00
RCA23_c09900	hypothetical protein			100.00	100.00
RCA23_c09910	hypothetical protein, DsrE/F-like	COG1553	P	100.00	100.00
RCA23_c09920	glutathione peroxidase Gpo	COG0386	O	100.00	100.00
RCA23_c09930	putative ion channel			100.00	100.00
RCA23_c09950	hypothetical protein			34.50	100.00
RCA23_c09940	ribosomal large subunit pseudouridine synthase D	COG0564	J	100.00	91.09
RCA23_c09960	RNA polymerase sigma-32 factor RpoH	COG0568	K	100.00	100.00

RCA23_c09970	putative oxidoreductase	COG0673	R	100.00	100.00
RCA23_c09980	oligoendopeptidase F	COG1164	E	100.00	97.87
RCA23_c09990	2-keto-3-deoxy-L-rhamnonate aldolase RhmA	COG3836	G	100.00	100.00
RCA23_c10000	hypothetical protein, lysophospholipase L2	COG2267	I	100.00	100.00
RCA23_c10010	putative sterol-binding protein	COG3255	I	100.00	100.00
RCA23_c10020	hypothetical protein			100.00	86.99
RCA23_c10030	ATP-dependent RNA helicase RhlB	COG4581	L	100.00	99.79
RCA23_c10040	putative heat shock protein	COG1188	J	100.00	100.00
RCA23_c10050	ferredoxin FdxA	COG1146	C	87.02	100.00
RCA23_c10060	putative transcriptional regulator, CarD family	COG1329	K	19.08	100.00
RCA23_c10070	cobalamin-5-phosphate synthase CobS	COG0368	H	100.00	100.00
RCA23_c10080	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase Cob	COG2038	H	100.00	100.00
RCA23_c10090	hypothetical protein			100.00	99.41
RCA23_c10100	glyoxalase/bleomycin resistance protein	COG3565	R	100.00	100.00
RCA23_c10110	hypothetical protein			100.00	100.00
RCA23_c10120	choline dehydrogenase BetA	COG2303	E	100.00	100.00
RCA23_c10140	23S rRNA (guanosine-2'-O-)-methyltransferase RlmB	COG0566	J	100.00	67.11
RCA23_c10150	phosphoribosyl-ATP pyrophosphatase HisE	COG0140	E	100.00	100.00
RCA23_c10160	imidazole glycerol phosphate synthase subunit HisF	COG0107	E	100.00	100.00
RCA23_c10170	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazo	COG0106	E	100.00	100.00
RCA23_c10180	imidazole glycerol phosphate synthase, glutamine amidotransferase subunit H	COG0118	E	100.00	87.64
RCA23_c10190	imidazoleglycerol-phosphate dehydratase HisB	COG0131	E	100.00	100.00
RCA23_c10210	hypothetical protein, calcineurin-like phosphoesterase-like	COG1409	R	100.00	100.00
RCA23_c10220	pyruvate carboxylase Pyc	COG1038	C	100.00	98.75
RCA23_c10230	L-lactate dehydrogenase IldD	COG1304	C	100.00	100.00
RCA23_c10240	putative DNA helicase II	COG0210	L	100.00	100.00
RCA23_c10250	gamma-glutamylputrescine oxidoreductase PuuB	COG0665	E	100.00	100.00
RCA23_c10260	putative cysteine/O-acetylserine efflux protein	COG1280	E	100.00	100.00
RCA23_c10270	hypothetical protein			100.00	100.00
RCA23_c10280	putative protein Mrp	COG0489	D	100.00	100.00
RCA23_c10290	hypothetical protein			100.00	100.00
RCA23_c10300	hypothetical protein, MraZ	COG2001	S	100.00	100.00
RCA23_c10310	S-adenosyl-L-methionine-dependent methyltransferase MraW	COG0275	M	100.00	100.00

RCA23_c10320	hypothetical protein	COG5462	S	100.00	100.00
RCA23_c10330	peptidoglycan synthase FtsI	COG0768	M	100.00	98.32
RCA23_c10340	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase Mu	COG0769	M	100.00	100.00
RCA23_c10350	putative UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase MurF	COG0770	M	100.00	100.00
RCA23_c10360	phospho-N-acetylmuramoyl-pentapeptide-transferase MraY	COG0472	M	100.00	100.00
RCA23_c10370	UDP-N-acetylmuramoylalanine--D-glutamate ligase MurD	COG0771	M	100.00	100.00
RCA23_c10380	putative glycosyltransferase, sugar binding region	COG3774	M	100.00	100.00
RCA23_c10390	putative galactoside 2-alpha-L-fucosyltransferase 1			100.00	100.00
RCA23_c10400	hypothetical protein			100.00	100.00
RCA23_c10410	hypothetical protein	COG3306	M	100.00	100.00
RCA23_c10420	hypothetical protein HI0933	COG2081	R	100.00	96.64
RCA23_c10430	cell division protein FtsW	COG0772	D	100.00	100.00
RCA23_c10440	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-	COG0707	M	100.00	100.00
RCA23_c10450	UDP-N-acetylmuramate--L-alanine ligase MurC	COG0773	M	100.00	100.00
RCA23_c10460	permease of the drug/metabolite transporter superfamily	COG0697	G	100.00	100.00
RCA23_c10470	UDP-N-acetylenolpyruvoylglucosamine reductase MurB	COG0812	M	100.00	85.96
RCA23_c10480	D-alanine--D-alanine ligase Ddl	COG1181	M	100.00	100.00
RCA23_c10490	putative cell division protein FtsQ			57.96	100.00
RCA23_c10500	cell division protein FtsA	COG0849	D	82.02	100.00
RCA23_c10510	cell division protein FtsZ	COG0206	D	100.00	100.00
RCA23_c10520	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase LpxC	COG0774	M	100.00	100.00
RCA23_c10530	outer membrane assembly lipoprotein	COG4105	R	100.00	100.00
RCA23_c10540	DNA repair protein RecN	COG0497	L	100.00	100.00
RCA23_c10550	hypothetical protein DUF427	COG2343	S	100.00	100.00
RCA23_c10560	putative Xaa-Pro aminopeptidase	COG0006	E	100.00	100.00
RCA23_c10570	aerobic cobaltochelatase subunit CobT	COG4547	H	100.00	100.00
RCA23_c10580	aerobic cobaltochelatase subunit CobS	COG0714	R	100.00	100.00
RCA23_c10590	hypothetical protein, DnaJ	COG0484	O	100.00	100.00
RCA23_c10600	putative stress-induced morphoprotein, BolA type	COG0271	T	54.17	100.00
RCA23_c10610	aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit GatB	COG0064	J	95.11	90.54
RCA23_c10620	aminopeptidase N	COG0308	E	95.06	95.89
RCA23_c10630	malate synthase GlcB	COG2225	C	100.00	97.29
RCA23_c10640	hypothetical protein DUF336	COG3193	R	100.00	100.00

RCA23_c10650	hypothetical protein				100.00	100.00
RCA23_c10660	gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD	COG2071	R		100.00	100.00
RCA23_c10670	putative D-beta-hydroxybutyrate dehydrogenase	COG1028	I		100.00	100.00
RCA23_c10680	NAD(P) transhydrogenase alpha subunit PntA	COG3288	C		100.00	100.00
RCA23_c10690	NAD(P) transhydrogenase beta subunit PntB	COG1282	C		100.00	100.00
RCA23_c10700	hypothetical protein	COG4949	S		100.00	100.00
RCA23_c10710	putative acetylornithine deacetylase ArgE	COG0624	E		100.00	100.00
RCA23_c10720	molybdenum cofactor biosynthesis protein MoaA	COG2896	H		100.00	98.21
RCA23_c10730	acetyl-coenzyme A synthase AcsA	COG0365	I		100.00	94.60
RCA23_c10740	hypothetical protein	COG4321	R		100.00	100.00
RCA23_c10750	hypothetical protein				100.00	100.00
RCA23_c10760	fumarate hydratase class II	COG0114	C		100.00	100.00
RCA23_c10770	hypothetical protein	COG3814	S		100.00	81.29
RCA23_c10780	putative chromate transport protein	COG2059	P		100.00	100.00
RCA23_c10790	hydantoinase / oxoprolinase family protein	COG0145	E		100.00	100.00
RCA23_c10800	hypothetical protein				100.00	100.00
RCA23_c10810	putative quinone oxidoreductase yhdH	COG0604	C		100.00	97.47
RCA23_c10820	cysteine desulfurase SufS	COG0520	E		100.00	100.00
RCA23_c10830	deoxyribodipyrimidine photo-lyase PhrB	COG0415	L		89.88	88.75
RCA23_c10840	cyclopropane-fatty-acyl-phospholipid synthase Cfa	COG2230	M		45.00	100.00
RCA23_c10850	ADP-ribose pyrophosphatase NudF	COG0494	L		100.00	99.46
RCA23_c10860	cysteine synthase CysK	COG0031	E		100.00	93.55
RCA23_c10870	mechanosensitive ion channel protein MscS	COG3264	M		100.00	100.00
RCA23_c10890	putative phage integrase	COG4974	L	GI 5	34.02	25.54
RCA23_c10900	hypothetical protein			GI 5	0.00	19.20
RCA23_c10910	hypothetical protein, tetratricopetide domain TPR-1	COG0457	R	GI 5	0.00	0.00
RCA23_c10920	hypothetical protein, resolvase-like	COG1961	L	GI 5	100.00	100.00
RCA23_c10930	hypothetical protein, DUF2924			GI 5	100.00	37.53
RCA23_c10940	putative TRAP transporter solute receptor	COG2358	R	GI 5	35.69	62.17
RCA23_c10950	TRAP transporter, 4TM/12TM fusion protein	COG4666	R	GI 5	100.00	57.74
RCA23_c10960	putative sulfatase YidJ	COG3119	P	GI 5	100.00	58.08
RCA23_c10970	arylsulfatase	COG1234	R	GI 5	100.00	74.80
RCA23_c10980	siderophore interactin protein, vibriobactin utilization protein-like	COG2375	P	GI 5	100.00	76.80

RCA23_c10990	hypothetical protein, metallo-beta-lactamase	COG2015	Q	GI 5	97.10	63.07
RCA23_c11000	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	GI 5	64.31	47.69
RCA23_c11010	glutathione S-transferase	COG0625	O	GI 5	99.55	62.01
RCA23_c11020	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	GI 5	100.00	19.67
RCA23_c11030	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	GI 5	100.00	20.78
RCA23_c11040	hypothetical protein, DUF6 transmembrane protein	COG0697	G	GI 5	100.00	34.32
RCA23_c11050	ferric reductase like transmembrane component family			GI 5	100.00	60.14
RCA23_c11060	hypothetical protein			GI 5	100.00	72.69
RCA23_c11070	hypothetical protein			GI 5	100.00	67.93
RCA23_c11080	sulfatase family protein	COG3119	P	GI 5	12.09	49.51
RCA23_c11090	short-chain dehydrogenase/reductase SDR	COG1028	I	GI 5	3.29	35.31
RCA23_c11100	maleylacetoacetate isomerase MaiA	COG0625	O	GI 5	100.00	8.30
RCA23_c11110	salicylaldehyde dehydrogenase	COG1012	C	GI 5	100.00	42.59
RCA23_c11120	transcriptional regulator, GntR family	COG1802	K	GI 5	100.00	47.91
RCA23_c11130	transposase	COG3316	L	GI 5	4.66	0.00
RCA23_c11140	putative phage integrase	COG4974	L	GI 5	0.00	0.00
RCA23_c11150	hypothetical protein, alginate lyase-lyase			GI 5	0.00	0.00
RCA23_c11160	HTH-type transcriptional regulator, GntR family	COG1802	K	GI 5	100.00	80.77
RCA23_c11170	branched-chain amino acid ABC transporter, periplasmic substrate-binding pr	COG0683	E	GI 5	5.22	76.60
RCA23_c11180	branched-chain amino acid ABC transporter, permease protein	COG0559	E	GI 5	91.81	90.14
RCA23_c11190	branched-chain amino acid ABC transporter, permease protein	COG4177	E	GI 5	100.00	71.25
RCA23_c11200	branched-chain amino acid ABC transporter, ATP-binding protein	COG0411	E	GI 5	77.12	44.58
RCA23_c11210	branched-chain amino acid ABC transporter, ATP-binding protein	COG0410	E	GI 5	0.00	43.98
RCA23_c11220	hydantoin racemase HyuE	COG4126	E	GI 5	0.00	62.57
RCA23_c11230	transposase	COG3316	L	GI 5	0.00	79.66
RCA23_c11240	glutamine transport ATP-binding protein GlnQ	COG1126	E	GI 5	0.00	0.00
RCA23_c11250	putative inner membrane amino-acid ABC transporter permease protein	COG0765	E	GI 5	0.00	0.00
RCA23_c11260	transposase	COG3316	L	GI 5	0.00	0.00
RCA23_c11270	transcriptional regulator CoxC	COG3300	T	GI 5	43.53	100.00
RCA23_c11280	carbon monoxide dehydrogenase medium chain CoxM	COG1319	C	GI 5	100.00	100.00
RCA23_c11290	carbon monoxide dehydrogenase small chain CoxS	COG2080	C	GI 5	100.00	91.11
RCA23_c11300	carbon monoxide dehydrogenase large chain CoxL	COG1529	C	GI 5	89.29	98.06
RCA23_c11310	AAA+ ATPase chaperone CoxD	COG0714	R	GI 5	95.43	92.47

RCA23_c11320	carbon monoxide dehydrogenase accessory protein CoxE	COG3552	R	GI 5	100.00	100.00
RCA23_c11330	carbon monoxide dehydrogenase accessory protein CoxF	COG1975	O	GI 5	100.00	91.92
RCA23_c11340	hypothetical protein	COG2068	R	GI 5	100.00	100.00
RCA23_c11350	carbon monoxide dehydrogenase protein CoxG	COG3427	S	GI 5	100.00	100.00
RCA23_c11360	carbon monoxide dehydrogenase accessory protein CoxI	COG1975	O	GI 5	100.00	86.99
RCA23_c11370	transposase	COG3316	L	GI 5	100.00	0.00
RCA23_c11380	dienelactone hydrolase	COG0412	Q	GI 5	0.00	96.25
RCA23_c11390	aldo/keto reductase	COG0667	C	GI 5	94.94	97.52
RCA23_c11400	mandelate racemase MdlA	COG4948	M	GI 5	100.00	91.07
RCA23_c11410	hypothetical protein DUF1498	COG3822	R	GI 5	100.00	100.00
RCA23_c11420	two-component system, sensor histidine kinase protein	COG0642	T	GI 5	100.00	100.00
RCA23_c11430	two-component system, response regulator protein	COG0745	T	GI 5	100.00	100.00
RCA23_c11440	hypothetical protein	COG3181	S	GI 5	100.00	100.00
RCA23_c11450	putative tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S	GI 5	100.00	99.40
RCA23_c11460	hypothetical protein	COG2828	S	GI 5	100.00	90.66
RCA23_c11470	D-isomer specific 2-hydroxyacid dehydrogenase	COG0111	H	GI 5	100.00	94.32
RCA23_c11480	putative ammonia monooxygenase	COG3180	R	GI 5	100.00	100.00
RCA23_c11490	transcriptional regulator	COG1802	K	GI 5	100.00	100.00
RCA23_c11500	fumarate reductase flavoprotein subunit FccA	COG1053	C	GI 5	100.00	100.00
RCA23_c11510	aspartate aminotransferase AspC	COG0436	E	GI 5	100.00	100.00
RCA23_c11520	hypothetical protein			GI 5	100.00	100.00
RCA23_c11530	aerobic glycerol-3-phosphate dehydrogenase GlpD	COG0578	C	GI 5	100.00	98.78
RCA23_c11540	aerobic glycerol-3-phosphate dehydrogenase GlpD	COG0578	C	GI 5	100.00	100.00
RCA23_c11550	hypothetical protein	COG1192	D	GI 5	0.00	0.00
RCA23_c11560	chemotaxis protein CheW	COG0835	N	GI 5	64.10	0.00
RCA23_c11570	methyl-accepting chemotaxis protein II	COG0840	N	GI 5	99.09	0.00
RCA23_c11580	chemotaxis histidine kinase CheA	COG0643	N	GI 5	0.00	0.00
RCA23_c11590	response regulator CheY	COG2114	T	GI 5	0.00	0.00
RCA23_c11600	chemotaxis protein methyltransferase CheR	COG1352	N	GI 5	0.00	0.00
RCA23_c11610	chemotaxis response regulator protein-glutamate methylesterase CheB	COG2201	N	GI 5	0.00	0.00
RCA23_c11620	metal-dependent hydrolase	COG1235	R	GI 5	0.00	0.00
RCA23_c11630	integrase	COG2801	L	GI 5	0.00	0.00
RCA23_c11640	methyltransferase, FkbM family			GI 5	0.00	0.00

RCA23_c11650	hypothetical protein, sugar transferase-like			GI 5	0.00	63.40
RCA23_c11660	aldo/keto reductase	COG0667	C	GI 5	0.00	35.84
RCA23_c11670	hypothetical protein			GI 5	0.00	92.39
RCA23_c11680	hypothetical protein	COG0790	R	GI 5	0.00	0.00
RCA23_c11690	hypothetical protein	COG4430	S	GI 5	0.00	73.65
RCA23_c11700	hypothetical protein	COG2755	E	GI 5	0.00	0.00
RCA23_c11710	phytanoyl-CoA dioxygenase family protein	COG5285	Q	GI 5	100.00	100.00
RCA23_c11720	putative bacterial extracellular solute binding protein, family 3	COG1791	S	GI 5	0.00	92.32
RCA23_c11730	hypothetical protein			GI 5	0.00	6.62
RCA23_c11740	SOUL heme-binding protein			GI 5	99.67	100.00
RCA23_c11750	arylsulfatase precursor	COG3119	P	GI 5	96.71	96.71
RCA23_c11760	hypothetical protein DUF583			GI 5	0.00	0.00
RCA23_c11770	integrase	COG2801	L	GI 5	0.00	0.00
RCA23_c11780	hypothetical protein	COG3616	E	GI 5	79.40	94.19
RCA23_c11790	glutamine synthase GlnA type I	COG0174	E	GI 5	0.00	0.00
RCA23_c11800	putative HTH-type transcriptional regulator, RpiR family	COG1737	K	GI 5	0.00	0.00
RCA23_c11810	N-formylglutamate amidohydrolase	COG3931	E	GI 5	0.00	0.00
RCA23_c11820	putative isochorismatase family protein	COG1335	Q	GI 5	0.00	0.00
RCA23_c11830	branched-chain amino acid ABC transporter, ATP-binding protein	COG0411	E	GI 5	70.26	0.00
RCA23_c11840	branched-chain amino acid ABC transporter, periplasmic binding protein	COG0683	E	GI 5	100.00	0.00
RCA23_c11850	branched-chain amino acid ABC transporter, ATP-binding protein	COG0410	E	GI 5	0.00	0.00
RCA23_c11860	branched-chain amino acid ABC transporter, permease protein	COG0559	E	GI 5	0.00	0.00
RCA23_c11870	branched-chain amino acid ABC transporter, permease protein	COG4177	E	GI 5	0.00	3.49
RCA23_c11880	acetamidase/formamidase family protein	COG2421	C	GI 5	0.00	18.96
RCA23_c11890	hypothetical protein	COG1028	I	GI 5	100.00	100.00
RCA23_c11910	putative prophage integrase	COG0582	L	GI 5	0.00	20.44
RCA23_c11920	hypothetical protein			GI 5	0.00	0.00
RCA23_c11930	hypothetical protein			GI 5	0.00	73.21
RCA23_c11940	putative AAA ATPase	COG3598	L	GI 5	0.00	73.00
RCA23_c11950	hypothetical protein			GI 5	0.00	0.00
RCA23_c11960	hypothetical protein			GI 5	0.00	0.00
RCA23_c11970	hypothetical protein			GI 5	0.00	100.00
RCA23_c11980	hypothetical protein			GI 5	0.00	32.25

RCA23_c11990	branched-chain amino acid ABC transporter, ATP-binding protein	COG0410	E	GI 5	100.00	75.68
RCA23_c12000	branched-chain amino acid ABC transporter, ATP-binding protein	COG0411	E	GI 5	100.00	0.00
RCA23_c12010	branched-chain amino acid ABC transporter, permease protein	COG4177	E	GI 5	100.00	46.14
RCA23_c12020	branched-chain amino acid ABC transporter, permease protein	COG0559	E	GI 5	92.06	14.54
RCA23_c12030	branched-chain amino acid ABC transporter, periplasmic binding protein	COG0683	E	GI 5	0.00	45.84
RCA23_c12040	enoyl-CoA hydratase/carnithine racemase	COG1024	I	GI 5	0.00	14.86
RCA23_c12050	acetyl-CoA synthase-like protein	COG0318	I	GI 5	77.41	47.70
RCA23_c12060	dehydrogenase iron-sulfur-binding subunit	COG2080	C	GI 5	100.00	23.73
RCA23_c12070	dehydrogenase FAD-binding subunit	COG1319	C	GI 5	92.00	25.70
RCA23_c12080	dehydrogenase molybdenum-binding subunit	COG1529	C	GI 5	0.00	12.86
RCA23_c12090	putative HTH-type transcriptional regulator, AraC family	COG2207	K	GI 5	0.00	0.00
RCA23_c12100	periplasmic binding protein-like	COG1879	G	GI 5	0.00	0.00
RCA23_c12110	ROK family transcriptional repressor	COG1940	K	GI 5	0.00	8.48
RCA23_c12120	xylose isomerase	COG1082	G	GI 5	0.00	0.00
RCA23_c12130	sugar ABC transporter, periplasmic binding protein	COG1879	G	GI 5	0.00	0.00
RCA23_c12140	putative oxidoreductase	COG0673	R	GI 5	0.00	0.00
RCA23_c12150	sugar ABC transporter, ATP-binding protein	COG1129	G	GI 5	0.00	6.98
RCA23_c12160	sugar ABC transporter, permease protein	COG1172	G	GI 5	0.00	0.00
RCA23_c12170	hypothetical protein, monoxygenase-like	COG1359	S	GI 5	0.00	0.00
RCA23_c12180	hypothetical protein	COG1082	G	GI 5	0.00	62.23
RCA23_c12190	putative oxidoreductase	COG0673	R	GI 5	0.00	11.02
RCA23_c12200	peptidase M20D, amidohydrolase	COG1473	R	GI 5	100.00	27.15
RCA23_c12210	X-Pro dipeptidase	COG0006	E	GI 5	46.86	29.93
RCA23_c12220	hypothetical protein			GI 5	0.00	0.00
RCA23_c12230	hypothetical protein, transmembrane			GI 5	0.00	0.00
RCA23_c12240	hypothetical protein, DNA breaking rejoining enzymes family protein-like			GI 5	88.64	0.00
RCA23_c12250	hypothetical protein			GI 5	0.00	0.00
RCA23_c12260	hypothetical protein			GI 5	0.00	17.82
RCA23_c12270	hypothetical protein			GI 5	0.00	0.00
RCA23_c12280	type I restriction-modification system, M subunit	COG0286	V	GI 5	0.00	0.00
RCA23_c12290	type I restriction-modification system, S subunit	COG0732	V	GI 5	0.00	0.00
RCA23_c12300	type I restriction-modification system, R subunit	COG0610	V	GI 5	0.00	3.78
RCA23_c12310	hypothetical protein	COG3012	S	GI 5	0.00	0.00

RCA23_c12320	hypothetical protein	COG2865	K	GI 5	0.00	0.00
RCA23_c12330	hypothetical protein			GI 5	0.00	0.00
RCA23_c12340	hypothetical protein, DNA helicase-like protein			GI 5	0.00	0.00
RCA23_c12350	hypothetical protein			GI 5	0.00	0.00
RCA23_c12360	integrase	COG0582	L	GI 5	0.00	0.00
RCA23_c12370	hypothetical protein			GI 5	0.00	0.00
RCA23_c12380	hypothetical protein			GI 5	0.00	0.00
RCA23_c12390	hypothetical protein			GI 5	0.00	0.00
RCA23_c12400	hypothetical protein			GI 5	0.00	0.00
RCA23_c12410	hypothetical protein, DUF1994			GI 5	100.00	47.10
RCA23_c12420	exonuclease I	COG2925	L	GI 5	100.00	70.01
RCA23_c12430	transpeptidase, penicillin binding protein	COG4953	M	GI 5	28.88	44.78
RCA23_c12440	hypothetical protein	COG2373	R	GI 5	22.49	26.91
RCA23_c12450	hypothetical protein			GI 5	0.00	24.93
RCA23_c12460	hypothetical protein			GI 5	0.00	0.00
RCA23_c12470	type III restriction enzyme, res subunit	COG1061	K	GI 5	0.00	31.71
RCA23_c12480	putative type III restriction system protein, mod subunit			GI 5	0.00	53.63
RCA23_c12490	putative serine/threonine protein kinase	COG0515	R	GI 5	0.00	20.27
RCA23_c12500	serine/threonine protein phosphatase PrpC	COG0631	T	GI 5	0.00	37.50
RCA23_c12510	DNA helicase, UvrD/REP type	COG0210	L	GI 5	0.00	0.00
RCA23_c12520	hypothetical protein	COG1463	Q	GI 5	0.00	0.00
RCA23_c12530	hypothetical protein, OmpA/MotB-like	COG1360	N	GI 5	0.00	0.00
RCA23_c12540	hypothetical protein			GI 5	0.00	0.00
RCA23_c12550	ATP-dependent helicase	COG0553	K	GI 5	0.00	0.00
RCA23_c12560	hypothetical protein			GI 5	0.00	0.00
RCA23_c12570	hypothetical protein			GI 5	0.00	0.00
RCA23_c12580	hypothetical protein			GI 5	0.00	0.00
RCA23_c12590	hypothetical protein			GI 5	49.19	0.00
RCA23_c12600	hypothetical protein			GI 5	89.56	0.00
RCA23_c12610	hypothetical protein			GI 5	0.00	0.00
RCA23_c12620	hypothetical protein			GI 5	0.00	0.00
RCA23_c12630	integrase	COG2801	L	GI 5	0.00	100.00
RCA23_c12640	hypothetical protein			GI 5	0.00	0.00

RCA23_c12650	hypothetical protein			GI 5	0.00	0.00
RCA23_c12660	DNA integration/recombination/invertion protein			GI 5	0.00	0.00
RCA23_c12690	putative beta-lactamase-like protein	COG0491	R		100.00	100.00
RCA23_c12680	osmolarity sensor protein EnvZ	COG0642	T		100.00	100.00
RCA23_c12700	hypothetical protein				100.00	100.00
RCA23_c12710	50S ribosomal protein L21	COG0261	J		100.00	100.00
RCA23_c12720	50S ribosomal protein L27	COG0211	J		100.00	100.00
RCA23_c12730	hypothetical protein, LysE type translocator	COG1280	E		100.00	100.00
RCA23_c12740	putative acetyltransferase	COG1670	J		100.00	100.00
RCA23_c12750	GCN5-like N-acetyltransferase	COG1670	J		100.00	100.00
RCA23_c12760	GTP-binding protein Obg	COG0536	R		100.00	100.00
RCA23_c12770	glutamate 5-kinase ProB	COG0263	E		85.73	99.01
RCA23_c12780	gamma-glutamyl phosphate reductase ProA	COG0014	E		80.35	88.94
RCA23_c12800	hypothetical protein	COG5385	S		100.00	100.00
RCA23_c12790	hypothetical protein				100.00	100.00
RCA23_c12810	hypothetical protein	COG3176	R		100.00	100.00
RCA23_c12820	putative phosphate acyltransferase	COG0204	I		100.00	98.29
RCA23_c12830	thiamine-phosphate pyrophosphorylase ThiE	COG0352	H		100.00	92.56
RCA23_c12840	tRNA (cytidine/uridine-2'-O-)-methyltransferase TrmJ	COG0565	J		100.00	100.00
RCA23_c12850	heme A synthase CtaA	COG1612	O		100.00	100.00
RCA23_c12860	thermostable carboxypeptidase 1	COG2317	E		100.00	97.64
RCA23_c12870	DNA gyrase subunit A	COG0188	L		100.00	98.76
RCA23_c12880	hypothetical protein	COG1495	O		100.00	100.00
RCA23_c12890	DedA family protein	COG1238	S		100.00	100.00
RCA23_c12900	hypothetical protein			GI 6	100.00	98.00
RCA23_c12910	putative snoaL-like polyketide cyclase	COG3631	R	GI 6	0.00	87.55
RCA23_c12920	hypothetical protein, DUF1330	COG5470	S	GI 6	0.00	100.00
RCA23_c12930	hypothetical protein, DUF3303			GI 6	0.00	100.00
RCA23_c12940	hypothetical protein, transmembrane			GI 6	100.00	0.00
RCA23_c12950	hypothetical protein, transmembrane			GI 6	100.00	0.00
RCA23_c12960	integrase	COG2801	L	GI 6	2.13	0.00
RCA23_c12970	hypothetical protein			GI 6	0.00	0.00
RCA23_c12980	site-specific recombinase, resolvase family protein	COG1961	L	GI 6	0.00	0.00

RCA23_c12990	hypothetical protein	COG4731	S	GI 6	0.00	100.00
RCA23_c13000	hypothetical protein			GI 6	0.00	0.00
RCA23_c13010	transposase	COG3415	L	GI 6	0.00	0.00
RCA23_c13020	ribsomal protein S21	COG0828	J	GI 6	100.00	100.00
RCA23_c13030	hypothetical protein			GI 6	100.00	45.83
RCA23_c13040	integrase	COG2801	L	GI 6	6.48	0.00
RCA23_c13050	hypothetical protein			GI 6	100.00	85.99
RCA23_c13060	putative phage integrase	COG0582	L	GI 6	100.00	100.00
RCA23_c13080	putative phage integrase	COG0582	L	GI 6	100.00	98.71
RCA23_c13090	hypothetical protein, GYD domain	COG4274	S	GI 6	0.00	0.00
RCA23_c13100	hypothetical protein			GI 6	0.00	12.95
RCA23_c13110	hypothetical protein			GI 6	0.00	100.00
RCA23_c13120	hypothetical protein			GI 6	73.28	65.53
RCA23_c13130	integrase	COG2801	L	GI 6	0.00	0.00
RCA23_c13140	hypothetical protein	COG5586	S	GI 6	90.17	100.00
RCA23_c13150	hypothetical protein			GI 6	100.00	84.93
RCA23_c13160	hypothetical protein			GI 6	100.00	100.00
RCA23_c13170	hypothetical protein	COG4274	S	GI 6	100.00	100.00
RCA23_c13180	hypothetical protein			GI 6	84.67	100.00
RCA23_c13190	hypothetical protein			GI 6	0.00	100.00
RCA23_c13210	phosphoribosylformylglycinamidine cyclo-ligase PurM	COG0150	F		90.87	100.00
RCA23_c13220	phosphoribosylglycinamide formyltransferase PurN	COG0299	F		100.00	100.00
RCA23_c13230	ribonuclease D	COG0349	J		100.00	100.00
RCA23_c13240	SufE-like protein	COG2166	R		100.00	81.51
RCA23_c13250	hypothetical protein				100.00	100.00
RCA23_c13260	putative methionine synthase (B12 dependent) subunit 2	COG5012	R		100.00	100.00
RCA23_c13270	putative methionine synthase (B12 dependent) subunit 1	COG0646	E		100.00	100.00
RCA23_c13280	phosphoribosylaminoimidazole-succinocarboxamide synthase PurC	COG0152	F		100.00	99.77
RCA23_c13290	phosphoribosylformylglycinamidine (FGAM) synthase PurS	COG1828	F		100.00	100.00
RCA23_c13300	phosphoribosylformylglycinamidine synthase PurQ	COG0047	F		100.00	100.00
RCA23_c13310	C4-dicarboxylate transport sensor protein DctB	COG4191	T		100.00	100.00
RCA23_c13320	C4-dicarboxylate transport transcriptional regulatory protein DctD	COG2204	T		100.00	100.00
RCA23_c13330	ribonuclease E	COG1530	J		100.00	99.28

RCA23_c13340	putative sulfurtransferase tusA	COG0425	O	100.00	100.00
RCA23_c13350	cytochrome C biogenesis protein transmembrane region	COG0785	O	100.00	100.00
RCA23_c13360	hypothetical protein			100.00	100.00
RCA23_c13370	putative cytochrome P450	COG2124	Q	100.00	100.00
RCA23_c13380	DNA alkylation repair enzyme	COG4912	L	100.00	100.00
RCA23_c13390	hypothetical protein	COG3153	R	100.00	100.00
RCA23_c13400	glucosamine--fructose-6-phosphate aminotransferase GlmS	COG0449	M	100.00	100.00
RCA23_c13410	bifunctional protein GlmU	COG1207	M	100.00	100.00
RCA23_c13420	putative HAD-superfamily hydrolase	COG0546	R	100.00	100.00
RCA23_c13440	MmgE/PrpD family protein	COG2079	R	100.00	99.16
RCA23_c13430	putative pyridoxal-phosphate-dependent aminotransferase	COG0399	M	94.55	87.12
RCA23_c13450	isovaleryl-CoA dehydrogenase	COG1960	I	100.00	92.78
RCA23_c13460	hypothetical protein			100.00	100.00
RCA23_c13470	methylcrotonoyl-CoA carboxylase beta subunit MccB	COG4799	I	100.00	100.00
RCA23_c13480	hypothetical protein			100.00	94.08
RCA23_c13490	methylcrotonoyl-CoA carboxylase alpha subunit MccA	COG4770	I	100.00	98.45
RCA23_c13500	hydroxymethylglutaryl-CoA lyase MvaB	COG0119	E	100.00	100.00
RCA23_c13510	putative methylglutaconyl-CoA hydratase	COG1024	I	100.00	100.00
RCA23_c13520	NADH-quinone oxidoreductase subunit A	COG0838	C	100.00	100.00
RCA23_c13530	NADH-quinone oxidoreductase subunit NuoB	COG0377	C	100.00	100.00
RCA23_c13540	hypothetical protein DUF2158	COG5475	S	100.00	100.00
RCA23_c13550	NADH-quinone oxidoreductase subunit C	COG0852	C	100.00	100.00
RCA23_c13560	hypothetical protein			100.00	100.00
RCA23_c13570	NADH-quinone oxidoreductase subunit D	COG0649	C	100.00	97.94
RCA23_c13580	NADH-quinone oxidoreductase subunit E	COG1905	C	100.00	84.88
RCA23_c13590	hypothetical protein			100.00	100.00
RCA23_c13600	NADH-quinone oxidoreductase subunit F	COG1894	C	100.00	95.14
RCA23_c13610	hypothetical protein			100.00	96.79
RCA23_c13620	NADH-quinone oxidoreductase subunit G	COG1034	C	100.00	98.27
RCA23_c13630	hypothetical protein			100.00	88.76
RCA23_c13640	NADH-quinone oxidoreductase subunit H	COG1005	C	100.00	91.62
RCA23_c13650	NADH-quinone oxidoreductase subunit I	COG1143	C	100.00	100.00
RCA23_c13660	NADH-quinone oxidoreductase subunit J	COG0839	C	100.00	100.00

RCA23_c13670	NADH-quinone oxidoreductase subunit K	COG0713	C	100.00	100.00
RCA23_c13680	NADH-quinone oxidoreductase subunit L	COG1009	C	100.00	100.00
RCA23_c13690	NADH-quinone oxidoreductase subunit M	COG1008	C	100.00	100.00
RCA23_c13700	NADH-quinone oxidoreductase subunit N	COG1007	C	100.00	100.00
RCA23_c13710	biotin-[acetyl-CoA-carboxylase] ligase	COG0340	H	100.00	100.00
RCA23_c13720	type III pantothenate kinase CoaX	COG1521	K	100.00	100.00
RCA23_c13730	putative ribonuclease	COG0595	R	100.00	100.00
RCA23_c13740	ATP-dependent RNA helicase RhlE	COG0513	L	100.00	95.14
RCA23_c13750	peptide chain release factor 3	COG4108	J	100.00	93.25
RCA23_c13760	hypothetical protein			100.00	89.10
RCA23_c13770	putative short chain dehydrogenase	COG1028	I	100.00	90.77
RCA23_c13780	arsenite methyltransferase	COG2226	H	100.00	100.00
RCA23_c13790	Fe(3+) ions import ATP-binding protein FbpC	COG3842	E	51.16	90.44
RCA23_c13800	putative helix-turn-helix protein	COG2378	K	76.75	100.00
RCA23_c13810	sec-independent protein translocase protein TatA	COG1826	U	100.00	100.00
RCA23_c13820	sec-independent protein translocase protein TatB	COG1826	U	36.47	100.00
RCA23_c13830	sec-independent protein translocase protein TatC	COG0805	U	4.94	97.98
RCA23_c13840	hypothetical protein DUF815	COG2607	R	100.00	100.00
RCA23_c13850	putative peptidoglycan-binding peptidase	COG0739	M	100.00	100.00
RCA23_c13860	protein-L-isoaspartate O-methyltransferase Pcm	COG2518	O	100.00	100.00
RCA23_c13870	5'-nucleotidase SurE	COG0496	R	100.00	100.00
RCA23_c13880	putative short chain dehydrogenase	COG1028	I	100.00	100.00
RCA23_c13890	amidophosphoribosyltransferase PurF	COG0034	F	100.00	98.57
RCA23_c13900	putative colicin V production protein			100.00	100.00
RCA23_c13910	DNA repair protein RadA	COG1066	O	100.00	100.00
RCA23_c13920	ABC transporter ATP-binding protein	COG1127	Q	100.00	100.00
RCA23_c13930	hypothetical protein DUF140	COG0767	Q	100.00	100.00
RCA23_c13940	alanine racemase, biosynthetic	COG0787	M	100.00	100.00
RCA23_c13950	replicative DNA helicase DnaB	COG0305	L	100.00	100.00
RCA23_c13960	orotate phosphoribosyltransferase PyrE	COG0461	F	100.00	100.00
RCA23_c13970	dihydroorotase PyrC	COG0418	F	100.00	95.58
RCA23_c13980	hypothetical protein			100.00	100.00
RCA23_c13990	hypothetical protein			100.00	63.96

RCA23_c14000	hypothetical protein				100.00	87.67
RCA23_c14010	malate dehydrogenase Mdh	COG0039	C		100.00	100.00
RCA23_c14020	citrate lyase beta subunit CitE	COG2301	G		100.00	100.00
RCA23_c14030	putative mesaconyl-CoA hydratase	COG2030	I		100.00	100.00
RCA23_c14040	succinate dehydrogenase cytochrome b556 subunit SdhC	COG2009	C		100.00	100.00
RCA23_c14050	succinate dehydrogenase hydrophobic membrane anchor subunit SdhD	COG2142	C		100.00	100.00
RCA23_c14060	succinate dehydrogenase flavoprotein subunit SdhA	COG1053	C		79.62	100.00
RCA23_c14070	hypothetical protein				100.00	100.00
RCA23_c14080	succinate dehydrogenase iron-sulfur subunit SdhB	COG0479	C		100.00	100.00
RCA23_c14090	purine nucleoside phosphorylase deoD-type	COG0813	F		100.00	91.91
RCA23_c14100	acetyltransferase	COG1246	E		100.00	94.56
RCA23_c14110	tryptophan synthase alpha chain TrpA	COG0159	E		100.00	95.03
RCA23_c14120	GTP-dependent nucleic acid-binding protein EngD	COG0012	J		84.61	100.00
RCA23_c14130	Non-specific ribonucleoside hydrolase rihC	COG1957	F		94.70	100.00
RCA23_c14150	hypothetical protein DUF583	COG1664	M	GI 7	100.00	61.64
RCA23_c14160	protease HtpX	COG0501	O	GI 7	100.00	82.72
RCA23_c14170	histone deacetylase	COG0123	B	GI 7	100.00	66.45
RCA23_c14180	peripheral-type benzodiazepine receptor/signal transduction protein TspO	COG3476	T	GI 7	100.00	48.84
RCA23_c14190	DNA-invertase Hin	COG1961	L	GI 7	60.90	12.61
RCA23_c14200	hypothetical protein			GI 7	0.00	0.00
RCA23_c14210	hypothetical protein			GI 7	0.00	100.00
RCA23_c14220	hypothetical protein, RmlC-like cupin family	COG3450	R	GI 7	0.00	77.25
RCA23_c14230	integrase	COG4974	L	GI 7	99.30	63.65
RCA23_c14240	hypothetical membrane protein			GI 7	100.00	89.26
RCA23_c14260	hypothetical protein	COG4530	S		100.00	100.00
RCA23_c14270	hypothetical protein DUF45	COG1451	R		100.00	100.00
RCA23_c14280	HTH-type transcriptional regulator, GntR family	COG1802	K		100.00	100.00
RCA23_c14290	dihydrolipoyl dehydrogenase Lpd	COG1249	C		100.00	100.00
RCA23_c14300	dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate (	COG0508	C		100.00	96.66
RCA23_c14310	2-oxoglutarate dehydrogenase E1 component SucA	COG0567	C		88.95	100.00
RCA23_c14320	succinyl-CoA ligase [ADP-forming] alpha subunit SucD	COG0074	C		100.00	100.00
RCA23_c14330	succinyl-CoA ligase [ADP-forming] beta subunit SucC	COG0045	C		100.00	100.00
RCA23_c14340	butyryl-CoA dehydrogenase	COG1960	I		100.00	100.00

RCA23_c14350	translation initiation factor IF-3	COG0290	J	100.00	100.00
RCA23_c14360	ferredoxin--NADP reductase Fpr	COG1018	C	100.00	100.00
RCA23_c14370	hypothetical protein	COG3749	S	100.00	85.29
RCA23_c14380	cysH'	COG0175	E	100.00	100.00
RCA23_c14390	cysI/sir: sulfite reductase (ferredoxin)	COG0155	P	100.00	100.00
RCA23_c14400	hypothetical protein			100.00	100.00
RCA23_c14410	siroheme synthase CysG	COG0007	H	100.00	100.00
RCA23_c14420	HTH-type transcriptional regulator, AsnC family	COG1522	K	100.00	100.00
RCA23_c14450	peptidase	COG0624	E	100.00	100.00
RCA23_c14460	hypothetical protein	COG1426	S	100.00	100.00
RCA23_c14470	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase IspG	COG0821	I	100.00	95.06
RCA23_c14480	TPR-repeat containing protein	COG4783	R	98.33	100.00
RCA23_c14490	aspartate aminotransferase	COG0436	E	100.00	89.56
RCA23_c14500	penicillin-binding protein 1A	COG5009	M	100.00	100.00
RCA23_c14510	peptide chain release factor 2	COG1186	J	41.96	93.96
RCA23_c14520	hypothetical protein DUF583	COG1664	M	0.00	100.00
RCA23_c14530	putative peptidase, M23 family	COG0739	M	55.06	100.00
RCA23_c14540	hypothetical protein DUF455	COG2833	S	100.00	100.00
RCA23_c14550	peroxiredoxin Bcp	COG1225	O	100.00	100.00
RCA23_c14560	hypothetical protein			100.00	95.00
RCA23_c14570	S-adenosylmethionine:tRNA ribosyltransferase-isomerase QueA	COG0809	J	100.00	100.00
RCA23_c14580	MFS-type transporter			100.00	100.00
RCA23_c14590	hypothetical protein DUF924	COG3803	S	100.00	100.00
RCA23_c14600	dihydrolipoyl dehydrogenase Lpd	COG1249	C	100.00	99.79
RCA23_c14610	uvrABC system protein A	COG0178	L	100.00	99.37
RCA23_c14620	MmgE/PrpD family protein	COG2079	R	100.00	100.00
RCA23_c14630	3-hydroxyisobutyrate dehydrogenase MmsB	COG2084	I	100.00	100.00
RCA23_c14640	3-hydroxyisobutyryl-CoA hydrolase	COG1024	I	100.00	100.00
RCA23_c14650	isobutyryl-CoA dehydrogenase	COG1960	I	100.00	95.32
RCA23_c14660	methylmalonate-semialdehyde dehydrogenase MmsA	COG1012	C	100.00	100.00
RCA23_c14670	putative HTH-type transcriptional regulator, LysR family	COG0583	K	100.00	100.00
RCA23_c14680	phosphopantetheine adenylyltransferase CoaD	COG0669	H	100.00	100.00
RCA23_c14690	glyceraldehyde-3-phosphate dehydrogenase Gap	COG0057	G	100.00	100.00

RCA23_c14700	glyceraldehyde-3-phosphate dehydrogenase Gap	COG0057	G	100.00	100.00
RCA23_c14710	transketolase TktA	COG0021	G	95.09	95.93
RCA23_c14720	hypothetical protein			100.00	100.00
RCA23_c14730	cell division protein ZapA	COG3027	S	100.00	100.00
RCA23_c14740	putative glutaredoxin	COG0278	O	100.00	100.00
RCA23_c14750	hypothetical protein			100.00	100.00
RCA23_c14760	hypothetical protein, BolA-like	COG0271	T	100.00	100.00
RCA23_c14770	phosphoribosylformylglycinamidine synthase PurL	COG0046	F	100.00	97.42
RCA23_c14780	HTH-type transcriptional regulator, LysR family	COG0583	K	100.00	100.00
RCA23_c14790	hypothetical protein, pyruvate ferredoxin/flavodoxin oxidoreductase	COG4231	C	100.00	97.99
RCA23_c14800	glutamate racemase MurI	COG0796	M	100.00	100.00
RCA23_c14810	N-acetyl-gamma-glutamyl-phosphate reductase ArgC	COG0002	E	86.88	98.93
RCA23_c14820	cytochrome c-type biogenesis protein CcmE	COG2332	O	54.53	100.00
RCA23_c14830	cytochrome c-type biogenesis protein CcmF	COG1138	O	100.00	99.59
RCA23_c14840	cytochrome c-type biogenesis protein CcmH	COG3088	O	100.00	100.00
RCA23_c14850	putative enoyl-CoA hydratase FadB	COG1024	I	100.00	100.00
RCA23_c14860	hypothetical protein			100.00	100.00
RCA23_c14870	citrate synthase GltA	COG0372	C	100.00	94.65
RCA23_c14880	glutamyl-tRNA synthase 2	COG0008	J	100.00	100.00
RCA23_c14890	hypothetical protein competence protein E	COG0658	R	100.00	100.00
RCA23_c14900	LexA repressor	COG1974	K	100.00	100.00
RCA23_c14910	molybdopterin biosynthesis protein MoeA	COG0303	H	100.00	95.24
RCA23_c14920	molybdenum cofactor biosynthesis protein MoaC	COG0315	H	100.00	100.00
RCA23_c14930	indole-3-glycerol phosphate synthase TrpC	COG0134	E	100.00	100.00
RCA23_c14940	anthranilate phosphoribosyltransferase TrpD	COG0547	E	100.00	91.67
RCA23_c14950	anthranilate synthase component TrpG	COG0512	E	100.00	99.48
RCA23_c14960	hypothetical protein, divergent polysaccharide deacetylase			100.00	100.00
RCA23_c14970	hypothetical protein			100.00	100.00
RCA23_c14990	hypothetical protein	COG3108	S	100.00	100.00
RCA23_c15000	putative L,D-transpeptidase YcbB	COG2989	S	100.00	100.00
RCA23_c15010	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase LpxD	COG1044	M	100.00	100.00
RCA23_c15020	acyl carrier protein	COG0236	I	100.00	100.00
RCA23_c15030	3-oxoacyl-[acyl-carrier-protein] synthase FabF	COG0304	I	100.00	100.00

RCA23_c15040	hypothetical protein, invasion protein B	COG5342	R	100.00	100.00
RCA23_c15050	inner membrane protein	COG4536	P	90.76	96.61
RCA23_c15060	tyrosine recombinase XerD	COG4974	L	100.00	100.00
RCA23_c15070	hypothetical protein			100.00	100.00
RCA23_c15080	hypothetical protein			100.00	100.00
RCA23_c15090	shikimate kinase AroK	COG0703	E	100.00	100.00
RCA23_c15100	3-dehydroquinate synthase AroB	COG0337	E	100.00	100.00
RCA23_c15110	single-stranded DNA-binding protein	COG0629	L	100.00	100.00
RCA23_c15120	glutathione import ATP-binding protein GsiA	COG0444	E	100.00	98.51
RCA23_c15130	putative ABC transporter inner membrane component	COG1173	E	100.00	94.69
RCA23_c15140	putative ABC transporter permease protein	COG0601	E	100.00	100.00
RCA23_c15150	ABC transporter extracellular solute-binding protein	COG0747	E	60.90	100.00
RCA23_c15160	tRNA delta(2)-isopentenylpyrophosphate transferase MiaA	COG0324	J	74.97	100.00
RCA23_c15170	uridylate kinase PyrH	COG0528	F	100.00	100.00
RCA23_c15180	ribosome recycling factor Frr	COG0233	J	100.00	100.00
RCA23_c15190	undecaprenyl pyrophosphate synthase UppS	COG0020	I	100.00	100.00
RCA23_c15200	putative cytidylyltransferase	COG0575	I	100.00	100.00
RCA23_c15210	1-deoxy-D-xylulose 5-phosphate reductoisomerase Dxr	COG0743	I	100.00	98.71
RCA23_c15220	RIP metalloprotease RseP	COG0750	M	100.00	100.00
RCA23_c15230	putative outer membrane assembly factor	COG4775	M	100.00	92.25
RCA23_c15240	putative outer tein			70.97	100.00
RCA23_c15250	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase FabZ	COG0764	I	0.00	100.00
RCA23_c15260	acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase LpxA	COG1043	M	83.33	100.00
RCA23_c15270	hypothetical protein DUF1009	COG3494	S	100.00	94.03
RCA23_c15280	lipid-A-disaccharide synthase LpxB	COG0763	M	100.00	100.00
RCA23_c15290	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase TrmU	COG0482	J	85.49	99.56
RCA23_c15300	hypothetical protein			100.00	92.98
RCA23_c15310	cell cycle transcriptional regulator	COG0745	T	77.73	100.00
RCA23_c15320	DNA ligase LigA	COG0272	L	100.00	89.77
RCA23_c15330	ATP-dependent DNA helicase RecG	COG1200	L	98.76	100.00
RCA23_c15340	hypothetical protein			100.00	100.00
RCA23_c15350	hypothetical protein	COG0822	C	100.00	100.00
RCA23_c15360	phosphoribosyl-AMP cyclohydrolase HisI	COG0139	E	100.00	100.00

RCA23_c15370	glutamyl-Q tRNA(Asp) synthase GluQ	COG0008	J	86.90	100.00
RCA23_c15380	hypothetical protein, methyltransferase	COG4976	R	100.00	100.00
RCA23_c15390	tRNA uridine 5-carboxymethylaminomethyl modification enzyme Gid	COG1206	J	100.00	96.56
RCA23_c15400	putative crotonase	COG1024	I	100.00	94.04
RCA23_c15410	putative thioesterase	COG2050	Q	69.98	100.00
RCA23_c15420	50S ribosomal protein L13	COG0102	J	0.00	100.00
RCA23_c15430	30S ribosomal protein S9	COG0103	J	0.00	100.00
RCA23_c15440	integrase			0.00	64.64
RCA23_c15460	choline dehydrogenase BetA	COG2303	E	100.00	89.43
RCA23_c15470	ABC transporter, permease protein	COG4662	H	100.00	100.00
RCA23_c15480	ABC transporter, ATP-binding protein	COG1131	V	100.00	100.00
RCA23_c15490	ABC transporter, periplasmic substrate-binding protein	COG2998	H	100.00	100.00
RCA23_c15500	hypothetical membrane protein	COG1238	S	100.00	100.00
RCA23_c15510	hypothetical protein DUF6 family, transmembrane	COG2962	R	100.00	100.00
RCA23_c15520	hypothetical protein	COG2068	R	100.00	94.33
RCA23_c15530	dimethylglycine dehydrogenase	COG0404	E	100.00	95.19
RCA23_c15540	dimethylglycine dehydrogenase	COG0404	E	100.00	98.03
RCA23_c15550	short-chain dehydrogenase/reductase SDR	COG1028	I	100.00	100.00
RCA23_c15560	hypothetical transmembrane protein	COG3788	R	100.00	100.00
RCA23_c15570	hypothetical protein, tetR regulator	COG1309	K	100.00	100.00
RCA23_c15580	homoserine O-succinyltransferase MetA	COG1897	E	100.00	100.00
RCA23_c15590	hypothetical protein			100.00	100.00
RCA23_c15600	putative integral membrane protein	COG2510	S	60.07	100.00
RCA23_c15610	Gcn5-like N-acetyltransferase	COG1670	J	100.00	84.62
RCA23_c15620	GMP synthase GuaA	COG0519	F	100.00	100.00
RCA23_c15630	trimethylamine methyltransferase MttB	COG5598	H	100.00	100.00
RCA23_c15640	hypothetical protein			100.00	100.00
RCA23_c15650	lipoyl synthase LipA	COG0320	H	100.00	100.00
RCA23_c15660	hypothetical protein			100.00	100.00
RCA23_c15670	hypoxanthine phosphoribosyltransferase Hpt	COG0634	F	100.00	100.00
RCA23_c15680	hypothetical protein	COG2867	I	100.00	100.00
RCA23_c15690	ammonium transporter AmtB	COG0004	P	100.00	100.00
RCA23_c15700	putative competence-damaged protein	COG1546	R	100.00	100.00

RCA23_c15710	phosphatidylglycerophosphatase A	COG1267	I	100.00	100.00
RCA23_c15720	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase IspF	COG0245	I	100.00	97.73
RCA23_c15730	tRNA-dihydrouridine synthase B	COG0042	J	100.00	95.22
RCA23_c15740	histidine kinase, nitrogen regulation protein NtrB	COG3852	T	100.00	100.00
RCA23_c15750	nitrogen regulation protein NtrC	COG2204	T	100.00	100.00
RCA23_c15760	histidine kinase, nitrogen regulation protein NtrY	COG5000	T	100.00	97.57
RCA23_c15770	nitrogen assimilation regulatory protein NtrX	COG2204	T	100.00	100.00
RCA23_c15780	Trk system potassium uptake protein TrkA	COG0569	P	100.00	96.80
RCA23_c15790	putative Trk system potassium uptake protein trkH	COG0168	P	100.00	99.73
RCA23_c15800	RNA-binding protein Hfq	COG1923	R	100.00	100.00
RCA23_c15810	GTP-binding protein HflX	COG2262	R	100.00	100.00
RCA23_c15820	acyl-homoserine lactone acylase QuiP	COG2366	R	100.00	96.81
RCA23_c15830	hypothetical protein	COG2930	S	100.00	93.91
RCA23_c15840	delta-aminolevulinic acid dehydratase HemB	COG0113	H	100.00	91.89
RCA23_c15850	hypothetical protein			100.00	100.00
RCA23_c15860	transcription-repair-coupling factor Mfd	COG1197	L	100.00	98.08
RCA23_c15870	putative DSBA-like thioredoxin family protein	COG2761	Q	100.00	66.67
RCA23_c15880	long-chain-fatty-acid-CoA ligase	COG0318	I	100.00	94.05
RCA23_c15890	HTH-type transcriptional regulator	COG1396	K	100.00	100.00
RCA23_c15900	aquaporin AqpZ	COG0580	G	100.00	100.00
RCA23_c15910	putative extracellular solute-binding protein	COG4166	E	100.00	100.00
RCA23_c15920	lysine exporter protein LysE	COG1279	R	100.00	100.00
RCA23_c15930	hypothetical protein DUF502	COG2928	S	100.00	99.43
RCA23_c15940	pseudouridine-5'-phosphate glycosidase PsuG	COG2313	Q	100.00	100.00
RCA23_c15950	hypothetical protein, pfkB family carbohydrate kinase	COG0524	G	100.00	100.00
RCA23_c15960	cold shock protein CspA	COG1278	K	100.00	100.00
RCA23_c15970	30S ribosomal protein S2	COG0052	J	100.00	100.00
RCA23_c15980	elongation factor Ts	COG0264	J	79.68	100.00
RCA23_c15990	hypothetical protein			95.07	79.32
RCA23_c16000	hypothetical protein, acetyltransferase-like	COG1670	J	100.00	48.43
RCA23_c16010	hypothetical protein, transcription factor NusA like			59.81	73.52
RCA23_c16020	putative phenylacetic acid degradation protein	COG0663	R	0.00	100.00
RCA23_c16030	guanylate kinase Gmk	COG0194	F	91.12	83.02

RCA23_c16040	hypothetical protein, YicC-like	COG1561	S	100.00	100.00
RCA23_c16050	hypothetical protein, DUF1457	COG5388	S	100.00	97.93
RCA23_c16060	phospho-2-dehydro-3-deoxyheptonate aldolase	COG3200	E	100.00	98.91
RCA23_c16070	putative HTH-type transcriptional regulator, AraC family	COG4977	K	100.00	98.52
RCA23_c16080	putative amino-acid binding protein	COG0683	E	100.00	80.88
RCA23_c16090	putative branched-chain amino acid transport ATP-binding protein	COG0411	E	84.49	93.46
RCA23_c16100	putative branched-chain amino acid transport ATP-binding protein	COG0410	E	100.00	79.13
RCA23_c16110	putative branched-chain amino acid transport system permease protein	COG0559	E	100.00	82.35
RCA23_c16120	putative branched-chain amino acid transport system permease protein	COG4177	E	100.00	97.53
RCA23_c16130	GTP-binding protein TypA	COG1217	T	91.58	100.00
RCA23_c16140	hypothetical protein DUF1330	COG5470	S	100.00	65.29
RCA23_c16150	alanyl-tRNA synthase AlaS	COG0013	J	77.92	100.00
RCA23_c16160	protein RecA	COG0468	L	100.00	91.02
RCA23_c16170	hypothetical protein			100.00	100.00
RCA23_c16180	sensor transduction histidine kianse	COG0642	T	71.23	100.00
RCA23_c16190	hypothetical protein			40.23	92.64
RCA23_c16200	hypothetical protein, NOL1/NOP2/sun family	COG0144	J	100.00	89.03
RCA23_c16210	inosine-5'-monophosphate dehydrogenase GuaB	COG0516	F	100.00	100.00
RCA23_c16220	hypothetical protein	COG3894	R	100.00	92.60
RCA23_c16240	cysteine desulfurase SufS	COG0520	E	100.00	100.00
RCA23_c16250	hypothetical protein			100.00	100.00
RCA23_c16260	hypothetical protein			100.00	95.60
RCA23_c16270	FeS assembly protein SufD	COG0719	O	100.00	100.00
RCA23_c16280	FeS assembly ATPase SufC	COG0396	O	100.00	100.00
RCA23_c16290	FeS assembly protein SufB	COG0719	O	59.19	100.00
RCA23_c16300	putative cysteine desulfurase	COG1104	E	100.00	100.00
RCA23_c16310	HTH-type transcriptional regulator Rrf2	COG1959	K	100.00	100.00
RCA23_c16320	hypothetical protein	COG2945	R	100.00	100.00
RCA23_c16330	putative oxidoreductase	COG0673	R	100.00	100.00
RCA23_c16340	inosose dehydratase IolE	COG1082	G	100.00	100.00
RCA23_c16350	putative oxidoreductase	COG0673	R	100.00	100.00
RCA23_c16360	HTH-type transcriptional repressor, LacI family	COG1609	K	100.00	100.00
RCA23_c16370	hypothetical protein			100.00	100.00

RCA23_c16380	sugar (ribose) ABC-transport system ATP binding protein	COG1129	G	100.00	100.00
RCA23_c16390	sugar (ribose) ABC transporter permease protein	COG1172	G	100.00	99.35
RCA23_c16400	sugar (ribose) ABC transporter periplasmic binding protein	COG1879	G	100.00	100.00
RCA23_c16410	beta-glucosidase BglA	COG2723	G	100.00	84.77
RCA23_c16420	maltose/maltodextrin import ATP-binding protein MalK	COG3839	G	100.00	100.00
RCA23_c16430	putative alpha-glucosidase AglA	COG0366	G	100.00	100.00
RCA23_c16440	alpha-glucoside transport system permease protein AglG	COG0395	G	100.00	100.00
RCA23_c16450	ABC transporter, membrane spanning protein	COG0395	G	100.00	100.00
RCA23_c16460	alpha-glucoside transport system permease protein AglF	COG1175	G	100.00	96.57
RCA23_c16470	alpha-glucosides-binding periplasmic protein AglE	COG1653	G	100.00	98.21
RCA23_c16480	HTH-type transcriptional regulator	COG1609	K	100.00	100.00
RCA23_c16490	oxidoreductase	COG4989	R	100.00	100.00
RCA23_c16500	alkanesulfonate monooxygenase SsuD	COG2141	C	100.00	87.13
RCA23_c16510	putative regulatory DNA binding protein	COG2188	K	100.00	94.56
RCA23_c16520	putative oxidoreductase	COG0673	R	100.00	100.00
RCA23_c16530	glucose-6-phosphate isomerase Pgi	COG0166	G	100.00	100.00
RCA23_c16540	6-phosphogluconolactonase Pgl	COG0363	G	100.00	100.00
RCA23_c16550	glucose-6-phosphate 1-dehydrogenase Zwf	COG0364	G	74.66	88.29
RCA23_c16560	hypothetical protein, radical SAM	COG0535	R	100.00	100.00
RCA23_c16570	putative 5-methylcytosine-specific restriction enzyme McrA	COG1403	V	100.00	100.00
RCA23_c16580	putative phospholipase/carboxylesterase	COG0400	R	94.37	98.52
RCA23_c16590	HhH-GPD superfamily base excision DNA repair protein	COG0122	L	98.24	100.00
RCA23_c16600	precorrin-6A reductase CobK	COG2099	H	100.00	97.68
RCA23_c16610	cobalamin (vitamin B12) biosynthesis CbiDprotein CbiD	COG1903	H	100.00	49.41
RCA23_c16620	uroporphyrinogen-III C-methyltransferase CobA	COG0007	H	100.00	100.00
RCA23_c16630	cobyrinic acid A,C-diamide synthase CobB	COG1797	H	100.00	100.00
RCA23_c16640	putative major facilitator superfamily transporter			100.00	95.59
RCA23_c16650	hypothetical protein	COG1562	I	100.00	100.00
RCA23_c16660	cysteinyl-tRNA synthase CysS	COG0215	J	100.00	100.00
RCA23_c16670	aspartate aminotransferase AspC	COG0436	E	100.00	100.00
RCA23_c16680	trimethylamine methyltransferase MttB	COG5598	H	100.00	100.00
RCA23_c16690	ABC transporter ATP-binding protein Uup	COG0488	R	97.63	100.00
RCA23_c16700	hypothetical protein	COG3176	R	100.00	85.60

RCA23_c16710	peptide methionine sulfoxide reductase MsrB	COG0229	O	100.00	89.22
RCA23_c16720	hypothetical protein, lipoprotein	COG2913	J	100.00	81.38
RCA23_c16730	hypothetical protein	COG1399	R	100.00	100.00
RCA23_c16740	fatty acid/phospholipid synthesis protein PlsX	COG0416	I	100.00	100.00
RCA23_c16750	3-oxoacyl-[acyl-carrier-protein] synthase FabH	COG0332	I	100.00	100.00
RCA23_c16760	integration host factor alpha subunit IhfA	COG0776	L	100.00	100.00
RCA23_c16770	hypothetical protein, MerR family regulatory protein	COG0789	K	100.00	95.91
RCA23_c16790	deoxyguanosinetriphosphate triphosphohydrolase-like protein Dcd	COG0717	F	88.73	100.00
RCA23_c16800	putative segregation and condesation protein B	COG1386	K	3.74	100.00
RCA23_c16810	hypothetical protein, segregation and condensation protein A	COG1354	S	100.00	84.61
RCA23_c16820	beta-hexosaminidase NagZ	COG1472	G	100.00	100.00
RCA23_c16830	hypothetical protein			100.00	100.00
RCA23_c16840	arginyl-tRNA synthase	COG0018	J	100.00	100.00
RCA23_c16850	deoxyguanosinetriphosphate triphosphohydrolase-like protein	COG0232	F	100.00	100.00
RCA23_c16860	putative iron-sulfur insertion protein erpA	COG0316	S	100.00	95.41
RCA23_c16870	exodeoxyribonuclease III	COG0708	L	57.16	93.92
RCA23_c16880	putative cytochrome B561	COG3038	C	100.00	99.14
RCA23_c16890	putative dnak suppressor protein dksA	COG1734	T	100.00	63.36
RCA23_c16900	hypothetical protein	COG0714	R	100.00	100.00
RCA23_c16910	hypothetical protein	COG3825	S	100.00	98.90
RCA23_c16920	putative peptidase, M48 family	COG0501	O	100.00	78.99
RCA23_c16930	hypothetical protein			100.00	100.00
RCA23_c16940	putative ribosomal RNA large subunit methyltransferase	COG1092	R	100.00	97.33
RCA23_c16950	phosphogluconate dehydratase Edd	COG0129	E	100.00	100.00
RCA23_c16960	KHG/KDPG aldolase Eda	COG0800	G	100.00	100.00
RCA23_c16970	hypothetical protein			100.00	100.00
RCA23_c16980	glutamate-ammonia-ligase adenylyltransferase GlnE	COG1391	O	100.00	100.00
RCA23_c16990	hypothetical protein	COG2606	S	100.00	100.00
RCA23_c17000	hypothetical protein			100.00	100.00
RCA23_c17010	hypothetical protein			100.00	92.89
RCA23_c17020	hypothetical protein			100.00	92.34
RCA23_c17030	putative aromatic-ring-hydroxylating dioxygenase	COG4638	P	77.49	100.00
RCA23_c17040	putative acetolactate synthase small subunit IlvH	COG0440	E	100.00	100.00

RCA23_c17050	acetolactate synthase isozyme large subunit IlvI	COG0028	E	100.00	92.98
RCA23_c17060	hypothetical membrane protein, porin-like			100.00	100.00
RCA23_c17070	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	100.00	94.50
RCA23_c17080	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	100.00	84.97
RCA23_c17090	TRAP dicarboxylate transporter, subunit DctM	COG4664	Q	100.00	100.00
RCA23_c17100	putative arginyl-tRNA--protein transferase Ate	COG2935	O	100.00	97.45
RCA23_c17110	vitamin B12-dependent ribonucleotide reductase NrdJ	COG0209	F	95.29	97.51
RCA23_c17120	hypothetical protein			100.00	100.00
RCA23_c17140	hypothetical protein DUF192	COG1430	S	100.00	100.00
RCA23_c17150	cold shock protein	COG1278	K	100.00	100.00
RCA23_c17160	pyridoxine/pyridoxamine 5'-phosphate oxidase PdxH	COG0259	H	100.00	100.00
RCA23_c17170	enoyl-[acyl-carrier-protein] reductase FabI	COG0623	I	100.00	99.11
RCA23_c17180	xanthine phosphoribosyltransferase Gpt	COG0503	F	100.00	100.00
RCA23_c17190	pyrimidine-specific ribonucleoside hydrolase RihA	COG1957	F	100.00	88.35
RCA23_c17200	hypothetical protein UPF0061	COG0397	S	100.00	100.00
RCA23_c17210	putative sodium/hydrogen exchanger	COG0025	P	100.00	100.00
RCA23_c17220	ferric uptake regulator protein Fur	COG0735	P	100.00	100.00
RCA23_c17230	putative S-adenosylmethionine uptake transporter	COG0697	G	100.00	94.63
RCA23_c17240	enolase Eno	COG0148	G	100.00	100.00
RCA23_c17250	anhydro-N-acetylmuramic acid kinase AnmK	COG2377	O	100.00	100.00
RCA23_c17260	tyrosyl-tRNA synthase TyrS	COG0162	J	100.00	100.00
RCA23_c17270	aspartate aminotransferase AspC	COG0436	E	100.00	95.63
RCA23_c17280	hypothetical protein	COG0760	O	100.00	99.35
RCA23_c17290	anthranilate synthase component TrpE	COG0147	E	100.00	99.74
RCA23_c17300	protein soxG	COG0491	R	100.00	100.00
RCA23_c17310	protein soxH	COG0491	R	100.00	96.72
RCA23_c17320	5-aminolevulinate synthase HemA	COG0156	H	100.00	72.73
RCA23_c17330	sulfide dehydrogenase flavoprotein chain SoxF	COG3439	S	100.00	100.00
RCA23_c17340	domain of unknown function DUF1791	COG1416	S	100.00	100.00
RCA23_c17350	sulfite oxidase cytochrome subunit SoxD	COG3474	C	100.00	100.00
RCA23_c17360	sulfite oxidase molybdopterin subunit SoxC	COG2041	R	100.00	100.00
RCA23_c17370	sulfur oxidation protein SoxB	COG0737	F	100.00	82.16
RCA23_c17380	diheme cytochrome c	COG3258	C	100.00	100.00

RCA23_c17390	protein SoxZ			100.00	89.09
RCA23_c17400	protein SoxY	COG5501	S	100.00	81.32
RCA23_c17410	cytochrome c	COG2010	C	100.00	64.22
RCA23_c17420	thioredoxin SoxW	COG2143	O	100.00	71.86
RCA23_c17430	cytochrome c-type biogenesis protein SoxV	COG0785	O	100.00	100.00
RCA23_c17440	protein SoxS			100.00	100.00
RCA23_c17450	HTH-type transcriptional regulator, ArsR family	COG0640	K	100.00	100.00
RCA23_c17460	putative soxT, transmembrane protein DUF395	COG2391	R	100.00	100.00
RCA23_c17470	hypothetical protein DUF395	COG2391	R	100.00	100.00
RCA23_c17480	NAD-binding protein	COG2084	I	100.00	100.00
RCA23_c17490	glyoxylate reductase GyaR	COG1052	C	100.00	96.82
RCA23_c17500	glutamyl-tRNA(Gln) amidotransferase subunit A	COG0154	J	100.00	99.25
RCA23_c17510	TRAP dicarboxylate transporter, subunit DctM	COG4664	Q	100.00	96.38
RCA23_c17520	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	100.00	100.00
RCA23_c17530	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	100.00	100.00
RCA23_c17540	hypothetical protein			100.00	100.00
RCA23_c17550	glycerol kinase GlpK	COG0554	C	100.00	100.00
RCA23_c17560	ribosomal RNA large subunit methyltransferase J	COG0293	J	100.00	100.00
RCA23_c17570	putative Ppx/GppA phosphatase family protein	COG0248	F	100.00	99.73
RCA23_c17580	hypothetical protein			100.00	100.00
RCA23_c17590	hypothetical protein	COG0685	E	100.00	100.00
RCA23_c17600	pterin domain containing enzyme	COG1410	E	100.00	97.75
RCA23_c17610	HpcH/HpaI aldolase family protein	COG3836	G	100.00	100.00
RCA23_c17620	glucokinase Glk	COG0837	G	100.00	100.00
RCA23_c17630	hypothetical protein DUF1006	COG3214	S	100.00	100.00
RCA23_c17640	reductive dehalogenase	COG1018	C	100.00	95.60
RCA23_c17650	hypothetical protein, XdhC and CoxI	COG1975	O	100.00	100.00
RCA23_c17660	2-hydroxy-3-oxopropionate reductase GlxR	COG2084	I	100.00	90.68
RCA23_c17670	3-hydroxybutyrate dehydrogenase	COG1028	I	100.00	97.71
RCA23_c17680	dihydroxy-acid dehydratase IlvD	COG0129	E	100.00	99.22
RCA23_c17690	fumarylacetoacetate hydrolase	COG0179	Q	100.00	100.00
RCA23_c17700	hypothetical protein			100.00	100.00
RCA23_c17710	nucleoside diphosphate kinase Ndk	COG0105	F	100.00	100.00

RCA23_c17720	hypothetical protein			59.55	55.34
RCA23_c17730	ABC transporter ATP-binding protein	COG0488	R	100.00	96.25
RCA23_c17740	MarC family integral membrane protein	COG2095	U	100.00	100.00
RCA23_c17750	hypothetical protein	COG3577	R	100.00	92.48
RCA23_c17760	putative DNA polymerase III chi subunit, HolC	COG2927	L	100.00	100.00
RCA23_c17770	cytosol aminopeptidase PepA	COG0260	E	100.00	100.00
RCA23_c17780	puative permease, YjgP/YjgQ family	COG0795	R	100.00	100.00
RCA23_c17790	puative permease, YjgP/YjgQ family	COG0795	R	100.00	100.00
RCA23_c17800	putative organic solvent tolerance protein	COG1452	M	91.88	100.00
RCA23_c17810	hypothetical protein, SurA	COG0760	O	100.00	92.83
RCA23_c17820	4-hydroxythreonine-4-phosphate dehydrogenase PdxA	COG1995	H	100.00	100.00
RCA23_c17830	dimethyladenosine transferase KsgA	COG0030	J	100.00	100.00
RCA23_c17840	hypothetical protein			100.00	100.00
RCA23_c17850	modification methylase, hemK family	COG2890	J	100.00	100.00
RCA23_c17860	peptide chain release factor 1	COG0216	J	100.00	100.00
RCA23_c17870	hypothetical protein	COG4446	S	100.00	100.00
RCA23_c17880	agmatinase SpeB	COG0010	E	100.00	100.00
RCA23_c17890	alpha/beta hydrolase			100.00	86.08
RCA23_c17900	agmatinase SpeB	COG0010	E	100.00	100.00
RCA23_c17910	hippurate hydrolase HipO	COG1473	R	100.00	100.00
RCA23_c17920	protein MazG	COG3956	R	100.00	96.06
RCA23_c17930	ABC transporter periplasmic iron-binding protein FutA	COG1840	P	93.77	95.75
RCA23_c17940	putative peptidyl-prolyl cis-trans isomerase Ppi	COG0652	O	100.00	78.69
RCA23_c17950	putative peptidyl-prolyl cis-trans isomerase Ppi	COG0652	O	100.00	100.00
RCA23_c17960	phosphoglycerate kinase Pgk	COG0126	G	100.00	100.00
RCA23_c17970	fructose-bisphosphate aldolase class 1	COG3588	G	100.00	100.00
RCA23_c17980	hypothetical protein, septum formation initiator	COG2919	D	100.00	100.00
RCA23_c17990	pyruvate dehydrogenase E1 component alpha subunit PdhA	COG1071	C	100.00	100.00
RCA23_c18000	pyruvate dehydrogenase E1 component beta subunit PdhB	COG0022	C	100.00	100.00
RCA23_c18010	dihydrolipoyllysine-residue acetyltransferase PdhC	COG0508	C	100.00	100.00
RCA23_c18020	serine acetyltransferase CysE	COG1045	E	100.00	97.37
RCA23_c18030	putative gene transfer agent protein			50.04	58.62
RCA23_c18040	putative gene transfer agent large terminase part 1	COG5323	S	10.00	3.73

RCA23_c18050	putative gene transfer agent large terminase part 2	COG5323	S	100.00	71.38
RCA23_c18060	aminodeoxychorismate lyase	COG1559	R	100.00	100.00
RCA23_c18070	3-oxoacyl-[acyl-carrier-protein] synthase FabF	COG0304	I	100.00	100.00
RCA23_c18080	acyl carrier protein AcpP	COG0236	I	100.00	100.00
RCA23_c18090	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	100.00	100.00
RCA23_c18100	malonyl CoA-acyl carrier protein transacylase	COG0331	I	100.00	100.00
RCA23_c18110	30S ribosomal protein S6	COG0360	J	100.00	94.81
RCA23_c18120	30S ribosomal protein S18	COG0238	J	100.00	100.00
RCA23_c18130	50S ribosomal protein L9	COG0359	J	100.00	100.00
RCA23_c18140	trigger factor (TF)	COG0544	O	100.00	100.00
RCA23_c18160	nitrogen regulatory protein P-II 1	COG0347	E	100.00	95.87
RCA23_c18170	glutamine synthase GlnA type I	COG0174	E	100.00	99.15
RCA23_c18190	hypothetical protein			100.00	100.00
RCA23_c18180	dimethylpropiothetin dethiomethylase DddP	COG0006	E	100.00	97.95
RCA23_c18200	Biotin transporter BioY	COG1268	R	100.00	87.00
RCA23_c18210	hypothetical protein			100.00	100.00
RCA23_c18220	adenylosuccinate lyase PurB	COG0015	F	100.00	100.00
RCA23_c18230	putative nitrile hydratase, beta subunit			100.00	100.00
RCA23_c18240	hypothetical protein			100.00	100.00
RCA23_c18250	nitrile hydratase alpha subunit NthA			100.00	100.00
RCA23_c18260	hypothetical protein DUF6 transmembrane	COG0697	G	100.00	100.00
RCA23_c18270	putative lipid A biosynthesis lauroyl acyltransferase	COG1560	M	100.00	100.00
RCA23_c18280	hypothetical protein	COG5429	S	100.00	100.00
RCA23_c18290	aconitate hydratase AcnA	COG1048	C	100.00	99.93
RCA23_c18300	cytochrome c biogenesis protein CcmG	COG0526	O	100.00	100.00
RCA23_c18310	cytochrome c-type biogenesis protein CcmC	COG0755	O	100.00	100.00
RCA23_c18320	cytochrome c-type biogenesis protein CcmB	COG2386	O	100.00	100.00
RCA23_c18330	cytochrome c biogenesis ATP-binding export protein CcmA	COG4133	O	100.00	100.00
RCA23_c18340	hypothetical protein	COG3737	S	100.00	100.00
RCA23_c18350	protein export membrane protein SecF	COG0341	U	100.00	100.00
RCA23_c18360	protein export membrane protein SecD	COG0342	U	100.00	96.14
RCA23_c18370	putative immunogenic membrane protein YajC	COG1862	U	100.00	100.00
RCA23_c18380	seryl-tRNA synthase SerS	COG0172	J	100.00	99.15

RCA23_c18390	hypothetical protein, alpha/beta hydrolase-like	COG2267	I	0.00	86.50
RCA23_c18400	GTP-binding protein EngA	COG1160	R	100.00	100.00
RCA23_c18410	putative quinoprotein	COG1520	S	100.00	100.00
RCA23_c18420	hypothetical protein DUF2133	COG4649	S	100.00	100.00
RCA23_c18430	RND efflux transporter, MFP subunit	COG0845	M	100.00	100.00
RCA23_c18440	RND efflux transporter, permease protein	COG0841	V	100.00	98.65
RCA23_c18450	ABC transporter ATP-binding protein	COG5265	O	100.00	97.90
RCA23_c18460	hypothetical protein, peptidoglycan-binding protein domain LysM	COG1652	S	100.00	100.00
RCA23_c18470	protein RarD	COG2962	R	100.00	100.00
RCA23_c18480	superoxide dismutase SodB	COG0605	P	100.00	100.00
RCA23_c18490	sarcosine oxidase subunit SoxG	COG4583	E	100.00	100.00
RCA23_c18500	sarcosine oxidase alpha subunit SoxA	COG0404	E	100.00	91.87
RCA23_c18510	sarcosine oxidase subunit SoxD	COG4311	E	100.00	100.00
RCA23_c18520	sarcosine oxidase beta subunit SoxB	COG0665	E	100.00	100.00
RCA23_c18530	cytochrome c-type biogenesis protein CycH	COG4235	O	100.00	100.00
RCA23_c18540	hypothetical protein	COG0816	L	100.00	100.00
RCA23_c18550	hypothetical protein DUF1289	COG3313	R	100.00	100.00
RCA23_c18560	hypothetical protein			100.00	100.00
RCA23_c18570	hypothetical protein DUF81	COG0730	R	100.00	100.00
RCA23_c18580	tRNA-dihydrouridine synthase Dus	COG0042	J	97.08	83.63
RCA23_c18590	putative peroxiredoxin (thioredoxin reductase)	COG0678	O	0.00	0.00
RCA23_c18600	benzaldehyde dehydrogenase	COG1012	C	0.00	0.00
RCA23_c18610	hypothetical protein			0.00	0.00
RCA23_c18620	hypothetical protein	COG2130	R	99.52	100.00
RCA23_c18630	MFS-type transporter	COG2814	G	100.00	100.00
RCA23_c18640	sorbitol dehydrogenase	COG1028	I	100.00	100.00
RCA23_c18650	6-hydroxynicotinate 3-monooxygenase	COG0654	H	100.00	85.61
RCA23_c18660	hypothetical protein			100.00	100.00
RCA23_c18670	glycerate kinase	COG2379	G	100.00	90.19
RCA23_c18680	putative carboxymuconolactone decarboxylase	COG0599	S	100.00	100.00
RCA23_c18690	2,4-dienoyl-CoA reductase [NADPH]	COG1902	C	100.00	97.17
RCA23_c18700	thiamine pyrophosphate protein	COG0028	E	100.00	100.00
RCA23_c18710	rffG/rfbB: dTDP-glucose 4,6-dehydratase	COG1088	M	91.97	100.00

RCA23_c18720	hypothetical protein			100.00	94.02
RCA23_c18730	choline dehydrogenase BetA	COG2303	E	100.00	94.44
RCA23_c18740	succinyl-CoA:3-ketoacid-coenzyme A transferase subunit ScoB	COG2057	I	100.00	100.00
RCA23_c18750	succinyl-CoA:3-ketoacid-coenzyme A transferase subunit ScoA	COG1788	I	100.00	100.00
RCA23_c18760	hypothetical membrane protein			100.00	100.00
RCA23_c18770	putative dimethyl sulfoniopropionate demethylase DmdA	COG0404	E	100.00	100.00
RCA23_c18780	hydantoin utilization protein A	COG0145	E	100.00	100.00
RCA23_c18790	hydantoin utilization protein B	COG0146	E	100.00	94.67
RCA23_c18800	FAD dependent oxidoreductase	COG0665	E	73.78	100.00
RCA23_c18810	hypothetical protein, 3-beta hydroxysteroid dehydrogenase	COG0451	M	100.00	100.00
RCA23_c18820	hypothetical protein, serine/threonine-protein kinase	COG1262	S	100.00	100.00
RCA23_c18830	hypothetical protein	COG0457	R	100.00	85.69
RCA23_c18840	hypothetical protein DUF1989	COG3665	S	100.00	100.00
RCA23_c18850	putative diaminopropionate ammonia-lyase	COG1171	E	70.99	100.00
RCA23_c18860	hypothetical protein, metallopeptidase M24	COG0006	E	100.00	91.49
RCA23_c18870	hypothetical protein, Asp/Glu/hydantoin racemase	COG3473	Q	100.00	100.00
RCA23_c18880	putative 3-oxoadipate enol-lactonase 2, alpha/beta hydrolase family	COG0596	R	100.00	100.00
RCA23_c18890	hypothetical protein DUF1185			100.00	100.00
RCA23_c18900	limonene 1,2-monooxygenase LimB	COG2141	C	100.00	100.00
RCA23_c18910	aldehyde dehydrogenase, cytosolic	COG1012	C	100.00	100.00
RCA23_c18920	ABC transporter, spermidine/putrescine import, permease protein PotC	COG1177	E	100.00	100.00
RCA23_c18930	ABC transporter, spermidine/putrescine import, permease protein PotB	COG1176	E	100.00	100.00
RCA23_c18940	ABC transporter, spermidine/putrescine import, substrate binding protein Pot	COG0687	E	100.00	92.44
RCA23_c18950	ABC transporter, spermidine/putrescine import, ATP-binding protein PotA	COG3842	E	100.00	84.06
RCA23_c18960	putative nitrilotriacetate monooxygenase component B	COG1853	R	100.00	100.00
RCA23_c18970	hypothetical protein, DUF268			70.60	99.44
RCA23_c18980	hypothetical protein DUF28	COG0217	S	100.00	100.00
RCA23_c18990	sodium/sulphate symporter	COG0471	P	100.00	96.98
RCA23_c19000	hypothetical protein	COG0697	G	95.82	99.45
RCA23_c19010	hypothetical protein	COG1692	S	33.09	100.00
RCA23_c19020	osmotically inducible OsmC-like protein	COG1764	O	100.00	100.00
RCA23_c19030	3-hydroxyacyl-CoA dehydrogenase, NAD-binding	COG1250	I	100.00	95.02
RCA23_c19040	putative transcriptional regulator	COG0703	E	100.00	97.50

RCA23_c19050	thioesterase-like protein	COG0824	R	31.07	100.00
RCA23_c19060	benzoyl-CoA oxygenase component A	COG0369	P	0.00	94.54
RCA23_c19070	benzoyl-CoA oxygenase component B	COG3396	S	86.65	98.01
RCA23_c19080	benzoyl-CoA-dihydrodiol lyase BoxC	COG1024	I	100.00	100.00
RCA23_c19090	hypothetical protein DUF309			100.00	100.00
RCA23_c19100	alpha/beta hydrolase	COG0596	R	100.00	100.00
RCA23_c19110	benzoate-coenzyme A ligase	COG0318	I	100.00	100.00
RCA23_c19120	putative 5-formyltetrahydrofolate cyclo-ligase family protein	COG0212	H	100.00	100.00
RCA23_c19130	magnesium transporter MgtE	COG2239	P	100.00	97.82
RCA23_c19140	guanine deaminase GuaD	COG0402	F	78.94	88.73
RCA23_c19150	putative hydroxydechloroatrazine ethylaminohydrolase	COG0402	F	100.00	100.00
RCA23_c19160	putative inositol monophosphatase family protein	COG0483	G	100.00	100.00
RCA23_c19170	putative helix-turn-helix protein	COG1396	K	100.00	100.00
RCA23_c19180	putative alcohol dehydrogenase	COG0604	C	100.00	100.00
RCA23_c19190	putative inner membrane protein	COG0670	R	100.00	100.00
RCA23_c19200	hypothetical protein			100.00	100.00
RCA23_c19210	putative N-acetylmuramoyl-L-alanine amidase amiD	COG3023	V	100.00	100.00
RCA23_c19220	hypothetical protein			100.00	100.00
RCA23_c19230	glutamyl-tRNA(Gln) amidotransferase subunit A	COG0154	J	100.00	100.00
RCA23_c19240	aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit GatC	COG0721	J	100.00	100.00
RCA23_c19250	putative deaminase	COG0590	F	100.00	66.90
RCA23_c19260	putative ribosomal large subunit pseudouridine synthase B	COG1187	J	100.00	100.00
RCA23_c19270	molybdate ABC transporter, ATP-binding protein ModC	COG4148	P	100.00	97.94
RCA23_c19280	molybdate ABC transporter, pemease protein ModB	COG4149	P	100.00	100.00
RCA23_c19290	molybdate ABC transporter, substrate binding protein ModA	COG0725	P	92.51	100.00
RCA23_c19300	hypothetical protein, NUDIX hydrolase	COG0494	L	100.00	100.00
RCA23_c19310	hypothetical protein DUF1178	COG5319	S	100.00	100.00
RCA23_c19320	aspartokinase LysC	COG0527	E	100.00	100.00
RCA23_c19330	phosphoenolpyruvate-protein phosphotransferase PstI	COG3605	T	100.00	99.02
RCA23_c19340	hypothetical protein, acetyltransferase-like	COG3153	R	100.00	100.00
RCA23_c19350	hypothetical protein	COG1853	R	100.00	100.00
RCA23_c19360	putative sulfate transporter	COG0659	P	100.00	100.00
RCA23_c19370	methylthioadenosine phosphorylase MtnP	COG0005	F	100.00	100.00

RCA23_c19380	adenine phosphoribosyltransferase Apt	COG0503	F	100.00	100.00
RCA23_c19390	ribosomal-protein-alanine acetyltransferase RimJ	COG1670	J	100.00	100.00
RCA23_c19400	uncharacterized zinc protease YmxG	COG0612	R	100.00	100.00
RCA23_c19410	threonine synthase ThrC	COG0498	E	100.00	100.00
RCA23_c19420	hypothetical protein, SURF1	COG3346	S	100.00	100.00
RCA23_c19430	cytochrome c oxidase subunit 3	COG1845	C	100.00	97.64
RCA23_c19440	cytochrome c oxidase assembly protein CtaG	COG3175	O	100.00	100.00
RCA23_c19450	protoheme IX farnesyltransferase CtaB	COG0109	O	100.00	100.00
RCA23_c19460	cytochrome c oxidase subunit 2 precursor	COG1622	C	100.00	100.00
RCA23_c19470	protein TldD	COG0312	R	100.00	96.41
RCA23_c19480	SMF family protein	COG0758	L	100.00	97.57
RCA23_c19490	DNA topoisomerase TopA	COG0550	L	66.94	100.00
RCA23_c19500	fructose-bisphosphate aldolase Fba	COG0191	G	100.00	100.00
RCA23_c19510	5-deoxy-glucuronate isomerase IolB	COG3718	G	100.00	100.00
RCA23_c19520	5-dehydro-2-deoxygluconokinase IolC	COG0524	G	100.00	95.76
RCA23_c19530	3D-(3,5/4)-trihydroxycyclohexane-1,2-dione hydrolase IolD	COG3962	E	100.00	100.00
RCA23_c19540	hypothetical protein DUF989	COG3748	S	100.00	100.00
RCA23_c19550	5-hydroxyisourate hydrolase UraH	COG2351	R	100.00	100.00
RCA23_c19560	uric acid degradation bifunctional protein PucL	COG3195	S	100.00	96.32
RCA23_c19570	putative allantoin catabolism protein YlbA	COG3257	R	100.00	97.94
RCA23_c19580	sarcosine oxidase beta subunit SoxB	COG0665	E	100.00	100.00
RCA23_c19590	sarcosine oxidase subunit SoxD	COG4311	E	100.00	100.00
RCA23_c19600	sarcosine oxidase alpha subunit SoxA	COG0404	E	99.56	99.42
RCA23_c19610	sarcosine oxidase subunit SoxG	COG4583	E	100.00	100.00
RCA23_c19620	putative HTH-type transcriptional regulator, AraC family	COG4977	K	100.00	98.81
RCA23_c19630	putative DNA-binding protein	COG1396	K	100.00	99.85
RCA23_c19640	putative membrane protein			100.00	100.00
RCA23_c19650	hypothetical protein			100.00	87.92
RCA23_c19660	hypothetical protein, thioesterase	COG2050	Q	15.78	100.00
RCA23_c19670	hypothetical protein, calcineurin-like phosphoesterase-like	COG1407	R	91.10	74.81
RCA23_c19680	putative DEAD/DEAH box helicase	COG1201	R	100.00	97.90
RCA23_c19690	bidunctional enzyme FolD	COG0190	H	100.00	100.00
RCA23_c19700	formate--tetrahydrofolate ligase Fhs	COG2759	F	100.00	98.45

RCA23_c19710	hypothetical protein			100.00	100.00
RCA23_c19720	cell division protease FtsH	COG0465	O	100.00	100.00
RCA23_c19730	tRNA(Ile)-lysidine synthase TilS	COG0037	D	100.00	100.00
RCA23_c19740	hypothetical protein	COG1729	S	100.00	100.00
RCA23_c19750	peptidoglycan-associated lipoprotein Pal	COG2885	M	100.00	90.42
RCA23_c19760	Tol-Pal system beta propeller repeat protein TolB	COG0823	U	100.00	100.00
RCA23_c19770	hypothetical protein, TolA-like			100.00	86.45
RCA23_c19780	biopolymer transport protein TolR	COG0848	U	100.00	100.00
RCA23_c19790	biopolymer transport protein TolQ	COG0811	U	100.00	100.00
RCA23_c19800	acyl-CoA thioester hydrolase	COG0824	R	100.00	100.00
RCA23_c19810	hypothetical protein			100.00	100.00
RCA23_c19820	holliday junction ATP-dependent DNA helicase RuvB	COG2255	L	100.00	100.00
RCA23_c19830	holliday junction ATP-dependent DNA helicase RuvA	COG0632	L	100.00	100.00
RCA23_c19840	crossover junction endodeoxyribonuclease RuvC	COG0817	L	100.00	100.00
RCA23_c19850	putative ribosomal protein L11 methyltransferase	COG2264	J	100.00	100.00
RCA23_c19860	peptide methionine sulfoxide reductase MsrA	COG0225	O	100.00	100.00
RCA23_c19870	hypothetical protein, major facilitator superfamily transporter			100.00	100.00
RCA23_c19880	L-lactate dehydrogenase IldD	COG1304	C	100.00	100.00
RCA23_c19890	BCCT family transporter involved in DMSP uptake	COG1292	M	100.00	91.01
RCA23_c19900	putative MFS-type transporter			100.00	87.26
RCA23_c19910	50S ribosomal protein L25	COG1825	J	100.00	100.00
RCA23_c19920	peptidyl-tRNA hydrolase Pth	COG0193	J	100.00	100.00
RCA23_c19930	hypothetical protein	COG4427	S	100.00	99.22
RCA23_c19940	hypothetical protein	COG3651	S	100.00	100.00
RCA23_c19950	tryptophan synthase beta chain TrpB	COG0133	E	100.00	100.00
RCA23_c19960	N-(5'-phosphoribosyl)anthranilate isomerase TrpF	COG0135	E	100.00	100.00
RCA23_c19970	hypothetical protein			100.00	100.00
RCA23_c19980	integration host factor beta subunit IhfB	COG0776	L	100.00	100.00
RCA23_c19990	30S ribosomal protein S1	COG0539	J	100.00	100.00
RCA23_c20000	hypothetical protein			100.00	100.00
RCA23_c20010	hypothetical protein, lacI family HTH-type regulatory protein	COG1609	K	0.88	100.00
RCA23_c20020	putative phytanoyl-CoA dioxygenase	COG5285	Q	33.68	100.00
RCA23_c20030	aldo/keto reductase	COG0667	C	100.00	100.00

RCA23_c20040	cytidylate kinase Cmk	COG0283	F	15.77	94.05
RCA23_c20050	3-phosphoshikimate 1-carboxyvinyltransferase AroA	COG0128	E	100.00	100.00
RCA23_c20060	tRNA (guanine-N(7)-)-methyltransferase TrmB	COG0220	R	100.00	100.00
RCA23_c20070	S-adenosylmethionine synthase MetK	COG0192	H	100.00	100.00
RCA23_c20080	apolipoprotein N-acyltransferase Lnt	COG0815	M	100.00	100.00
RCA23_c20090	putative magnesium and cobalt efflux protein	COG1253	R	100.00	100.00
RCA23_c20100	hypothetical protein UPF0054	COG0319	R	100.00	100.00
RCA23_c20110	PhoH-like protein	COG1702	T	100.00	100.00
RCA23_c20120	RNA modification enzyme, MiaB family	COG0621	J	100.00	96.92
RCA23_c20130	hypothetical protein			100.00	100.00
RCA23_c20140	putative ferric uptake regulator family protein	COG0735	P	100.00	100.00
RCA23_c20150	3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase FabA	COG0764	I	100.00	100.00
RCA23_c20160	3-oxoacyl-[acyl-carrier-protein] synthase FabB	COG0304	I	100.00	100.00
RCA23_c20170	enoyl-[acyl-carrier-protein] reductase FabI	COG0623	I	100.00	97.47
RCA23_c20180	threo-3-hydroxyaspartate ammonia-lyase SRY	COG1171	E	100.00	100.00
RCA23_c20190	2-haloalkanoic acid dehalogenase	COG1011	R	100.00	97.51
RCA23_c20200	3-oxoadipate enol-lactonase CatD	COG0596	R	100.00	100.00
RCA23_c20210	putative tetracycline resistance protein, class C	COG2814	G	100.00	97.88
RCA23_c20220	NAD/mycothiol-dependent formaldehyde dehydrogenase	COG1062	C	100.00	100.00
RCA23_c20230	D-galactonate dehydratase DgoD	COG4948	M	100.00	100.00
RCA23_c20240	dimethylglycine dehydrogenase	COG0404	E	100.00	99.88
RCA23_c20250	trimethylamine methyltransferase MttB	COG5598	H	100.00	99.42
RCA23_c20260	hypothetical protein	COG0607	P	100.00	100.00
RCA23_c20270	uncharacterized aminotransferase	COG0161	H	100.00	100.00
RCA23_c20280	aminomethyltransferase, mitochondrial	COG0404	E	100.00	98.18
RCA23_c20290	glycine cleavage system protein GcvH	COG0509	E	100.00	100.00
RCA23_c20300	glycine dehydrogenase GcvP	COG1003	E	100.00	99.96
RCA23_c20310	D-amino acid dehydrogenase small subunit DadA	COG0665	E	100.00	100.00
RCA23_c20320	putrescine ABC transport system putrescine-binding periplasmic protein PotF	COG0687	E	100.00	100.00
RCA23_c20330	putrescine ABC transport system ATP-binding protein PotG	COG3842	E	100.00	100.00
RCA23_c20340	putrescine ABC transport system permease protein PotI	COG1177	E	100.00	100.00
RCA23_c20350	putrescine ABC transport system permease protein PotB	COG1176	E	100.00	100.00
RCA23_c20360	rieske 2Fe-2S domain protein	COG4638	P	100.00	100.00

RCA23_c20370	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	100.00	100.00
RCA23_c20380	TRAP dicarboxylate transporter, subunit DctQ	COG3090	G	100.00	100.00
RCA23_c20390	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	100.00	100.00
RCA23_c20400	putative phytanoyl-CoA dioxygenase	COG5285	Q	100.00	100.00
RCA23_c20410	putative phenylacetic acid degradation NADH oxidoreductase PaaE	COG1018	C	100.00	82.96
RCA23_c20420	phenylacetate-CoA oxygenase, subunit PaaD	COG2151	R	100.00	100.00
RCA23_c20430	phenylacetate-CoA oxygenase, subunit PaaC	COG3396	S	100.00	100.00
RCA23_c20440	phenylacetate-CoA oxygenase, subunit PaaB	COG3460	Q	100.00	100.00
RCA23_c20450	phenylacetate-CoA oxygenase, subunit PaaA	COG3396	S	100.00	100.00
RCA23_c20460	phenylacetic acid degradation operon negative regulatory protein PaaX	COG3327	K	100.00	100.00
RCA23_c20470	phenylacetic acid degradation protein PaaZ	COG1012	C	100.00	100.00
RCA23_c20480	putative enoyl-CoA hydratase PaaG	COG1024	I	100.00	100.00
RCA23_c20490	acyl-coenzyme A thioesterase PaaI	COG2050	Q	100.00	100.00
RCA23_c20500	beta-ketoadipyl CoA thiolase PaaJ	COG0183	I	100.00	100.00
RCA23_c20510	phenylacetate-CoA ligase PaaK	COG1541	H	100.00	100.00
RCA23_c20520	putative HTH-type transcriptional regulator, TetR family	COG1309	K	100.00	100.00
RCA23_c20530	3-dehydroquinate dehydratase AroQ	COG0757	E	100.00	100.00
RCA23_c20540	tyrosine decarboxylase	COG0076	E	100.00	100.00
RCA23_c20550	putative short chain dehydrogenase	COG0300	R	100.00	100.00
RCA23_c20560	hypothetical protein, acetyltransferase-like	COG1247	M	100.00	100.00
RCA23_c20570	protein CsaA	COG0073	R	100.00	100.00
RCA23_c20580	pyrroline-5-carboxylate reductase ProC	COG0345	E	100.00	88.85
RCA23_c20590	hypothetical protein DUF1790	COG5465	S	100.00	100.00
RCA23_c20600	hypothetical protein DUF526	COG2960	S	100.00	82.81
RCA23_c20610	prolipoprotein diacylglyceryl transferase Lgt	COG0682	M	100.00	100.00
RCA23_c20620	hypothetical protein DUF185	COG1565	S	100.00	100.00
RCA23_c20630	multi-copper polyphenol oxidoreductase laccase	COG1496	S	100.00	100.00
RCA23_c20640	hypothetical protein			100.00	85.94
RCA23_c20650	leucine-responsive regulatory protein Lrp	COG1522	K	100.00	93.74
RCA23_c20660	thioredoxin reductase TrxB	COG0492	O	100.00	100.00
RCA23_c20670	bifunctional sulfate adenylyltransferase / adenylyl-sulfate kinase CysC	COG2046	P	100.00	100.00
RCA23_c20680	hypothetical protein DUF1150			100.00	100.00
RCA23_c20690	small heat shock protein IbpA	COG0071	O	100.00	100.00

RCA23_c20700	hypothetical protein DUF465	COG5481	S	100.00	100.00
RCA23_c20710	N5-carboxyaminoimidazole ribonucleotide mutase PurE	COG0041	F	100.00	100.00
RCA23_c20720	N5-carboxyaminoimidazole ribonucleotide synthase PurK	COG0026	F	100.00	100.00
RCA23_c20730	hypothetical protein	COG1075	R	100.00	84.51
RCA23_c20740	chaperonin GroEL	COG0459	O	77.09	97.64
RCA23_c20750	chaperonin GroS	COG0234	O	3.82	100.00
RCA23_c20760	hypothetical protein			100.00	99.68
RCA23_c20770	hypothetical protein			100.00	86.83
RCA23_c20780	hypothetical protein			100.00	100.00
RCA23_c20790	creatinase	COG0006	E	100.00	100.00
RCA23_c20800	putative membrane lipoprotein			100.00	100.00
RCA23_c20810	hypothetical protein DUF261	COG5482	S	100.00	100.00
RCA23_c20820	putative manganese-dependent inorganic pyrophosphatase PpaC	COG1227	C	100.00	100.00
RCA23_c20830	hypothetical protein, HAD family hydrolase	COG0647	G	100.00	100.00
RCA23_c20840	putative MaoC-like dehydratase	COG2030	I	100.00	100.00
RCA23_c20850	riboflavin biosynthesis protein RibF	COG0196	H	100.00	100.00
RCA23_c20860	hypothetical protein UPF0260	COG2983	S	100.00	100.00
RCA23_c20870	low specificity L-threonine aldolase ItaE	COG2008	E	100.00	100.00
RCA23_c20880	hypothetical protein, 2-hydroxychromene-2-carboxylate isomerase-like	COG3917	Q	100.00	100.00
RCA23_c20890	ribose-phosphate pyrophosphokinase Prs	COG0462	F	79.31	100.00
RCA23_c20900	hypothetical protein			100.00	80.37
RCA23_c20910	ATP synthase epsilon chain AtpC	COG0355	C	100.00	100.00
RCA23_c20920	ATP synthase beta subunit AtpD	COG0055	C	100.00	95.93
RCA23_c20930	ATP synthase gamma chain AtpG	COG0224	C	100.00	95.86
RCA23_c20940	ATP synthase alpha subunit AtpA	COG0056	C	100.00	100.00
RCA23_c20950	ATP synthase delta chain AtpH	COG0712	C	100.00	98.24
RCA23_c20960	hypothetical protein			88.98	100.00
RCA23_c20970	hydroxyacylglutathione hydrolase GloB	COG0491	R	100.00	100.00
RCA23_c20980	ATP-dependent Clp protease ATP-binding subunit ClpA	COG0542	O	100.00	100.00
RCA23_c20990	hypothetical protein, OmpA	COG1360	N	100.00	100.00
RCA23_c21000	hypothetical protein			100.00	100.00
RCA23_c21010	putative cation transport protein ChaC	COG3703	P	100.00	93.41
RCA23_c21020	hypothetical protein	COG1357	S	88.91	90.70

RCA23_c21030	protein TyrC	COG0287	E	100.00	100.00
RCA23_c21040	histidinol-phosphate aminotransferase HisC	COG0079	E	100.00	100.00
RCA23_c21050	30S ribosomal protein S4	COG0522	J	100.00	91.17
RCA23_c21060	cold-shock DEAD box protein A	COG0513	L	100.00	100.00
RCA23_c21070	hypothetical protein DUF393	COG3011	S	100.00	100.00
RCA23_c21080	vitamin B12-dependent ribonucleoside-diphosphate reductase	COG0209	F	100.00	100.00
RCA23_c21090	hypothetical protein	COG5458	S	100.00	100.00
RCA23_c21100	ATP-phosphoribosyltransferase HisG	COG0040	E	100.00	100.00
RCA23_c21110	ATP phosphoribosyltransferase regulatory subunit HisZ	COG3705	E	98.19	95.42
RCA23_c21120	histidyl-tRNA synthase HisS	COG0124	J	73.57	100.00
RCA23_c21130	hypothetical protein	COG2900	S	100.00	100.00
RCA23_c21140	DNA polymerase III 1 alpha subunit DnaE	COG0587	L	100.00	98.86
RCA23_c21150	xanthine dehydrogenase XdhA	COG4630	F	100.00	100.00
RCA23_c21160	xanthine dehydrogenase XdhB	COG4631	F	95.14	98.87
RCA23_c21170	xanthine dehydrogenase accessory protein XdhC	COG1975	O	100.00	100.00
RCA23_c21180	sugar ABC transporter, ATP-binding protein	COG3845	R	100.00	100.00
RCA23_c21190	sugar ABC transporter, permease protein	COG4603	R	100.00	91.90
RCA23_c21200	ABC transporter permease protein	COG1079	R	100.00	100.00
RCA23_c21210	putative basic membrane protein	COG1744	R	100.00	100.00
RCA23_c21220	putative FAD dependent oxidoreductase	COG0665	E	86.30	100.00
RCA23_c21230	dimethylglycine dehydrogenase	COG0404	E	99.44	100.00
RCA23_c21240	pantothenate synthase PanC	COG0414	H	100.00	80.00
RCA23_c21250	3-methyl-2-oxobutanoate hydroxymethyltransferase PanB	COG0413	H	100.00	100.00
RCA23_c21260	hypothetical protein	COG3807	S	100.00	100.00
RCA23_c21270	peptidyl-dipeptidase Dcp	COG0339	E	58.17	99.85
RCA23_c21280	molybdopterin biosynthesis protein MoeB	COG0476	H	100.00	100.00
RCA23_c21290	deoxyuridine 5'-triphosphate nucleotidohydrolase Dut	COG0756	F	100.00	100.00
RCA23_c21300	coenzyme A biosynthesis bifunctional protein CoaBC	COG0452	H	100.00	100.00
RCA23_c21310	RNA polymerase sigma-32 factor RpoH	COG0568	K	100.00	98.63
RCA23_c21320	bifunctional adenosylcobalamin biosynthesis protein CobP	COG2087	H	100.00	100.00
RCA23_c21330	putative phosphoglycerate mutase family protein	COG0406	G	100.00	100.00
RCA23_c21340	hypothetical protein, glutathione S-transferase	COG0625	O	100.00	100.00
RCA23_c21350	glutathione S-transferase family protein			100.00	100.00

RCA23_c21360	magnesium chelatase-like protein	COG0606	O	100.00	100.00
RCA23_c21370	magnesium chelatase	COG0606	O	100.00	91.27
RCA23_c21380	putative transmembrane protein			43.10	100.00
RCA23_c21390	putative sulfite exporter TauE/SafE	COG0730	R	20.92	100.00
RCA23_c21400	branched-chain amino acid transporter, permease component	COG1079	R	100.00	100.00
RCA23_c21410	branched-chain amino acid transporter, permease component	COG4603	R	100.00	100.00
RCA23_c21420	ABC transporter ATP-binding protein	COG3845	R	100.00	97.79
RCA23_c21430	putative ABC transporter, periplasmic substrate-binding protein	COG1744	R	100.00	100.00
RCA23_c21440	hypothetical protein, acetyltransferase-like	COG0456	R	100.00	100.00
RCA23_c21450	NifU-like protein	COG0694	O	100.00	100.00
RCA23_c21460	putative universal stress protein	COG0589	T	100.00	94.35
RCA23_c21470	tryptophanyl-tRNA synthase TrpS	COG0180	J	100.00	100.00
RCA23_c21480	hypothetical protein, rhomboid protease	COG0705	R	100.00	87.56
RCA23_c21490	virulence factor MviN homolg	COG0728	R	94.92	100.00
RCA23_c21500	uridylyltransferase GlnD	COG2844	O	100.00	100.00
RCA23_c21510	hypothetical protein	COG0683	E	100.00	100.00
RCA23_c21520	putative tetrapyrrole methylase	COG0313	R	100.00	100.00
RCA23_c21530	hypothetical protein UPF0102	COG0792	L	100.00	100.00
RCA23_c21540	glutathione synthase GshB	COG0189	H	100.00	100.00
RCA23_c21550	putative esterase/lipase	COG0657	I	100.00	97.46
RCA23_c21560	NusB-like protein NusB	COG0781	K	100.00	100.00
RCA23_c21570	6,7-dimethyl-8-ribityllumazine synthase RibH	COG0054	H	100.00	100.00
RCA23_c21580	riboflavin biosynthesis protein RibB	COG0108	H	100.00	100.00
RCA23_c21590	riboflavin synthase alpha chain RibE	COG0307	H	100.00	89.97
RCA23_c21600	putative capsule polysaccharide export protein	COG3562	M	100.00	98.39
RCA23_c21610	putative polysaccharide export protein	COG1596	M	100.00	89.38
RCA23_c21620	putative capsule polysaccharide export protein	COG3563	M	90.91	100.00
RCA23_c21630	riboflavin biosynthesis protein RibD	COG0117	H	100.00	100.00
RCA23_c21640	transcriptional repressor NrdR	COG1327	K	100.00	100.00
RCA23_c21650	RNA polymerase sigma factor RpoD	COG0568	K	100.00	100.00
RCA23_c21660	DNA primase DnaG	COG0358	L	100.00	100.00
RCA23_c21670	sarcosine oxidase subunit SoxG	COG4583	E	100.00	100.00
RCA23_c21680	sarcosine oxidase alpha subunit SoxA	COG0404	E	100.00	98.16

RCA23_c21690	sarcosine oxidase subunit SoxD	COG4311	E	100.00	100.00
RCA23_c21700	sarcosine oxidase beta subunit SoxB	COG0665	E	100.00	100.00
RCA23_c21710	phosphatidylserine decarboxylase proenzyme Psd	COG0688	I	100.00	100.00
RCA23_c21720	diacylglycerol kinase	COG0818	M	100.00	100.00
RCA23_c21730	putative sulfatase	COG2194	R	100.00	100.00
RCA23_c21740	hypothetical protein, CDP-alcohol phosphatidyltransferase-like	COG1183	I	100.00	100.00
RCA23_c21750	hypothetical protein	COG0671	I	100.00	100.00
RCA23_c21760	hypothetical protein	COG2194	R	100.00	100.00
RCA23_c21770	hypothetical protein, DUF1705	COG2194	R	100.00	100.00
RCA23_c21780	glutamate dehydrogenase GluD	COG0334	E	96.73	99.23
RCA23_c21790	hypothetical protein			100.00	100.00
RCA23_c21800	2-hydroxy-3-oxopropionate reductase GarR	COG2084	I	90.37	100.00
RCA23_c21810	reductive dehalogenase	COG1600	C	26.15	100.00
RCA23_c21820	pirin	COG1741	R	100.00	99.67
RCA23_c21830	transcriptional regulator	COG1309	K	100.00	100.00
RCA23_c21840	hypothetical protein			100.00	100.00
RCA23_c21850	homoserine dehydrogenase Hom	COG0460	E	91.55	100.00
RCA23_c21860	fructose-1,6-bisphosphatase class II	COG1494	G	100.00	100.00
RCA23_c21870	putative single-stranded-DNA-specific exonuclease	COG0608	L	100.00	100.00
RCA23_c21890	sulfopyruvate decarboxylase beta subunit ComE	COG0028	E	100.00	100.00
RCA23_c21900	sulfopyruvate decarboxylase alpha subunit ComD	COG4032	R	100.00	100.00
RCA23_c21910	zinc-type alcohol dehydrogenase	COG1063	E	100.00	100.00
RCA23_c21920	gluconate 5-dehydrogenase Gno	COG1028	I	100.00	100.00
RCA23_c21930	histidinol dehydrogenase HisD	COG0141	E	100.00	100.00
RCA23_c21940	HTH-type transcriptional regulator, LacI family	COG1609	K	100.00	100.00
RCA23_c21950	HTH-type transcriptional regulator, IclR family	COG1414	K	100.00	100.00
RCA23_c21960	alcohol dehydrogenase AlkJ	COG2303	E	100.00	95.01
RCA23_c21970	integral membrane protein TerC	COG0861	P	100.00	86.40
RCA23_c21980	putative TRAP transporter, DctP subunit	COG4663	Q	33.69	100.00
RCA23_c21990	TRAP transporter, DctM/DctQ subunit	COG4664	Q	89.37	100.00
RCA23_c22000	trimethylamine methyltransferase	COG5598	H	99.44	100.00
RCA23_c22010	putative HAD-family hydrolase	COG0546	R	98.57	77.40
RCA23_c22020	hypothetical protein			0.00	100.00

RCA23_c22030	hypothetical protein				0.00	100.00
RCA23_c22040	hypothetical protein				72.85	88.66
RCA23_c22050	hypothetical protein	COG5349	S		100.00	100.00
RCA23_c22060	hypothetical protein, NUDIX hydrolase	COG0494	L		100.00	100.00
RCA23_c22070	hypothetical protein				100.00	100.00
RCA23_c22080	putative fatty acid desaturase	COG3239	I		100.00	100.00
RCA23_c22090	putative helix-turn-helix protein	COG1396	K		100.00	100.00
RCA23_c22100	short chain dehydrogenase	COG1028	I		100.00	100.00
RCA23_c22110	putative 3-hydroxyacyl-CoA dehydrogenase	COG1250	I		100.00	100.00
RCA23_c22130	hypothetical protein, DUF849	COG3246	S		100.00	100.00
RCA23_c22120	transcriptional regulator, AraC family	COG4977	K		100.00	98.15
RCA23_c22140	protein Tas	COG0667	C		100.00	100.00
RCA23_c22150	hypothetical protein	COG4544	S		100.00	100.00
RCA23_c22160	putative MFS-type transporter	COG2211	G		100.00	99.35
RCA23_c22170	methionyl-tRNA synthase MetG	COG0143	J		100.00	100.00
RCA23_c22190	integrase			GI 8	65.69	100.00
RCA23_c22200	integrase	COG4974	L	GI 8	100.00	93.84
RCA23_c22210	putative nucleotidyl transferase	COG1210	M	GI 8	42.77	100.00
RCA23_c22220	hypothetical protein			GI 8	0.00	92.08
RCA23_c22230	hypothetical protein			GI 8	7.79	76.78
RCA23_c22240	putative replication factor C, small subunit	COG2256	L	GI 8	100.00	100.00
RCA23_c22250	hypothetical protein			GI 8	100.00	100.00
RCA23_c22260	hypothetical protein			GI 8	100.00	98.44
RCA23_c22270	ParB-like nuclease	COG1475	K	GI 8	100.00	97.83
RCA23_c22280	cadmium-transporting ATPase CadA	COG2217	P	GI 8	94.20	92.03
RCA23_c22290	hypothetical protein	COG3034	S	GI 8	0.00	100.00
RCA23_c22300	putative lipoprotein signal peptidase	COG0597	M	GI 8	0.00	100.00
RCA23_c22310	putative ZIP zinc transporter	COG0428	P	GI 8	3.41	100.00
RCA23_c22320	hypothetical protein DUF411	COG3019	R	GI 8	100.00	100.00
RCA23_c22330	SCO-like protein	COG1999	R	GI 8	100.00	100.00
RCA23_c22340	hypothetical protein	COG2847	S	GI 8	100.00	100.00
RCA23_c22350	hypothetical protein	COG1651	O	GI 8	100.00	100.00
RCA23_c22360	hypothetical protein, disulfide bond formation protein	COG1495	O	GI 8	100.00	100.00

RCA23_c22370	SCO-like protein	COG1999	R	GI 8	100.00	100.00
RCA23_c22380	putative HTH-type transcriptional regulator	COG0789	K	GI 8	100.00	100.00
RCA23_c22390	hypothetical protein UPF0060	COG1742	S	GI 8	26.61	100.00
RCA23_c22400	hypothetical protein			GI 8	100.00	100.00
RCA23_c22410	hypothetical protein			GI 8	90.61	88.57
RCA23_c22420	transposase	COG3316	L	GI 8	1.98	0.00
RCA23_c22430	MORN motif precursor	COG4642	S	GI 8	92.77	87.96
RCA23_c22440	transposase	COG3316	L	GI 8	87.01	0.00
RCA23_c22450	hypothetical protein			GI 8	100.00	41.98
RCA23_c22460	hypothetical protein			GI 8	0.00	100.00
RCA23_c22470	hypothetical protein	COG3295	S	GI 8	0.00	0.00
RCA23_c22480	hypothetical protein			GI 8	0.00	0.00
RCA23_c22490	hypothetical protein			GI 8	0.00	0.00
RCA23_c22500	transcriptional regulator, LuxR family	COG2197	T	GI 8	0.00	0.00
RCA23_c22510	hypothetical protein	COG3316	L	GI 8	0.00	0.00
RCA23_c22520	hypothetical protein, HTH-type transcriptional regulator, LuxR family	COG5616	S	GI 8	0.00	43.46
RCA23_c22530	cysteine synthase CysK	COG0031	E	GI 8	0.00	23.80
RCA23_c22540	hypothetical protein			GI 8	0.00	80.00
RCA23_c22550	hypothetical protein			GI 8	0.00	0.00
RCA23_c22560	hypothetical protein			GI 8	43.51	100.00
RCA23_c22570	hypothetical protein			GI 8	100.00	98.38
RCA23_c22580	hypothetical protein	COG5470	S	GI 8	100.00	15.00
RCA23_c22590	coniferyl aldehyde dehydrogenase CalB	COG1012	C	GI 8	92.50	91.22
RCA23_c22600	regulatory protein NocR	COG0583	K	GI 8	97.88	47.60
RCA23_c22610	sn-glycerol-3-phosphate-binding periplasmic protein UgpB	COG1653	G	GI 8	100.00	71.23
RCA23_c22620	sn-glycerol-3-phosphate transport system permease protein UgpA	COG1175	G	GI 8	100.00	100.00
RCA23_c22630	sn-glycerol-3-phosphate transport system permease protein UgpE	COG0395	G	GI 8	100.00	60.57
RCA23_c22640	sn-glycerol-3-phosphate import ATP-bindingprotein UgpC	COG3839	G	GI 8	100.00	30.18
RCA23_c22650	putative glycerophosphoryl diester phosphodiesterase	COG0584	C	GI 8	86.31	77.10
RCA23_c22660	aerobic glycerol-3-phosphate dehydrogenase GlpD	COG0578	C	GI 8	100.00	36.21
RCA23_c22670	hypothetical protein, Na+/Pi-cotransporter	COG1283	P	GI 8	90.48	16.94
RCA23_c22680	hypothetical protein, calcineurin-like phosphoesterase-like	COG1409	R	GI 8	100.00	27.78
RCA23_c22690	ABC transporter permease protein	COG0395	G	GI 8	95.06	30.25

RCA23_c22700	ABC transporter permease protein	COG1175	G	GI 8	21.59	0.00
RCA23_c22710	ABC transporter extracellular solute-binding protein	COG1653	G	GI 8	100.00	0.00
RCA23_c22720	ABC transporter ATP-binding protein	COG3839	G	GI 8	100.00	0.00
RCA23_c22730	HTH-type transcriptional regulator, DeoR family	COG1349	K	GI 8	100.00	28.92
RCA23_c22740	HAD-superfamily hydrolase, subfamily IIB	COG0561	R	GI 8	100.00	33.58
RCA23_c22750	putative inner membrane transporter	COG0697	G	GI 8	100.00	100.00
RCA23_c22760	soluble aldose sugar dehydrogenase YliI	COG2133	G	GI 8	100.00	97.17
RCA23_c22770	putative FKBP-type peptidyl-prolyl cis-trans isomerase	COG0545	O	GI 8	75.38	43.59
RCA23_c22780	putative NnrU family protein	COG4094	S	GI 8	81.42	81.42
RCA23_c22790	hypothetical protein			GI 8	98.28	44.44
RCA23_c22800	putative adenine methyltransferase	COG0863	L	GI 8	52.56	49.53
RCA23_c22810	hypothetical protein			GI 8	100.00	0.00
RCA23_c22820	phage uncharacterised protein	COG5410	S	GI 8	40.93	0.00
RCA23_c22830	hypothetical protein, resolvase-like	COG1961	L	GI 8	0.00	0.00
RCA23_c22840	putative prophage integrase	COG0582	L	GI 8	0.00	0.00
RCA23_c22850	hypothetical protein, putative phage-like protein			GI 8	0.00	0.00
RCA23_c22860	hypothetical protein	COG3311	K	GI 8	0.00	0.00
RCA23_c22870	hypothetical protein			GI 8	0.00	0.00
RCA23_c22880	hypothetical protein			GI 8	54.14	46.50
RCA23_c22890	DNA polymerase III	COG5545	R	GI 8	100.00	89.39
RCA23_c22900	hypothetical protein, periplasmic binding protein-like	COG2998	H	GI 8	0.00	0.00
RCA23_c22910	hypothetical protein			GI 8	0.00	81.13
RCA23_c22920	transposase	COG3316	L	GI 8	0.00	5.37
RCA23_c22930	molybdopterin biosynthesis protein MoeA	COG0303	H	GI 8	0.00	75.38
RCA23_c22940	molybdopterin-guanine dinucleotide biosynthesis protein MobB	COG1763	H	GI 8	0.00	97.39
RCA23_c22950	hypothetical protein, MobA-like	COG0746	H	GI 8	0.00	75.08
RCA23_c22960	formate dehydrogenase family accessory protein FdhD	COG1526	C	GI 8	0.00	96.01
RCA23_c22970	molybdopterin-converting factor subunit MoaE	COG0314	H	GI 8	0.00	98.42
RCA23_c22980	molybdopterin-converting factor subunit MoaD	COG1977	H	GI 8	0.00	92.68
RCA23_c22990	transposase	COG3328	L	GI 8	0.00	100.00
RCA23_c23000	molybdenum cofactor biosynthesis protein MoaA	COG2896	H	GI 8	0.00	76.98
RCA23_c23010	transposase			GI 8	0.00	89.01
RCA23_c23020	ABC transporter, ATP-binding cassette protein	COG3839	G	GI 8	0.00	76.75

RCA23_c23030	ABC transporter, permease protein	COG4662	H	GI 8	0.00	68.37
RCA23_c23040	formate dehydrogenase, gamma subunit FdhI	COG2864	C	GI 8	0.00	97.17
RCA23_c23050	formate dehydrogenase, iron-sulfur subunit FdhB	COG0437	C	GI 8	0.00	100.00
RCA23_c23060	formate dehydrogenase alpha subunit FdhA	COG0243	C	GI 8	7.36	98.65
RCA23_c23070	putative twin-arginine translocation pathway signal sequence domain			GI 8	100.00	93.23
RCA23_c23080	hypothetical protein, cytoplasmic chaperon TorD	COG3381	R	GI 8	100.00	43.89
RCA23_c23090	hypothetical protein			GI 8	100.00	81.11
RCA23_c23100	putative molybdopterin-guanine dinucleotide biosynthesis protein A			GI 8	47.65	51.60
RCA23_c23110	hypothetical protein, 4Fe-4S ferredoxin-like	COG1148	C	GI 8	53.82	95.73
RCA23_c23120	hypothetical protein			GI 8	0.00	82.64
RCA23_c23130	hypothetical protein			GI 8	0.00	100.00
RCA23_c23140	hypothetical protein, Mrp/NBP35 family protein	COG0489	D	GI 8	0.00	56.63
RCA23_c23150	hypothetical protein	COG0340	H	GI 8	0.00	100.00
RCA23_c23160	hypothetical protein			GI 8	0.00	100.00
RCA23_c23170	putative phage terminase			GI 8	0.00	0.00
RCA23_c23180	transposase	COG3316	L	GI 8	0.00	0.00
RCA23_c23190	putative phage helicase	COG3378	R	GI 8	0.00	58.97
RCA23_c23200	hypothetical protein			GI 8	0.00	100.00
RCA23_c23210	putative pyridoxal 4-dehydrogenase	COG0667	C	GI 8	0.00	0.00
RCA23_c23220	2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase KduD	COG1028	I	GI 8	0.00	0.00
RCA23_c23230	putative L-rhamnonate dehydratase rhamD	COG4948	M	GI 8	0.00	0.00
RCA23_c23240	ureidoglycolate lyase	COG0179	Q	GI 8	0.00	0.00
RCA23_c23250	D-3-phosphoglycerate dehydrogenase	COG1052	C	GI 8	0.00	0.00
RCA23_c23260	uncharacterized oxidoreductase	COG1028	I	GI 8	0.00	0.00
RCA23_c23270	putative rhamnose mutarotase RhaM	COG3254	S	GI 8	0.00	0.00
RCA23_c23280	putative acetoacetate decarboxylase Adc	COG4689	Q	GI 8	0.00	0.00
RCA23_c23290	uncharacterized oxidoreductase	COG0673	R	GI 8	0.00	0.00
RCA23_c23300	transcriptional regulator, GntR family	COG2186	K	GI 8	0.00	0.00
RCA23_c23310	putative ribose ABC transport system, permease protein RbsC	COG1172	G	GI 8	0.00	0.00
RCA23_c23320	putative ribose ABC transport system, permease protein RbsC	COG1172	G	GI 8	0.00	0.00
RCA23_c23330	ribose import ABC transporter, ATP-binding protein RbsA	COG1129	G	GI 8	0.00	0.00
RCA23_c23340	putative rhamnose ABC transport system, substrate-binding protein RhaS	COG1879	G	GI 8	0.00	0.00
RCA23_c23350	galactonate dehydratase	COG4948	M	GI 8	0.00	0.00

RCA23_c23360	xylose isomerase-like	COG1082	G	GI 8	0.00	0.00
RCA23_c23370	putative aldo/keto reductase	COG0667	C	GI 8	0.00	0.00
RCA23_c23380	amidohydrolase	COG3618	R	GI 8	0.00	0.00
RCA23_c23390	hypothetical protein			GI 8	0.00	0.00
RCA23_c23400	transposase	COG3316	L	GI 8	0.00	5.79
RCA23_c23410	L-rhamnose mutarotase RhaM	COG3254	S	GI 8	0.00	31.45
RCA23_c23420	ureidoglycolate lyase	COG0179	Q	GI 8	0.00	27.06
RCA23_c23430	dihydrodipicolinate synthase DapA	COG0329	E	GI 8	0.00	0.00
RCA23_c23440	amidohydrolase	COG3618	R	GI 8	0.00	0.00
RCA23_c23450	3-ketoacyl-(acyl-carrier-protein) reductase FabG	COG1028	I	GI 8	0.00	0.00
RCA23_c23460	L-rhamnonate dehydratase RhmD	COG4948	M	GI 8	0.00	0.00
RCA23_c23470	polyamine ABC transporter, ATP-binding protein PotA	COG3842	E	GI 8	0.00	0.00
RCA23_c23480	polyamine ABC transporter, permease protein	COG1177	E	GI 8	0.00	0.00
RCA23_c23490	polyamine ABC transporter, permease protein	COG1176	E	GI 8	0.00	0.00
RCA23_c23500	polyamine ABC transporter, substrate binding protein	COG0687	E	GI 8	0.00	0.00
RCA23_c23510	hypothetical protein			GI 8	0.00	0.00
RCA23_c23520	hypothetical protein			GI 8	0.00	0.00
RCA23_c23530	uncharacterized oxidoreductase	COG1028	I	GI 8	55.29	0.00
RCA23_c23540	L-rhamnonate dehydratase	COG4948	M	GI 8	100.00	7.73
RCA23_c23550	ribose ABC transport system, permease protein RbsC	COG1172	G	GI 8	9.20	0.00
RCA23_c23560	ribose import ABC transporter, ATP-binding protein RbsA	COG1129	G	GI 8	0.00	29.44
RCA23_c23570	ribose import ABC transporter, substrate binding protein RbsC	COG1879	G	GI 8	1.80	0.00
RCA23_c23580	HTH-type transcriptional regulator, GntR family	COG1609	K	GI 8	100.00	0.00
RCA23_c23590	hypothetical protein	COG3311	K	GI 8	100.00	0.00
RCA23_c23600	putative DNA-binding protein	COG1396	K	GI 8	0.00	100.00
RCA23_c23610	glutamine amidotransferase-like protein GlxB	COG0067	E	GI 8	0.00	97.88
RCA23_c23620	glutamate synthase alpha subunit GlxC	COG0070	E	GI 8	0.00	80.68
RCA23_c23630	glutamate synthase large subunit GlxD	COG0069	E	GI 8	0.00	96.35
RCA23_c23640	glutamine synthase GlnA type III	COG0174	E	GI 8	0.00	75.58
RCA23_c23650	hypothetical protein			GI 8	0.00	0.00
RCA23_c23660	methyltrasnferase	COG0863	L	GI 8	42.40	20.00
RCA23_c23670	hypothetical protein			GI 8	100.00	4.43
RCA23_c23680	putative terminase, large subunit	COG5410	S	GI 8	94.28	41.63

RCA23_c23690	hypothetical protein, DUF2924			GI 8	0.00	0.00
RCA23_c23700	hypothetical protein, resolvase-like	COG1961	L	GI 8	9.15	0.00
RCA23_c23720	signal transduction histidine kinase	COG0642	T		100.00	97.31
RCA23_c23730	mercuric reductase MerA	COG1249	C		100.00	99.22
RCA23_c23740	hypothetical protein	COG0398	S		100.00	93.31
RCA23_c23750	putative ribonuclease P	COG0594	J		100.00	100.00
RCA23_c23760	hypothetical protein DUF37	COG0759	S		100.00	100.00
RCA23_c23770	hypothetical protein				100.00	100.00
RCA23_c23780	tRNA 2-thiocytidine biosynthesis protein TtcA	COG0037	D		100.00	100.00
RCA23_c23790	putative diguanylate phosphodiesterase	COG2200	T		100.00	100.00
RCA23_c23800	inner membrane protein OxaA	COG0706	U		100.00	99.50
RCA23_c23810	hypothetical protein, molybdenum cofactor sulfurase	COG3217	R		100.00	100.00
RCA23_c23820	putative GTP-binding protein EngB	COG0218	R		100.00	100.00
RCA23_c23830	acetylglutamate kinase ArgB	COG0548	E		100.00	100.00
RCA23_c23840	putative fatty acid hydroxylase	COG3000	I		100.00	97.35
RCA23_c23850	hypothetical protein	COG1148	C		100.00	94.52
RCA23_c23860	hypothetical protein, phosphogylcerate mutase	COG2062	T		100.00	100.00
RCA23_c23870	glutamate/glutamine/aspartate/asparagine transport ATP-binding protein Bzt	COG1126	E		100.00	100.00
RCA23_c23880	glutamate/glutamine/aspartate/asparagine transport system permease prote	COG0765	E		100.00	100.00
RCA23_c23890	glutamate/glutamine/aspartate/asparagine transport system permease prote	COG4597	E		62.65	78.89
RCA23_c23900	glutamate/glutamine/aspartate/asparagine-binding protein BztA	COG0834	E		100.00	98.19
RCA23_c23910	ATP chaperone protein	COG5387	O		93.48	100.00
RCA23_c23920	putative phosphoglycerate phosphatase	COG0546	R		100.00	100.00
RCA23_c23930	ribosomal large subunit pseudouridine synthase C	COG0564	J		100.00	100.00
RCA23_c23940	putative ccrb-like protein	COG0239	D		100.00	100.00
RCA23_c23950	replication-associated recombination protein A	COG2256	L		100.00	100.00
RCA23_c23960	HTH-type transcriptional regulator, LuxR family	COG2197	T		100.00	100.00
RCA23_c23970	50S ribosomal protein L17	COG0203	J		100.00	100.00
RCA23_c23980	DNA-directed RNA polymerase alpha subunit RpoA	COG0202	K		100.00	100.00
RCA23_c23990	30S ribosomal protein S11	COG0100	J		100.00	100.00
RCA23_c24000	30S ribosomal protein S13	COG0099	J		100.00	100.00
RCA23_c24010	adenylate kinase Adk	COG0563	F		100.00	99.38
RCA23_c24020	preprotein translocase subunit SecY	COG0201	U		100.00	100.00

RCA23_c24030	50S ribosomal protein L15	COG0200	J	100.00	100.00
RCA23_c24040	50S ribosomal protein L30	COG1841	J	100.00	100.00
RCA23_c24050	30S ribosomal protein S5	COG0098	J	100.00	100.00
RCA23_c24060	50S ribosomal protein L18	COG0256	J	100.00	100.00
RCA23_c24070	50S ribosomal protein L6	COG0097	J	100.00	100.00
RCA23_c24080	30S ribosomal protein S8	COG0096	J	100.00	100.00
RCA23_c24090	30S ribosomal protein S14	COG0199	J	100.00	100.00
RCA23_c24100	50S ribosomal protein L5	COG0094	J	100.00	100.00
RCA23_c24110	50S ribosomal protein L24	COG0198	J	100.00	100.00
RCA23_c24120	50S ribosomal protein L14	COG0093	J	100.00	100.00
RCA23_c24130	30S ribosomal protein S17	COG0186	J	100.00	100.00
RCA23_c24140	50S ribosomal protein L29	COG0255	J	70.53	100.00
RCA23_c24150	hypothetical protein			71.98	95.65
RCA23_c24160	50S ribosomal protein L23	COG0089	J	100.00	100.00
RCA23_c24170	50S ribosomal protein L4	COG0088	J	100.00	100.00
RCA23_c24180	50S ribosomal protein L3	COG0087	J	100.00	100.00
RCA23_c24190	30S ribosomal protein S10	COG0051	J	100.00	100.00
RCA23_c24200	elongation factor Tu (EF-Tu)	COG0050	J	100.00	31.12
RCA23_c24210	elongation factor FusA	COG0480	J	100.00	100.00
RCA23_c24220	30S ribosomal protein S7	COG0049	J	100.00	100.00
RCA23_c24230	30S ribosomal protein S12	COG0048	J	100.00	100.00
RCA23_c24240	hypothetical protein	COG1072	H	100.00	51.54
RCA23_c24250	ABC transporter ATP-binding protein	COG1129	G	100.00	94.47
RCA23_c24260	ABC tansporter permease protein	COG1172	G	100.00	100.00
RCA23_c24270	putative ABC transporter periplasmic binding protein	COG1879	G	100.00	96.82
RCA23_c24280	putative transcriptional repressor	COG1940	K	100.00	86.31
RCA23_c24290	hypothetical protein, acyl-CoA thioesterase-like	COG2755	E	100.00	100.00
RCA23_c24300	ABC transporter ATP-binding protein	COG1136	V	100.00	100.00
RCA23_c24310	putative ABC transporter permease protein	COG3127	Q	100.00	100.00
RCA23_c24320	hypothetical protein, gylcosyl transferase family 2	COG0463	M	95.78	100.00
RCA23_c24330	aminotransferase class-III	COG0161	H	100.00	100.00
RCA23_c24340	putative ectoine utilization protein EutD	COG0006	E	100.00	100.00
RCA23_c24350	putative integral membrane proein DUF6			100.00	100.00

RCA23_c24360	4-coumarate--CoA ligase	COG0318	I		100.00	100.00
RCA23_c24370	hypothetical protein	COG0679	R		100.00	100.00
RCA23_c24380	hypothetical protein DUF125	COG1814	S		100.00	100.00
RCA23_c24390	transcriptional regulator, AsnC family	COG1522	K		100.00	86.95
RCA23_c24400	arginase ArcA	COG0010	E		100.00	100.00
RCA23_c24410	ornithine cyclodeaminase ArcB	COG2423	E		100.00	100.00
RCA23_c24420	bifunctional protein PutA	COG4230	C		100.00	100.00
RCA23_c24430	hypothetical protein				0.00	100.00
RCA23_c24440	hypothetical protein				0.00	100.00
RCA23_c24460	hypothetical protein				100.00	100.00
RCA23_c24450	creatinase	COG0006	E		100.00	100.00
RCA23_c24470	putative hydrolase	COG0596	R		95.99	0.00
RCA23_c24480	glycine betaine/L-proline transport system permease protein ProW	COG4176	E		100.00	46.78
RCA23_c24490	glycine betaine/L-proline transport ATP-binding protein ProV	COG4175	E		19.41	5.59
RCA23_c24500	glycine betaine-binding periplasmic protein ProX	COG2113	E		13.46	13.87
RCA23_c24510	putative HTH-type transcriptional regluator, LysR family	COG0583	K		100.00	85.20
RCA23_c24520	D-cysteine desulfhydrase DcyD	COG2515	E		100.00	100.00
RCA23_c24530	hypothetical protein				100.00	86.27
RCA23_c24540	high-affinity zinc uptake system protein ZnuA	COG4531	P		100.00	100.00
RCA23_c24550	zinc uptake regulator	COG0735	P		100.00	100.00
RCA23_c24560	zinc import ATP-binding protein ZnuC	COG1121	P		100.00	100.00
RCA23_c24570	high-affinity zinc uptake system membrane protein ZnuB	COG1108	P		100.00	100.00
RCA23_c24580	4-hydroxyphenylpyruvate dioxygenase Hpd	COG3185	E		100.00	97.53
RCA23_c24590	hypothetical protein			GI 9	0.00	83.48
RCA23_c24600	hypothetical protein			GI 9	0.00	64.23
RCA23_c24610	hypothetical protein	COG1028	I	GI 9	21.14	100.00
RCA23_c24620	hypothetical protein	COG1402	R	GI 9	100.00	77.28
RCA23_c24630	putative gluconolactonase	COG3386	G	GI 9	100.00	95.04
RCA23_c24640	ribokinase RbsK	COG0524	G	GI 9	100.00	100.00
RCA23_c24650	ribose ABC transporter protein RbsD	COG4154	G	GI 9	100.00	79.69
RCA23_c24660	putative ribose ABC transporter, ATP-binding protein RbsA	COG3839	G	GI 9	100.00	96.72
RCA23_c24670	putative ribose ABC transporter, substrate binding protein RbsB	COG1653	G	GI 9	100.00	92.16
RCA23_c24680	putative ribose ABC transporter, permease protein RbsC	COG1175	G	GI 9	100.00	100.00

RCA23_c24690	putative ribose ABC transporter, permease protein RbsC	COG0395	G	GI 9	100.00	80.61
RCA23_c24700	hypothetical protein, L-fucose isomerase-like			GI 9	100.00	99.66
RCA23_c24710	sugar isomerase	COG0794	M	GI 9	100.00	81.72
RCA23_c24720	xylulose kinase XylB	COG1070	G	GI 9	24.61	100.00
RCA23_c24730	aminopeptidase SgcX	COG1363	G	GI 9	0.00	86.83
RCA23_c24740	photosystem I biogenesis protein BtpA	COG0434	R	GI 9	0.00	100.00
RCA23_c24750	pyridoxal 4-dehydrogenase Pld	COG0667	C	GI 9	17.69	100.00
RCA23_c24760	putative amidohydrolase	COG3618	R	GI 9	100.00	100.00
RCA23_c24770	putative carbohydrate kinase, pfkB family	COG0524	G	GI 9	100.00	100.00
RCA23_c24780	xylose isomerase family protein	COG4952	M	GI 9	100.00	98.89
RCA23_c24790	putative HTH-type transcriptional regulator, GntR family	COG1802	K	GI 9	100.00	100.00
RCA23_c24800	mandelate racemase/muconate lactonizing enzyme	COG4948	M	GI 9	100.00	100.00
RCA23_c24810	fatty acid desaturase	COG3239	I	GI 9	0.00	100.00
RCA23_c24820	hypothetical protein			GI 9	0.00	41.32
RCA23_c24830	transposase	COG3316	L	GI 9	0.00	0.00
RCA23_c24840	transporter, LysE family	COG1280	E	GI 9	0.00	0.00
RCA23_c24850	hypothetical protein			GI 9	0.00	68.85
RCA23_c24860	putative transporter, periplasmic binding protein	COG2358	R	GI 9	0.00	0.00
RCA23_c24870	ribitol 2-dehydrogenase RbtD	COG4221	R	GI 9	0.00	0.00
RCA23_c24880	hypothetical protein			GI 9	0.00	16.00
RCA23_c24890	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa s	COG0457	R	GI 9	0.00	24.22
RCA23_c24900	putative phage integrase	COG4974	L	GI 9	100.00	50.82
RCA23_c24920	D-alanyl-D-alanine carboxypeptidase DacC	COG1686	M		69.53	100.00
RCA23_c24930	thymidylate kinase Tmk	COG0125	F		100.00	98.54
RCA23_c24940	hypothetical protein, DNA polymerase III delta subunit	COG2812	L		100.00	100.00
RCA23_c24950	TatD family deoxyribonuclease	COG0084	L		100.00	100.00
RCA23_c24960	hypothetical protein, metallo-beta-lactamase	COG1235	R		100.00	100.00
RCA23_c24970	membrane transport protein	COG0679	R		100.00	88.92
RCA23_c24980	FAD dependent oxidoreductase	COG0665	E		100.00	100.00
RCA23_c24990	UvrABC system protein B	COG0556	L		100.00	94.66
RCA23_c25000	hypothetical protein				100.00	89.12
RCA23_c25010	hypothetical protein				100.00	72.28
RCA23_c25020	hypothetical protein				100.00	100.00

RCA23_c25030	lysyl-tRNA synthase LysS	COG1384	J	100.00	93.21
RCA23_c25040	hypothetical protein			100.00	100.00
RCA23_c25050	putative D-Ala-D-Ala carboxypeptidase 3 (S13) family protein	COG2027	M	100.00	100.00
RCA23_c25060	hypothetical protein			100.00	100.00
RCA23_c25070	nicotinic acid mononucleotide adenyltransferase NadD	COG1057	H	100.00	100.00
RCA23_c25080	ABC transporter, ATP-binding cassette protein, ChvD family	COG0488	R	100.00	100.00
RCA23_c25090	cold shock protein CspA	COG1278	K	100.00	100.00
RCA23_c25100	glutamyl-tRNA synthase 1	COG0008	J	100.00	100.00
RCA23_c25110	glutamine-dependent NAD(+) synthase NadE	COG0171	H	100.00	98.47
RCA23_c25120	putative phosphatidylinositol-4-phosphate 5-kinase	COG4642	S	69.37	100.00
RCA23_c25130	hypothetical protein	COG1028	I	100.00	100.00
RCA23_c25140	2-isopropylmalate synthase LeuA	COG0119	E	100.00	100.00
RCA23_c25150	rod shape-determining protein MreB	COG1077	D	100.00	99.33
RCA23_c25160	rod shape-determining protein MreC	COG1792	M	100.00	100.00
RCA23_c25170	hypothetical protein			100.00	96.48
RCA23_c25180	penicillin-binding protein 2	COG0768	M	100.00	100.00
RCA23_c25190	rod shape-determining protein RodA	COG0772	D	100.00	100.00
RCA23_c25200	glyoxylate/hydroxypyruvate reductase GhrA	COG0111	H	100.00	100.00
RCA23_c25210	hypothetical protein			100.00	100.00
RCA23_c25220	hypothetical protein	COG1573	L	100.00	100.00
RCA23_c25230	Na+/H+ antiporter NhaA	COG3004	P	100.00	100.00
RCA23_c25240	capsule polysaccharide export protein KpsT	COG1134	G	100.00	96.36
RCA23_c25250	capsule polysaccharide export protein KpsE	COG3524	M	100.00	100.00
RCA23_c25270	hypothetical protein	COG1832	R	100.00	100.00
RCA23_c25280	ferredoxin PetF	COG0633	C	100.00	100.00
RCA23_c25290	selenium binding protein			100.00	100.00
RCA23_c25300	aldehyde dehydrogenase	COG1012	C	100.00	100.00
RCA23_c25310	CoA-transferase family III protein involved in DMSP degradation	COG1804	C	100.00	100.00
RCA23_c25320	nitrogen regulatory protein P-II 2	COG0347	E	100.00	100.00
RCA23_c25330	ammonium transporter	COG0004	P	100.00	100.00
RCA23_c25340	homoprotocatechuate degradative operon repressor	COG1846	K	100.00	100.00
RCA23_c25350	5-carboxymethyl-2-hydroxymuconic semialdehyde dehydrogenase HpaE	COG1012	C	100.00	100.00
RCA23_c25360	3,4-dihydroxyphenylacetate 2,3-dioxygenase HpaD	COG0346	E	100.00	100.00

RCA23_c25370	3-carboxy-cis,cis-muconate cycloisomerase PcaB	COG0015	F	100.00	100.00
RCA23_c25380	hypothetical protein DUF849	COG3246	S	100.00	100.00
RCA23_c25390	protocatechuate 3,4-dioxygenase alpha chain PcaG	COG3485	Q	100.00	100.00
RCA23_c25400	protocatechuate 3,4-dioxygenase beta chain PcaH	COG3485	Q	100.00	100.00
RCA23_c25410	4-carboxymuconolactone decarboxylase PcaC	COG0599	S	100.00	100.00
RCA23_c25420	hypothetical protein			100.00	100.00
RCA23_c25430	transcriptional regulatory protein	COG0745	T	100.00	99.42
RCA23_c25440	two component signal transduction histidine kinase ChvG	COG0642	T	100.00	100.00
RCA23_c25450	hypothetical protein			54.01	100.00
RCA23_c25460	Na(+)-phosphate symporter Pit	COG0306	P	100.00	100.00
RCA23_c25470	hypothetical protein			100.00	100.00
RCA23_c25480	2-dehydro-3-deoxyphosphooctonate aldolase KdsA	COG2877	M	100.00	100.00
RCA23_c25490	arabinose 5-phosphate isomerase KdsD	COG0794	M	100.00	95.21
RCA23_c25500	3-deoxy-manno-octulosonate cytidylyltransferase KdsB	COG1212	M	100.00	100.00
RCA23_c25510	hypothetical protein			51.69	66.95
RCA23_c25520	bacterial surface protein			19.09	34.25
RCA23_c25530	catalase-peroxidase KatG	COG0376	P	100.00	100.00
RCA23_c25540	hypothetical protein	COG2067	I	100.00	74.54
RCA23_c25550	thioesterase/thiol ester dehydrase-isomerase	COG0824	R	100.00	73.66
RCA23_c25560	iron-containing alcohol dehydrogenase	COG1454	C	100.00	100.00
RCA23_c25570	hypothetical protein	COG1960	I	100.00	100.00
RCA23_c25580	nodulation protein N-like protein	COG2030	I	100.00	97.22
RCA23_c25590	putative phosphotransferase, eukaryotic acyl-CoA dehydrogenase	COG3173	R	100.00	94.81
RCA23_c25600	putative HTH-type transcriptional regulator, TetR family	COG1309	K	100.00	100.00
RCA23_c25610	fatty acid oxidation complex alpha subunit FadJ	COG1250	I	100.00	100.00
RCA23_c25620	acyl-CoA dehydrogenase	COG1960	I	100.00	100.00
RCA23_c25630	acyl-CoA dehydrogenase	COG1960	I	100.00	100.00
RCA23_c25640	hypothetical protein, MarR family			100.00	95.44
RCA23_c25650	acyl-CoA dehydrogenase	COG1960	I	100.00	100.00
RCA23_c25660	quinone oxidoreductase (1.6.5.5.)	COG0604	C	100.00	100.00
RCA23_c25670	hypothetical protein			100.00	100.00
RCA23_c25680	putative DNA repair protein RecO	COG1381	L	100.00	92.32
RCA23_c25690	hypothetical protein	COG5447	S	100.00	100.00

RCA23_c25700	GTP-binding protein Era	COG1159	R	100.00	100.00
RCA23_c25710	ribonuclease 3	COG0571	K	100.00	100.00
RCA23_c25720	signal peptidase I	COG0681	U	100.00	100.00
RCA23_c25730	holo-[acyl-carrier-protein] synthase AcpS	COG0736	I	100.00	100.00
RCA23_c25740	pyridoxine 5'-phosphate synthase PdxJ	COG0854	H	89.25	100.00
RCA23_c25750	hypothetical protein DUF2062	COG3216	S	0.00	100.00
RCA23_c25760	guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase SpoT	COG0317	T	97.23	96.31
RCA23_c25770	DNA-directed RNA polymerase subunit omega	COG1758	K	100.00	100.00
RCA23_c25780	putative 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphoki	COG0801	H	100.00	96.41
RCA23_c25790	hypothetical protein DUF88	COG1432	S	100.00	100.00
RCA23_c25800	4-hydroxy-3-methylbut-2-enyl diphosphate reductase IspH	COG0761	I	100.00	100.00
RCA23_c25810	hypothetical protein, methyltransferase	COG2227	H	100.00	100.00
RCA23_c25820	ribonuclease H	COG0328	L	100.00	100.00
RCA23_c25830	hypothetical protein			100.00	100.00
RCA23_c25840	methionyl-tRNA formyltransferase Fmt	COG0223	J	100.00	100.00
RCA23_c25850	peptide deformylase Def	COG0242	J	100.00	100.00
RCA23_c25860	peptide deformylase Def	COG0242	J	100.00	97.46
RCA23_c25870	aminotransferase	COG1168	E	100.00	95.20
RCA23_c25880	hypothetical protein			100.00	100.00
RCA23_c25890	precorrin-4 C(11)-methyltransferase CobM	COG2875	H	100.00	100.00
RCA23_c25900	precorrin-3B C(17)-methyltransferase CobJ	COG1010	H	100.00	100.00
RCA23_c25910	precorrin-2 C(20)-methyltransferase CobI	COG2243	H	100.00	100.00
RCA23_c25920	precorrin-6Y C(5,15)-methyltransferase CobL	COG2242	H	100.00	100.00
RCA23_c25930	precorrin-8X methylmutase CobH	COG2082	H	100.00	100.00
RCA23_c25940	sirohydrochlorin cobaltochelatase CbiX	COG2138	S	100.00	100.00
RCA23_c25950	methyltransferase, FkbM family	COG3774	M	45.28	7.19
RCA23_c25960	hypothetical protein, glutathione S-transferase	COG0435	O	100.00	100.00
RCA23_c25970	hypothetical protein DUF1636	COG5469	S	100.00	100.00
RCA23_c25980	threonine-phosphate decarboxylase CobC	COG0079	E	100.00	100.00
RCA23_c25990	cobalamin biosynthesis protein CobD	COG1270	H	100.00	100.00
RCA23_c26000	putative peptidoglycan-binding lytic murein transglycosylase	COG2951	M	100.00	100.00
RCA23_c26010	putative chromosome partition protein smc	COG1196	D	100.00	98.50
RCA23_c26020	hypothetical protein			100.00	100.00

RCA23_c26030	hypothetical protein	COG1028	I		100.00	95.56
RCA23_c26040	hypothetical protein DUF81	COG0730	R		100.00	100.00
RCA23_c26050	50S ribosomal protein L2	COG0090	J		100.00	98.31
RCA23_c26060	30S ribosomal protein S19	COG0185	J		100.00	100.00
RCA23_c26070	50S ribosomal protein L22	COG0091	J		100.00	100.00
RCA23_c26080	30S ribosomal protein S3	COG0092	J		100.00	100.00
RCA23_c26090	50S ribosomal protein L16	COG0197	J		100.00	100.00
RCA23_c26100	hypothetical protein			GI 10	100.00	100.00
RCA23_c26120	hypothetcial protein	COG0666	R	GI 10	0.00	0.00
RCA23_c26130	transcriptional regulator			GI 10	0.00	61.06
RCA23_c26140	hypothetical protein			GI 10	0.00	54.41
RCA23_c26150	hypothetical protein, UPF0311			GI 10	52.81	100.00
RCA23_c26160	transcriptional regulator, MarR family	COG1846	K	GI 10	100.00	100.00
RCA23_c26170	long-chain-fatty-acid-CoA-ligase	COG0318	I	GI 10	100.00	100.00
RCA23_c26180	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	GI 10	100.00	89.01
RCA23_c26190	TRAP dicarboxylate transporter, subunit DctQ			GI 10	100.00	55.75
RCA23_c26200	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	GI 10	100.00	98.36
RCA23_c26210	3-hydroxybutyryl-CoA dehydratase Crt	COG1024	I	GI 10	100.00	92.11
RCA23_c26220	metal-dependent hydrolase	COG2159	R	GI 10	100.00	94.61
RCA23_c26230	FAD dependent monooxygenase	COG0654	H	GI 10	100.00	85.76
RCA23_c26240	arylsulfatase	COG3119	P	GI 10	100.00	82.21
RCA23_c26250	transcriptional regulator, MarR family	COG1846	K	GI 10	100.00	58.78
RCA23_c26260	putative transporter, periplasmic binding protein	COG2358	R	GI 10	100.00	86.08
RCA23_c26270	TRAP transporter, 4TM/12TM fusion protein	COG4666	R	GI 10	100.00	70.89
RCA23_c26280	hypothetical protein, UPF0261	COG5441	S	GI 10	100.00	88.29
RCA23_c26290	gamma-glutamylputrescine synthase PuuA	COG0174	E	GI 10	100.00	100.00
RCA23_c26300	cytochrome P450	COG2124	Q	GI 10	77.02	64.97
RCA23_c26310	glutamine amidotransferase class-I	COG0518	F	GI 10	0.00	90.85
RCA23_c26320	tripartite tricarboxylate transporter (TTT) protein TctC	COG3181	S	GI 10	41.28	95.46
RCA23_c26330	tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S	GI 10	47.59	100.00
RCA23_c26340	hypothetical protein, transmembrane			GI 10	0.00	64.08
RCA23_c26350	gamma-glutamyl-gamma-aminobutyraldehyde dehydrogenase PuuC	COG1012	C	GI 10	51.89	100.00
RCA23_c26360	hypothetical protein			GI 10	0.00	0.00

RCA23_c26370	hypothetical protein			GI 10	0.00	0.00
RCA23_c26380	hypothetical protein			GI 10	0.00	0.00
RCA23_c26390	putative fucosyltransferase			GI 10	0.00	0.00
RCA23_c26400	integrase	COG2801	L	GI 10	0.00	91.19
RCA23_c26410	putative extracellular solute-binding protein	COG1840	P	GI 10	96.80	37.21
RCA23_c26420	two-component system, sensor histidine kinase protein	COG0642	T	GI 10	100.00	32.90
RCA23_c26430	two-component system, response regulator protein	COG0745	T	GI 10	100.00	47.68
RCA23_c26440	tripartite tricarboxylate transporter (TTT) protein TctC	COG3181	S	GI 10	80.50	100.00
RCA23_c26450	tripartite tricarboxylate transporter (TTT) protein TctB			GI 10	100.00	100.00
RCA23_c26460	tripartite tricarboxylate transporter (TTT) protein TctA	COG3333	S	GI 10	100.00	99.07
RCA23_c26470	hypothetical protein			GI 10	100.00	100.00
RCA23_c26480	hypothetical protein			GI 10	100.00	100.00
RCA23_c26490	transposase A			GI 10	7.56	86.83
RCA23_c26510	hypothetical protein			GI 10	0.00	0.00
RCA23_c26520	hypothetical protein	COG1709	K	GI 10	0.00	0.00
RCA23_c26530	hypothetical protein			GI 10	0.00	0.00
RCA23_c26540	putative RTX toxin and hemolysin-type calcium binding protein			GI 10	0.00	0.00
RCA23_c26550	putative RTX toxin and hemolysin-type calcium binding protein			GI 10	7.47	5.66
RCA23_c26560	type I secretion outer membrane protein, TolC family	COG1538	M	GI 10	94.05	90.27
RCA23_c26570	type I secretion system ATP-binding component	COG2274	V	GI 10	84.42	96.96
RCA23_c26580	type I RTX secretion system membrane fusion protein, HlyD family	COG1566	V	GI 10	100.00	100.00
RCA23_c26590	hypothetical protein			GI 10	100.00	100.00
RCA23_c26600	hypothetical protein			GI 10	100.00	99.21
RCA23_c26610	hypothetical protein			GI 10	100.00	100.00
RCA23_c26620	hypothetical protein			GI 10	100.00	100.00
RCA23_c26630	hypothetical protein DUF2125			GI 10	100.00	100.00
RCA23_c26640	Asp/Glu racemase	COG3473	Q	GI 10	88.47	100.00
RCA23_c26650	cystathionine beta-lyase	COG0626	E	GI 10	100.00	100.00
RCA23_c26660	deoxyribodipyrimidine photo-lyase PhrB	COG0415	L	GI 10	100.00	100.00
RCA23_c26670	hypothetical protein	COG3380	R	GI 10	100.00	100.00
RCA23_c26680	O-sialoglycoprotein endopeptidase Gcp	COG0533	O	GI 10	100.00	100.00
RCA23_c26690	Sua5/YciO/YrdC/YwlC family protein	COG0009	J	GI 10	100.00	100.00
RCA23_c26700	putative glycoprotease family protein	COG1214	O	GI 10	100.00	100.00

RCA23_c26710	hypothetical protein UPF0079	COG0802	R	GI 10	100.00	100.00
RCA23_c26720	xylose isomerase family protein	COG3622	G	GI 10	31.51	100.00
RCA23_c26730	putative phage integrase	COG0582	L	GI 10	0.00	79.94
RCA23_c26740	hypothetical protein			GI 10	0.00	74.23
RCA23_c26750	putative prophage regulatory protein	COG3311	K	GI 10	0.00	0.00
RCA23_c26760	hypothetical protein			GI 10	0.00	0.00
RCA23_c26770	DNA polymerase	COG0749	L	GI 10	0.00	67.34
RCA23_c26780	hypothetical protein			GI 10	0.00	0.00
RCA23_c26790	hypothetical protein, snoaL-like polyketide cyclase	COG5485	R	GI 10	0.00	0.00
RCA23_c26800	hypothetical protein			GI 10	0.00	56.96
RCA23_c26810	glyoxalase/bleomycin resistance protein/dioxygenase superfamily protein			GI 10	0.00	0.00
RCA23_c26820	hypothetical protein			GI 10	0.00	16.37
RCA23_c26830	putative ABC transporter permease protein	COG0395	G	GI 10	0.00	0.00
RCA23_c26840	putative ABC transporter permease protein	COG1175	G	GI 10	0.00	0.00
RCA23_c26850	putative ABC transporter extracellular solute binding protein	COG1653	G	GI 10	0.00	0.00
RCA23_c26860	ABC transporter ATP-binding protein	COG3839	G	GI 10	0.00	0.00
RCA23_c26870	histidinol dehydrogenase HisD	COG0141	E	GI 10	0.00	0.00
RCA23_c26880	hypothetical protein, snoaL-like polyketide cyclase	COG5485	R	GI 10	0.00	0.00
RCA23_c26890	short chain dehydrogenase	COG1028	I	GI 10	32.12	34.95
RCA23_c26900	putative HTH-type transcriptional regulator LacI family	COG1609	K	GI 10	100.00	0.00
RCA23_c26910	hypothetical protein, snoaL-like polyketide cyclase			GI 10	30.64	1.00
RCA23_c26920	hypothetical protein			GI 10	80.64	23.91
RCA23_c26930	hypothetical protein, 2-hydroxypropyl-CoM lyase-like	COG0620	E	GI 10	73.76	0.00
RCA23_c26940	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	GI 10	0.00	0.00
RCA23_c26950	hypothetical protein	COG0684	H	GI 10	0.00	0.00
RCA23_c26960	3-hydroxyisobutyrate dehydrogenase MmsB	COG2084	I	GI 10	33.55	11.59
RCA23_c26970	dihydroxy-acid dehydratase IlvD	COG0129	E	GI 10	100.00	36.50
RCA23_c26980	hypothetical protein	COG4091	E	GI 10	33.78	51.45
RCA23_c26990	putative choline (or alcohol) dehydrogenase	COG2303	E	GI 10	4.21	29.85
RCA23_c27000	aldehyde dehydrogenase	COG1012	C	GI 10	100.00	76.74
RCA23_c27010	transporter, LysE family	COG1280	E	GI 10	100.00	0.17
RCA23_c27020	TRAP dicarboxylate transporter, subunit DctM	COG4664	Q	GI 10	100.00	24.34
RCA23_c27030	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	GI 10	25.53	41.49

RCA23_c27040	TRAP dicarboxylate transporter, subunit DctP	COG4663	Q	GI 10	0.00	0.00
RCA23_c27050	transcriptional regulator, LysR family	COG0583	K	GI 10	0.00	66.37
RCA23_c27060	mandelate racemase/muconate lactonizing enzyme	COG4948	M	GI 10	100.00	100.00
RCA23_c27070	Zn-dependant oxidoreductase	COG0604	C	GI 10	56.06	97.66
RCA23_c27090	HTH-type transcriptional regulator, LysR family	COG0583	K	GI 10	0.00	70.65
RCA23_c27100	class II aldolase	COG0235	G	GI 10	0.00	29.10
RCA23_c27110	fatty acid desaturase	COG3239	I	GI 10	0.00	69.78
RCA23_c27120	3-methyl-2-oxobutanoate hydroxymethyltransferase PanB	COG0413	H	GI 10	0.00	30.86
RCA23_c27130	putative rieske [2Fe-2S] protein	COG2146	P	GI 10	0.00	89.00
RCA23_c27150	hypothetical protein	COG1396	K		100.00	100.00
RCA23_c27160	arginase family protein	COG0010	E		100.00	100.00
RCA23_c27170	hypothetical protein	COG1082	G		100.00	100.00
RCA23_c27180	putative HTH-type transcriptional regulator, AraC family	COG2169	F		100.00	100.00
RCA23_c27190	hypothetical protein. fatty acid desaturase	COG3239	I		100.00	100.00
RCA23_c27200	AMP-dependent synthase / ligase	COG0318	I		100.00	100.00
RCA23_c27210	ABC transporter ATP-binding protein	COG1132	V		100.00	100.00
RCA23_c27220	spore coat family protein				100.00	100.00
RCA23_c27230	PapD-like chaperone involved in fimbrial biogenesis	COG3121	N		100.00	100.00
RCA23_c27240	fimbrial biogenesis outer membrane usher protein	COG3188	N		100.00	100.00
RCA23_c27250	fimbrial biogenesis outer membrane usher protein	COG3188	N		100.00	100.00
RCA23_c27260	hypothetical protein				100.00	100.00
RCA23_c27280	hypothetical protein	COG0697	G		100.00	100.00
RCA23_c27290	DNA-directed RNA polymerase beta subunit RpoC	COG0086	K		86.33	100.00
RCA23_c27300	DNA-directed RNA polymerase beta subunit RpoB	COG0085	K		100.00	100.00
RCA23_c27310	50S ribosomal protein L7/L12	COG0222	J		100.00	100.00
RCA23_c27320	50S ribosomal protein L10	COG0244	J		100.00	100.00
RCA23_c27330	50S ribosomal protein L1	COG0081	J		100.00	100.00
RCA23_c27340	50S ribosomal protein L11	COG0080	J		100.00	100.00
RCA23_c27350	transcription antitermination protein NusG	COG0250	K		100.00	100.00
RCA23_c27360	preprotein translocase, subunit SecE	COG0690	U		100.00	100.00
RCA23_c27370	hypothetical protein				100.00	100.00
RCA23_c27430	hypothetical protein				100.00	100.00
RCA23_c27440	hypothetical protein, transmembrane				100.00	100.00

RCA23_c27450	hypothetical protein			74.48	100.00
RCA23_c27460	putative acetolactate synthase isozyme 1 large subunit	COG0028	E	100.00	100.00
RCA23_c27470	putative cation transporter	COG0168	P	100.00	100.00
RCA23_c27480	GTP cyclohydrolase FolE	COG1469	S	100.00	100.00
RCA23_c27490	O-succinylhomoserine sulfhydrylase MetZ	COG0626	E	100.00	100.00
RCA23_c27500	hypothetical protein	COG0625	O	100.00	100.00
RCA23_c27510	putative intracellular septation protein	COG2917	D	100.00	99.00
RCA23_c27520	hypothetical protein	COG0697	G	100.00	100.00
RCA23_c27530	putative cell division protein	COG0552	U	100.00	100.00
RCA23_c27540	NADH-quinone oxidoreductase subunits E/F (fused)	COG1894	C	100.00	100.00
RCA23_c27550	exodeoxyribonuclease 7 large subunit XseA	COG1570	L	100.00	100.00
RCA23_c27560	phosphoribosylamine--glycine ligase PurD	COG0151	F	100.00	93.19
RCA23_c27570	hypothetical protein, ferredoxin	COG0633	C	100.00	100.00
RCA23_c27580	hypothetical protein	COG0587	L	100.00	100.00
RCA23_c27590	periplasmic serine protease DO-like	COG0265	O	100.00	100.00
RCA23_c27600	putative protein hflC	COG0330	O	100.00	100.00
RCA23_c27610	protein HflK	COG0330	O	100.00	100.00
RCA23_c27620	glutathione reductase Gor	COG1249	C	100.00	100.00
RCA23_c27630	ribose-5-phosphate isomerase A	COG0120	G	100.00	100.00
RCA23_c27640	hypothetical protein			100.00	100.00
RCA23_c27650	L-serine dehydratase SdaB	COG1760	E	100.00	99.49
RCA23_c27660	putative integral membrane protein	COG0697	G	100.00	100.00
RCA23_c27670	hypothetical protein, thiamine pyrophosphokinase	COG1564	H	100.00	100.00
RCA23_c27680	hypothetical protein			100.00	100.00
RCA23_c27690	hypothetical protein			100.00	89.18
RCA23_c27700	hypothetical protein			100.00	100.00
RCA23_c27710	adenylosuccinate synthase PurA	COG0104	F	100.00	100.00
RCA23_c27720	preprotein translocase, SecG subunit SecG			100.00	100.00
RCA23_c27730	CTP synthase	COG0504	F	100.00	100.00
RCA23_c27740	hypothetical protein, DUF1330	COG5470	S	100.00	94.90
RCA23_c27750	hypothetical protein DUF1332	COG4103	S	100.00	100.00
RCA23_c27760	hypothetical protein	COG3221	P	100.00	100.00
RCA23_c27770	amino acid transport ATP-binding protein	COG4598	E	100.00	100.00

RCA23_c27780	putative amino acid transport extracellular solute binding protein	COG0834	E	100.00	100.00
RCA23_c27790	putative ABC transporter inner membrane component	COG4215	E	100.00	100.00
RCA23_c27800	putative amino acid transport permease protein	COG4160	E	100.00	100.00
RCA23_c27810	gamma-glutamylputrescine synthase PuuA	COG0174	E	100.00	98.51
RCA23_c27820	hypothetical protein, glutamine amidotransferase class I	COG0518	F	100.00	100.00
RCA23_c27830	gamma-glutamylputrescine synthase PuuA	COG0174	E	100.00	100.00
RCA23_c27840	gamma-glutamylputrescine oxidoreductase PuuB	COG0665	E	100.00	97.62
RCA23_c27850	aminotransferase class III	COG0161	H	100.00	100.00
RCA23_c27860	putative PHB synthesis repressor protein	COG5394	S	100.00	100.00
RCA23_c27870	hypothetical protein			100.00	100.00
RCA23_c27880	poly(R)-hydroxyalkanoic acid synthase PhaC	COG3243	I	100.00	100.00
RCA23_c27890	polyhydroxyalkanoate depolymerase PhaZ	COG4553	I	98.97	99.60
RCA23_c27900	putative ribosome biogenesis GTPase RsgA	COG1162	R	100.00	100.00
RCA23_c27910	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	100.00	100.00
RCA23_c27920	hypothetical protein			100.00	100.00
RCA23_c27930	hypothetical protein, alpha/beta hydrolase-like	COG0596	R	100.00	100.00
RCA23_c27940	threonyl-tRNA synthase ThrS	COG0441	J	100.00	95.74
RCA23_c27950	hypothetical protein, ArsC	COG1393	P	100.00	100.00
RCA23_c27960	thymidylate synthase	COG1351	F	100.00	100.00
RCA23_c27970	putative glyoxalase/bleomycin resistance protein/dioxygenase superfamily pr	COG0346	E	100.00	100.00
RCA23_c27980	hypothetical protein			100.00	100.00
RCA23_c27990	putative transcriptional regulatorm marR family	COG1846	K	100.00	100.00
RCA23_c28000	hypothetical protein DUF339	COG2938	S	100.00	100.00
RCA23_c28010	hypothetical protein, DNA binding	COG1813	K	100.00	100.00
RCA23_c28020	aspartate aminotransferase AatA	COG0436	E	100.00	100.00
RCA23_c28030	hypothetical protein			100.00	100.00
RCA23_c28040	multidrug resistance protein NorM	COG0534	V	100.00	100.00
RCA23_c28050	DNA topoisomerase 4 subunit B	COG0187	L	100.00	100.00
RCA23_c28060	hypothetical protein	COG3238	S	100.00	100.00
RCA23_c28070	putative malate/L-lactate dehydrogenase	COG2055	C	100.00	99.90
RCA23_c28080	lipoprotein-releasing system ATP-binding protein LolD	COG1136	V	100.00	100.00
RCA23_c28090	lipoprotein-releasing system transmembrane protein, lolC/E family	COG4591	M	100.00	100.00
RCA23_c28100	prolyl-tRNA synthase ProS	COG0442	J	100.00	100.00

RCA23_c28110	hypothetical protein DUF20	COG0628	R	100.00	100.00
RCA23_c28120	hypothetical protein	COG0593	L	100.00	100.00
RCA23_c28130	polyphosphate kinase Ppk	COG0855	P	100.00	100.00
RCA23_c28140	putative phosphatase	COG0248	F	100.00	99.16
RCA23_c28150	hypothetical protein, DnaJ	COG1076	O	100.00	100.00
RCA23_c28160	hypothetical protein			100.00	100.00
RCA23_c28170	methylmalonyl-CoA mutase McmA	COG1884	I	100.00	100.00
RCA23_c28180	hypothetical protein			100.00	100.00
RCA23_c28190	biotin carboxylase AccC	COG4770	I	100.00	100.00
RCA23_c28200	hypothetical protein			100.00	100.00
RCA23_c28210	propionyl-CoA carboxylase beta chain, mitochondrial precursor	COG4799	I	100.00	100.00
RCA23_c28220	putative major facilitator superfamily transporter	COG2814	G	100.00	100.00
RCA23_c28230	hypothetical protein	COG3800	R	100.00	100.00
RCA23_c28240	betaine aldehyde dehydrogenase BetB	COG1012	C	100.00	100.00
RCA23_c28250	hypothetical protein	COG0697	G	100.00	100.00
RCA23_c28260	TRAP dicarboxylate transporter, subunit DctM	COG1593	G	100.00	98.72
RCA23_c28270	TRAP dicarboxylate transporter, subunit DctQ	COG4665	Q	100.00	87.46
RCA23_c28280	TRAP dicarboxylate transporter, subunit DctP	COG1638	G	100.00	100.00
RCA23_c28290	transcriptional regulator, LacI family	COG1609	K	100.00	100.00
RCA23_c28300	hypothetical protein, DUF81 family	COG0730	R	100.00	100.00
RCA23_c28310	serine--glyoxylate aminotransferase SgaA	COG0075	E	100.00	99.66
RCA23_c28320	D-lactate dehydrogenase Dld	COG0277	C	100.00	90.09
RCA23_c28330	hypothetical protein			100.00	79.22
RCA23_c28340	AMP-binding enzyme	COG0318	I	100.00	80.01
RCA23_c28350	succinate-semialdehyde dehydrogenase SucD	COG1012	C	100.00	100.00
RCA23_c28360	hypothetical protein	COG3333	S	100.00	100.00
RCA23_c28370	hypothetical protein			100.00	100.00
RCA23_c28380	putative tripartite tricarboxylate transporter family receptor	COG3181	S	18.15	89.84
RCA23_c28390	putative ETC complex I subunit			100.00	100.00
RCA23_c28450	sulfoxide reductase heme-binding subunit YedZ	COG2717	S	100.00	100.00
RCA23_c28460	sulfoxide reductase catalytic subunit YedY	COG2041	R	100.00	100.00
RCA23_c28470	3-oxoacyl-[acyl-carrier-protein] reductase FabG	COG1028	I	100.00	100.00
RCA23_c28480	putative haloacid dehalogenase-like hydrolase	COG0637	R	100.00	100.00

RCA23_c28490	chaperone protein ClpB	COG0542	O	100.00	100.00
RCA23_c28500	orotidine 5'-phosphate decarboxylase PyrF	COG0284	F	100.00	100.00
RCA23_c28510	hypothetical protein			100.00	100.00
RCA23_c28520	SPFH domain/band 7 family protein	COG0330	O	100.00	100.00
RCA23_c28530	DNA polymerase IV	COG0389	L	100.00	100.00
RCA23_c28540	putative N-formylglutamate amidohydrolase	COG3741	E	100.00	100.00
RCA23_c28560	hypothetical protein, phenylacetate-coenzyme A ligase	COG1541	H	100.00	100.00
RCA23_c28570	high-affinity branched-chain amino acid transporter ATP-binding protein	COG0410	E	100.00	99.64
RCA23_c28580	hypothetical protein	COG0683	E	100.00	57.30
RCA23_c28590	high-affinity branched-chain amino acid transporter permease protein	COG4177	E	100.00	71.22
RCA23_c28600	high-affinity branched-chain amino acid transporter permease protein	COG0559	E	100.00	90.82
RCA23_c28610	high-affinity branched-chain amino acid transporter ATP-binding protein	COG0411	E	100.00	100.00
RCA23_c28620	putative long-chain-fatty-acid-CoA ligase	COG1022	I	100.00	100.00
RCA23_c28630	tRNA pseudouridine synthase B	COG0130	J	100.00	100.00
RCA23_c28650	ribosome-binding factor A	COG0858	J	100.00	100.00
RCA23_c28640	dihydrodipicolinate synthase DapA	COG0289	E	100.00	100.00
RCA23_c28660	hypothetical protein DUF1674	COG5508	S	100.00	100.00
RCA23_c28670	putative ribosomal RNA small subunit methyltransferase B	COG0144	J	100.00	100.00
RCA23_c28680	hypothetical protein, heparinase II/III	COG5360	S	79.28	100.00
RCA23_c28690	bifunctional purine biosynthesis protein PurH	COG0138	F	100.00	93.76
RCA23_c28700	signal peptidase II	COG0597	M	100.00	100.00
RCA23_c28710	uncharacterized zinc protease y4wA	COG0612	R	100.00	100.00
RCA23_c28720	uncharacterized zinc protease y4wB	COG0612	R	100.00	100.00
RCA23_c28730	DNA mismatch repair protein MutL	COG0323	L	100.00	100.00
RCA23_c28740	putative RmuC family protein	COG1322	S	100.00	100.00
RCA23_c28750	transcriptional activator ChrR	COG3806	T	100.00	100.00
RCA23_c28760	RNA polymerase sigma factor SigK	COG1595	K	100.00	100.00
RCA23_c28770	hypothetical protein	COG2907	R	100.00	100.00
RCA23_c28780	hypothetical protein DUF1365	COG3496	S	100.00	100.00
RCA23_c28790	hypothetical protein	COG2211	G	100.00	100.00
RCA23_c28800	hypothetical protein			100.00	100.00
RCA23_c28810	putative short chain dehydrogenase	COG4221	R	100.00	100.00
RCA23_c28820	saccharopine dehydrogenase (NAD+,L-lysine-forming)	COG3288	C	100.00	99.34

RCA23_c28830	putative glutathione S-transferase	COG0625	O	100.00	100.00
RCA23_c28840	putative phosphoglycerate mutase family protein	COG0406	G	100.00	100.00
RCA23_c28850	formate dehydrogenase Fdh	COG3383	R	100.00	100.00
RCA23_c28860	hypothetical protein	COG1357	S	75.52	54.92
RCA23_c28870	hypothetical protein DUF1643	COG4333	S	100.00	100.00
RCA23_c28880	haloacetate dehalogenase DehH	COG0596	R	100.00	100.00
RCA23_c28890	30S ribosomal protein S15	COG0184	J	100.00	100.00
RCA23_c28900	hypothetical protein, DUF6 transmembrane protein	COG0697	G	100.00	100.00
RCA23_c28910	polyribonucleotide nucleotidyltransferase Pnp	COG1185	J	100.00	100.00
RCA23_c28920	aldehyde dehydrogenase	COG1012	C	97.85	100.00
RCA23_c28930	ribosomal large subunit pseudouridine synthase A	COG0564	J	100.00	96.43
RCA23_c28940	protein OtnG			100.00	99.37
RCA23_c28950	hypothetical protein, DUF940 putative lipoprotein			27.08	100.00
RCA23_c28960	hypothetical protein			100.00	100.00
RCA23_c28970	UDP-glucose 4-epimerase GalE	COG1087	M	100.00	100.00
RCA23_c28980	putative FAD linked oxidase	COG0277	C	100.00	100.00
RCA23_c28990	signal recognition particle protein Ffh	COG0541	U	100.00	96.66
RCA23_c29000	putative chorismate mutase type II	COG1605	E	100.00	100.00
RCA23_c29010	30S ribosomal protein S16	COG0228	J	100.00	100.00
RCA23_c29020	ribosome maturation factor RimM	COG0806	J	100.00	100.00
RCA23_c29030	tRNA (guanine-N(1)-)-methyltransferase TrmD	COG0336	J	100.00	100.00
RCA23_c29040	50S ribosomal protein L19	COG0335	J	100.00	100.00
RCA23_c29050	50S ribosomal protein L31	COG0254	J	100.00	100.00
RCA23_c29060	ATPase MipZ	COG1192	D	100.00	100.00
RCA23_c29070	fructokinase	COG0524	G	100.00	100.00
RCA23_c29080	P-hydroxybenzoate hydroxylase PobA	COG0654	H	100.00	98.64
RCA23_c29090	kynureninase KynU	COG3844	E	100.00	100.00
RCA23_c29100	HTH-type transcriptional regulator, GntR family	COG1802	K	100.00	100.00
RCA23_c29110	hypothetical protein, glutathione S-transferase	COG0625	O	100.00	100.00
RCA23_c29120	L-threonine 3-dehydrogenase	COG0451	M	100.00	100.00
RCA23_c29130	enoyl-CoA hydratase	COG1024	I	100.00	100.00
RCA23_c29140	enoyl-CoA hydratase	COG1250	I	100.00	98.59
RCA23_c29150	3-hydroxyacyl-CoA dehydrogenase FadN	COG1028	I	100.00	100.00

RCA23_c29160	hypothetical protein			90.85	100.00
RCA23_c29170	integrase	COG4974	L	100.00	100.00
RCA23_c29180	hypothetical protein, porin-like			100.00	73.20
RCA23_c29190	leucyl-tRNA synthase LeuS	COG0495	J	100.00	100.00
RCA23_c29200	hypothetical protein DUF2159			100.00	100.00
RCA23_c29210	hypothetical protein			100.00	100.00
RCA23_c29220	hypothetical protein, glutathione S-transferase	COG0625	O	100.00	100.00
RCA23_c29230	hypothetical protein	COG2081	R	100.00	100.00
RCA23_c29240	hypothetical protein, porin			100.00	59.27
RCA23_c29250	hypothetical protein, alanine racemase-like	COG0325	R	100.00	100.00
RCA23_c29260	hypothetical protein, YkuD	COG3786	S	100.00	100.00
RCA23_c29270	GTP cyclohydrolase II RibA	COG0807	H	100.00	100.00
RCA23_c29280	response regulator receiver protein	COG0745	T	100.00	100.00
RCA23_c29290	2-oxo-hepta-3-ene-1,7-dioic acid hydratase HpcG	COG3971	Q	100.00	100.00
RCA23_c29300	fumarylacetoacetate hydrolase family protein	COG0179	Q	100.00	100.00
RCA23_c29310	5-carboxymethyl-2-hydroxymuconate isomerase	COG3232	E	100.00	100.00
RCA23_c29320	homoprotocatechuate degradation operon regulator HpaR	COG1846	K	100.00	100.00
RCA23_c29330	4-hydroxyphenylacetate 3-monooxygenase oxygenase component HpaH	COG2368	Q	100.00	98.22
RCA23_c29340	ABC sugar transporter, ATPase subunit	COG3839	G	100.00	100.00
RCA23_c29350	mannitol 2-dehydrogenase MtlK	COG0246	G	100.00	99.22
RCA23_c29360	transcriptional regulator, AraC family	COG2207	K	100.00	100.00
RCA23_c29370	ABC sugar transporter, permease protein	COG0395	G	100.00	100.00
RCA23_c29380	ABC sugar transporter, permease protein	COG1175	G	100.00	72.95
RCA23_c29390	ABC transporter, periplasmic substrate-binding protein	COG1653	G	100.00	95.16
RCA23_c29400	putative flavohemoglobin / bacterial hemoglobin	COG1017	C	0.00	73.22
RCA23_c29410	hypothetical protein			0.00	100.00
RCA23_c29420	Na+/H+ antiporter NhaA	COG3004	P	0.00	98.75
RCA23_c29430	hypothetical protein			3.54	100.00
RCA23_c29440	hypothetical protein DUF55	COG2947	S	100.00	100.00
RCA23_c29450	protein YciI	COG2350	S	100.00	100.00
RCA23_c29460	glycerol-3-phosphate dehydrogenase GpsA	COG0240	C	100.00	100.00
RCA23_c29470	hypothetical protein, uroporophyrinogen-III synthase HemD	COG1587	H	100.00	99.86
RCA23_c29480	hypothetical protein	COG3264	M	100.00	100.00

RCA23_c29490	hypothetical protein, HemY	COG3898	S	100.00	100.00
RCA23_c29500	transcriptional regulatory protein, Ars family	COG0640	K	100.00	100.00
RCA23_c29510	magnesium-chelatase subunit BchO	COG0596	R	100.00	100.00
RCA23_c29520	magnesium-chelatase subunit BchD	COG1240	H	100.00	98.79
RCA23_c29530	magnesium-chelatase subunit BchI	COG1239	H	100.00	100.00
RCA23_c29540	spheroidene monooxygenase CrtA			100.00	100.00
RCA23_c29550	phytoene dehydrogenase CrtI	COG1233	Q	100.00	100.00
RCA23_c29560	phytoene synthase CrtB	COG1562	I	100.00	100.00
RCA23_c29570	hydroxyneurosporene dehydrogenase CrtC			100.00	84.93
RCA23_c29580	methoxyneurosporene dehydrogenase CrtD	COG1233	Q	100.00	100.00
RCA23_c29590	geranylgeranyl pyrophosphate synthase CrtE	COG0142	H	100.00	91.37
RCA23_c29600	hydroxyneurosporene methyltransferase CrtF			100.00	100.00
RCA23_c29610	2-desacetyl-2-hydroxyethyl bacteriochlorophyllide A dehydrogenase BchC	COG1063	E	100.00	100.00
RCA23_c29620	chlorophyllide reductase BchX	COG1348	P	100.00	100.00
RCA23_c29630	chlorophyllide reductase BchY			100.00	100.00
RCA23_c29640	chlorophyllide reductase subunit BchZ	COG2710	C	100.00	99.38
RCA23_c29650	protein PufQ			100.00	100.00
RCA23_c29660	light-harvesting protein B-870 beta chain PufB			100.00	100.00
RCA23_c29670	light-harvesting protein B-870 alpha chain PufA			100.00	100.00
RCA23_c29680	reaction center protein L chain PufL			100.00	100.00
RCA23_c29690	reaction center protein M chain PufM			100.00	100.00
RCA23_c29700	protein PufX			100.00	100.00
RCA23_c29710	1-deoxy-D-xylulose-5-phosphate synthase Dxs	COG1154	H	100.00	100.00
RCA23_c29720	isopentenyl-diphosphate delta-isomerase Idi	COG1443	I	100.00	100.00
RCA23_c29730	geranylgeranyl reductase BchP	COG0644	C	100.00	100.00
RCA23_c29740	bacteriochlorophyll synthase 44.5 kDa chain			100.00	100.00
RCA23_c29750	bacteriochlorophyll synthase BchG	COG0382	H	100.00	100.00
RCA23_c29760	cytochrome c-551	COG3474	C	100.00	100.00
RCA23_c29770	uroporphyrinogen decarboxylase HemE	COG0407	H	100.00	100.00
RCA23_c29780	porphobilinogen deaminase HemC	COG0181	H	100.00	100.00
RCA23_c29790	hypothetical protein, NmrA-like	COG0702	M	100.00	90.83
RCA23_c29800	5-aminolevulinate synthase HemA	COG0156	H	99.67	100.00
RCA23_c29810	hypothetical protein			43.99	100.00

RCA23_c29820	aerobic Mg-protoporphyrin IX monomethyl ester oxidative cyclase AcsF			100.00	100.00
RCA23_c29830	hypothetical protein			100.00	100.00
RCA23_c29840	hypothetical protein			100.00	100.00
RCA23_c29850	hypothetical protein			90.92	100.00
RCA23_c29860	reaction center protein PuhA			22.87	100.00
RCA23_c29870	protein PucC			100.00	100.00
RCA23_c29880	magnesium-protoporphyrin O-methyltransferase BchM	COG2227	H	100.00	100.00
RCA23_c29890	light-independent protochlorophyllide reductase iron-sulfur ATP-binding prote	COG1348	P	100.00	100.00
RCA23_c29900	magnesium-chelatase subunit BchH	COG1429	H	100.00	100.00
RCA23_c29910	light-independent protochlorophyllide reductase subunit BchB	COG2710	C	100.00	100.00
RCA23_c29920	light-independent protochlorophyllide reductase subunit BchN	COG2710	C	100.00	100.00
RCA23_c29930	2-vinyl bacteriochlorophyllide hydratase BchF			100.00	100.00
RCA23_c29940	putative transcriptional regulator PpaA			100.00	100.00
RCA23_c29950	transcriptional regulator PpsR	COG3829	K	100.00	100.00
RCA23_c29960	hypothetical protein	COG5012	R	100.00	100.00
RCA23_c29970	transcriptional regulator protein FixJ	COG4566	T	100.00	100.00
RCA23_c29980	peripheral-type benzodiazepine receptor/signal transduction protein TspO	COG3476	T	100.00	100.00
RCA23_c29990	hypothetical protein, integral membrane proteins YeeE/YedE	COG2391	R	100.00	100.00
RCA23_c30000	hypothetical protein, integral membrane proteins YeeE/YedE	COG2391	R	100.00	100.00
RCA23_c30010	putative beta-lactamase hydrolase-like protein	COG0491	R	100.00	100.00
RCA23_c30020	hypothetical protein	COG3453	S	100.00	100.00
RCA23_c30030	sulphate transporter	COG0659	P	100.00	100.00
RCA23_c30040	amidase	COG0154	J	100.00	100.00
RCA23_c30050	hypothetical protein			100.00	100.00
RCA23_c30060	aminoglycoside phosphotransferase Aph			100.00	100.00
RCA23_c30070	hypothetical protein	COG2885	M	100.00	100.00
RCA23_c30080	chaperone protein DnaJ	COG0484	O	100.00	100.00
RCA23_c30090	chaperone protein DnaK	COG0443	O	100.00	100.00
RCA23_c30100	putative alpha-ketoglutarate-dependent dioxygenase AlkB	COG3145	L	100.00	100.00
RCA23_c30110	putative ABC-2 type transporter	COG1682	G	100.00	100.00
RCA23_c30120	3'(2'),5'-bisphosphate nucleotidase CysQ	COG1218	P	100.00	100.00
RCA23_c30130	UTP-glucose-1-phosphate uridylyltransferase GalU	COG1210	M	100.00	26.43
RCA23_c30140	hypothetical protein	COG0463	M	4.59	18.36

RCA23_c30150	hypothetical protein			0.00	53.82
RCA23_c30160	hypothetical protein beta-1,6-N-acetylglucosaminyltransferases			0.00	6.85
RCA23_c30170	hypothetical protein			14.99	36.27
RCA23_c30180	nitrogen regulatory protein PtsN	COG1762	G	100.00	100.00
RCA23_c30190	putative sigma 54 modulation protein	COG1544	J	100.00	100.00
RCA23_c30200	lipopolysaccharide export system ATP-binding protein LptB	COG1137	R	100.00	100.00
RCA23_c30210	putative lipopolysaccharide export system protein LptA	COG1934	S	100.00	100.00
RCA23_c30220	hypothetical protein			100.00	100.00
RCA23_c30230	putative 3'-5'-exonuclease	COG0349	J	100.00	94.63
RCA23_c30240	glycine amidinotransferase	COG1834	E	100.00	100.00
RCA23_c30250	hypothetical protein	COG4123	R	100.00	100.00
RCA23_c30260	hypothetical protein, metal-dependent phosphohydrolase	COG1896	R	100.00	100.00
RCA23_c30270	S-adenosyl-L-homocysteine hydrolase AhcY	COG0499	H	100.00	100.00
RCA23_c30280	photosynthetic apparatus regulatory protein RegA	COG4567	T	100.00	100.00
RCA23_c30290	protein SenC	COG1999	R	100.00	100.00
RCA23_c30300	sensor histidine kinase RegB	COG0642	T	100.00	100.00
RCA23_c30310	hypothetical protein			100.00	100.00
RCA23_c30320	hypothetical protein, aminoglycoside phosphotransferase	COG3178	R	100.00	98.19
RCA23_c30330	hypothetical protein, nucleotidyl transferase	COG1208	M	100.00	100.00
RCA23_c30340	double-strand break repair protein AddB	COG3893	L	100.00	100.00
RCA23_c30350	double-strand break repair helicase AddA	COG1074	L	97.97	100.00
RCA23_c30360	thioredoxin TrxA	COG3118	O	100.00	100.00
RCA23_c30370	ATP-dependent protease HslV	COG5405	O	100.00	100.00
RCA23_c30380	ATP-dependent hsl protease ATP-binding subunit HslU	COG1220	O	100.00	96.87
RCA23_c30390	MFS-type transporter	COG2814	G	100.00	100.00
RCA23_c30400	hypothetical protein, Smr protein/MutS2	COG2840	S	100.00	100.00
RCA23_c30410	putative lytic murein transglycosylase	COG2821	M	100.00	100.00
RCA23_c30420	hypothetical protein, TIM44	COG4395	S	100.00	100.00
RCA23_c30430	putative cytoplasmic membrane protein FxsA	COG3030	R	100.00	100.00
RCA23_c30440	protein-export protein SecB	COG1952	U	100.00	100.00
RCA23_c30450	DNA polymerase III subunit epsilon	COG0847	L	100.00	100.00
RCA23_c30460	dephospho-CoA kinase CoaE	COG0237	H	100.00	100.00
RCA23_c30470	shikimate 5-dehydrogenase AroE	COG0169	E	100.00	100.00

RCA23_c30480	transcription termination factor Rho	COG1158	K	100.00	100.00
RCA23_c30490	tRNA modification GTPase MnmE	COG0486	R	100.00	100.00
RCA23_c30500	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	COG0445	D	100.00	99.84
RCA23_c30510	ribosomal RNA small subunit methyltransferase G	COG0357	M	100.00	100.00
RCA23_c30520	chromosome-partitioning protein ParA	COG1192	D	100.00	100.00
RCA23_c30530	chromosome-partitioning protein ParB	COG1475	K	100.00	100.00
RCA23_c30540	oxygen-independent coproporphyrinogen III oxidase HemN	COG0635	H	100.00	100.00
RCA23_c30550	nucleoside-triphosphatase RdgB	COG0127	F	100.00	100.00
RCA23_c30560	ribonuclease PH	COG0689	J	100.00	100.00
RCA23_c30570	heat-inducible transcription repressor HrcA	COG1420	K	100.00	100.00
RCA23_c30580	protein GrpE	COG0576	O	100.00	100.00
RCA23_c30590	DNA mismatch repair protein MutS	COG0249	L	100.00	100.00
RCA23_c30600	NADP-dependent malic enzyme MaeB	COG0281	C	100.00	98.05
RCA23_c30610	pfkB family carbohydrate kinase	COG0524	G	100.00	100.00
RCA23_c30620	putative aminotransferase class IV	COG0115	E	100.00	100.00
RCA23_c30630	hypothetical protein, branched-chain-amino-acid aminotransferase-like			100.00	100.00
RCA23_c30640	argininosuccinate synthase ArgG	COG0137	E	100.00	100.00
RCA23_c30650	threonine dehydratase, biosynthetic	COG1171	E	100.00	100.00
RCA23_c30660	putative NUDIX hydrolase	COG0494	L	100.00	100.00
RCA23_c30670	heat shock protein 33	COG1281	O	100.00	100.00
RCA23_c30680	putative NUDIX hydrolase	COG0494	L	100.00	100.00
RCA23_c30690	hypothetical protein, poly A polymerase	COG0617	J	100.00	100.00
RCA23_c30700	ABC transporter ATP-binding/permease protein	COG1132	V	100.00	100.00
RCA23_c30710	hypothetical protein	COG2265	J	95.06	100.00
RCA23_c30720	ion transport protein			2.89	100.00
RCA23_c30730	hypothetical protein	COG3034	S	100.00	100.00
RCA23_c30740	hypothetical protein, SCP-like extracellular protein	COG2340	S	100.00	100.00
RCA23_c30750	hypothetical protein	COG1376	S	100.00	100.00
RCA23_c30760	ferrochelatase HemH	COG0276	H	100.00	100.00
RCA23_c30770	hypothetical protein			100.00	100.00
RCA23_c30780	hypothetical protein, ComF/GntX family	COG1040	R	100.00	100.00
RCA23_c30790	glutaredoxin GrxC	COG0695	O	100.00	100.00
RCA23_c30800	putative carbon-nitrogen hydrolase	COG0388	R	100.00	100.00

RCA23_c30810	putative HTH-type transcriptional regulator, MarR family	COG1846	K	100.00	100.00
RCA23_c30820	3-demethylubiquinone-9 3-O-methyltransferase UbiG	COG2227	H	100.00	100.00
RCA23_c30830	proline iminopeptidase Pip	COG0596	R	100.00	94.74
RCA23_c30840	putative peptide transport system permease protein	COG0601	E	100.00	100.00
RCA23_c30850	putative peptide transport system permease protein	COG1173	E	100.00	100.00
RCA23_c30860	peptide transport system ATP-binding protein	COG1123	R	100.00	100.00
RCA23_c30870	putative periplasmic peptide-binding protein	COG0747	E	100.00	97.48
RCA23_c30880	hypothetical protein UPF0090	COG0779	S	100.00	100.00
RCA23_c30890	transcription elongation protein NusA	COG0195	K	100.00	100.00
RCA23_c30900	hypothetical protein DUF448	COG2740	K	100.00	100.00
RCA23_c30910	translation initiation factor IF-2	COG0532	J	100.00	100.00
RCA23_c30920	putative mutator MutT protein	COG1051	F	100.00	100.00
RCA23_c30930	arginine biosynthesis bifunctional protein ArgJ	COG1364	E	100.00	100.00
RCA23_c30940	putative peptidylprolyl isomerase	COG0760	O	100.00	100.00
RCA23_c30950	protein translocase subunit SecA	COG0653	U	100.00	100.00
RCA23_c30960	putative O-acetyltransferase OatA	COG1835	I	87.06	100.00
RCA23_c30970	hypothetical protein	COG4922	S	100.00	100.00
RCA23_c30980	hypothetical protein, transcriptional regulator-like	COG1733	K	100.00	100.00
RCA23_c30990	UDP-glucoronate 5'-epimerase LspL	COG0451	M	62.96	96.49
RCA23_c31000	hypothetical protein			100.00	98.60
RCA23_c31010	phosphoglucosamine mutase GlmM	COG1109	G	100.00	100.00
RCA23_c31020	putative ubiquinone biosynthesis protein UbiB	COG0661	R	100.00	100.00
RCA23_c31030	ubiquinone/menaquinone biosynthesis methyltransferase UbiE	COG2226	H	100.00	100.00
RCA23_c31040	formamidopyrimidine-DNA glycosylase MutM	COG0266	L	100.00	100.00
RCA23_c31050	enoyl-CoA hydratase/isomerase	COG1024	I	100.00	100.00
RCA23_c31060	30S ribosomal protein S20	COG0268	J	100.00	100.00

STUDY 6:

COMPOSITION AND ACTIVITY OF THE BACTERIOPLANKTON COMMUNITIES IN THE DRAKE PASSAGE AND ANTARCTIC PENINSULA REGION WITH A SPECIAL EMPHASIS ON THE *ROSEOBACTER* CLADE AND DISSOLVED ORGANIC MATTER

SIMON M[1], BILLERBECK S[1], BRINKHOFF T[1], DOGS M[1], MÜLLENMEISTER S[1], SEIBT M[1], SEIDEL M[1], SMITS M[2], WAGNER-DÖBLER I[3], WANG H[3], **WEMHEUER B**[2], AND WURST M[1]

REPORTS ON POLAR AND MARINE RESEARCH 653: 49-54.
(hdl: 10013/epic.4037)

[1]INSTITUTE FOR CHEMISTRY AND BIOLOGY OF THE MARINE ENVIRONMENT (ICBM), CARL-VON-OSSIETZKY-UNIVERSITY OF OLDENBURG, CARL-VON-OSSIETZKY-STR. 9-11, D-26111 OLDENBURG, GERMANY; [2]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY; [3]HELMHOLTZ CENTRE FOR INFECTION RESEARCH, INHOFFENSTRAßE 7, D-38124 BRAUNSCHWEIG, GERMANY

Author contributions to the work: all

Contributed data: all

Wrote the publication: MS

Berichte
zur Polar-
und Meeresforschung

652
2012

Reports
on Polar and Marine Research

The Expedition of the Research Vessel "Polarstern" to the Antarctic in 2012 (ANT-XXVIII/4)

Edited by
Magnus Lucassen
with contributions of the participants

HELMHOLTZ GEMEINSCHAFT

ALFRED-WEGENER-INSTITUT FÜR
POLAR- UND MEERESFORSCHUNG
in der Helmholtz-Gemeinschaft
D-27570 BREMERHAVEN
Bundesrepublik Deutschland

ISSN 1866-3192

13. COMPOSITION AND ACTIVITY OF THE BACTERIOPLANKTON COMMUNITIES IN THE DRAKE PASSAGE AND ANTARCTIC PENINSULA REGION WITH A SPECIAL EMPHASIS ON THE *ROSEOBACTER* CLADE AND DISSOLVED ORGANIC MATTER

Meinhard Simon[1], Sara Billerbeck[1], Thorsten Brinkhoff[1], Marco Dogs[1], Swaantje Müllenmeister[1], Maren Seibt[1], Michael Seidel[1], Maike Smits[3], Irene Wagner-Döbler[2], Hui Wang[2], Bernd Wemheuer[3], Mascha Wurst1

[1]ICBM
[2]HZI
[3]University of Göttingen

Objectives

We aim at a comprehensive assessment of the bacterioplankton community in the Drake Passage and the Peninsula region of the Southern Ocean with a special emphasis on the *Roseobacter* clade and its major bacterioplankton subclusters. This project is part of a key work package of the Transregional Collaborative Research Center *Ecology, Physiology and Molecular Biology of the Roseobacter clade*: Towards a Systems Biology Understanding of a Globally Important Clade of Marine Bacteria (TRR 51). The work includes investigations of the biogeography, growth and population dynamics, the genomic potential (metagenomics) and actively expressed genes (metatranscriptomics) and the impact on the decomposition of dissolved organic matter (DOM) and cycling by the bacterioplankton communities. Our investigations can only be done in a concerted action in which also the other members of the bacterioplankton communities are considered, as well as bulk parameters of the entire bacterioplankton communities and relevant biogeochemical parameters such as chlorophyll and the composition and concentration of DOM. Previous studies dealing with some of these aspects have been carried out in this region before (Manganelli et al. 2009, Straza et al. 2010). However, one particular group of marine bacteria and in particular the *Roseobacter* clade have not been investigated before in such great detail.

A list of investigated parameters is given in table 13.1.

Tab. 13.1: List of parameters studied for assessing bacterioplankton communities during cruise ANT-XXVIII/4.

Parameter	Water depths (m)								
	20	40	60	100	200	500	1000	2000	>3000
POC	+	+	+	+		+	+	+	+
Chlorophyll	+	+	+	+	+				
Phytoplankton	+	+	+	+					
Inorganic nutrients	+	+	+	+	+	+	+	+	+
Bacterial abundance	+	+	+	+	+	+	+	+	+
Bacterial production	+	+	+	+	+	+	+		
Glucose turnover rate	+	+	+	+	+	+	+		
Amino acid turnover rate	+	+	+	+	+				
FISH	+		+	+					
Aerobic anoxygenic bacteria	+	+	+	+	+				
DGGE	+	+	+	+	+	+	+	+	+
16S rRNA gene Pyrosequencing	+	+	+	+	+	+	+	+	+
Metagenomics	+		+						+
Metatranscriptomics	+	+	+	+	+	+	+	+	+
DOC	+	+	+	+	+	+	+	+	+
DOM	+	+	+	+	+	+	+	+	+
Vitamins	+	+	+	+	+	+	+	+	+

Work at sea

Our main work on shipboard was the collection and processing of water samples from depths between 20 and 3500 m. Samples were collected with Niskin bottles mounted on a CTD rosette from the mixed layer, the mesopelagic and bathypelagic zones (for details on the CTD see Badewien et al., chapter 18, this volume). Our sampling scheme included fixed depths between 20 and 200 m and at deep stations also at 500 and 1000 m. Below, sampling depths were identified according to the temperature and salinity profile. In total 16 stations were visited, five deep stations on the continental slope and in the Drake Passage and 11 stations on the continental shelf. For exact locations and further details see table 13.2 and Fig. 13.1.

Tab. 13.2: List of stations and depths investigated for studies of the bacterioplankton and dissolved organic matter during cruise ANT-XXVIII/4.

Station	Date	Ship time	Depth (m)	Sampled Depths (m)	
PS79/178	16.03.12	1:30	3994	20-200	
PS79/179	16.03.12	15:19	4210	20-200	500, 1000
PS79/193	18.03.12	22:10	2095	20-200	500, 1000, 2000
PS79/205	20.03.12	20:30	3540	20-200	500, 1000, 2000, 3500
PS79/217	22.03.12	20:30	1845	20-200	
PS79/232	24.03.12	20:30	331	20-200	

Station	Date	Ship time	Depth (m)	Sampled Depths (m)	
PS79/235	25.03.12	13:00	316	20-200	
PS79/241	26.03.12	13:30	1845	20-200	500, 1000, 1800
PS79/244	27.03.12	12:30	394	20-200	
PS79/247	28.03.12	10:30	316	20-200	
PS79/251	29.03.12	08:30	436	20-200	
PS79/262	30.03.12	20:30	4420	20-200	500, 1000, 2000
PS79/267	01.04.12	04:00	2190	20-200	
PS79/277	02.04.12	21:00	1484	20-200	
PS79/281	03.04.12	23:00	740	20-200	350, 600
PS79/287	04.04.12	22:00	3644	20-200	500, 1000, 2500, 3600

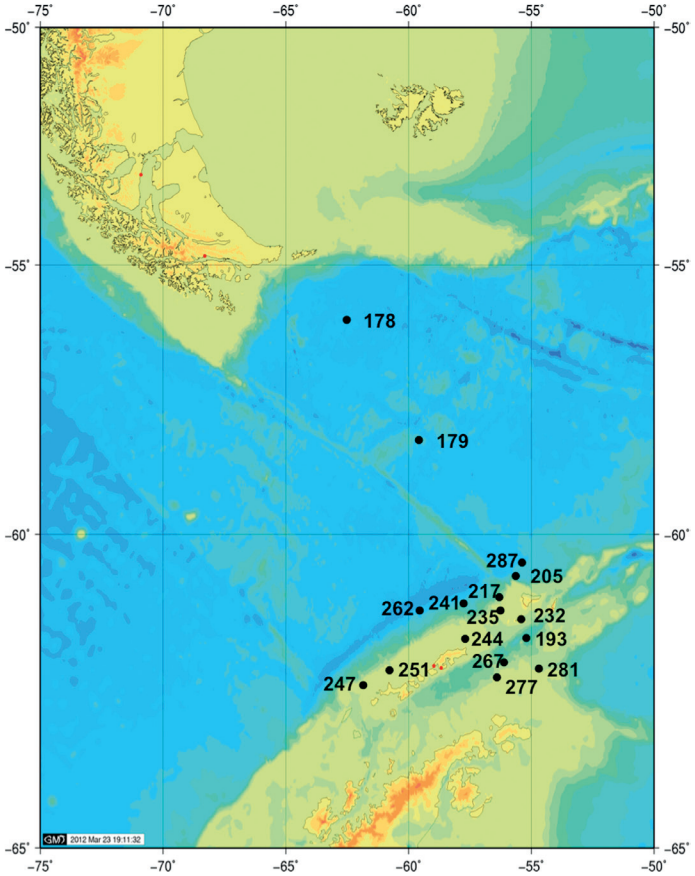


Fig. 13.1: Map with stations where samples for this study were collected during cruise ANT-XXVIII/4

In most cases, samples were filtered onto membrane filters of various type and size, frozen at -20 or -80 °C. Further processing will be done in the home laboratories. Samples for phytoplankton and nutrient analyses were fixed with Lugol's solution and HgCl, respectively, and will be analyzed further at home. Samples for the analyses of dissolved amino acids, carbohydrates, vitamins, DOC and DOM were prefiltered and frozen or acidified (DOC). Samples for bacterial abundance, production and turnover of dissolved free amino acids and glucose were analyzed on shipboard. Bacterial abundance was assessed by flow cytometry and bacterial production and substrate turnover by radiotracer techniques and applying 14C-leucine, 3H-leucine, -glucose and –amino acids. For details on the methods see Simon and Azam (1989) and Simon and Rosenstock (2007).

Preliminary results

The analysis of the water masses from the CTD and sigma-T data showed that we sampled quite different water masses (Badewien et al., chapter 18 this volume). The data on bacterial abundance, production and turnover rates of amino acids and glucose in most cases did not reflect these water masses, but rather differences between the stations on the continental shelf, the slope and Drake Passage. The more productive stations on the shelf exhibited consistently higher rates of bacterial biomass production and also shorter generation times of the bacterioplankton communities than those on the continental slope and the Drake Passage (Fig. 13.2 and 13.3). Whereas the values of bacterial biomass production on the continental slope and in the Drake Passage are similar to data from a previous study in this region (Manganelli et al. (2009) the values on the shelf are on average significantly higher than those of previous studies (Manganelli et al.2009, Straza et al. 2010). Bacterioplankton bulk generation times of as fast as 1 - 2 days in the mixed layer of this cold, but obviously very productive region were most remarkable. These findings are surprising, considering that the study was carried out in the late austral fall and at *in-situ* temperatures of 0 - 1 °C and including stations close to the ice edge. We might have encountered a particularly productive situation, reflected also by a surprisingly high abundance of fin wales and krill in this region (chapter 3, 16 this volume). Below 100 m, bacterial activity, as shown in the form of generation time, was much lower than above as a result of the much lower substrate supply below the mixed layer. The data on concentrations of chlorophyll, dissolved amino acids and carbohydrates and DOC will provide us with further details on the availability of substrates to the bacterioplankton.

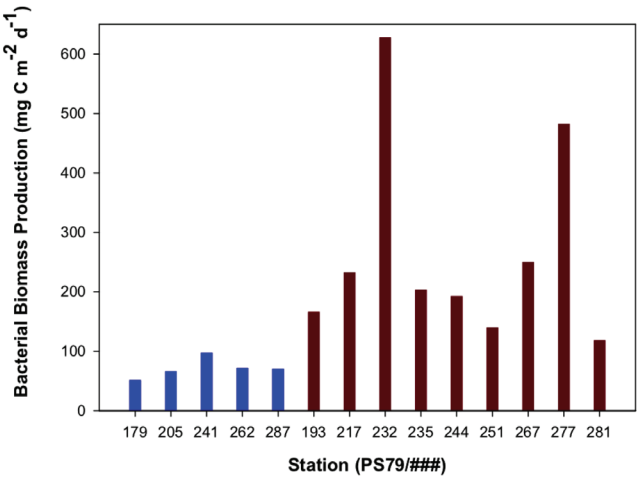


Fig. 13.2: Bacterial biomass production integrated from 0 to 200 m at all stations visited during cruise ANT-XXVIII/4. Stations on the continental slope and the Drake Passage (179-287) are shown on the left in blue and those on the continental shelf on the right in dark red (193-281).

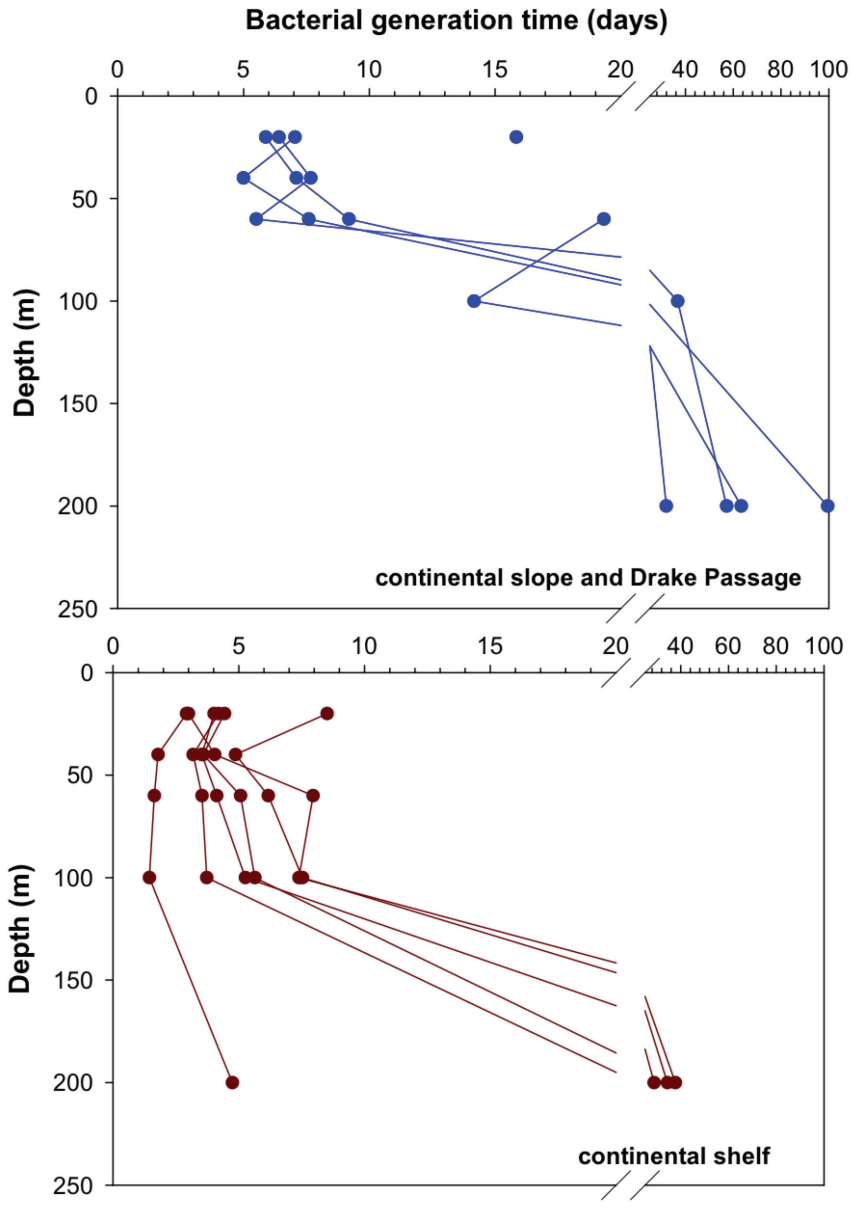


Fig. 13.3: Bacterioplankton bulk generation times of various profiles on the continental slope and the Drake Passage (upper panel) and on the continental shelf (lower panel) during cruise ANT-XXVIII/4.

The most interesting question is whether and how these differences will be reflected in the composition and metabolic activities of the bacterioplankton communities at these stations. Our analyses hopefully will shed light on this issue.

These results and those of all other parameters for which samples were collected, however, will become available only after processing of the samples in the home laboratories after several months to years.

Data management

All finally processed data will be stored on a server at ICBM and of TRR 51 and will be available on request if not otherwise mentioned. Data on the pyrosequencing and metagenomics and metatranscriptomics will be processed and stored on a server of the Göttingen Genomics lab at the University of Göttingen and at HZI. The final data of the metagenomics and metatranscriptomics and on sequenced genes will be made publicly available via GenBank and NCBI. Most of the data will be published in international peer-reviewed journals.

References

Manganelli M, Malfatti F, Samo TJ, Mitchell BG, Wang H, Azam F (2009) Major role of microbes in carbon fluxes during Austral winter in the Southern Drake Passage. PLoS ONE 4, e6941.

Simon M, Azam F (1989) Protein content and protein synthesis rates of planktonic marine bacteria. Marine Ecology Progress Series 51, 201-213.

Simon M, Rosenstock B (2007) Different coupling of dissolved amino acid, protein, and carbohydrate turnover to heterotrophic picoplankton production in the Southern Ocean in austral summer and fall. Limnology and Oceanography 52, 85-95.

Straza TRA, Ducklow HW, Murray AE, Kirchman DL (2010) Abundance and single-cell activity of bacterial groups in Antarctic coastal waters. Limnology and Oceanography 55, 2526–2536.

STUDY 7:

HIGH ABUNDANCE OF HETEROTROPHIC PROKARYOTES IN HYDROTHERMAL SPRINGS OF THE AZORES AS REVEALED BY A NETWORK OF 16S RRNA GENE-BASED METHODS

SAHM K[1], JOHN P[1], NACKE H[2], **WEMHEUER B**[2], GROTE R[1], DANIEL R[2], AND ANTRANIKIAN G[1]

EXTREMOPHILES 17(4): 649-62.

INSTITUTE FOR TECHNICAL MICROBIOLOGY, HAMBURG UNIVERSITY OF TECHNOLOGY, KASERNENSTR. 12, 21073 HAMBURG, GERMANY; [2]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY;

Author contributions to the work:

Performed the experiments: PJ

Analyzed data: KS, HN, BW

Contributed data on environmental properties and analysis of these data: PJ

Contributed to sampling: RG

Wrote the publication: KS

Conceived and designed the experiments: KS, GA

Extremophiles (2013) 17:649–662
DOI 10.1007/s00792-013-0548-2

ORIGINAL PAPER

High abundance of heterotrophic prokaryotes in hydrothermal springs of the Azores as revealed by a network of 16S rRNA gene-based methods

Kerstin Sahm · Patrick John · Heiko Nacke · Bernd Wemheuer · Ralf Grote · Rolf Daniel · Garabed Antranikian

Received: 21 March 2013 / Accepted: 12 May 2013 / Published online: 26 May 2013

Abstract Two hydrothermal springs (AI: 51 °C, pH 3; AIV: 92 °C, pH 8) were analysed to determine prokaryotic community composition. Using pyrosequencing, 93,576 partial 16S rRNA gene sequences amplified with V2/V3-specific primers for Bacteria and Archaea were investigated and compared to 16S rRNA gene sequences from direct metagenome sequencing without prior amplification. The results were evaluated by fluorescence in situ hybridization (FISH). While in site AIV Bacteria and Archaea were detected in similar relative abundances (Bacteria 40 %, Archaea 35 %), the acidic spring AI was dominated by Bacteria (68 %). In spring AIV the combination of 16S rRNA gene sequence analysis and FISH revealed high abundance (>50 %) of heterotrophic bacterial genera like *Caldicellulosiruptor*, *Dictyoglomus*, and *Fervidobacterium*. In addition, chemolithoautotrophic Aquificales were detected in the bacterial community with *Sulfurihydrogenibium* being the dominant genus. Regarding Archaea, only Crenarchaeota, were detected, dominated by the family Desulfurococcaceae (>50 %). In addition, Thermoproteaceae made up almost 25 %. In the acidic spring (AI) prokaryotic diversity was lower than in the hot, slightly alkaline spring AIV. The bacterial community of site AI was dominated by organisms related to the chemolithoautotrophic genus *Acidithiobacillus* (43 %), to the heterotrophic *Acidicaldus* (38 %) and to *Anoxybacillus* (7.8 %). This study reveals differences in the relative abundance of heterotrophic versus autotrophic microorganisms as compared to other hydrothermal habitats. Furthermore, it shows how different methods to analyse prokaryotic communities in complex ecosystems can complement each other to obtain an in-depth picture of the taxonomic composition and diversity within these hydrothermal springs.

Keywords Biodiversity · Thermoacidophile ecology · Thermophile ecology · Metagenome

Communicated by M. da Costa.

K. Sahm (✉) · P. John · R. Grote · G. Antranikian
Institute for Technical Microbiology, Hamburg University of Technology, Kasernenstr. 12, 21073 Hamburg, Germany
e-mail: sahm@tuhh.de

H. Nacke · B. Wemheuer · R. Daniel
Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University Göttingen, Göttingen, Germany

Introduction

With regard to their physicochemical conditions, terrestrial hydrothermal springs are highly diverse habitats, offering a wide range of ecological niches. These niches exhibit extreme conditions, such as high temperature, low or high pH, and the presence of toxic ions. Their extreme features are expected to lead to limited biodiversity, making hydrothermal habitats ideal model systems to study principles of community structure and function. Prokaryotic diversity in hydrothermal ecosystems has been extensively studied, in particular at Yellowstone National Park repeatedly revealing characteristic taxonomic groups like Aquificales and specific groups of Crenarchaeota, in particular (among others Hamamura et al. 2012; Meyer-Dombard et al. 2005; Spear et al. 2005).

The abundant, often predominant occurence of Aquificales has led to the hypothesis that H_2-based chemolithoautotrophic processes play a dominant role in hot-temperature ecosystems, in which photosynthesis is no longer possible (Spear

et al. 2005; Inskeep et al. 2010; Hugenholtz et al. 1998). Aquificales often form the basis for chemolithotrophic biofilm or mat structures, e.g. the pink filaments of *Thermocrinis ruber* and the white sausage-shaped filaments found in sulphur-turf mats of Japanese hot springs (Blank et al. 2002; Yamamoto et al. 1998).

Estimates on community composition, however, have been performed mainly on the basis of clone libraries with the known bias of cloning-based technologies and relatively low sample size taken into account (among others Blank et al. 2002; Hugenholtz et al. 1998; Reysenbach et al. 1994; Swingley et al. 2012). With next generation sequencing methods the analysis of larger 16S rRNA gene-data sets has become feasible, making it possible to increase the extent of sequence-based biodiversity studies, leading to a higher degree of reliability. Analysis of large sets of short rRNA gene fragments comprising variable regions has been applied to infer community composition and biodiversity estimates in terrestrial and marine habitats (Huse et al. 2008; Miller et al. 2009; Nacke et al. 2011; Roesch et al. 2007; Will et al. 2010; Youssef et al. 2009). However, the analysis is usually based on PCR-amplified DNA fragments, which still suffers from potential primer bias. Direct sequencing of metagenomic DNA can circumvent this bias, reducing, however, the amount of data available for analysis since non-16S rRNA gene sequences are also generated (Inskeep et al. 2010). To obtain an in-depth picture of prokaryotic hydrothermal spring communities, we combined both sequencing of short variable rRNA gene fragments and phylogenetic analysis of the metagenome, evaluating it with fluorescence in situ hybridization (FISH), which allows direct quantification of specific microbial lineages in the sample, directly. Our study was conducted with samples from the Furnas Valley, situated on São Miguel, Azores.

The Azores are a group of geographically isolated islands in the Atlantic Ocean. They are almost exclusively of volcanic origin, with the main hydrothermal area being the Furnas Valley on the Island of Sao Miguel. However, compared to other hydrothermal zones like Yellowstone National Park and Iceland the overall amount of data available is still limited. Brock and Brock (1967) gave a detailed description of the hot springs of Furnas Valley pointing out the thermal and chemical complexity of the Furnas springs and a major difference with regard to other terrestrial hydrothermal areas: in Furnas "… the largest spring is at the highest elevation and is alkaline, whereas some of the lower springs are smaller and more acidic. This is completely opposite to the situation found in Yellowstone National Park and Iceland… (Allen and Day 1935; Barth 1950)". Here the higher springs are small and acidic and the lower ones large and alkaline. The hot springs of Furnas Valley have been a valuable source for thermophilic microorganisms and thermostable enzymes with useful features for biotechnological applications (among others Albuquerque et al. 2008, 2010, 2012; Franca et al. 2006; Riessen and Antranikian 2001; Friedrich and Antranikian 1996). However, only little data are available on overall diversity (Hamamura et al. 2012).

Applying different 16S rRNA gene-based and metagenomic approaches, we provide an in-depth study of bacterial and archaeal diversity as a quantitative basis for understanding ecological interactions within the prokaryotic community. The study reveals the wealth of biocatalytic potential for enzymes from heterotrophic organisms still waiting to be recovered from these challenging extreme habitats.

Materials and methods

Study site and sampling

Sediment, biofilm, and water samples were collected from various hydrothermal springs at Furnas Valley, São Miguel, Azores, in September 2010. All necessary permits were obtained for the described field studies from the regional government of the Azores (LICENCA No 107/2010/DRA). A description of the different sampling sites is provided in Table 1. Several samples were collected at a solfataric field (sampling site A) (Fig. 1). Additional samples were collected from the hydrothermal vents Caldeirão (sampling site B) and Caldeira do Asmodeu (sampling site C). With the exception of the solfatara sample AIV (water sample) mainly sediments, muds or biofilms were collected and supplemented with water from the respective hydrothermal spring. Sediments and mud were characterized by brownish, ochre, grey, black or green appearance. The biofilms were black, white or light-grey, and the microbial mats encountered along the effluent stream of one hydrothermal vent exhibited green, orange, yellowish or brownish colours (Fig. 1).

Each sample was transferred to a sterile serum bottle, preventing exposure to air as far as possible. After sampling, the bottles were closed immediately with a butyl rubber stopper and a screw cap and 0.01 % w/v (final concentration) sodium sulphide was added by means of a syringe to counteract potential oxygen intake. Within 4 h after sampling, aliquots of 500 µL per sample were fixed for 1 h with 3 % (final concentration) paraformaldehyde for subsequent FISH analyses as described by Ravenschlag et al. (2000) and stored at −20 °C. The original samples were stored at 4 °C.

Table 1 Sampling site characteristics

Site	*T* (°C)	pH	Biotope	Sample
A	51	3	Small shallow hydrothermal spring at solfataric field near Caldeira do Esguicho	AI
	84	2.5–3	Small shallow hydrothermal spring at solfataric field near Caldeira do Esguicho	AII
	85	8	Small shallow hydrothermal spring at solfataric field near Caldeira do Esguicho	AIII
	92	8	Solfatara (Caldeira do Esguicho), effluent stream	AIV
B	60	6	Hydrothermal spring (Caldeirão)	BI
	65	7	Small stream near Caldeirão (Chalet Quente)	BII
	70	6	Hydrothermal spring (Caldeirão, paved basin)	BIII
C	55	8	Effluent of Caldeira do Asmodeu	CI
	76	8	Hydrothermal spring (Caldeira do Asmodeu), edge zone	CII

Fig. 1 Solfataric field (study site A) with sulphur deposits and various hydrothermal springs of temperatures up to 97 °C and pH values varying between 2.5 and 8. The rising gases (mainly water vapour and hydrogen sulphide) originate from a solfatara, Caldeira do Esguicho (*red arrow*, sample AIV). Several sediment and water samples were collected at this site. The *green arrow* depicts site AI

Determination of chemical parameters

Aliquots of approx. 120 mL of the original samples were centrifuged at 3,800×*g* for 40 min. The supernatants were filtered (pore size of the filters <0.45 μm) and analysed for chemical parameters applying the following methods/instruments: optical emission spectroscopy with inductively coupled plasma (Optima 7000 DV OES/ICP, PerkinElmer, Inc.) (Na, K, Mg, Ca, Fe, S), ultraviolet–visible spectrophotometry (Lambda 25 UV–VIS, PerkinElmer, Inc.) (NH_4^+, Fe^{2+}, NO^{2-}), ion chromatography (Metrohm AG, Germany) (Cl^-, NO_3^-, SO_3^{2-}, SO_4^{2-}, PO_4^{3-}) and TOC analyser (highTOC, Elementar Analysensysteme GmbH, Germany) (DOC, TIC).

Isolation of DNA

Total community DNA was isolated as described by Zhou et al. (1996), based on lysis with high-salt extraction buffer and extended heating in the presence of sodium dodecyl sulphate (SDS). The samples were centrifuged at 3,939×*g* for 1 h and 2–3 g of the pellet was resuspended in DNA extraction buffer, containing 1 % hexadecylmethylammonium bromide (CTAB). Prior to DNA isolation, three freeze-and-thaw cycles in liquid nitrogen and a water bath at 65 °C were conducted. The precipitated DNA was dissolved in 120 μL sterile water and analysed spectrophotometrically at 260/280 nm and by agarose gel electrophoresis (0.4 % agarose). DNA solutions derived from samples with high proportions of sediments appeared intensively orange or yellowish, indicating the co-extraction of humic substances which interfere with UV-photometric measurement of DNA concentration. Thus, quantification of DNA extracted from these samples was merely done by agarose gel electrophoresis.

In addition, we applied the PowerMax™ Soil DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, USA) as an alternative method for isolation of genomic DNA from all samples. Microorganisms were lysed by a combination of SDS and further disruption agents, and mechanical force using beads. Approx. 2 g of sediment was used for DNA extraction. The DNA was eluted in 2 mL sterile water and quantified as described above. Pooled DNA from both methods was used for metagenomic pyrosequencing or PCR amplification.

Denaturing gradient gel electrophoresis

PCR amplification and DGGE were performed as described by Muyzer et al. (1998). We routinely added 3 μg/μL (final concentration) bovine serum albumin to the PCR reaction mixture to prevent inhibition of the DNA polymerase by humic substances. The sequences of the primers used for amplification of 16S rRNA gene-fragments were as follows: Arc334F: 5′-CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCGCCCGACGGGGYGCAGCAGGCGCGA-3′, Arc915R: 5′-GTGCTCCCCCGCCAATTCCT-3′ (Archaea); 314F: 5′-CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCGCCCGCCTACGGGAGGCAGCAG-3′, 907R: 5′-CCGTCAATTCMTTTGAGTTT-3′ (Bacteria). *Taq* DNA polymerase (Fermentas GmbH, St. Leon-Rot, Germany) was used in PCR. The touchdown PCR protocol by Muyzer (annealing temperatures ranging from 65 to 55 °C) was applied for both primer pairs. We used the Dcode™ System (Bio-Rad Laboratories GmbH, Munich, Germany) for DGGE. Electrophoresis was carried out using a 6 % polyacrylamide gel with a denaturing gradient of 20–80 %, at 55 °C, 200 V for 5 h (Bacteria-DGGE), and a denaturing gradient of 30–80 % at 60 °C, 150 V for 2 h (Archaea-DGGE). After electrophoresis, the gels were stained in Roti®-Gel Stain (Carl Roth GmbH + Co. KG, Karlsruhe, Germany) and photographed on a UV transilluminator.

Amplification of 16S rRNA gene fragments and pyrosequencing

DNA preparations from both isolation approaches were used as template for individual PCR amplification of 16S rRNA gene fragments. The PCR products derived from either DNA preparation were combined for sequencing in equal amounts. The sequences of the bacterial and archaeal V2/V3 region-specific primers used in respective PCRs are as follows: Bacteria, V2f: 5′-AGTGGCGGACGGGTGAGTAA-3′; V3r, 5′-CCGCGGCTGCTGGCAC-3′ [(Will et al. 2010) V3r modified]; Archaea, Arc113f: 5′-ACKGCTSAGTAACACGTGG-3′; Arc520r, 5′-TACCGCGGCKGCTGGCA-3′ [Arc113f, (Baker and Cowan 2004); Arc520r, this study]. The template-specific primer sequences were complemented with a GS FLX Titanium primer sequence including the sequence adaptor, a four-base library key sequence, and a MID (multiplex identifier) sequence at the 5′ end of the specific primer. The Amplicon Fusion Primers were commercially synthesized (Fermentas GmbH, St. Leon-Rot, Germany).

The PCR mixtures (final volume 50 μL) contained 10 μL fivefold Phusion HF buffer (New England Biolabs GmbH, Frankfurt am Main, Germany), 300 μg BSA, 0.2 mM final concentration of each dNTP, 0.4 μM final concentration of each primer, 0.5 U Phusion High-Fidelity DNA polymerase (New England Biolabs GmbH, Frankfurt am Main, Germany), and 30–70 ng DNA. For Archaea-specific PCR the following steps were conducted: initial denaturation at 98 °C for 30 s and 28 cycles of denaturation at 98 °C for 10 s, primer annealing for 20 s using a temperature gradient ranging from 68 °C to 55 °C (1 °C touchdown every two cycles) and extension at 72 °C for 15 s, followed by a final extension period at 72 °C for 10 min. For Bacteria-specific PCR the cycling scheme deviated from the one described above by the annealing temperature gradient which was from 68 to 58 °C over 22 cycles and six additional cycles at a constant annealing temperature of 58 °C. Subsequently, the PCR products were purified by employing the GeneJet™ Gel Extraction Kit (Fermentas GmbH, St. Leon-Rot, Germany) and quantified spectrophotometrically.

The isolated metagenomic DNA was used to create a 454-shotgun library following the GS Rapid library protocol (Roche). The library was sequenced with the Genome Sequencer FLX (Roche, Mannheim, Germany) using Titanium chemistry. Sequencing was performed by the Göttingen Genomics Laboratory. In total, 2,044,797 shotgun reads were generated by two complete sequencing runs.

The partial 16S rRNA gene amplicons were sequenced using the unidirectional sequencing protocol (Roche) and Titanium chemistry. One medium lane of a Titanium Picotiter plate was used for sequencing, resulting in 125,702 shotgun sequences.

The sequences have been deposited at NCBI database under the accession number: SRA059339.

Analysis of pyrosequencing-derived data

All pyrosequencing reads obtained from partial 16S rRNA gene amplicons were reassigned to samples AI and AIV based on the unique MIDs. Different scripts of the QIIME software pipeline (Caporaso et al. 2010) and additional programs were applied to perform pyrosequencing data preprocessing and downstream analyses. Removal of sequences <200 bp, and sequences containing more than two primer mismatches or long homopolymers (>8 bp) was performed by applying the script "split_libraries.py".

Pyrosequencing noise was removed by using the scripts "denoise_wrapper.py" and "inflate_denoiser.py". Subsequently, the programs cutadapt (Martin 2011) and UCHIME (Edgar et al. 2011) (reference database: Greengenes gold database) were applied to remove remaining primer sequences and chimeras, respectively.

Operational taxonomic units (OTUs) were identified at genetic distances of 3 and 20 %, representing the species and phylum level, respectively, according to Schloss and Handelsman (2005) by applying the QIIME script "pick_otus.py" (Schloss and Handelsman 2005). In addition, taxonomic classification of OTUs as well as calculation of rarefaction curves, the Shannon index and the Chao1 index were performed using QIIME (Shanon and Weaver 1949; Chao and Bunge 2002). With respect to the script "assign_taxonomy.py", related to taxonomic classification of OTUs, preprocessed sequences were compared to the SILVA ribosomal RNA database (release 108) using BLASTN (Pruesse et al. 2007). Furthermore, a customized script to remove all OTUs from the OTU table that have been classified as chloroplasts was applied.

Metagenomic 16SrRNA gene sequences with a minimum length of 100 bp were analysed in the same way by comparing them to the SILVA ribosomal RNA database (release 108) using BLASTN (Pruesse et al. 2007).

Fluorescence in situ hybridization

An 80 µL aliquot of the fixed sample was diluted 1:10 in PBS/ethanol and treated by sonication using Branson Sonifier 450 (Branson Ultrasonics Corporation, Danbury, USA) at a setting of 1 min, duty cycle 20 and output control 2, three times with 60 s lapse between each sonication step while cooled on ice. Subsequently, the sample was resuspended in 400 µL PBS/ethanol. In situ hybridization was conducted following the protocol by Snaidr et al. (1997). An aliquot of 2.5 µL was applied to a poly-L-lysine-coated microscopic slide. After dehydration, hybridization of the respective 5′-cyanine 3-labelled oligonucleotide probe to the ribosomal sequence in the target cells was carried out at 46 °C for 100–120 min. The slide was washed and stained with 2 µg/mL 4′,6′-diamidino-2-phenylindol (DAPI) for 5 min on ice. Cells were counted using a fluorescent microscope (Zeiss Axioplan) and means were calculated by counting 10–20 randomly chosen fields for each analysis, which corresponded to at least 1,000 Dapi-stained cells.

The FISH probes applied in this study were Eub338 (5′-GCTGCCTCCCGTAGGAGT-3′, specificity: domain Bacteria, 35 % formamide in hybridization buffer) (Amann et al. 1990), Arc915 (5′-GTGCTCCCCCGCCAATTCCT-3′, specificity: domain Archaea, 35 % formamide in hybridization buffer) (Stahl and Amann 1991), Cal450 (5′- CTCCCCGTCCAAAGAGGT-3′, specificity: genus *Caldicellulosiruptor*, 30 % formamide in hybridization buffer) (O-Thong et al. 2007), SHyd540 (5′-TCGCGCAACGTTCGGGACC-3′, specificity: genus *Sulfurihydrogenibium*, 60 % formamide in hybridization buffer), and Fer660 (5′-GTTCCGTCTGCCTCTGCC-3′, specificity: genus *Fervidobacterium*, 20 % formamide in hybridization buffer). Probes were commercially synthesized as 5′-indocyanine 3-labelled oligonucleotides (Fermentas GmbH, St. Leon-Rot, Germany). Specificity for the newly designed oligonucleotide probes SHyd540 and Fer660 was verified by hybridization to control strains that possess one mismatch position within the ribosomal target sequence (SHyd540: *Rhodothermus marinus* (DSMZ 4252); Fer660: *Fervidobacterium gondwanense* (DSMZ 13020) and *F. nodosum* (DSMZ 5306)). Under the given stringency discrimination of one mismatch was possible.

Results

Sampling site and sample characteristics

Samples were collected in September 2010 in the Valley of Furnas from nine sites with a wide variety of physico-chemical characteristics. In situ temperatures ranged from 51 to 92 °C, whereas pH values varied between 8 and 2.5 (Table 1). Two sampling sites (AI and AIV) were selected for in-depth analysis, including the determination of several chemical parameters (Table 2). Site AI had a temperature of 51 °C and an acidic pH of 3. The greyish muddy hot spring exhibited oxidized conditions with high concentrations of sulphate (12.5 mM) and iron, which was mainly present as Fe^{3+} (1,400 µM). Site AIV, with a temperature of 92 °C and a pH of 8, exhibited reduced conditions, indicated by the proportion of Fe^{2+} to Fe^{3+}, which can be taken as a proxy for redox conditions (Mathur et al. 2007). The concentrations of sulphate and iron were low (100 and 20 µM) and iron was mainly present as Fe^{2+}. Organic carbon content was high with 0.006 % (60 mg/L) at site AI and 0.03 % (300 mg/L) at site AIV.

Community fingerprinting using DGGE

Prokaryotic diversity was estimated using denaturing gradient gel electrophoresis (DGGE) based on 16S rRNA gene fragments amplified with Bacteria- and Archaea-specific primers. Results are shown in Figs. 2 and 3. Taking the number of bands in the DGGE-profile as diversity indicator, the highest bacterial diversity was detected at sites with temperatures between 55 and 85 °C and pH values between 7 and 8. Analysis of the DGGE-profiles revealed a

Table 2 Chemical analysis of sampling sites AI and AIV

	Na (mM)	K (mM)	Mg (mM)	Ca (mM)	Fe^{2+} (μM)	Fe^{3+} [(1)] (μM)	Fe (μM)	NH_4^+ (μM)	Cl^- (mM)
AI	6.0	1.5	0.7	1.1	152	1391	1543	388	1.0
AIV	19.7	0.7	0.1	0.3	22	2	25	26	5.4

	NO_2^- (μM)	NO_3^- (μM)	SO_3^{2-} (μM)	SO_4^{2-} (mM)	S (mM)	PO_4^{2-} (mM)	TIC (%)	DOC (%)
AI	<200	<16	<13	12.5	12.6	0.01	<0.0005	0.006
AIV	<200	<16	<13	0.1	0.3	0.1	0.009	0.028

Fe^{3+} was calculated as the difference of Fe and Fe^{2+}

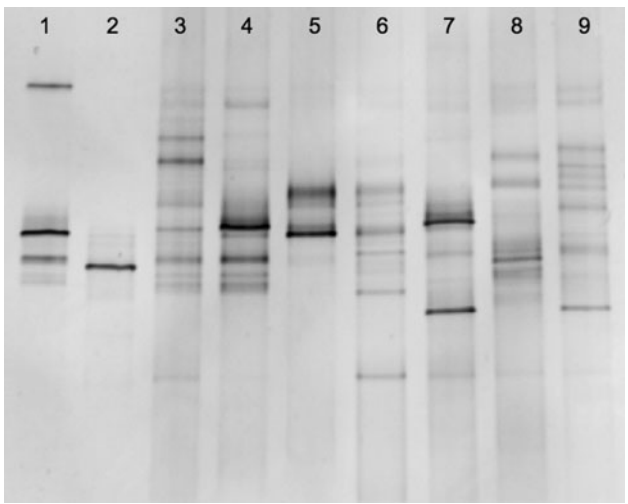


Fig. 2 DGGE analysis of 16S rRNA gene fragments amplified with Bacteria-specific primers. PCR products were analysed on a denaturing gradient from 20 to 80 %. The samples were grouped according to sampling site (A, B, C) and increasing temperature. The *lanes* can be attributed to the following sites: *1* AI (51 °C, pH 3), *2* AII (84 °C, pH 2.5–3), *3* AIII (85 °C, pH 8), *4* AIV (92 °C, pH 8), *5* BI (60 °C, pH 6), *6* BII (65 °C, pH 7), *7* BIII (70 °C, pH 6), *8* CI (55 °C, pH 8), *9* CII (76 °C, pH 8)

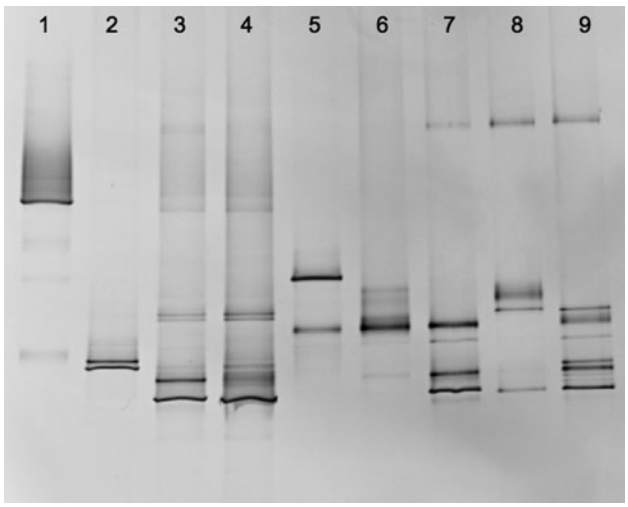


Fig. 3 DGGE analysis of 16S rRNA gene fragments amplified with Archaea-specific primers. PCR products were analysed on a denaturing gradient from 30 to 80 %. The samples were grouped according to sampling site (A, B, C) and increasing temperature. The *lanes* can be attributed to the following sites: *1* AI (51 °C, pH 3), *2* AII (84 °C, pH 2.5–3), *3* AIII (85 °C, pH 8), *4* AIV (92 °C, pH 8), *5* BI (60 °C, pH 6), *6* BII (65 °C, pH 7), *7* BIII (70 °C, pH 6), *8* CI (55 °C, pH 8), *9* CII (76 °C, pH 8)

prominent effect of pH. In particular, the combination of acidic pH and high temperature (pH 2.5–3 and 84 °C) led to a very low bacterial diversity (Fig. 2, lane 2) while at moderate temperature of 51 °C and pH 3 more bacterial bands were detected (Fig. 2, lane 1). Analysis of archaeal 16S rRNA genes using DGGE revealed an overall lower number of bands (Fig. 3). Sites AIII and AIV exhibit an almost identical band-pattern despite the difference in temperature (85 vs. 92 °C).

Based on site-specific characteristics and the DGGE results we selected two sites (AI and AIV) for further analysis. Both sites are connected, with the outflow of AIV running into AI. Furthermore, they exhibit medium diversity and particularly challenging conditions with regard to pH (AI: pH 3) or temperature (AIV: 92 °C).

Community composition determined by fluorescence in situ hybridization

To quantify the relative contributions of Bacteria and Archaea to the communities of sites AI and AIV, we applied FISH using domain-specific probes (Fig. 4). The acidic hydrothermal spring AI was dominated by Bacteria (68 %). Archaea could not be detected at all and 32 % of the cells could only be stained with DAPI. At the high-temperature site AIV (92 °C) 75 % of all cells could be detected by FISH. The relative contribution of the specific domains was almost even with 35 % Archaea and 40 % Bacteria. Total cell number determined by DAPI-staining was 4.6×10^7/ml for AI and 2.6×10^8/ml for AIV.

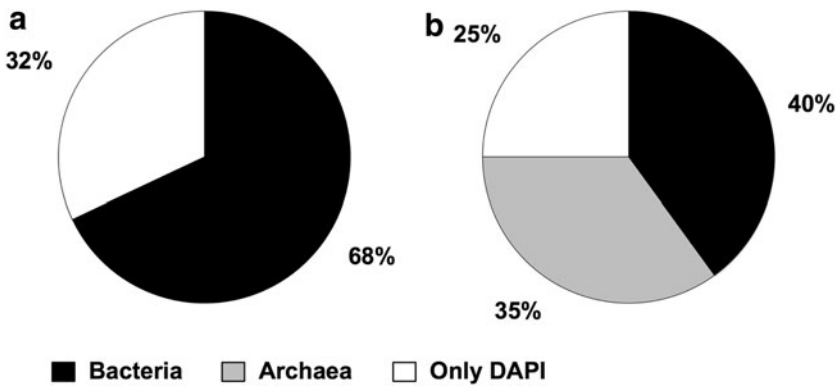


Fig. 4 Domain-level composition of microbial communities AI and AIV as determined by FISH. For Bacteria EUB333 was used and for Archaea probe Arch915. Relative abundance is based on the total cell number determined by DAPI-staining. **a** Depicts the results for site AI and **b** for site AIV

Phylogenetic analyses based on 16S rRNA gene sequences

To evaluate prokaryotic diversity of the selected sites, partial 16S rRNA genes were amplified from metagenomic DNA and sequenced by pyrosequencing. Based on former studies, we decided to amplify the V2/V3 region for both, Bacteria and Archaea, comprising 435 bp (Nacke et al. 2011; Yu et al. 2008; Will et al. 2010; Kysela et al. 2005). While primers for Bacteria were already available, new primers to amplify the corresponding region in the archaeal 16S rRNA gene had to be designed. The forward primer was based on the data of Baker and Cowan with slight modifications (Baker and Cowan 2004) while the reverse primer was designed de-novo. A total number of 125,702 sequences were generated. After preprocessing including quality filtering and denoising, 93,576 sequences with an average length of 390 bases could be obtained for further analyses. The number of total reads of every sample and remaining numbers of reads after quality filtering and denoising are depicted in Table 3. Amplification was specific, yielding the expected domain-specific sequences apart from archaeal amplification of DNA from site AIV.

Furthermore, metagenomic DNA derived from site AIV was analysed using direct pyrosequencing without prior amplification. Of the data generated 0.1 % represented partial 16S rRNA genes. The 16S rRNA gene fragments $f \geq 100$ bp were extracted from the dataset. The resulting 725 partial 16S rRNA gene sequences were analysed by comparing them to the SILVA ribosomal RNA database using BLASTN.

Comparison of the sequencing approaches

Combined results of the two sequencing approaches are shown in Tables 4, 5 and 6. The acidic spring AI was dominated by Proteobacteria (>80 %) and Firmicutes (10 %) (Table 4) while site AIV was dominated by bacteria belonging to 4 different phyla: Thermotogae, Firmicutes, Dictyoglomi, and Aquificae (Table 5). Each of these phylum-groups mainly consisted of one dominating genus. The acidic spring AI was dominated by Proteobacteria known for their acidophilic properties: the heterotrophic genus *Acidicaldus* (38 %) and the chemolithoautotrophic *Acidithiobacillus* (43 %). In addition, almost 8 % of the sequences was related to *Anoxybacillus* (Table 4). Within site AIV (92 °C, pH 8) members of *Dictyoglomus*, *Caldicellulosiruptor*, and *Fervidobacterium* constituted up to 61 % of the bacterial community according to data resulting from non-amplified pyrosequencing of metagenomic DNA, whereas analysis of the amplified V2/V3 16S rRNA gene region indicated an even higher abundance of these heterotrophic genera (88 %) (Table 5). Data from direct metagenome sequencing also revealed high abundance of two genera from the phylum Aquificae, with the chemolithoautotrophic *Sulfurihydrogenibium* being the dominant genus (22 %).

Table 3 Evaluation of V2/V3 pyrosequencing data

	Total sequences	Preprocessed sequences	Average read length	Assigned to Bacteria (%)	Assigned to Archaea (%)
AI—Bact.	37,044	28,565	385	97.0	3
AI—Arch.	49,573	35,607	386	1.3	98.7
AIV—Bact.	39,085	29,404	400	99.9	0.1

The archaeal community of the hot spring AIV (92 °C) was almost exclusively composed of Crenarchaeota while in the acidic spring AI only Euryarchaeota were detected. Table 6 depicts the detailed composition. The archaeal community of site AI mainly harboured *Thermoplasma* (89 %), while clone group BLSdp215 represented the second largest group (9 % of all sequences). Site AIV was dominated by Desulfurococcaceae (75 %) with the genera *Sulfophobococcus*, *Desulfurococcus* being most abundant within this family. Furthermore, the genus *Pyrobaculum* of the Thermoprotaceae made up 25 % of the sequences.

Table 4 Composition of the bacterial community of AI determined by V2/V3 sequencing

Only genera that occur at an abundances of above 0.5 % are shown

Dominant taxa are given in bold

Phylum	Family	Genus	Rel. abundance (%) AI V2/V3
Proteobacteria (α)	**Acetobacteraceae**	***Acidicaldus***	**38**
		Uncultured	0.6
Proteobacteria (γ)	**Acidithiobacillaceae**	***Acidithiobacillus***	**43**
Firmicutes	**Bacillaceae**	***Anoxybacillus***	**7.8**
	Clostridiaceae	*Caloramator*	1.2
		Clostridium	0.5
Nitrospirae	Nitrospiraceae	*Leptospirillum*	0.9
Actinobacteria	Acidimicrobiaceae	*Acidimicrobium*	0.7

Table 5 Detailed comparison of the bacterial community composition of AIV determined by V2/V3 or metagenome sequencing

Phylum	Family	Genus	Rel. abundance (%)	
			AIV V2/V3	AIV metagenome
Aquificae	**Hydrogenothermaceae**	***Sulfurihydrogenibium***	**0.4**	**22**
	Aquificaceae	*Hydrogenobacter*	0.1	6
Dictyoglomi	**Dictyoglomaceae**	***Dictyoglomus***	**22**	**27**
Deinococcus-Thermus	Thermaceae	*Thermus*	0.8	2
Firmicutes	**Thermoanaerobacteraceae**	***Caldicellulosiruptor***	**19**	**21**
	Clostridiaceae	*Clostridium*	4.9	0.6
	Peptococcaceae	*Desulfosporosinus*	ND	1
	Ruminococcaceae	*Incertae Sedis*	1.63	ND
Thermotogae	**Thermotogaceae**	***Fervidobacterium***	**47**	**13**
		Thermotoga	0.7	0.8
Thermodesulfobacteria	Thermodesulfobacteriaceae	*Thermodesulfobacterium*	ND	1
Nitrospirae	Nitrospiraceae	*Thermodesulfovibrio*	0.72	0.4

Only genera that occur at an abundances of above 0.5 % in at least one of the two approaches are shown

Dominant taxa are given in bold

ND not detected

Table 6 Detailed comparison of the archaeal community composition of AI and AIV determined by metagenome sequencing and V2/V3 sequencing, respectively

Phylum	Family	Genus	Rel. abundance of archaea (%)	
			AI	AIV[a]
Crenarchaeota	**Desulfurococcaceae**	***Sulfophobococcus***	**ND**	**55**
		Desulfurococcus	**ND**	**19**
		Staphylothermus	ND	0.5
		Stetteria	ND	0.5
	Thermoproteaceae	***Pyrobaculum***	**ND**	**25**
Euryarchaeota	**Thermoplasmataceae**	***Thermoplasma***	**89**	**ND**
		BSLdp215 uncultured	9	ND
		A10 uncultured	1	ND

Only genera that occur at an abundances of above 0.5 % are shown

Dominant taxa are given in bold

ND not detected

[a] Data are based on the metagenome analysis

Microbial diversity indices

Hot springs are usually referred to as low-diversity habitats. Diversity indices are well established measures to quantify diversity on a parametric (Shanon) or non-parametric (Chao1) basis. While the Shanon index takes species richness and species evenness into account, the Chao index is based on species richness, with the obtained value giving the number of expected operational taxonomic units (OTUs). To calculate rarefaction curves, richness, and diversity indices, OTUs at genetic distances of 3 and 20 % were determined. With respect to the Bacteria, OTU-based analyses were performed at the same level of surveying effort by using 22,100 randomly selected sequences per sample to enable accurate OTU-based comparisons between site AI and AIV.

Comparison of the rarefaction analyses with the number of OTUs determined by Chao1 richness estimator revealed that 76.2–90.6 % (20 % genetic distance) and 54.5–77.7 % (3 % genetic distance) of the taxonomic richness were covered by the surveying effort (Fig. 5; Table 7). Thus, we did not survey the full extent of taxonomic diversity at these genetic distances, but especially at the phylum level a substantial fraction of the estimated richness was assessed by the surveying effort. The bacterial diversity at sites AI and AIV exhibits an interesting difference: on the phylum level the acidic spring is more diverse than the slightly alkaline high-temperature site AIV, with 50 versus 30 OTUs (Table 7), whereas on the species level the opposite pattern was discovered (AI: 178 OTUs, AIV: 236 OTUs).

Using Chao1 richness estimator, the maximum number of OTUs for the acidic spring bacterial community predicted at 3 % genetic distance was 229 belonging to 59 predicted phyla. With respect to site AIV (92 °C, pH8), 371 bacterial OTUs on species level and 33 OTUs on the phylum level were predicted by Chao1 richness estimator. The Shanon index (H'), determined for both habitats, indicated a slightly lower diversity in site AI than in site AIV (3 % genetic distance: AI: 1.8, AIV: 2.3; 20 % genetic distance: AI: 0.8, AIV: 1.0).

Archaeal diversity could only be estimated for site AI and was very low (H': 0.54 at 3 % and 0.09 at 20 % genetic distance) as expected due to the high dominance of *Thermoplasma* (Table 6). At 3 % sequence divergence a maximum number of 99 OTUs were predicted while 21

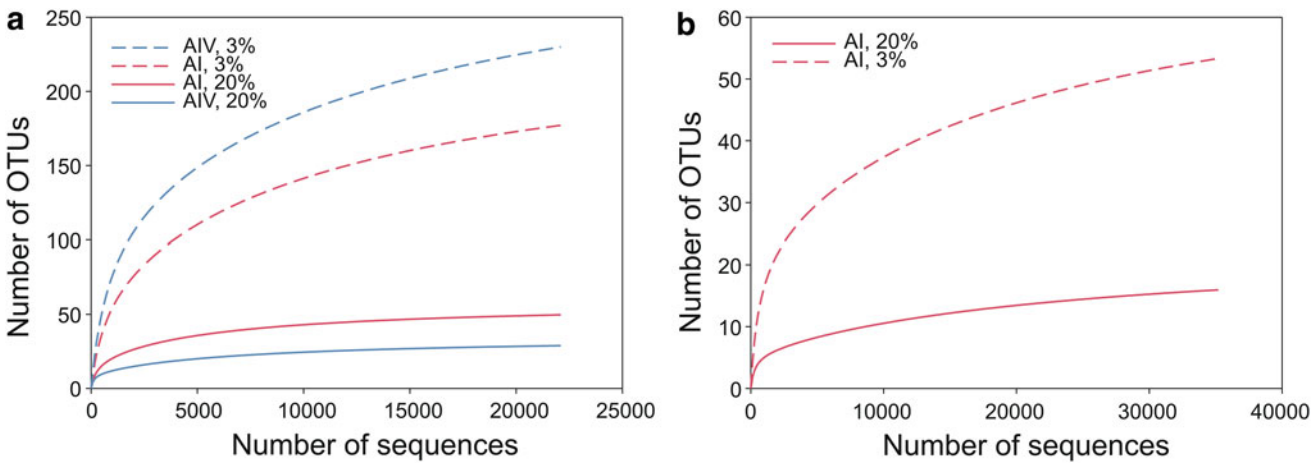


Fig. 5 Rarefaction analysis of V2/V3 sequencing results for Bacteria and Archaea. Rarefaction curves were calculated on the basis of 22,100 sequences for Bacteria (**a**) and based on 35,200 sequences for Archaea (**b**). The *curves* indicate the observed number of OTUs within the two samples AI (51 °C, pH 3) and AIV (92 °C, pH 8). OTUs are shown at the level of 3 and 20 % genetic distance

Table 7 Diversity estimates based on V2/V3 sequencing

Distance	Shannon index (H')		Rarefaction (no. of OTUs)		Chao1 (no. of OTUs)		Coverage (%)	
	3 %	20 %	3 %	20 %	3 %	20 %	3 %	20 %
AI Bacteria[a]	1.80	0.83	178	50	229	59	77.7	84.7
AI Archaea[b]	0.54	0.09	54	16	99	21	54.5	76.2
AIV Bacteria[a]	2.29	1.04	231	29	337	32	68.5	90.6

[a] No. of sequences analysed 22,100

[b] No. of sequences analysed 35,200

OTUs were predicted at 20 % sequence divergence using Chao1 richness estimator.

Evaluation of the pyrosequencing-based analyses by group-specific fluorescence in situ hybridization

At site AIV relative abundance of *Sulfurihydrogenibium* and *Fervidobacterium* differed depending on the sequencing approach (Table 5). While sequencing of non-amplified metagenomic DNA indicated high abundance of Aquificae (approx. 28 %), in particular *Sulfurihydrogenibium*, the sequencing of partial 16S rRNA gene amplicons hardly detected any. On the other hand, amplicon sequencing indicated dominance of *Fervidobacterium* (approx. 47 %). We applied FISH to quantify and further analyse these specific groups including also *Caldicellulosiruptor*, which represented approx. 21 % of the bacterial community according to data derived from direct sequencing and pyrosequencing of the V2/V3 16S rRNA gene region. *Dictyoglomus*, the fourth dominant genus in sample AIV could not be evaluated because no specific probe could be designed.

Two probes had to be designed de-novo. The probes specific for *Sulfurihydrogenibium* and *Fervidobacterium* were developed on the basis of signature sequences described by Harmsen et al. (1997), which had been specific for *Hydrogenobacter* and *Thermotoga*. An exchange of one and two bases, respectively, allowed to change the specificity to the genera in question. Primer specificity was confirmed in-silico using ProbeMatch of the RDP Database Project and hybridization conditions were experimentally verified using controls with one mismatch. At the stringency conditions applied, no signal could be detected in cells with 1 mismatch to the probe. A comparison of all three approaches to quantify dominant bacterial genera is shown in Fig. 6. Quantification by FISH underlined the high abundance of *Caldicellulosiruptor*, revealing that this genus represented 28 % of all Bacteria (detected with EUB-specific probe) as compared to 19 and 21 % detected by sequencing of partial 16S rRNA gene amplicons and direct sequencing of metagenomic DNA, respectively. Furthermore, FISH indicated that *Fervidobacterium* accounted for 22 % of the bacterial community and thus had been overestimated by analysis of the V2/V3 16S rRNA region (47 %) and underestimated according to data derived from direct sequencing of metagenomic DNA (14 %). While *Sulfurihydrogenibium* was almost undetected by sequencing of the V2/V3 16S rRNA gene region (0.4 %), the genus was overestimated by direct sequencing of metagenomic DNA (17 %) since FISH revealed it to account for 9 % of the Bacteria. Altogether, almost 60 % of the bacterial cells could be assigned to the three genera

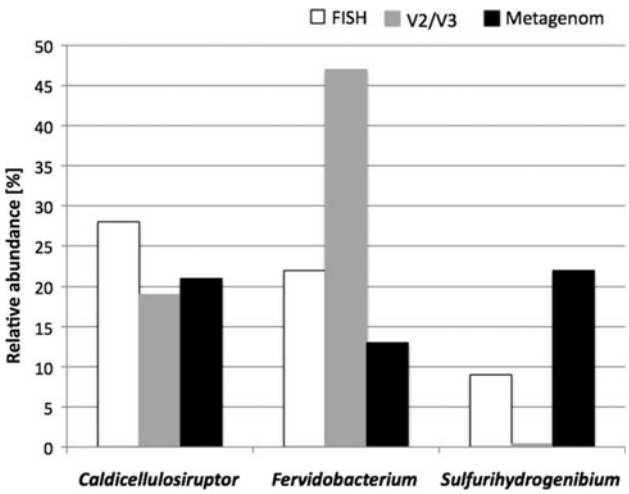


Fig. 6 Comparative quantification using FISH, V2/V3 and metagenome sequencing. The relative abundance relates to % of bacterial cells for FISH, % of bacterial rRNA gene sequences for V2/V3 and metagenome sequences

Caldicellulosiruptor, *Fervidobacterium*, and *Sulfurihydrogenibium* using FISH.

Discussion

Prokaryotic diversity

Hot springs are considered low-diversity habitats due to the high physicochemical constraints and the high degree of geographical isolation (Papke et al. 2003). The effect of temperature on prokaryotic diversity has been studied and results are contradictory. Yim et al. (2006) did not find a linear correlation between prokaryotic diversity and temperature in mats and streamers at temperatures between 52 and 69 °C in alkaline geothermal springs from Tibet. On the other hand, Miller et al. suggested that bacterial diversity along a temperature gradient from 39 to 72 °C in two alkaline hot springs from Yellowstone National Park is primarily controlled by temperature in photosynthetically active microbial mats which has also been shown by Hiraishi et al. in hot springs in Japan (Yim et al. 2006; Miller et al. 2009; Hiraishi et al. 1999). Furthermore, Skirnisdottir et al. (2000) found the lowest bacterial diversity at the highest temperature (80 °C) and highest sulphide concentration when comparing hot springs in Iceland with low and high concentration of sulphide. Based on DGGE analysis of nine sites, we found high bacterial diversity over a range of temperature between 55 and 85 °C as long as the pH was neutral to slightly alkaline. With respect to Archaea, the temperature range was slightly shifted upwards from 70 to 92 °C. In the

hydrothermal springs investigated here, lowest bacterial diversity was detected at acidic pH and high temperature.

If phylogenetic distance is taken into account, sites AI and AIV display a distinctive difference. OTU-based analyses showed that bacterial diversity was more phylum rich in site AI compared to site AIV. However, site AIV was more species rich than site AI. Similar observations have been published previously (Roesch et al. 2007; Yim et al. 2006). Yim et al. suggested that extreme stress might inhibit the survival of large numbers of closely related guild taxa in microbial population from hot spring sediments and streamers leading to higher phylum numbers (Yim et al. 2006). Roesch et al. (2007) found that agricultural soil was phylum poor and species rich, whereas undisturbed forest soil showed the opposite pattern. However, the authors could not find a clear explanation for this. The factors causing the differences between the two habitats AI and AIV investigated here remain elusive. It is possible that coping with low pH is a property more widespread in the phylogenetic tree on phylum level than coping with temperatures above 90 °C. On the other hand it could also mean, that a habitat with temperature above 90 °C allows faster speciation than the low pH habitat. As long as there is no additional data available to show that this is not a coincidental observation, we can only speculate. Further comparative studies will be needed to evaluate this and to establish, whether it also holds for archaeal populations.

Diversity indices help to compare habitats. However, only few data are available on diversity indices from hot springs. For circumneutral springs at temperatures between 73 and 86 °C Costa et al. (2009) calculated Shanon indices between 2.02 and 2.37 at 3 % genetic distance for Bacteria and between 1.18 and 2.7 for Archaea. At a genetic distance of 20 %, the values were between 1.82 and 2.29 for Bacteria and between 0.95 and 2.16 for Archaea. For slightly alkaline springs along a temperature gradient from 52 to 69 °C Lau et al. (2009) calculated slightly higher Shanon indices between 3 and 4 at 3 % genetic distance for Bacteria. In acidic springs at temperatures between 60 and 75 °C Mathur et al. (2007) calculated Shanon indices for bacteria at 3 % genetic distance between 0.83 and 3.28 at springs with high sulphide and low iron concentration and between 2.1 and 3.6 in springs with high iron and low sulphide concentration. Shanon indices for bacteria calculated for sites AI and AIV fall within the same range. The acidic spring AI, which contained high iron and low sulphide concentrations, exhibited a Shanon index of 1.8 for Bacteria and 0.54 for Archaea at 3 % sequence divergence and values of 0.83 and 0.09 at 20 % sequence divergence. The slightly alkaline high-temperature (92 °C) spring AIV shows a Shanon index of 2.29 at 3 % and 1.04 at 20 % sequence divergence for Bacteria. These values are, as expected, lower than in terrestrial soils, which are considered the most complex microbial ecosystems on Earth, with Shanon indices of 5–7 at 3 % sequence divergence and 2.5–4.5 at 20 % sequence divergence (Will et al. 2010; Nacke et al. 2011). Shanon indices calculated for another extreme environment, glacier ice (3.4 and 1.7 for 3 and 20 % sequence divergence, respectively) (Simon et al. 2009), were in the same range as in our study.

Dominant metabolic pathways in the hot springs of Furnas

Inferring physiology from 16S rRNA gene phylogeny is inherently biased (Jaspers and Overmann 2004), in particular if the closest relatives are sequences from uncultured organisms. Keeping this in mind we suggest putative physiology on the basis of the closest cultivated 16S rRNA gene relatives for our habitat, in particular for AIV because a large proportion of the partial sequences is more than 99 % identical to sequences from cultivated organisms. For *Fervidobacterium* 91 % of the sequences were more than 99 % identical to *F. islandicum*, for *Dictyoglomus* 88 % of the sequences were more than 99 % identical to *D. thermophilum*, and for *Caldicellulosiruptor* 85 % of the sequences were more than 99 % identical to *C. lactoaceticus*. Altogether only the rare members of the sequence pool amounting to 11 % were <97 % identical to cultivated or related to uncultivated species. The situation is different for AI and here in particular for *Acidicaldus*. 94 % of the *Acidicaldus*-specific sequences were related to uncultivated organisms. However, 99 % of the *Acidithiobacillus*-sequences were 99 % identical to *Acidithiobacillus caldus*. This is in contrast to the majority of other environmental studies, in which the largest proportion of sequences does not affiliate with cultivated species (Meyer-Dombard et al. 2005; Hugenholtz et al. 1998; Ivanova et al. 2011; Pagaling et al. 2012).

Although differing for two groups (*Sulfurihydrogenibium* and *Fervidobacterium*), both 16S rRNA gene sequence-based approaches indicated the dominance of heterotrophic bacterial genera in both springs. This is in contrast to many other studies in hot spring-environments where Aquificales have been repeatedly found to dominate 16S rRNA gene-libraries, e.g. from Yellowstone National Park, Iceland, New Zealand, and Japan (Hall et al. 2008; Meyer-Dombard et al. 2005; Blank et al. 2002; Jackson et al. 2001; Skirnisdottir et al. 2000; Hetzer et al. 2007; Yamamoto et al. 1998). These findings led to the hypothesis that chemolithotrophic physiology probably based on the oxidation of H_2 or reduced sulphur compounds is the major metabolic pathway in hot springs. Skirnisdottir et al. compared results from different studies on hot spring microbial mats and sediments and found that, depending on the chemical characteristics of the

spring, different subgroups of Aquificales dominated (Yamamoto et al. 1998; Reysenbach et al. 2000; Hugenholtz et al. 1998). In high-sulphide and sometimes iron-rich habitats groups, J and S were dominating and the related sequences can today be attributed to *Sulfurihydrogenibium*. The high abundance (21 %) of *Sulfurihydrogenibium*-related sequences within the metagenome from site AIV fitted into the picture. However, evaluation with FISH revealed that *Sulfurihydrogenibium* had been overestimated and was only accounting for 9 % of the bacterial cells.

The comparatively low abundance of chemolithoautotrophic organisms at site AIV might be related to the high concentration of dissolved organic carbon (DOC) of 284 mg/L (0.028 %). Only little data are available on the DOC content of hot springs. Yamamoto et al. measured a DOC content of 0.41–0.72 mg/L within the microbial sulphur-turf mat, while Hetzer et al. determined 0.7 mg/L in Champagne Pool, New Zealand, and Hall et al. described a DOC content of 10 mg/L in Coffee Pot hot spring (YNP) (Yamamoto et al. 1998; Hetzer et al. 2007; Hall et al. 2008). All sites showed clear dominance of Aquificales. The more than 20-fold or even 400-fold higher concentration of DOC in the Furnas spring could well be a reason for the abundant occurence of heterotrophic bacteria. Regarding autotrophic processes at site AIV, however, also Archaea have to be taken into account, since FISH detected 35 % of the cells to be of archaeal origin. Metagenomic rRNA data suggest that the heterotrophic genera *Sulfophobococcus* spp. and *Desulfurococcus* spp. account for approx. 74 % of the sequences while the remaining 26 % may well be related to chemolithotrophic genera. Further studies will show, if high abundance of heterotrophic microorganisms is a common feature for springs with higher organic carbon content.

The dominant genera in site AIV (*Caldicellulosiruptor*, *Dictyoglomus*, and *Fervidobacterium*) have also been detected in in situ enrichment cultures in hot springs from Kamchatka (Kublanov et al. 2009). While *Caldicellulosiruptor* sp. and *Dictyoglomus* sp. grew on polysaccharides like cellulose and chitin, *Fervidobacterium* sp. was enriched on proteinaceous substrates. Although dominated by one OTU, all groups show additional sequence diversity in particular *Dictyoglomus* with 44 OTUs at 3 % divergence. For *Fervidobacterium* and *Caldicellulosiruptor* the number of specific OTUs is 18 and 14, respectively. These results suggest that the hot springs of Furnas are a valuable source to retrieve new polymer-degrading organisms.

Conclusions

The microbial community of two hot springs of Furnas Valley was characterized on the basis of different rRNA gene-based methods. The combination of different primer-dependent and independent methods revealed major differences, emphasizing the importance of combining different methodological approaches. A large dataset of short rRNA gene sequences and metagenome analysis revealed that the habitats were dominated by three to four genera, both for the bacterial and the archaeal population. Archaeal diversity was particularly limited in the acidic spring AI (pH 3, 51 °C), in which *Thermoplasma* represented almost 90 %. However, overall abundance of Archaea in this acidic environment was low as revealed by FISH, while the hot slightly alkaline spring AIV (pH 8, 92 °C) had almost even abundances of Archaea and Bacteria. Unlike other hydrothermal habitats, the high-temperature slightly alkaline low-sulphate spring AIV was dominated by heterotrophic bacteria, probably due to a high content of dissolved organic carbon, suggesting that not all hydrothermal habitats are dominated by chemolithoautotrophs and that a wide range of chemical characteristics needs to be taken into account when analysing hot springs. This supports the efforts of the Earth Microbiome Project which states the importance of careful and comprehensive metadata collection for each and in particular for metagenomic environmental studies in order to improve comparability, so that general patterns can be identified on the basis of different studies (Knight et al. 2012; Yilmaz et al. 2011).

Our results suggest a natural enrichment of heterotrophic, polymer-degrading genera in the Furnas springs which make them particularly promising for the search of novel thermostable enzymes for application in biotechnology and biorefinery of the second generation.

Acknowledgments The authors gratefully acknowledge Milton da Costa for supporting their sampling on a field trip to the Valley of Furnas.

References

Albuquerque L, Rainey FA, Nobre MF, da Costa MS (2008) *Elioraea tepidiphila* gen. nov., sp. nov., a slightly thermophilic member of the Alphaproteobacteria. Int J Syst Evol Microbiol 58:773–778

Albuquerque L, Rainey FA, Nobre MF, da Costa MS (2010) *Meiothermus granaticius* sp. nov., a new slightly thermophilic red-pigmented species from the Azores. Syst Appl Microbiol 33:243–246

Albuquerque L, Rainey FA, Nobre MF, da Costa MS (2012) *Hydrotalea sandarakina* sp. nov., isolated from a hot spring runoff and emended description of the genus *Hydrotalea* and the species *Hydrotalea flava*. Int J Syst Evol Microbiol 62:1603–1608

Allen ET, Day AL (1935) Hot springs of Yellowstone National Park. Carnegie Institution of Washington Publication no. 466

Amann RI, Krumholz L, Stahl DA (1990) Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. J Bacteriol 172:762–770

Baker GC, Cowan DA (2004) 16 S rDNA primers and the unbiased assessment of thermophile diversity. Biochem Soc Trans 32:218–221

Barth TFW (1950) Volcanic geology, hot springs and geysers of Iceland. Carnegie Institution of Washington Publication no. 587

Blank CE, Cady SL, Pace NR (2002) Microbial composition of near-boiling silica-depositing thermal springs throughout Yellowstone National Park. Appl Environ Microbiol 68:5123–5135

Brock TD, Brock ML (1967) The hot springs of Furnas Valley, Azores. Int Revue ges Hydrobiol 52:545–558

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336

Chao A, Bunge J (2002) Estimating the number of species in a stochastic abundance model. Biometrics 58:531–539

Costa KC, Navarro JB, Shock EL, Zhang CL, Soukup D, Hedlund BP (2009) Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. Extremophiles 13:447–459

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200

Franca L, Rainey FA, Nobre MF, da Costa MS (2006) *Tepidicella xavieri* gen. nov., sp. nov., a betaproteobacterium isolated from a hot spring runoff. Int J Syst Evol Microbiol 56:907–912

Friedrich AB, Antranikian G (1996) Keratin Degradation by *Fervidobacterium pennavorans*, a novel thermophilic anaerobic species of the order Thermotogales. Appl Environ Microbiol 62:2875–2882

Hall JR, Mitchell KR, Jackson-Weaver O, Kooser AS, Cron BR, Crossey LJ, Takacs-Vesbach CD (2008) Molecular characterization of the diversity and distribution of a thermal spring microbial community by using rRNA and metabolic genes. Appl Environ Microbiol 74:4910–4922

Hamamura N, Meneghin J, Reysenbach AL (2012) Comparative community gene expression analysis of Aquificales-dominated geothermal springs. Environ Microbiol

Harmsen H, Prieur D, Jeanthon C (1997) Group-specific 16S rRNA-targeted oligonucleotide probes to identify thermophilic bacteria in marine hydrothermal vents. Appl Environ Microbiol 63:4061–4068

Hetzer A, Morgan HW, McDonald IR, Daughney CJ (2007) Microbial life in Champagne Pool, a geothermal spring in Waiotapu, New Zealand. Extremophiles 11:605–614

Hiraishi A, Umezawa T, Yamamoto H, Kato K, Maki Y (1999) Changes in quinone profiles of hot spring microbial mats with a thermal gradient. Appl Environ Microbiol 65:198–205

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR (1998) Novel division level bacterial diversity in a Yellowstone hot spring. J Bacteriol 180:366–376

Huse SM, Dethlefsen L, Huber JA, Mark Welch D, Relman DA, Sogin ML (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. PLoS Genet 4:e1000255

Inskeep WP, Rusch DB, Jay ZJ, Herrgard MJ, Kozubal MA, Richardson TH, Macur RE, Hamamura N, Jennings R, Fouke BW, Reysenbach AL, Roberto F, Young M, Schwartz A, Boyd ES, Badger JH, Mathur EJ, Ortmann AC, Bateson M, Geesey G, Frazier M (2010) Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. PLoS ONE 5:e9773

Ivanova I, Atanassov I, Lyutskanova D, Stoilova-Disheva M, Dimitrova D, Tomova I, Derekova A, Radeva G, Buchvarova V, Kambourova M (2011) High Archaea diversity in Varvara hot spring, Bulgaria. J Basic Microbiol 51:163–172

Jackson CR, Langner HW, Donahoe-Christiansen J, Inskeep WP, McDermott TR (2001) Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. Environ Microbiol 3:532–542

Jaspers E, Overmann J (2004) Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysiologies. Appl Environ Microbiol 70:4831–4839

Knight R, Jansson J, Field D, Fierer N, Desai N, Fuhrman JA, Hugenholtz P, van der Lelie D, Meyer F, Stevens R, Bailey MJ, Gordon JI, Kowalchuk GA, Gilbert JA (2012) Unlocking the potential of metagenomics through replicated experimental design. Nat Biotechnol 30:513–520

Kublanov IV, Perevalova AA, Slobodkina GB, Lebedinsky AV, Bidzhieva SK, Kolganova TV, Kaliberda EN, Rumsh LD, Haertle T, Bonch-Osmolovskaya EA (2009) Biodiversity of thermophilic prokaryotes with hydrolytic activities in hot springs of Uzon Caldera, Kamchatka (Russia). Appl Environ Microbiol 75:286–291

Kysela DT, Palacios C, Sogin ML (2005) Serial analysis of V6 ribosomal sequence tags (SARST-V6): a method for efficient, high-throughput analysis of microbial community composition. Environ Microbiol 7:356–364

Lau MC, Aitchison JC, Pointing SB (2009) Bacterial community composition in thermophilic microbial mats from five hot springs in central Tibet. Extremophiles 13:139–149

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17:10–12

Mathur J, Bizzoco RW, Ellis DG, Lipson DA, Poole AW, Levine R, Kelley ST (2007) Effects of abiotic factors on the phylogenetic diversity of bacterial communities in acidic thermal springs. Appl Environ Microbiol 73:2612–2623

Meyer-Dombard DR, Shock EL, Amend JP (2005) Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. Geobiology 3:211–227

Miller SR, Strong AL, Jones KL, Ungerer MC (2009) Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. Appl Environ Microbiol 75: 4565–4572

Muyzer G, Brinkhoff T, Nübel U, Santegoeds C, Schäfer H, Wawer C (1998) Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In: Akkermans ADL, van Elsas JD, de Bruijn FJ (eds) Molecular microbial ecology manual. Kluwer Academic Publishers, London, pp 3.4.4.1–3.4.4.27

Nacke H, Thurmer A, Wollherr A, Will C, Hodac L, Herold N, Schoning I, Schrumpf M, Daniel R (2011) Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils. PLoS ONE 6:e17000

O-Thong S, Prasertsan P, Karakashev D, Angelidaki I (2007) Specific detection of *Thermoanaerobacterium* spp., *Thermoanaerobacterium thermosaccharolyticum* and *Caldicellulosiruptor* spp. in thermophilic biohydrogen reactor using fluorescent in situ hybridization (FISH). Int J Hydrogen Energy 33:6082–6091

Pagaling E, Grant WD, Cowan DA, Jones BE, Ma Y, Ventosa A, Heaphy S (2012) Bacterial and archaeal diversity in two hot spring microbial mats from the geothermal region of Tengchong, China. Extremophiles 16:607–618

Papke RT, Ramsing NB, Bateson MM, Ward DM (2003) Geographical isolation in hot spring cyanobacteria. Environ Microbiol 5:650–659

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196

Ravenschlag K, Sahm K, Knoblauch C, Jorgensen BB, Amann R (2000) Community structure, cellular rRNA content, and activity of sulfate-reducing bacteria in marine arctic sediments. Appl Environ Microbiol 66:3592–3602

Reysenbach AL, Wickham GS, Pace NR (1994) Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. Appl Environ Microbiol 60:2113–2119

Reysenbach AL, Ehringer M, Hershberger K (2000) Microbial diversity at 83 °C in Calcite Springs, Yellowstone National Park: another environment where the Aquificales and "Korarchaeota" coexist. Extremophiles 4:61–67

Riessen S, Antranikian G (2001) Isolation of *Thermoanaerobacter keratinophilus* sp. nov., a novel thermophilic, anaerobic bacterium with keratinolytic activity. Extremophiles 5:399–408

Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AK, Kent AD, Daroub SH, Camargo FA, Farmerie WG, Triplett EW (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J 1:283–290

Schloss PD, Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol 71:1501–1506

Shanon WE, Weaver W (1949) The mathematical theory of communication. MD Comput 14:306–317

Simon C, Wiezer A, Strittmatter AW, Daniel R (2009) Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. Appl Environ Microbiol 75:7519–7526

Skirnisdottir S, Hreggvidsson GO, Hjorleifsdottir S, Marteinsson VT, Petursdottir SK, Holst O, Kristjansson JK (2000) Influence of sulfide and temperature on species composition and community structure of hot spring microbial mats. Appl Environ Microbiol 66:2835–2841

Snaidr J, Amann R, Huber I, Ludwig W, Schleifer KH (1997) Phylogenetic analysis and in situ identification of bacteria in activated sludge. Appl Environ Microbiol 63:2884–28896

Spear JR, Walker JJ, McCollom TM, Pace NR (2005) Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. Proc Natl Acad Sci USA 102:2555–2560

Stahl DA, Amann RI (1991) Development and application of nucleic acid probes. In: Stackebrandt E, Goodfellow M (eds) Nucleic acid techniques in bacterial systematics. Wiley, Chichester, pp 205–248

Swingley WD, Meyer-Dombard DR, Shock EL, Alsop EB, Falenski HD, Havig JR, Raymond J (2012) Coordinating environmental genomics and geochemistry reveals metabolic transitions in a hot spring ecosystem. PLoS ONE 7:e38108

Will C, Thürmer A, Wollherr A, Nacke H, Herold N, Schrumpf M, Gutknecht J, Wubet T, Buscot F, Daniel R (2010) Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. Appl Environ Microbiol 76:6751–6759

Yamamoto H, Hiraishi A, Kato K, Chiura HX, Maki Y, Shimizu A (1998) Phylogenetic evidence for the existence of novel thermophilic bacteria in hot spring sulfur-turf microbial mats in Japan. Appl Environ Microbiol 64:1680–1687

Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, Vaughan R, Hunter C, Park J, Morrison N, Rocca-Serra P, Sterk P, Arumugam M, Bailey M, Baumgartner L, Birren BW, Blaser MJ, Bonazzi V, Booth T, Bork P, Bushman FD, Buttigieg PL, Chain PS, Charlson E, Costello EK, Huot-Creasy H, Dawyndt P, DeSantis T, Fierer N, Fuhrman JA, Gallery RE, Gevers D, Gibbs RA, San Gil I, Gonzalez A, Gordon JI, Guralnick R, Hankeln W, Highlander S, Hugenholtz P, Jansson J, Kau AL, Kelley ST, Kennedy J, Knights D, Koren O, Kuczynski J, Kyrpides N, Larsen R, Lauber CL, Legg T, Ley RE, Lozupone CA, Ludwig W, Lyons D, Maguire E, Methe BA, Meyer F, Muegge B, Nakielny S, Nelson KE, Nemergut D, Neufeld JD, Newbold LK, Oliver AE, Pace NR, Palanisamy G, Peplies J, Petrosino J, Proctor L, Pruesse E, Quast C, Raes J, Ratnasingham S, Ravel J, Relman DA, Assunta-Sansone S, Schloss PD, Schriml L, Sinha R, Smith MI, Sodergren E, Spo A, Stombaugh J, Tiedje JM, Ward DV, Weinstock GM, Wendel D, White O, Whiteley A, Wilke A, Wortman JR, Yatsunenko T, Glockner FO (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. Nat Biotechnol 29:415–420

Yim LC, Hongmei J, Aitchison JC, Pointing SB (2006) Highly diverse community structure in a remote central Tibetan geothermal spring does not display monotonic variation to thermal stress. FEMS Microbiol Ecol 57:80–91

Youssef N, Sheik CS, Krumholz LR, Najar FZ, Roe BA, Elshahed MS (2009) Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. Appl Environ Microbiol 75:5227–5236

Yu Z, Garcia-Gonzalez R, Schanbacher FL, Morrison M (2008) Evaluations of different hypervariable regions of archaeal 16S rRNA genes in profiling of methanogens by Archaea-specific PCR and denaturing gradient gel electrophoresis. Appl Environ Microbiol 74:889–893

Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62:316–322

STUDY 8:

METAGENOME SURVEY OF A MULTISPECIES AND ALGAE-ASSOCIATED BIOFILM REVEALS KEY ELEMENTS OF BACTERIAL-ALGAE INTERACTIONS IN PHOTOBIOREACTORS

KROHN-MOLT I[1,2], **WEMHEUER B**[3], ALAWI M[4,5], POEHLEIN A[3], GÜLLERT S[1], SCHMEISSER C[1], POMMERENING-RÖSER A, GRUNDHOFF A[4], DANIEL R[3], HANELT D[2], AND STREIT WR[1]

APPLIED AND EINVIRONMENTAL MICROBIOLOGY 79 (20): 6196-6206

[1]MICROBIOLOGY AND BIOTECHNOLOGY, BIOCENTER KLEIN FLOTTBEK, UNIVERSITY OF HAMBURG, HAMBURG, GERMANY; [2]CELL BIOLOGY AND PHYCOLOGY, BIOCENTER KLEIN FLOTTBEK, UNIVERSITY OF HAMBURG, HAMBURG, GERMANY; [3]GENOMIC AND APPLIED MICROBIOLOGY AND GOETTINGEN GENOMICS LABROTORY, INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GOETTINGEN, GOETTINGEN, GERMANY; [4]VIRUS GENOMICS LEIBNIZ, INSTITUTE FOR EXPERIMENTAL VIROLOGY, HEINRICH-PETTE-INSTITUTE, HAMBURG, GERMANY; [5]UNIVERSITY MEDICAL CENTER HAMBURG-EPPENDORF, BIOINFORMATICS SERVICE FACILITY, HAMBURG, GERMANY

Author contributions to the work:

Performed the experiments: IKM

Analyzed data: IKM, BW, MA

Contributed data on bioreactor properties and analysis of these data: IKM, APR, CS, WRS

Wrote the publication: IKM, BW, WRS

Conceived and designed the experiments: IKM, WRS

AEM
Journals.ASM.org

Metagenome Survey of a Multispecies and Alga-Associated Biofilm Revealed Key Elements of Bacterial-Algal Interactions in Photobioreactors

Ines Krohn-Molt,[a,b] Bernd Wemheuer,[c] Malik Alawi,[d,e] Anja Poehlein,[c] Simon Güllert,[a] Christel Schmeisser,[a] Andreas Pommerening-Röser,[a] Adam Grundhoff,[d] Rolf Daniel,[c] Dieter Hanelt,[b] Wolfgang R. Streit[a]

University of Hamburg, Biocenter Klein Flottbek, Microbiology and Biotechnology, Hamburg, Germany[a]; University of Hamburg, Biocenter Klein Flottbek, Cell Biology and Phycology, Hamburg, Germany[b]; Georg-August-University Goettingen, Institute of Microbiology and Genetics, Goettingen, Germany[c]; Heinrich-Pette-Institute, Leibniz Institute for Experimental Virology, Virus Genomics, Hamburg, Germany[d]; University Medical Center Hamburg-Eppendorf, Bioinformatics Service Facility, Hamburg, Germany[e]

Photobioreactors (PBRs) are very attractive for sunlight-driven production of biofuels and capturing of anthropogenic CO_2. One major problem associated with PBRs however, is that the bacteria usually associated with microalgae in nonaxenic cultures can lead to biofouling and thereby affect algal productivity. Here, we report on a phylogenetic, metagenome, and functional analysis of a mixed-species bacterial biofilm associated with the microalgae *Chlorella vulgaris* and *Scenedesmus obliquus* in a PBR. The biofilm diversity and population dynamics were examined through 16S rRNA phylogeny. Overall, the diversity was rather limited, with approximately 30 bacterial species associated with the algae. The majority of the observed microorganisms were affiliated with *Alphaproteobacteria*, *Betaproteobacteria*, and *Bacteroidetes*. A combined approach of sequencing via GS FLX Titanium from Roche and HiSeq 2000 from Illumina resulted in the overall production of 350 Mbp of sequenced DNA, 165 Mbp of which was assembled in larger contigs with a maximum size of 0.2 Mbp. A KEGG pathway analysis suggested high metabolic diversity with respect to the use of polymers and aromatic and nonaromatic compounds. Genes associated with the biosynthesis of essential B vitamins were highly redundant and functional. Moreover, a relatively high number of predicted and functional lipase and esterase genes indicated that the alga-associated bacteria are possibly a major sink for lipids and fatty acids produced by the microalgae. This is the first metagenome study of microalga- and PBR-associated biofilm bacteria, and it gives new clues for improved biofuel production in PBRs.

The increasing global CO_2 concentrations and the resulting effects of global warming have led to the development of novel and alternative technologies for the production of non-fossil-based biofuels and methods of CO_2 capture. Since photobioreactors (PBRs) allow the continuous cultivation of microalgae under relatively controlled conditions, they have been the focus of many research projects, often linked to large-scale biofuel production (1–3). The microalgae in the PBRs produce significant amounts of lipids that are the basis of biofuel production (1, 2, 4, 5). Unfortunately the technology is not yet sustainable (6), and certainly, more research is needed, not only to improve the technology, but also to better understand the complex biology and interactions of microalgae with their environment.

The term microalgae comprises a phylogenetically very heterogeneous group of prokaryotic and eukaryotic microorganisms. They all employ oxygenic photosynthesis and thereby convert atmospheric CO_2 to biomass (7–9).

PBRs are usually vessels with varying volumes (3, 10, 11). Despite their advantages, however, PBRs suffer from several drawbacks, and biodiesel production with microalgae in PBRs faces many challenges (11, 12). One major problem associated with PBRs is linked to biofouling and the growth of microbial biofilms within the reactor. Bacterial growth in running PBRs results in disrupted flow, increased pressure, and reduced light intensities and ultimately affects algal productivity (13). While it is well known that microalgae are often associated with diverse bacteria in their natural aquatic ecosystems and in experimental cultures (14–16), only very few studies have addressed a detailed phylogenetic characterization of the associated microbes (17–19). Bacteria observed in these nonaxenic cultures belong to the classes *Alphaproteobacteria*, *Betaproteobacteria*, and *Gammaproteobacteria*, but bacteria affiliated with the phylum *Bacteroidetes* are also often observed. Additionally, many uncultivated bacteria are noted. However, no archaeal species have been detected, and only one study has reported the presence of an uncultivated fungus (18). Two very recent studies have addressed a partial phylogenetic characterization of the alga-associated microbial community within PBRs (16, 20). The authors of these studies show that the microalga *Chlorella vulgaris* is associated mainly with the *Alphaproteobacteria* (i.e., the genus *Sphinogomonas*) and *Dunaliella tertiolecta* is associated with the *Alphaproteobacteria* and *Gammaproteobacteria*.

Microalgae are in general believed to be auxotrophic for vitamin B_{12}, and a significant fraction of microalgal species appear to be auxotrophic for thiamine, while some are also auxotrophic for biotin (21–24). Within this framework, it is also possible that

Received 20 May 2013 **Accepted** 19 July 2013
Published ahead of print 2 August 2013
Address correspondence to Wolfgang R. Streit, wolfgang.streit@uni-hamburg.de.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.01641-13.

doi:10.1128/AEM.01641-13

other bacterial factors or metabolic activities have a positive influence on algal growth (25). Only very few studies have suggested that there is actually a negative impact on the population dynamics of the algae (26).

Metagenome technologies have proven to be powerful tools for the analysis of complex microbial communities and have led to a tremendous increase of knowledge about the functions, protein families, biotechnology, and ecology of microbial communities within the last decade (27–29). Recent examples from metagenome studies have given us a first insight into the complex interactions of microorganisms with their environments and hosts (29–33). Surprisingly, only very few studies have focused on metagenomes associated with algae (34, 35), and to date, no study is available that has addressed aquatic microalgae and their bacterial communities in PBRs.

Therefore, here, we have analyzed the phylogenetic and metagenomic diversity of a mixed-species biofilm grown in a PBR. The data suggest that the microbial community is composed of approximately 30 different microbial species and is capable of supporting algal growth by the production of B vitamers (cobalamin [B_{12}] and biotin [B_7]) and possibly other cofactors not yet identified. Remarkably, our data also suggest that at least some of the bacteria within the PBR biofilm cannot be cultivated without the microalgae and that the bacteria encode the potential to affect biofuel production.

MATERIALS AND METHODS

Microorganisms used in this study and cultivation media. *Scenedesmus obliquus* strain U169 and *Chlorella vulgaris* strain U126 were obtained from the collection of algae at the University of Hamburg, Hamburg, Germany (http://www.biologie.uni-hamburg.de/bzf/zeph/zephsvcke.htm). The defined medium for cultivation of algae was composed of 2 g liter^{-1} Flory Basis Fertilizer 1 (Euflor, Germany) and supplemented with 3.22 g liter^{-1} KNO_3. Flory Basis Fertilizer 1 is solely based on mineral compounds and was not supplemented with any vitamins. The pH was adjusted to 7.0. The algae were cultivated at 17°C in liquid medium at a natural light intensity in polyethyleneterephthalate flat-panel photobioreactors. The culture medium circulated at 1 liter min^{-1} in the PBR system. The PBR was aerated with compressed air and flue gas obtained from a combined block heat and power station. Physical parameters and culture conditions were monitored continuously (WTW IQ Sensor Net; System 2020 XT, Germany). The light intensity was determined with a LI-190 sensor (Li-Cor, USA). For a detailed description of the reactor, see Fig. S1 in the supplemental material and reference 36.

Escherichia coli strains were grown at 37°C in lysogenic broth (LB) medium (1% tryptone, 0.5% yeast extract, 1% NaCl) (37) supplemented with appropriate antibiotics. *Lactobacillus plantarum* ATCC 8014 and *Lactobacillus delbrueckii* subsp. *lactis* DSM 20335 for B vitamin detection were obtained from the DSMZ, Braunschweig, Germany. *Lactobacillus* strains were grown on MRS medium (38) under anaerobic conditions.

Media for cultivation of individual bacterial isolates derived from the PBR biofilm community were prepared as follows. R2A medium was prepared as described previously (39), and M9, TSB, and NB media were prepared according to the method of Sambrook and Russell (37). To stimulate microbial growth, the media were in part supplemented with algal culture extracts ranging from 5% to 50% (vol/vol). The inoculated plates were incubated for 5 to 7 days at 22°C under aerobic and anaerobic conditions.

Scanning electron microscopy. For scanning electron microscopy (SEM), biofilm samples were fixed in paraformaldehyde (1%) and glutaraldehyde (0.25%), dehydrated by ascending alcohol series, and dried at the critical point with Balzers CPD 030 Critical Point Dryer (Bal-Tec, Schalksmühle, Germany). After coating samples with gold using an SCD 050 sputter coater (Bal-Tec), scanning electron micrographs were taken with a Leo 1525 (Zeiss, Germany).

PBR biofilm DNA extraction and molecular technologies. Total nucleic acids were extracted from the biofilm samples using a previously published enzymatic cell lysis protocol with some modifications (40). The samples were stirred (200 rpm) overnight in 10 ml of extraction buffer (100 mM Tris-HCl, pH 8.0, 100 mM sodium EDTA, pH 8.0, 1.5 M NaCl, 0.1% Tween 80) with the addition of 5 mg lysozyme and 0.5 mg proteinase K at 37°C. Subsequently, SDS (1 ml; 20%) was added and incubated at 65°C for 90 min. The sample was centrifuged for 10 min at room temperature and 6,000 × *g*. The pellet was resuspended in 10 ml extraction buffer, incubated for 10 min at 65°C, and centrifuged again. The supernatant was collected and mixed with 1 volume polyethylene glycol (PEG) 6000 (30%) and 1.6 M sodium chloride. After 2 h of incubation at room temperature, the mixture was centrifuged again for 20 min at 10,000 × *g*. The resulting pellet was resuspended in 2 ml TE buffer (10 mM Tris-HCl, 1 mM sodium EDTA, pH 8.0). For separating proteins and polysaccharides, 7.5 M potassium acetate was added to the mixture, achieving a final concentration of 0.75 M. After 5 min of incubation on ice, proteins and polysaccharides were centrifuged at 16,000 × *g* and 4°C for 30 min. The DNA was finally extracted with phenol-chloroform–isoamyl alcohol and precipitated overnight at −20°C after adding 0.7 volume isopropanol with 1/10 volume of 3 M sodium acetate. The isolated DNA was used for PCR amplifications, as well as metagenomic analyses.

For the phylogenetic characterization, 16S rRNA genes were amplified using the standard primers 27f and 1492r (41). The amplified genes were ligated into the pDrive cloning vector (Qiagen, Hilden, Germany) and transformed into chemically competent TOP10 *E. coli* cells (Invitrogen, Karlsruhe, Germany). The 16S rRNA gene was sequenced with automated ABI377 technology following the manufacturer's instructions.

The large-insert fosmid library was constructed according to the Copy Control fosmid library production kit manual (Epicentre Biotechnologies, Madison, WI, USA). A total of 14,976 fosmid clones harboring inserts with an average size of 35 kb were generated. The insert rate was approximately 98%.

Sequencing of metagenomic DNA. For sequencing of metagenomic DNA on the GS FLX platform (Roche Applied Science), libraries were constructed by applying the GS Rapid Library Prep kit (Roche Applied Science) according to the manufacturer's protocol. A single library was sequenced in half a PicoTiter Plate on the GS FLX system using the titanium sequencing chemistry (Roche Applied Science, Mannheim, Germany). Raw data were processed by employing the analysis pipeline for whole-genome shotgun sequence reads and applying GS FLX System Software (version 2.3). Illumina sequencing was done using a single lane of a HiSeq 2000 paired-end run (2 sets of 100 cycles). *De novo* assembly of these short reads was performed using the Velvet assembly program version 1.2.08 (42).

For the sequences' functional characterization, we used the Integrated Microbial Genomes (IMG) pipeline. To further analyze the biological processes linked to the individual genes and open reading frames (ORFs), we mainly employed the KEGG (43), COG (44), and Pfam (45) databases, using a cutoff of 10^{-5}.

Phylogenetic analysis of bacteria associated with algae. 16S rRNA genes were amplified using oligonucleotide primers (27f, 5′-AGA GTT GAT CMT GGC TCA G-3′, and 1492r, 5′-GGY TAC CTT GTT ACG ACT T-3′) (41). PCR mixtures contained 100 ng of template DNA per μl, 0.2 mM each of the four deoxynucleoside triphosphates, 1.5 mM $MgCl_2$, 1 μM (each) primer, and 2.5 U of *Taq* DNA polymerase. The thermocycling conditions were 45 s of denaturation at 94°C, 45 s of primer annealing at 50°C, and 1 min 30 s of primer extension at 72°C. This cycle was repeated 30 times.

Denaturing gradient gel electrophoresis (DGGE) analyses were carried out with a DCode system (Bio-Rad, Munich, Germany) using denaturing gradients of 40 to 70% denaturants. For the amplification of the 16S rRNA gene fragments, the primers 341F-GC (5′-GCA CGG GGG CCT ACG

GGA GGC AGC AG-3′, containing a 30-bp GC-rich sequence at the 5′ end) and 907R (5′-CCG TCA ATT CCT TTR AGT TT-3′) were used (46, 47). The thermocycling conditions were 40 s of denaturation at 96°C, 40 s of primer annealing at 56°C, and 1 min of primer extension at 72°C. This cycle was repeated 30 times. DGGE bands were sampled for reamplification and sequencing by punching into individual bands with sterile pipette tips and were transferred to 1.5-ml microcentrifuge tubes that contained 50 μl of distilled water. The reamplified PCR products were cleaned up using the Gel/PCR DNA Fragments Extraction kit (Qiagen, Hilden, Germany) and finally sequenced using standard technologies.

For restriction fragment length polymorphism (RFLP) analysis, the amplified (27f and 1492r) 16S rRNA genes were cloned into pDrive and transformed in *E. coli* cells. The insert DNA of 100 clones of each sample (T0 to T5) was amplified using oligonucleotide primers (M13-20-f, 5′-GTA AAA CGA CGG CCA GT-3′, and M13-r, 5′-CAG GAA ACA GCT ATG ACC-3′). The thermocycling conditions were 45 s of denaturation at 94°C, 45 s of primer annealing at 57°C, and 1 min 45 s of primer extension at 72°C. This cycle was repeated 30 times. The fragments were digested by the restriction enzyme HpaII. Clones that showed different patterns were sequenced using standard technologies.

Furthermore, to investigate the composition of the microbial community in the photobioreactor, rRNA sequences of the small subunit (SSU) were extracted from the unassembled FLX 454 and Illumina data using SortMeRna (48). Low-quality (<20 Ns) and short (<80-bp) reads were removed prior to further analysis. To gain insight into the community composition, sequences were directly mapped onto the most recent Silva database (SSURef 111 NR) using Bowtie2 (49). The Silva taxonomy of the best hit in the database was affiliated to the respective query sequence. For deeper analysis of the bacterial community structure, the hypervariable regions V3 and V6 of the bacterial 16S rRNA genes were extracted using V-Xtractor (50) and were also mapped on the most recent Silva database.

EPS characterization. Exopolymeric substances (EPS) from the biofilm sample were extracted according to the method of Wingender et al. (51) with minor modifications. For this purpose, the biofilm samples were carefully removed from the PBR surface, weighed, and subsequently suspended in a ratio of 1:16 (wt/vol) in 0.14 M NaCl solution and homogenized by stirring at room temperature for 60 min. The suspension was centrifuged for 30 min at 20,000 × *g* and 10°C. The supernatants were twice sterile filtered using membrane filters (cellulose acetate; pore size, 0.20 μm). The filtrates were either processed immediately or aliquoted and stored until use at −20°C.

The lipid content was measured by extraction with *n*-hexane (52). To 0.5 ml of the EPS sample, 0.5 ml *n*-hexane was added and homogenized for 30 min by shaking at room temperature following a centrifugation step at 12,000 × *g* for 90 s. The upper phase (hexane and lipids) was transferred to a preweighed glass tube while avoiding the interphase. The solution was dried, and the glass vessel was weighed again.

The total carbohydrate content was determined using a phenol-sulfuric acid method (53) with dextran (neutral polysaccharides) and alginate (acidic polysaccharides) standards. To 0.5 ml of sample, 0.5 ml of phenol solution was added and mixed. Following this, 2.5 ml of concentrated sulfuric acid (H_2SO_4) was added, and the sample was mixed again following an incubation step at room temperature (10 min) and incubation at 30°C for 15 min. After a further 5-min incubation at room temperature, the absorbance of the samples was measured spectrophotometrically as the optical density at 480 nm (OD_{480}) for acidic polysaccharides (alginate) and the OD_{490} for neutral polysaccharides (dextran).

Uronic acids were determined according to the method of Filisetti-Cozzi and Carpita (54) with glucuronic acid as a standard. To 0.4 ml of each sample, 40 μl solution 1 (4 M sulfamic acid [H_3SO_3N], pH 1.6) was added and mixed. After this step, 2.4 ml of solution 2 [0.075 M $Na_2(B_4O_5(OH)_4) \times 8H_2O$ dissolved in H_2SO_4] was added, mixed, and then incubated for 20 min at 100°C. The samples were cooled in an ice bath for 5 min. Then, 80 μl of solution 3 (0.15% [wt/vol] *m*-hydroxybiphenyl [$C_{12}H_{10}O$] dissolved in 0.5% [wt/vol] NaOH) was added. Finally, the samples were thoroughly mixed and incubated for 10 min at room temperature, and the absorbance was recorded as OD_{525}.

The protein content was measured using a modified Lowry assay (55, 56).

Furthermore, the total DNA within the EPS solution was detected spectrophotometrically after extraction from the EPS using the UltraClean Microbial DNA Kit (Dianova, Hamburg, Germany).

Lipase and esterase activities were assayed using various *p*-nitrophenyl (pNP) substrates and following previously published protocols (57).

Vitamin B_{12} and biotin detection. For the determination of cobalamin the Difco B_{12} assay medium was used, and vitamin B_{12} was extracted and assayed according to the protocol published by Denter and Bisping (58) employing the *L. delbrueckii* subsp. *lactis* DSM 20335 indicator strain.

To determine the biotin content of the sample, the culture supernatant was centrifuged and sterile filtered. The biotin detection was done as described previously using the *Lactobacillus* growth assay and *L. plantarum* strain ATCC 8014 (59).

Nucleotide sequence accession numbers. This project has been assigned the GenBank BioProject number PRJNA197241. The sequences derived from Illumina and FLX454 sequencing were deposited in the NCBI Short Read Archive under the study accession number SRP021903. The genome assembly, together with predicted gene models and annotations, is available from www.jgi.doe.gov (DOE Joint Genome Institute) under the IMG Project ID 9992.

The 16S rRNA gene sequences obtained have been deposited in the GenBank database under accession numbers KC994651 to KC994663, KC994664 to KC994680, KC994681 to KC994881, and KC994882 to KC994889.

RESULTS AND DISCUSSION

Biofilm development and chemical analysis. PBRs are commonly inoculated with nonaxenic cultures of eukaryotic microalgae (14–16). In the current study, we examined a microbial biofilm in a PBR, which was inoculated with an *S. obliquus* and *C. vulgaris* culture. We monitored biofilm development over a period of 12 weeks by analyzing a total of six time points (T0 to T5) (Fig. 1). As early as 14 days after the initial transfer of the starter culture into the reactor, extensive biofilm formation was observed. The biofilm was tightly attached to the surface of the reactor and spread over the entire reactor within a period of 10 to 12 weeks, thereby strongly affecting the overall photosynthesis rates of the algae. This strong biofilm development resulted in shutdown of the PBR. Microscopic examinations of biofilm samples using SEM suggested that the bacteria were tightly attached to the algae (Fig. 2A to D). Additionally, EPS and nanowire-like filamentous structures were visible in the SEM images (Fig. 2). It is well known that EPS in biofilms mainly consists of polysaccharides, proteins, nucleic acids, and lipids (60). The concentrations of the individual components, however, vary and depend to a large extent on the different microorganisms and their medium and environmental conditions (61–63). EPS has diverse functions during biofilm development: it provides mechanical stability, builds up a three-dimensional network of polymers, and plays a role in cell-cell communication and defense against biocides (60, 64). A detailed chemical analysis of the PBR biofilm exopolymeric matrix suggested that it consisted mainly of fatty acids (10.4 mg/g), acidic and neutral polysaccharides (11.6 mg/g), and proteins (3.8 mg/g). However, only small amounts of uronic acid (<1%) and only traces of DNA were observed (Fig. 3). Since microalgae are well known to produce and release lipids

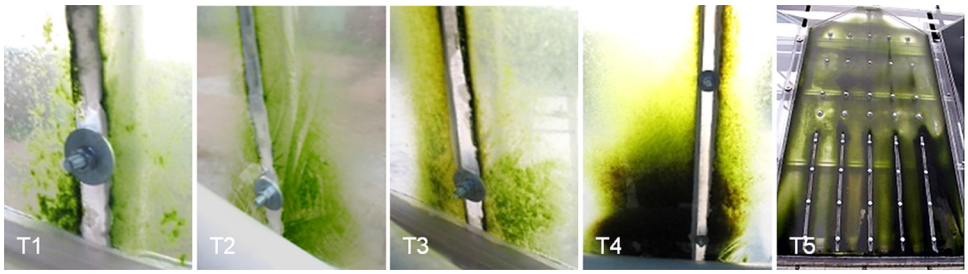


FIG 1 Development of a microbial biofilm over 12 weeks in a flat-panel PBR. T1, after 2 weeks; T2, after 4 weeks; T3, after 6 weeks; T4, after 8 weeks; T5, after 12 weeks of incubation. No microbial biofilm was observed at T0 (data not shown). Shortly after inoculation of the PBR at T0, a light intensity of approximately 150 μmol/m^{2}s was measured at the backside of the reactor. The light intensities with a fully developed biofilm were in the range of 10 μmol/m^{2}s or lower (T5).

(2, 5), it is not surprising that the EPS consisted largely of fatty acids.

Population structure of the alga-associated bacterial community. For phylogenetic analysis, DNA from the mature biofilm (T5), the different biofilm developmental stages (T1 to T4), and the nonaxenic starter culture (T0) was extracted. The quality of the isolated DNA was sufficiently high for the construction of a large-insert metagenome library and DNA-sequencing analysis using pyrosequencing and Illumina-based technologies.

We used different strategies to estimate the microbial diversity within the developing biofilm samples and the established biofilm. We amplified 16S rRNA genes using standard primers (27f and 1492r [41]) and cloned the amplified DNA fragments. Inserts of these clones were completely sequenced. After removal of potential chimeras and duplicates, we analyzed a total of 201 unique 16S rRNA gene sequences from the mature biofilm (T5) (KC994681 to KC994881). In order to identify unique phylotypes and to estimate the bacterial richness, a rarefaction analysis using QIIME (65) was performed. This analysis suggested that the 16S rRNA gene sequences represented 28 operational taxonomic units (OTUs) based on a >99% identity cutoff for bacterial 16S rRNA genes. Rarefaction curves reached saturation at distance levels of 20% (phylum level) and 3% (species level) (Fig. 4). This analysis also suggested that the overall diversity within the biofilm community was rather limited. In addition, a Shannon-Weaver index (66) of 3.01 at a genetic distance of 1% was observed. These results, together with abundance-based coverage estimators (ACE) (67) and Chao1 richness estimates (68), indicated that a significant fraction of the bacterial species diversity was assessed. The analyzed bacterial 16S rRNA gene sequences were mainly affiliated with members of the classes *Alphaproteobacteria*, *Betaproteobacteria*, and *Gammaproteobacteria* (Fig. 5; see Table S1 in the supplemental material). Additionally, a significant number of 16S rRNA gene hits were associated with the phylum *Bacteroidetes*. *Bacteroidetes* are among the main colonizers of the guts of humans and animals, but they are also found in many other habitats and are

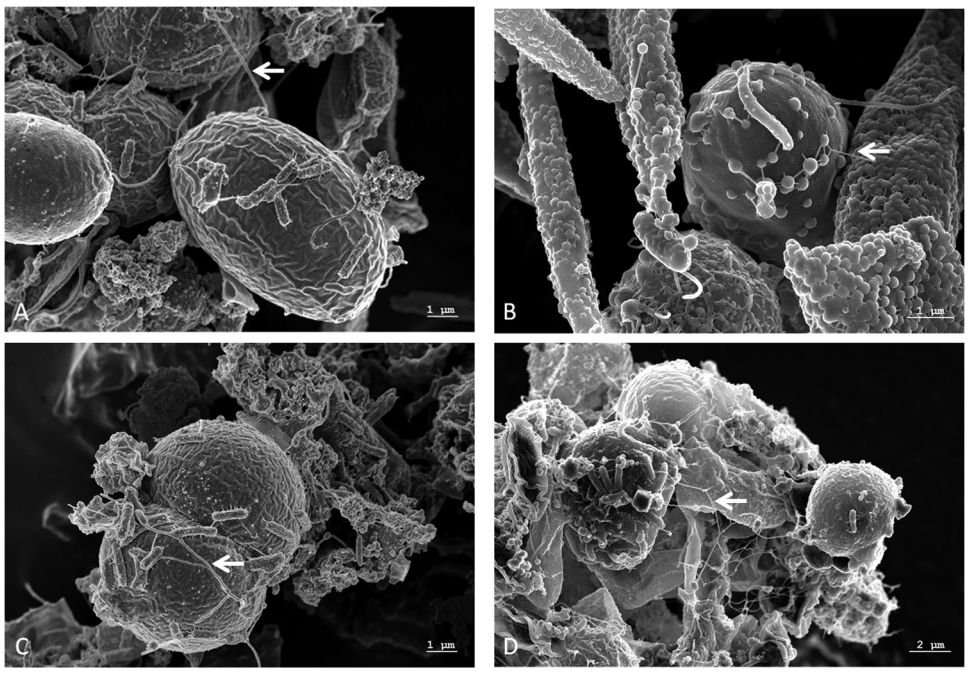


FIG 2 Scanning electron micrographs of a network of microalgae and bacteria in a PBR. Scale bars are indicated in the images (REM Leo 1525; 5.00 kV). The arrows point to nanowire-like structures that were identified. The large round cells (5 to 6 μm) are *C. vulgaris*, and the lemon-shaped cells (8 to 9 μm) are *S. obliquus*.

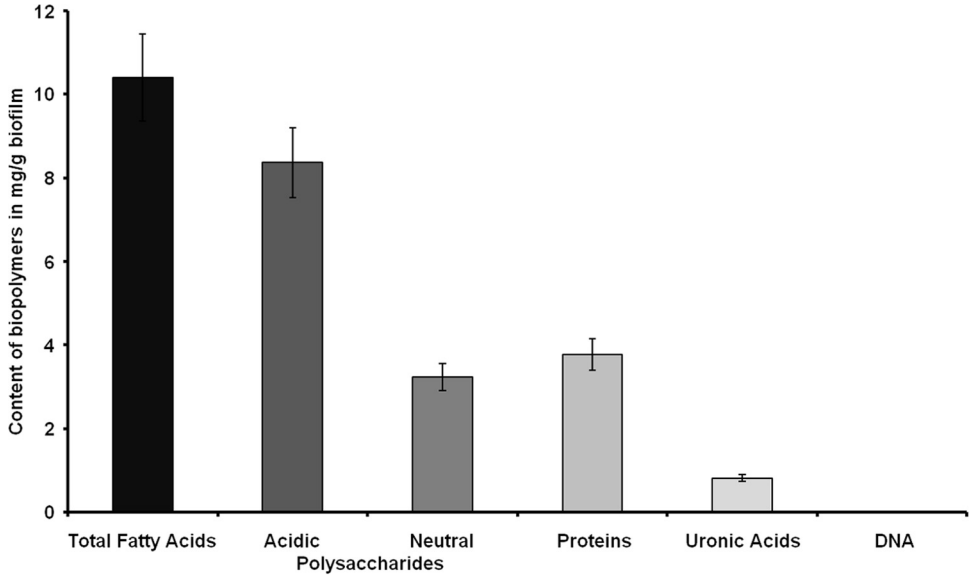


FIG 3 Contents of biopolymers of the extracellular polymeric substances from biofilm samples harvested at T5. The DNA content was <4 ng/mg wet weight. The data are mean values of at least three replicates. The error bars indicate simple standard deviations.

regarded as specialists in the degradation of high-molecular-weight organic matter (34). Overall, the most abundant bacterial organisms were members of the families *Sphingomonadaceae*, *Caulobacteraceae*, *Rhizobiaceae*, *Comamonadaceae*, *Xanthomonadaceae*, *Sphingobacteriaceae*, and *Flavobacteriaceae*.

Additional data from DGGE and RFLP analysis from the different time points during biofilm development and the nonaxenic starter culture supported these findings (see Fig. S2 and S3 in the supplemental material). The DGGE profiles and RFLP analysis, in combination with 16S rRNA gene sequences (DGGE, KC994651 to KC994663; RFLP, KC994664 to KC994680) derived from T0 to T5 further suggested that the phylogenetic structure of the biofilm population was rather stable over the entire time. The data also indicated that the microbes observed within the PBR biofilm were mostly already present in the starter culture used to inoculate the PBR at T0.

To further verify these data, we analyzed the DNA sequences of the mature PBR biofilm obtained by pyrosequencing and Illumina-based sequencing for the presence of hypervariable regions of rRNA gene fragments. Analysis of the V3 and V6 regions has been applied to study bacterial communities in diverse ecosystems (69–71). In total, 226 sequences from the FLX 454 (see Fig. S4A in the supplemental material) and 137,506 sequences from the Illumina (see Fig. S4C in the supplemental material) data sets were directly mapped onto the Silva database. For the analysis of the bacterial population, the hypervariable region V6 of the bacterial

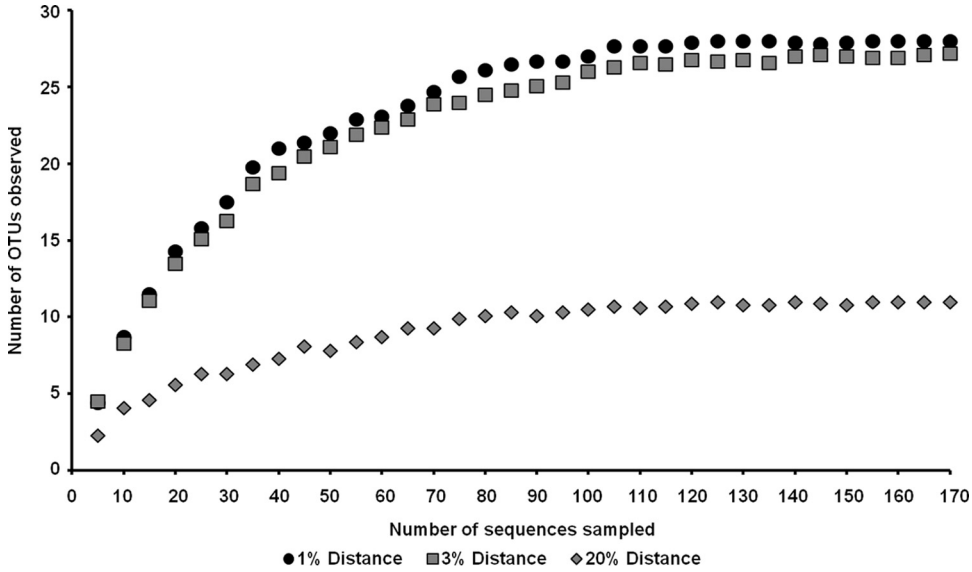


FIG 4 Rarefaction curves indicating the observed numbers of OTUs within the 16S rRNA gene libraries derived from the PBR biofilm at T5 after filtering of the chloroplast sequences. The curves were calculated with QIIME (65). OTUs are shown at 1, 3, and 20% genetic-distance levels.

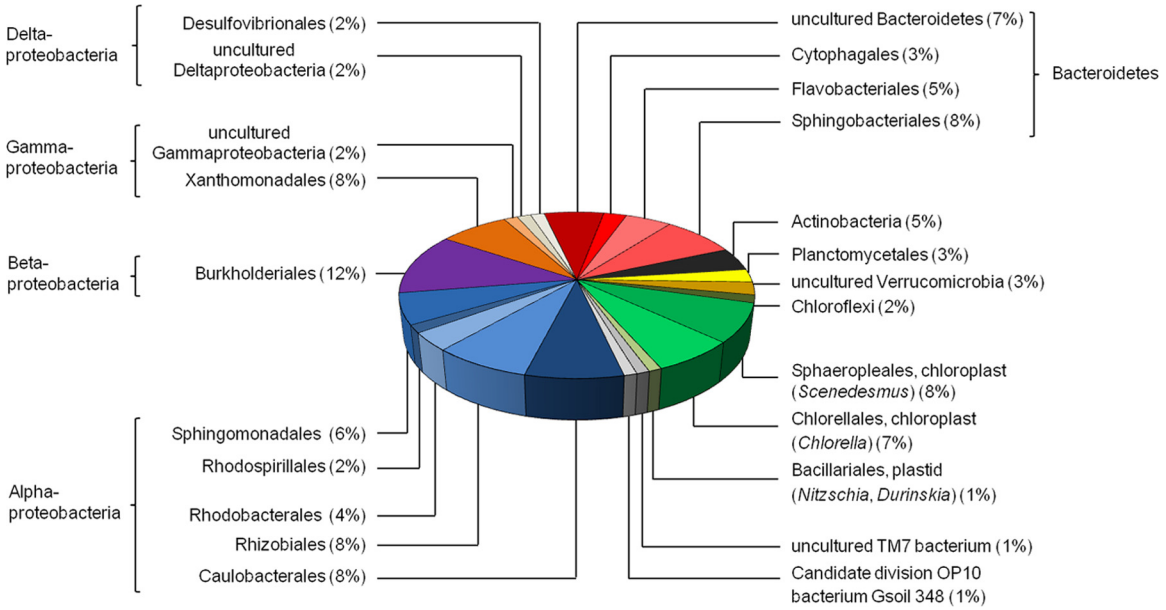


FIG 5 Phylogenetic assignment of the 16S rRNA clone library derived from the PBR biofilm by using the reference sequences stored in the NCBI database. The 16S rRNA gene sequences were analyzed with Finch TV (Geospiza) and BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html).

16S rRNA gene was analyzed. Altogether, 41 sequences from the FLX 454 (see Fig. S4B in the supplemental material) and 558 sequences from the Illumina (see Fig. S4D in the supplemental material) data sets were included in the analysis.

In total, the results from the analysis (see Fig. S4 in the supplemental material) largely confirmed the results and the phylogenic composition estimated by using the 16S rRNA gene clone library.

However, it is notable that primer biases occurring during amplification of 16S rRNA genes (72) and the relatively high copy number of the microalgal plastid genomes, including 16S rRNA genes, resulted in slightly different species counts. The observed fungi, diatoms, and golden algae most likely are contaminants resulting from previous use of the PBRs.

Cultivation of microorganisms from the PBR. Initial attempts to isolate bacteria associated with the algae by plating on different solid media (LB, tryptone-yeast [TY], Reasoner's 2A [R2A], TSB, and NB) under aerobic and anaerobic conditions, resulted in only two pure cultures. 16S rRNA gene sequencing suggested that the two isolates were affiliated with the genera *Brevundimonas* (JQ661035.1; isolate A) and *Paracoccus* (JQ404485.1; isolate B) (see Table S2 in the supplemental material). Members of these genera have been frequently isolated from soil and aquatic environments, but various clinical isolates are also known. Interestingly, tests in which we added various amounts of algal culture supernatant, including living cells, to the solid media did stimulate growth of another six bacterial organisms. These isolates were phylotyped by 16S rRNA gene sequencing (see Table S2 in the supplemental material). The 16S rRNA gene sequence of isolate C is most similar to that of *Chryseobacterium taichungense* YNB68 (JQ071521.1). *C. taichungense* belongs to the *Flavobacteriales* and was only recently described as a novel species able to hydrolyze many chromogenic substrates (73). Isolate D could be associated with a *Brevibacterium* species that was not further characterized (GQ199748.1). Brevibacteria are strictly aerobic Gram-positive microorganisms. Both isolates C and D were isolated on R2A agar plates supplemented with 50% (vol/vol) algal culture. The closest relative of isolate E is a member of the genus *Roseomonas*. The 16S rRNA gene sequence revealed a similarity of 78% to *Roseomonas* sp. (HQ588850.1). Isolate E was cultivated in the presence of 25% (vol/vol) algal culture on NB agar plates. 16S rRNA analysis of isolate F indicated that this strain is associated with the yet-uncultured *Cytophaga-Flavobacteria-Bacteroides* (CFB) group bacterial clone S15D-MN13 (AJ583211.1). The 16S rRNA gene sequence was 96% similar to the S15D-MN13 GenBank entry. Isolate G was similar to a hitherto uncultured gammaproteobacterium, clone L4 (EU887990.1). 16S rRNA sequencing of isolate H suggested that the isolate is affiliated with *Rhodococcus* sp. strain 3/2 (EU041710.1). Isolates F to H were grown on NB medium supplemented with 50% (vol/vol) algal culture.

Altogether, these data suggest that only a minor fraction of the bacteria within the PBR are easily cultivable in laboratory media (KC994882 to KC994889).

Analysis of the metabolic potential encoded by the PBR bacterial metagenome. Today, our knowledge of the physiology and metabolism of complex microbial communities associated with eukaryotic microalgae is very limited. Therefore, we analyzed the partial metagenome sequences of the PBR biofilm microbial community with respect to possible gene functions. We established and assembled 164.7 Mbp of DNA sequences for the PBR biofilm community in contigs ranging from 165 bp to 203,650 bp (Table 1). Approximately 347,000 protein-coding genes could be assigned. Of all possible ORFs predicted, a total of 55,767 were similar to ORFs and genes in the KEGG database; 125,999 ORFs matched the COG database, and 35,935 ORFs could be analyzed by a MetaCyc pathway analysis (74). In addition, it should be noted that up to 1.5% of all ORFs were derived from eukaryotic DNA in the sample. Within this framework, it should be taken into account that most likely a small but significant number of ORFs are not correctly predicted due to overlapping genes/ORFs and noncorrected frame shifts. Based on the assumption that most

TABLE 1 Overall numbers of sequences and contigs generated for PBR biofilm analysis

Parameter	Value
Reads FLX 454 (filtered)	
No.	569,233
Total length (bp)	322,201,969
Mean length (bp)	566
Range (bp)	20–1,637
GC mean (%)	54
Reads Illumina (filtered)	
No.	217,902,594
Total length (bp)	21,136,551,618
Mean length (bp)	97
Range (bp)	97
GC mean (%)	51
Contigs–assembly (velvet)	
No.	165,176
Total length (bp)	164,776,921
Mean length (bp)	997 ± 2,780
No. ≥400 bp	86,050
N50 size (bp)	2,410
Largest (bp)	203,650
GC mean (%)	52 ± 10

TABLE 2 Key features observed in the PBR biofilm metagenome using a KEGG-based analysis

Trait	% of all hits
Amino acid metabolism	16.6
Polysaccharide degradation and metabolism	16.1
Energy metabolism	8.1
Vitamin and cofactor biosynthesis	6.7
Secondary-metabolite synthesis	4.9
Lipid metabolism	4.8
Transport mechanisms	4.7
Xenobiotic- and aromatic-compound degradation	4.2
Exopolysaccharide biosynthesis and modification	1.9
Cell appendages and motility	1.8
Secretion systems	1.6
Others	14.2
Total	100

of these 30 different bacterial species have genomes ranging from 3 to 8 Mbp, it is reasonable to speculate that the overall bacterial biofilm metagenome ranges from 90 Mbp to 240 Mbp. Therefore, the 164.7-Mbp assembled DNA represents a significant fraction of the overall bacterial biofilm metagenome. Although the available sequences do not allow a complete analysis of the physiological and metabolic functions within this bacterial community, the sequences give a good estimate of the biofilm genome structure and its metabolic potential.

The KEGG analysis suggested that the metabolic and catabolic potential of the microbes living in the PBR biofilm is highly diverse and flexible (Table 2). Genes for many of the classical pathways linked to the degradation of polysaccharides, proteins, and cellulose, but also for the degradation of aromatic compounds, were identified. With respect to the degradation of biopolymers, at least 64 amylases, 57 cellulases (mainly GH5 and M-family cellulose), 22 genes linked to the depolymerization of poly-β-hydroxybutyrate (PHB), and a high number of proteases (2,419) were observed.

Many genes involved in lipid and fatty acid catabolism could be identified. Altogether, at least 1,137 genes and ORFs encoded esterolytic and/or lipolytic enzymes, suggesting a high hydrolytic activity for fatty acids within the analyzed microbial community. This observation was supported by data from a metagenome fosmid library screening. Altogether, 14,976 clones with an average insert size of 35 kbp were screened in tributyrin-containing indicator plates. This equals about 524 Mbp of screened DNA. From this screen, a surprisingly large number of lipolytic clones were identified. A total of 68 clones were identified in a first screen using the outlined screening procedures. Further tests with cell extracts of 16 of these clones confirmed that a wide range of fatty acids with various chain lengths were turned over by extracts of the fosmid clones (Fig. 6). The majority of the clones were active on the pNP-laurate (C_{12}) but had almost negligible activities on the pNP-palmitate (C_{16}) and pNP-stearate (C_{18}) and thus most likely encoded esterases acting preferentially on shorter-chain fatty acids. However, very few clones also hydrolyzed the longer-chain fatty acid substrates and are perhaps more likely to encode true lipases (i.e., pFOS73A2, pFOS108D6, and pFOS131A4). Additional tests with culture supernatants supported this hypothesis. In these tests, we observed small but significant hydrolytic activities ranging from 0.0732 to 0.3373 milliunits per ml in the cell-free culture supernatant of the PBR. Lipolytic enzymes, including lipases (EC 3.1.1.3) and carboxylesterases (EC 3.1.1.1), are of general cell importance for membrane and fatty acid metabolism, and they are well known to be important during various pathogenic and nonpathogenic host-microbe interactions (75, 76). Thus, we speculate that the relatively high number of esterolytic and lipolytic genes is perhaps a result of the known production of lipids and fatty acids by the algae and thus has possibly resulted in an enrichment of highly lipolytic and esterolytic microbes. It is further notable that many short-chain fatty acid oxidoreductases and oxidoreductases without an assigned substrate (>600) were identified within the PBR metagenome. Some of these will most likely be important for the oxidation/reduction of the lipids derived from algae.

The metagenome sequences further suggest that the bacteria associated within the PBR were mainly heterotrophs metabolizing a wide range of carbon and energy sources provided by the algae.

The microbial lifestyle in the PBR biofilm most likely included aerobic, as well as microaerobic, growth. Altogether, 1,844 cytochrome-encoding genes were counted, suggesting that highly diverse and branched electron transport chains are encoded by the bacteria living within the PBR biofilm. Branched electron transport chains are an adaptation to varying oxygen concentrations (77) and are observed in the majority of all bacterial genomes. In addition, all the genes needed for nitrate respiration were identified. The occurrence of these genes will certainly allow a wide range of metabolic activities under varying oxygen concentrations and could be an adaptation to life in biofilms that are known to include gradients of oxygen and often even have anoxic pockets (78, 79).

Biosynthesis genes of B-group vitamins. Vitamin B_{12} is one of the most complex cofactors and an essential vitamer that is required for molecular rearrangements of hydrogen or methyl groups in prokaryotic and eukaryotic cells, and it is involved in the

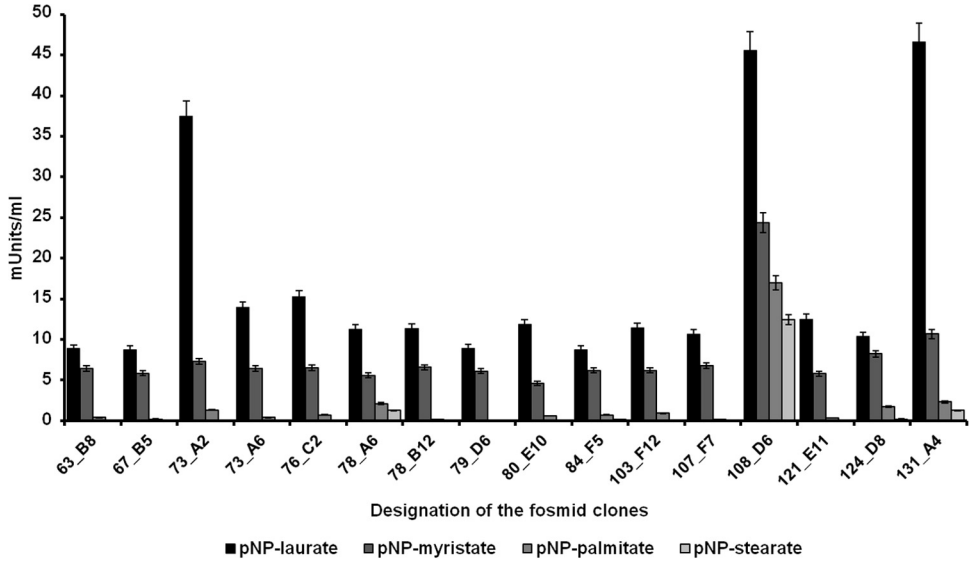


FIG 6 Lipolytic and esterolytic activities of 16 PBR- and metagenome-derived fosmid clones. Activities were assayed in the presence of various pNP substrates (pNP-laurate, pNP-myristate, pNP-palmitate, and pNP-stearate) at a final concentration of 1 mM. Extinction was measured at 405 nm against an enzyme-free blank after 30 min of incubation at 37°C. The data are mean values of at least three replicates. The error bars indicate simple standard deviations.

reduction of ribonucleotide triphosphate to 2′-deoxyribonucleotide triphosphate (80, 81). Higher eukaryotes (i.e., animals, humans, and protists) require B_{12} but do not synthesize the vitamer, and plants and fungi neither synthesize it nor use it. While plants have evolved a B_{12}-independent methionine synthase, algae require the vitamer for the function of B_{12}-dependent methionine synthase (22–24, 82, 83). Vitamin B_{12} is synthesized by bacteria and archaea, which have evolved two distinct routes for biosynthesis under aerobic and anaerobic conditions (22). More than 20 different steps are required for the biosynthesis of the corrinoid ring structure from uroporphyrinogen III (80, 81).

The KEGG-based analysis of the predicted genes and ORFs in the PBR biofilm metagenome suggested that all the genes with relevance to B_{12} vitamin biosynthesis were present, and many of them in multiple copies (see Fig. S5 in the supplemental material). Within this framework, it is notable that many algae are auxotrophic with respect to the production of thiamine and biotin, as well (21). Thiamine (vitamin B_1) is essential for some of the key enzymes in the primary carbohydrate and branched-chain amino acid metabolism. Altogether, we have identified at least 218 genes in the biofilm metagenome that are involved in thiamine biosynthesis (see Fig. S6 in the supplemental material) (84, 85). Biotin is a cofactor of biotin-dependent carboxylase reactions, among them acetyl coenzyme A (CoA) carboxylase, which is essential for fatty acid synthesis (86, 87). For all major genes involved in biotin biosynthesis, multiple copies were detected (see Fig. S7 in the supplemental material). Based on the KEGG analysis, almost 6.7% of all ORFs and genes encoded metabolism of cofactors and vitamin biosynthesis. Additional tests in an experimental test reactor confirmed that small amounts of the two vitamers B_{12} and biotin could already be detected in the PBR culture supernatant 48 h after inoculation (Fig. 7 and 8). The pico- and nanomolar amounts observed were certainly sufficient to support algal growth.

***N*-Acyl-homoserine lactone-dependent signaling plays a minor role in PBR biofilm formation.** Microbial biofilm formation depends to a large extent on the attachment of bacteria to surfaces, microcolony formation, and maturing of the biofilm structure (64, 88, 89). Within the PBR metagenome, many genes were identified that are essential to biofilm formation. Biofilm formation is regulated by a wide range of factors, including signaling molecules, such as autoinducer involved in quorum-sensing-dependent gene regulation. Interestingly, we have identified only a small number of genes that are involved in the synthesis of autoinducer I-type (*N*-acyl-homoserine lactone) molecules. *N*-Acyl-homoserine lactones are the key signaling molecules in the cell density-dependent system of gene regulation in many Gram-negative bacteria, and they are usually synthesized through a LuxI-like protein (90, 91). Employing the amino acid sequences of various LuxI homologues for Blastp searches, we could not detect more than a few homologues (<10) in the PBR metagenome. This finding is intriguing, and it might suggest that autoinducer I-dependent sig-

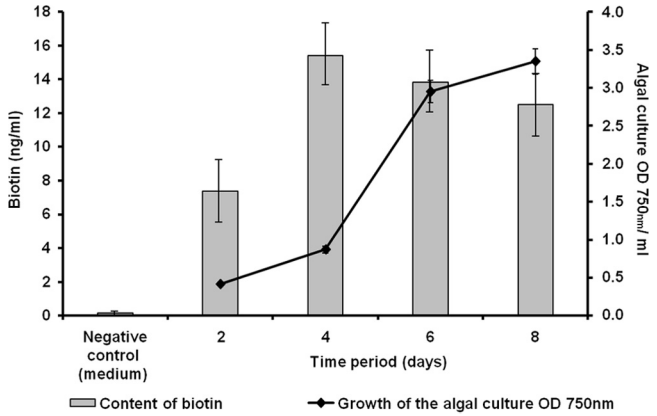


FIG 7 Biotin (ng/ml) detected in a PBR culture supernatant. Biotin was determined at different periods of growth of the algal culture. The data represent values of at least three replicates. The error bars indicate simple standard deviations.

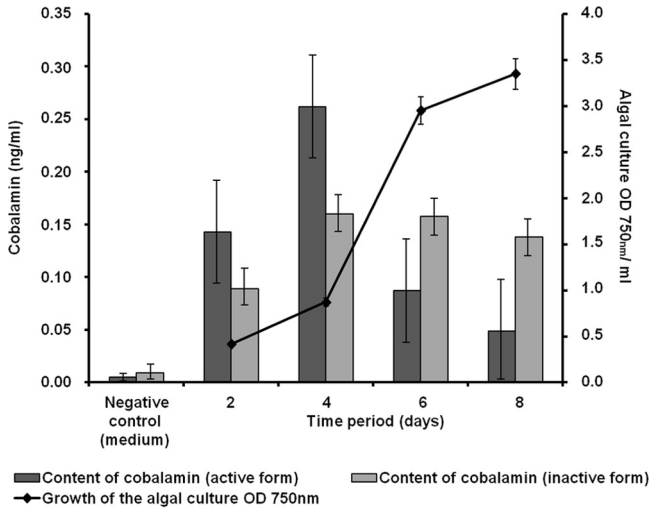


FIG 8 Content of cobalamin (ng/ml) in a model PBR. The cobalamin content (active and inactive forms) of the algal culture supernatant was determined at different time points during algal growth. The data represent values of at least three replicates. The error bars indicate simple standard deviations.

naling may play only a minor role within the observed ecological niche.

Summary and conclusions. Sequencing-based metagenome analyses in combination with function-based studies have in general significantly increased our knowledge of microbial communities and their phylogenetic makeup and genetic potential. Finally, these studies provided insight into the ecology of the analyzed communities and thereby enriched our understanding of biodiversity and biochemical functions. Many of these studies have been summarized previously (29, 92–94). Surprisingly, only a few studies have addressed metagenome technologies for the analysis of mixed-species biofilms in nonhuman interactions. The earliest studies dealt with the acid mine drainage biofilm (95) and a drinking water biofilm (96). More recently, studies have addressed microbial biofilm metagenomes of hydrothermal vent chimneys, biofilms on concrete water pipes, and an extremely acidophilic sulfur-oxidizing biofilm (97–99). Although not all studies have been listed here, it is evident that many different biotic and abiotic factors shape each of these communities and that each community clearly differs in its phylogenetic structure and genomic-information content.

In summary, our data give detailed insight into the microbial community and the metagenome of a PBR-associated microbial biofilm. Altogether, we observed 28 OTUs within the analyzed biofilm, and DGGE and RFLP analyses suggested that the community was rather stable over time. The sequence-based data and the laboratory measurement data clearly support earlier findings with respect to the bacterially produced B vitamins. Our data also suggest that bacterial growth depends in part on compounds released by the algae or bacteria within the community.

Finally, it is interesting that the bacteria in the PBR studied here encode many esterolytic and lipolytic enzymes. This finding was supported by extracting active lipolytic clones and by measuring the hydrolytic activities in the culture supernatants. Future work will have to use these data to further improve the efficiencies of PBRs with respect to CO_2 fixation and downstream processing of biofuels and/or other valuable compounds produced and released by the algae.

ACKNOWLEDGMENTS

This work was supported by the research and development project Technologies for Exploiting the Resource Microalgae (TERM) and the BMBF Project ChemBiofilm and was funded by the Federal Graduate School "C1-Chemistry in Resource and Energy Management" (C1-REM) (Hamburg, Germany).

REFERENCES

1. **Chisti Y.** 2007. Biodiesel from microalgae. Biotechnol. Adv. **25:**294–306.
2. **Williams PJLB, Laurens LML.** 2010. Microalgae as biodiesel and biomass feedstocks: review and analysis of the biochemistry, energetics and economics. Energy Environ. Sci. **3:**554–590.
3. **Ugwu CU, Aoyagi H, Uchiyama H.** 2008. Photobioreactors for mass cultivation of algae. Bioresour. Technol. **99:**4021–4028.
4. **Gouveia L, Oliveira A.** 2009. Microalgae as a raw material for biofuels production. J. Ind. Microbiol. **36:**269–274.
5. **Abomohra AF, Wagner M, El-Sheekh M, Hanelt D.** 2012. Lipid and total fatty acid productivity in photoautotrophic fresh water microalgae: screening studies towards biodiesel production. J. Appl. Phycol. **2012:**1–6.
6. **Waltz E.** 2013. Algal biofuels questioned. Nat. Biotechnol. **31:**12.
7. **Andersen RA.** 2004. Biology and systematics of heterokont and haptophyte algae. Am. J. Bot. **91:**1508–1522.
8. **Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJ.** 2004. The evolution of modern eukaryotic phytoplankton. Science **305:**354–360.
9. **Falkowski PG, Raven JA.** 2007. Aquatic photosynthesis, 2nd ed. Princeton University Press, Princeton, NJ.
10. **Lee YK.** 2001. Microalgal mass culture systems and methods: their limitation and potential. J. Appl. Phycol. **13:**307–315.
11. **Lehr F, Posten C.** 2009. Closed photo-bioreactors as tools for biofuel production. Curr. Opin. Biotechnol. **20:**280–285.
12. **Scott SA, Davey MP, Dennis JS, Horst I, Howe CJ, Lea-Smith DJ, Smith AG.** 2010. Biodiesel from algae: challenges and prospects. Curr. Opin. Biotechnol. **21:**277–286.
13. **Carvalho AP, Meireles LA, Malcata FX.** 2006. Microalgal reactors: a review of enclosed system designs and performances. Biotechnol. Prog. **22:**1490–1506.
14. **Cole JJ.** 1982. Interactions between bacteria and algae in aquatic ecosystems. Annu. Rev. Ecol. Syst. **13:**291–314.
15. **Lakaniemi AM, Hulatt CJ, Wakeman KD, Thomas DN, Puhakka JA.** 2012. Eukaryotic and prokaryotic microbial communities during microalgal biomass production. Bioresour. Technol. **124:**387–393.
16. **Lakaniemi AM, Intihar VM, Tuovinen OH, Puhakka JA.** 2012. Growth of Chlorella vulgaris and associated bacteria in photobioreactors. Microb. Biotechnol. **5:**69–78.
17. **Otsuka S, Abe Y, Fukui R, Nishiyama M, Sendoo K.** 2008. Presence of previously undescribed bacterial taxa in non-axenic Chlorella cultures. J. Gen. Appl. Microbiol. **54:**187–193.
18. **Watanabe K, Takihana N, Aoyagi H, Hanada S, Watanabe Y, Ohmura N, Saiki H, Tanaka H.** 2005. Symbiotic association in Chlorella culture. FEMS Microbiol. Ecol. **51:**187–196.
19. **Ueda H, Otsuka S, Senoo K.** 2009. Community composition of bacteria co-cultivated with microalgae in non-axenic algal cultures. Microbiol. Cult. Coll. **25:**21–25.
20. **Lakaniemi AM, Intihar VM, Tuovinen OH, Puhakka JA.** 2012. Growth of Dunaliella tertiolecta and associated bacteria in photobioreactors. J. Ind. Microbiol. Biotechnol. **39:**1357–1365.
21. **Croft MT, Warren MJ, Smith AG.** 2006. Algae need their vitamins. Eukaryot. Cell **5:**1175–1183.
22. **Warren MJ, Raux E, Schubert HL, Escalante-Semerena JC.** 2002. The biosynthesis of adenosylcobalamin (vitamin B12). Nat. Prod. Rep. **19:** 390–412.
23. **Helliwell KE, Wheeler GL, Leptos KC, Goldstein RE, Smith AG.** 2011. Insights into the evolution of vitamin B12 auxotrophy from sequenced algal genomes. Mol. Biol. Evol. **28:**2921–2933.
24. **Giovannoni SJ.** 2012. Vitamins in the sea. Proc. Natl. Acad. Sci. U. S. A. **109:**13888–13889.
25. **Mouget JL, Dakhama A, Lavoie MC, de la Noüe J.** 1995. Algal growth

enhancement by bacteria: is consumption of photosynthetic oxygen involved? FEMS Microbiol. Ecol. **18:**35–43.
26. **Berger PS, Rho J, Gunner H.** 1979. Bacterial suppression of Chlorella by hydroxylamine production. Water Res. **13:**267–273.
27. **Handelsman J.** 2004. Metagenomics: application of genomics to uncultured microorganisms. Microbiol. Mol. Biol. Rev. **68:**669–685.
28. **Simon C, Daniel R.** 2011. Metagenomic analyses: past and future trends. Appl. Environ. Microbiol. **77:**1153–1161.
29. **Streit WR, Schmitz RA.** 2004. Metagenomics—the key to the uncultured microbes. Curr. Opin. Microbiol. **7:**492–498.
30. **Simon C, Daniel R.** 2009. Achievements and new knowledge unraveled by metagenomic approaches. Appl. Microbiol. Biotechnol. **85:**265–276.
31. **Tremaroli V, Bäckhed F.** 2012. Functional interactions between the gut microbiota and host metabolism. Nature **489:**242–249.
32. **Weinstock GM.** 2012. Genomic approaches to studying the human microbiota. Nature **489:**250–256.
33. **Lewin A, Wentzel A, Valla S.** 2013. Metagenomics of microbial life in extreme temperature environments. Curr. Opin. Biotechnol. **24:**516–525.
34. **Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T.** 2011. Bacterial community assembly based on functional genes rather than species. Proc. Natl. Acad. Sci. U. S. A. **108:**14288–14293.
35. **Williams TJ, Wilkins D, Long E, Evans F, Demaere MZ, Raftery MJ, Cavicchioli R.** 2013. The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. Environ. Microbiol. **15:**1302–1317.
36. **Hindersin S, Leupold M, Kerner M, Hanelt D.** 2013. Irradiance optimization of outdoor microalgal cultures using solar tracked photobioreactors. Bioprocess Biosyst. Eng. **36:**345–355.
37. **Sambrook J, Russell DW.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
38. **De Man J, Rogosa M, Sharpe M.** 1960. A medium for the cultivation of lactobacilli. J. Appl. Bacteriol. **23:**130–135.
39. **Reasoner DJ, Geldreich EE.** 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl. Environ. Microbiol. **49:**1–7.
40. **Yeates C, Gillings MR, Davison AD, Altavilla N, Veal DA.** 1998. Methods for microbial DNA extraction from soil for PCR amplification. Biol. Proced. Online **1:**40–47.
41. **Lane DJ.** 1991. 16S/23S rRNA sequencing, p 115–175. *In* Stackebrandt E, Goodfellow MD (ed), Nucleic acid techniques in bacterial systematics. John Wiley and Sons, New York, NY.
42. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18:**821–829.
43. **Nakaya A, Katayama T, Itoh M, Hiranuka K, Kawashima S, Moriya Y, Okuda S, Tanaka M, Tokimatsu T, Yamanishi Y, Yoshizawa AC, Kanehisa M, Goto S.** 2013. KEGG OC: a large-scale automatic construction of taxonomy-based ortholog clusters. Nucleic Acids Res. **41:**D353–D357.
44. **Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV.** 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. **29:**22–28.
45. **Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer ELL, Bateman A.** 2008. The Pfam protein families database. Nucleic Acids Res. **36:**281–288.
46. **Buchholz-Cleven BEE, Rattunde B, Straub KL.** 1997. Screening for genetic diversity of isolates of anaerobic Fe(II)-oxidizing bacteria using DGGE and whole-cell hybridization. Syst. Appl. Microbiol. **20:**301–309.
47. **Muyzer G, de Waal EC, Uitterlinden AG.** 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. **59:**695–700.
48. **Kopylova E, Noe L, Touzet H.** 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics **28:**3211–3217.
49. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods **9:**357–359.
50. **Hartmann M, Howes CG, Abarenkov K, Mohn WW, Nilsson RH.** 2010. V-Xtractor: an open-source, high-throughput software tool to identify and extract hypervariable regions of small subunit (16S/18S) ribosomal RNA gene sequences. J. Microbiol. Methods **83:**250–253.
51. **Wingender J, Strathmann M, Rode A, Leis A, Flemming HC.** 2001. Isolation and biochemical characterization of extracellular polymeric substances from Pseudomonas aeruginosa. Methods Enzymol. **336:**302–314.
52. **Hara A, Radin NS.** 1978. Lipid extraction of tissues with a low-toxicity solvent. Anal. Biochem. **90:**420–426.
53. **Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F.** 1956. Colorimetric method for determination of sugars and related substances. Anal. Chem. **28:**350–356.
54. **Filisetti-Cozzi TM, Carpita NC.** 1991. Measurement of uronic acids without interference from neutral sugars. Anal. Biochem. **197:**157–162.
55. **Lowry OH, Rosebrough NJ, Farr AL, Randall RJ.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193:**265–275.
56. **Frølund B, Palmgren R, Keiding K, Nielsen PH.** 1996. Extraction of extracellular polymers from activated sludge using a cation exchange resin. Water Res. **30:**1749–1758.
57. **Elend C, Schmeisser C, Leggewie C, Babiak P, Carballeira JD, Steele HL, Reymond JL, Jaeger KE, Streit WR.** 2006. Isolation and biochemical characterization of two novel metagenome-derived esterases. Appl. Environ. Microbiol. **72:**3637–3645.
58. **Denter J, Bisping B.** 1994. Formation of B-vitamins by bacteria during the soaking process of soybeans for tempe fermentation. Int. J. Food. Microbiol. **22:**23–31.
59. **Entcheva P, Liebl W, Johann A, Hartsch T, Streit WR.** 2001. Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. Appl. Environ. Microbiol. **67:**89–99.
60. **Flemming HC, Wingender J.** 2010. The biofilm matrix. Nat. Rev. Microbiol. **8:**623–633.
61. **Branda SS, Vik S, Friedman L, Kolter R.** 2005. Biofilms: the matrix revisited. Trends Microbiol. **13:**20–26.
62. **Sutherland IW.** 2001. The biofilm matrix: an immobilized but dynamic microbial environment. Trends Microbiol. **9:**222–227.
63. **Flemming HC, Neu TR, Wozniak DJ.** 2007. The EPS matrix: the house of biofilm cells. J. Bacteriol. **189:**7945–7947.
64. **Stoodley P, Sauer K, Davies DG, Costerton JW.** 2002. Biofilms as complex differentiated communities. Annu. Rev. Microbiol. **56:**187–209.
65. **Kuczynski J, Stombaugh J, Walters WA, Gonzalez A, Caporaso JG, Knight R.** 2011. Using QIIME to analyze 16S rRNA gene sequences from microbial communities. Curr. Protoc. Bioinformatics **10:**10.7. doi:10.1002/0471250953.bi1007s36.
66. **Shannon CE, Weaver W.** 1949. The mathematical theory of communication, p 379–423. University of Illinois Press, Urbana, IL.
67. **Chao A, Lee SM.** 1992. Estimating the number of classes via sample coverage. J. Am. Stat. Assoc. **87:**210–217.
68. **Chao A.** 1984. Nonparametric estimation of the number of classes in a population. Scan. J. Statist. **11:**265–270.
69. **Christen R.** 2008. Global sequencing: a review of current molecular data and new methods available to assess microbial diversity. Microbes Environ. **23:**253–268.
70. **Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhee R, Reid G.** 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. PLoS One **5:**e15406. doi:10.1371/journal.pone.0015406.
71. **Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman DA, Sogin ML.** 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. PLoS Genet. **4:**e1000255. doi:10.1371/journal.pgen.1000255.
72. **Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glockner FO.** 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. **41:**e1. doi:10.1093/nar/gks808.
73. **Shen FT, Kämpfer P, Young CC, Lai WA, Arun AB.** 2005. Chryseobacterium taichungense sp. nov., isolated from contaminated soil. Int. J. Syst. Evol. Microbiol. **55:**1301–1304.
74. **Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Pujar A, Shearer AG, Travers M, Weerasinghe D, Zhang P, Karp PD.** 2012. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. **40:**D742–D753.
75. **Arpigny JL, Jaeger KE.** 1999. Bacterial lipolytic enzymes: classification and properties. Biochem. J. **343:**177–183.
76. **Jaeger KE, Ransac S, Dijkstra BW, Colson C, van Heuvel M, Misset O.** 1994. Bacterial lipases. FEMS Microbiol. Rev. **15:**29–63.
77. **Anraku Y.** 1988. Bacterial electron transport chains. Annu. Rev. Biochem. **57:**101–132.

78. **Stewart PS, Franklin MJ.** 2008. Physiological heterogeneity in biofilms. Nat. Rev. Microbiol. **6:**199–210.
79. **Williamson KS, Richards LA, Perez-Osorio AC, Pitts B, McInnerney K, Stewart P, Franklin MJ.** 2012. Heterogeneity in Pseudomonas aeruginosa biofilms includes expression of ribosome hibernation factors in the antibiotic-tolerant subpopulation and hypoxia-induced stress response in the metabolically active population. J. Bacteriol. **194:**2062–2073.
80. **Martens J, Barg H, Warren MJ, Jahn D.** 2002. Microbial production of vitamin B12. Appl. Microbiol. Biotechnol. **58:**275–285.
81. **Roth J, Lawrence J, Bobik T.** 1996. Cobalamin (coenzyme B12): synthesis and biological significance. Annu. Rev. Microbiol. **50:**137–181.
82. **Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG.** 2005. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature **438:**90–93.
83. **Duda J, Pedziwilk Z, Zodrow K.** 1957. Studies on the vitamin B12 content of the leguminous plants. Acta Microbiol. Pol. **6:**233–238.
84. **Begley TP, Ealick SE, McLafferty FW.** 2012. Thiamin biosynthesis: still yielding fascinating biological chemistry. Biochem. Soc. Trans. **40:**555–560.
85. **Jurgenson CT, Begley TP, Ealick SE.** 2009. The structural and biochemical foundations of thiamin biosynthesis. Annu. Rev. Biochem. **78:**569–603.
86. **Streit WR, Entcheva P.** 2003. Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. Appl. Microbiol. Biotechnol. **61:**21–31.
87. **Lin S, Cronan JE.** 2011. Closing in on complete pathways of biotin biosynthesis. Mol. Biosyst. **7:**1811–1821.
88. **Monds RD, O'Toole GA.** 2009. The developmental model of microbial biofilms: ten years of a paradigm up for review. Trends Microbiol. **17:**73–87.
89. **Klausen M, Gjermansen M, Kreft JU, Tolker-Nielsen T.** 2006. Dynamics of development and dispersal in sessile microbial communities: examples from Pseudomonas aeruginosa and Pseudomonas putida model biofilms. FEMS Microbiol. Lett. **261:**1–11.
90. **LaSarre B, Federle MJ.** 2013. Exploiting quorum sensing to confuse bacterial pathogens. Microbiol. Mol. Biol. Rev. **77:**73–111.
91. **Galloway WR, Hodgkinson JT, Bowden SD, Welch M, Spring DR.** 2011. Quorum sensing in Gram-negative bacteria: small-molecule modulation of AHL and AI-2 quorum sensing pathways. Chem. Rev. **111:**28–67.
92. **Gilbert JA, Dupont CL.** 2011. Microbial metagenomics: beyond the genome. Annu. Rev. Mar. Sci. **3:**347–371.
93. **Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, Bork P, Hugenholtz P, Rubin EM.** 2005. Comparative metagenomics of microbial communities. Science **308:**554–557.
94. **Simon C, Wiezer A, Strittmatter AW, Daniel R.** 2009. Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. Appl. Environ. Microbiol. **75:**7519–7526.
95. **Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature **428:**37–43.
96. **Schmeisser C, Stöckigt C, Raasch C, Wingender J, Timmis KN, Wenderoth DF, Flemming HC, Liesegang H, Schmitz RA, Jaeger KE, Streit WR.** 2003. Metagenome survey of biofilms in drinking-water networks. Appl. Environ. Microbiol. **69:**7298–7309.
97. **Gomez-Alvarez V, Revetta RP, Santo Domingo JW.** 2012. Metagenome analyses of corroded concrete wastewater pipe biofilms reveal a complex microbial system. BMC Microbiol. **12:**122.
98. **Brazelton WJ, Baross JA.** 2009. Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. ISME J. **3:**1420–1424.
99. **Jones DS, Albrecht HL, Dawson KS, Schaperdoth I, Freeman KH, Pi Y, Pearson A, Macalady JL.** 2012. Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. ISME J. **6:**158–170.

SUPPORTING INFORMATION FOR STUDY 8

CONTENTS:

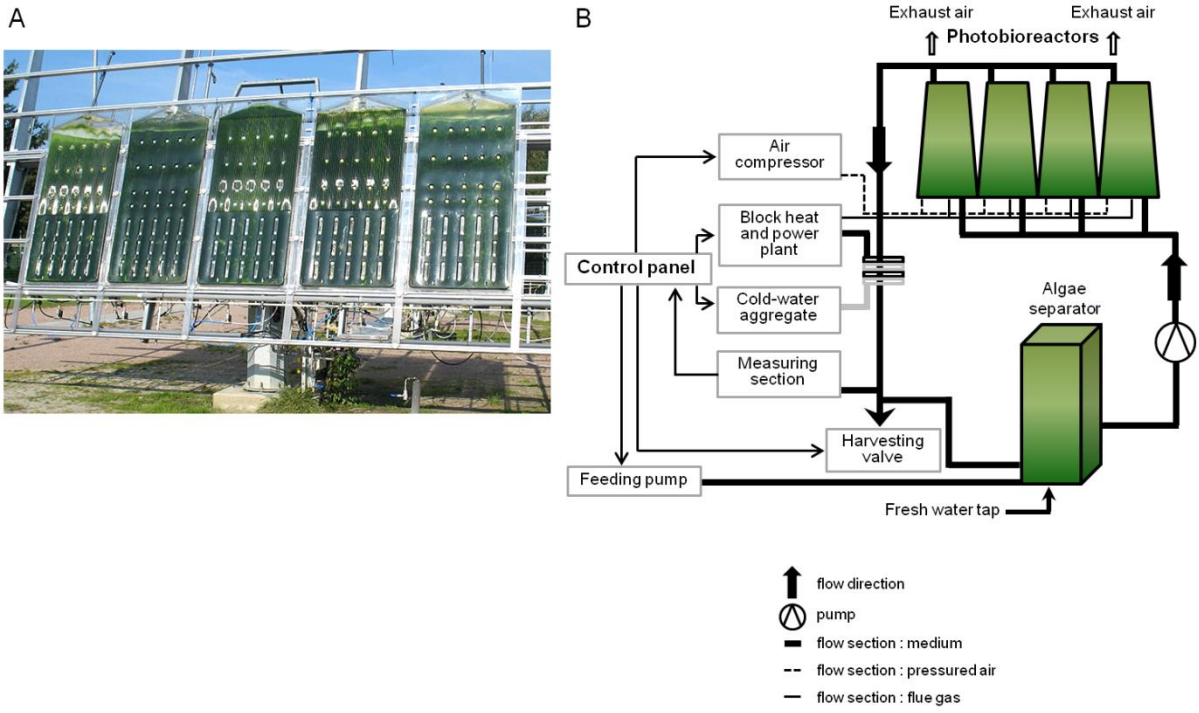


FIGURE S1, Krohn-Molt et al 2013,

Fig. S1:A: Photobioreactors mounted on a solar tracker.**B:**Flow chart of the microalgae pilot plant's function(modified from S. Hindersin, M. Leupold, M. Kerner, and D. Hanelt, "Irradiance optimization of outdoor microalgal cultures using solar tracked photobioreactors," Bioprocess Biosyst. Eng. 36:345-355, 2013, Fig. 2, with kind permission from Springer Science and Business Media [copyright 2013 Springer-Verlag]). For the duration of cultivation, part of the culture medium circulates through an external circuit. The algal biomass was harvesting continuously, which adjusts cell density to a constant value. The process parameters were monitored by measuring thetemperature, turbidity, pH, O_2, NH_4^+, NO_3^-, and K^+.Above a set value, nutrients were added automatically.A heat exchanger within the circular flow allowed temperature control.

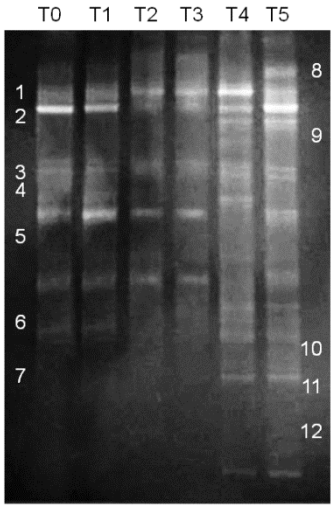


FIGURE S2, Krohn-Molt et al 2013

Fig S2: DGGE analysis of PCR-amplified 16S rRNA gene fragments obtained from six biofilm samples. Lanes 1-6 represent the different time points (T0-T5) of sampling, which were taken in 14 day intervals. The different DGGE bands (1-13) represent the 16S rRNA fragments of associated microorganisms, which could be assigned to sequences stored in the NCBI database. **(1)** *Pedobacter* sp. (EU585748.1), **(2)** *Flavobacterium* sp. (JX827624.1), **(3)** *Acidovorax* sp. (KC464816.1), **(4)** *Rhodobacter* sp. (KC157045.1), **(5)** plastid gene of *Nitzschiafrustulum* (AY221721.1), **(6)** chloroplast of *Scenedesmusobliquus* (DQ396875.1), **(7)** chloroplast of *Chlorella vulgaris* (AB001684.1), **(8)** Bacteroidetes bacterium (HM205113.1), **(9)** *Sinorhizobium* sp. (JQ316267.1), **(10)** *Caulobacter* sp. (EF020225.1), **(11)** uncultured Bacteroidetes bacterium (DQ463716.2), **(12)** uncultured Bacteroidetes bacterium (AY874003.1), **(13)** Gemmatimonadetes bacterium (JQ177808.1).

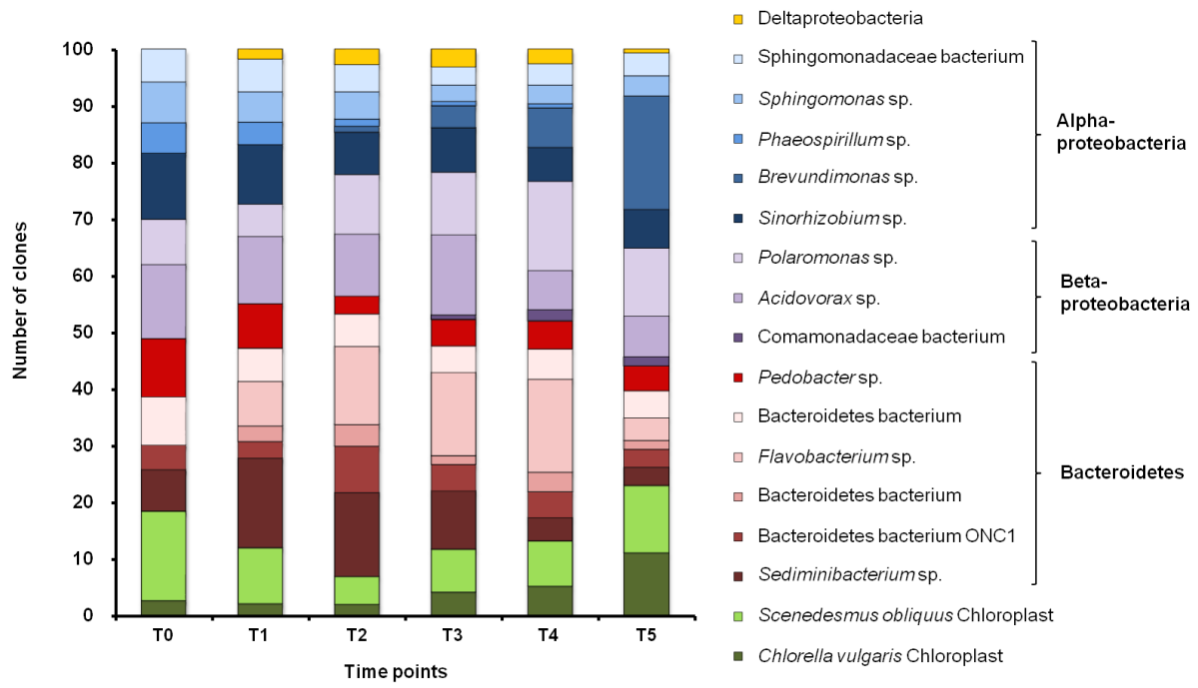


FIGURE S3, Krohn-Molt et al 2013

Fig. S3: Assignment of phylogenetic classes and orders of 16S rRNA sequences from the RFLP analysis. Per sample (T0-T5), 100 clones were analysed.

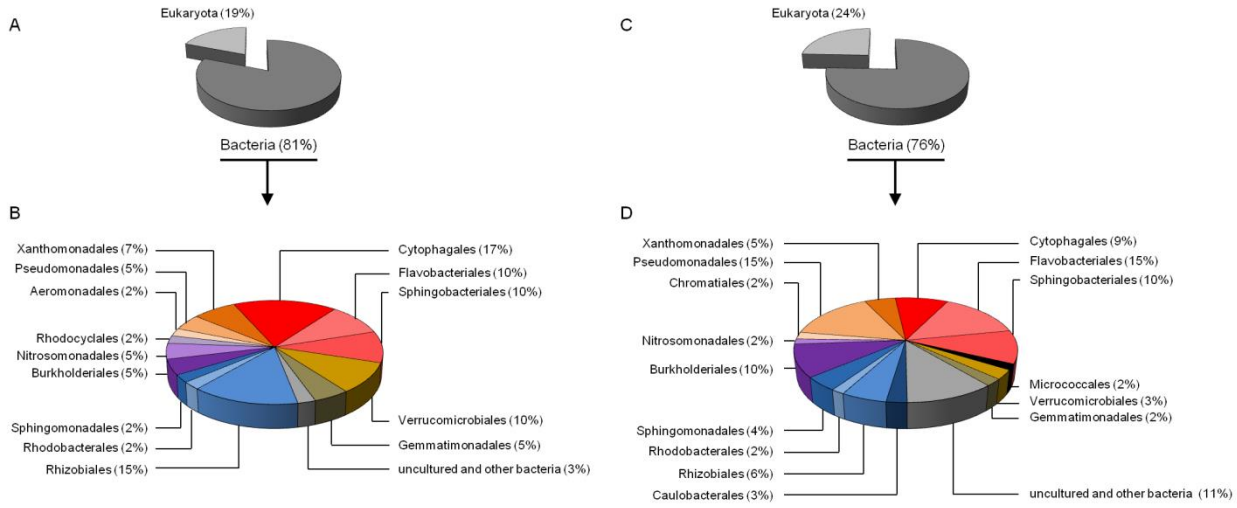


FIGURE S4, Krohn-Molt et al 2013

Fig. S4. Distribution of phylogenetic classes and orders of the metagenome-derived sequences in the PBR biofilm community. **A:**Overall diversity of 226 sequences from the FLX 454 dataset (Eukaryota 19%, Bacteria 81%), **B:** Phylogenetic classification of the V6 region of the 16S rRNAfrom the FLX 454 dataset (41 sequences). Alphaproteobacteria are represented by Sphingomonadales (2%), Rhodobacterales (2%), and Rhizobiales (15%).Betaprotebacteria were composed of Rhodocyclales (2%), Nitrosomonadales (5%), and Burkholderiales (5%).Gammaproteobacteriaconsistedof Xanthomonadales (7%), Pseudomonadales (5%), and Aeromonadales (2%). Bacteroidetes included Cytophagales (17%), Flavobacteriales (10%), and Sphingobacteriales (10%). 10% could be assigned to Verrucomicrobia(Verrucomicrobiales), and 5% could be classified asGemmatimonadetes(Gemmatimonadales). **C:**Overall diversity of 137,506 sequences from the Illumina dataset (Eukaryota 24%, Bacteria 76%),**D:** Phylogenetic classification of the V6 region of the 16S rRNA from the Illumina dataset (558 sequences).Alphaproteobacteria are represented by Sphingomonadales (4%), Rhodobacterales (2%), Rhizobiales (6%), and Caulobacterales (3%). Betaprotebacteria were composed of Nitrosomonadales (2%), and Burkholderiales (10%). GammaproteobacteriaconsistedofXanthomonadales (5%), Pseudomonadales (15%), and Chromatiales (2%). Bacteroidetes included Cytophagales (9%), Flavobacteriales (15%), and Sphingobacteriales (10%). 2% were represented by Actinobacteria(Micrococcales). 3% could be assigned to Verrucomicrobia(Verrucomicrobiales), and 2% could be classified asGemmatimonadetes(Gemmatimonadales).Phylogenetic groups accounting for <1% of the sequences and sequences which were not assigned to a specific taxonomic group were summarized in the artificial groups "uncultured and other bacteria".

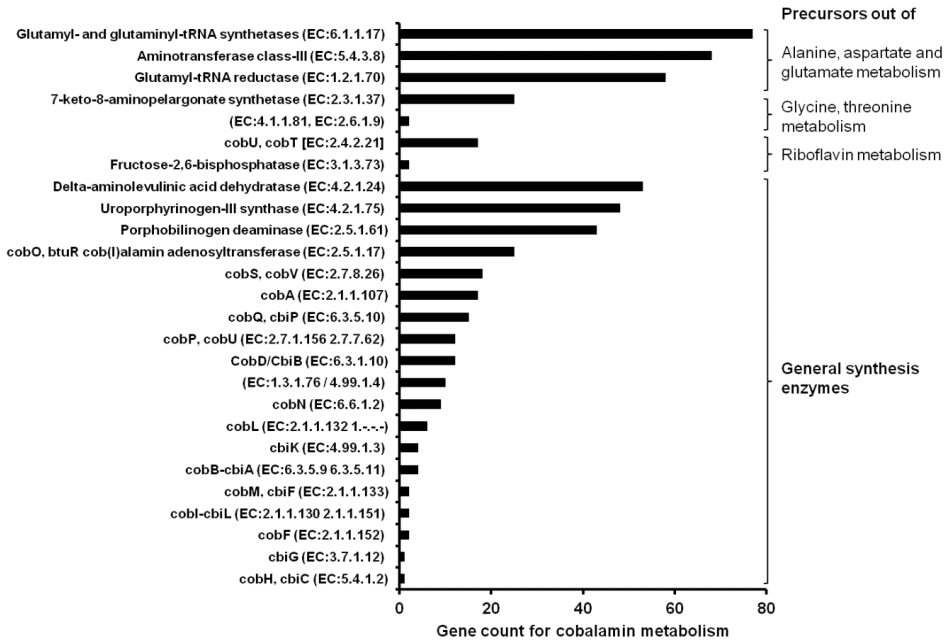


FIGURE S5, Krohn-Molt et al 2013

Fig. S5. Gene count for cobalamin metabolism. Analysis of the metagenome-derived sequences was done using the KEGG database. In total 533 sequences were assigned to the metabolism of cobalamin.

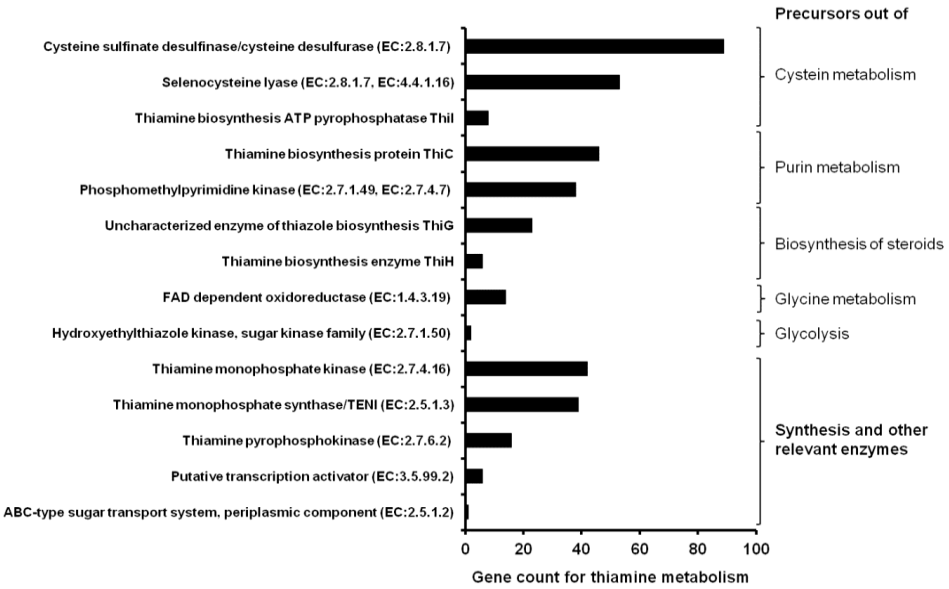


FIGURE S6, Krohn-Molt et al 2013

Fig. S6: Gene count for thiamine metabolism. Analysis of the metagenome-derived sequences was done using the KEGG database. In total 383 sequences were assigned to the metabolism of thiamine.

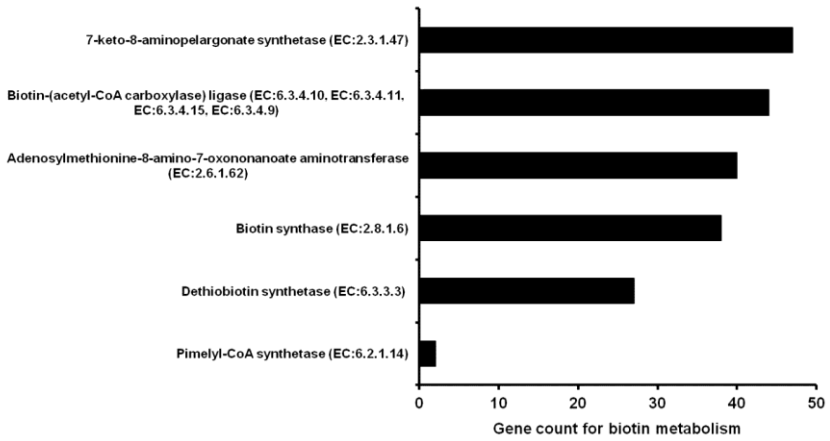


FIGURE S7, Krohn-Molt et al 2013

Fig. S7: Gene count for biotin metabolism. Analysis of the metagenome-derived sequences was done using the KEGG database. In total 198 sequences were assigned to the various biosynthesis steps of the vitamer biotin.

SUPPLEMENTAL TABLES

TABLE S1. Phylogenetic analysis of 171 16S rRNA clones sequences derived from thebiofilm's fully established microbial community (T5).

Phyla/Class	Sequence count	Order/Family	Sequence count
Alphaproteobacteria	48	Sphingomonadales	12
		Rhodobacterales	7
		Rhodospirillales	3
		Rhizobiales	15
		Caulobacterales	11
Betaproteobacteria	26	Burkholderiales	15
		uncultured Comamonadaceae	11
Deltaproteobacteria	6	Desulfovibrionales	3
		uncultured Deltaproteobacteria	3
Gammaproteobacteria	19	Xanthomonadales	16
		uncultured Gammaproteobacteria	3
		Flavobacteriales	10
		Cytophagales	5
		uncultured Bacteroidetes	13
Actinobacteria	9	Acidimicrobiales	2
		Actinomycetales	3
		uncultured Actionobacteria	4
Planctomycetes	7	Planctomycetales	7
Verrucomicrobia	5	uncultured Verrucomicrobia	5
Chloroflexi	3	uncultured Chloroflexi bacterium	3
unknown Bacteria	4	Candidate division OP10 bacterium	2
		uncultured TM7 bacterium	2

TABLE S2. Bacteria cultivated from the PBR microbial community on solid media.

Bacterial Isolate	**Closest relatives**	**16S rRNA gene identity (%)**
A	*Brevundimonas*sp. (JQ661035.1)	94
B	*Paracoccus*sp. (JQ404485.1)	96
C	*Chryseobacteriumtaichungense*(JQ071521.1)	99
D	*Brevibacterium*sp. / *Arthrobacter* sp. (GQ199748.1)	99
E	uncultivated / *Roseomonas*sp. (HQ588850.1)	78
F	uncultivated CFB affiliated bacterium (AJ583211.1)	96
G	*Xanthomonas*sp. (EU887990.1)	99
H	*Rhodococcus*sp. (EU041710.1)	99

STUDY 9:

IMPACT OF GRASSLAND MANAGEMENT REGIMES ON BACTERIAL ENDOPHYTE DIVERSITY DIFFERS WITH GRASS SPECIES AND SEASON

WEMHEUER F[1], KRETZSCHMAR D[1], **WEMHEUER B**[2], DANIEL R[2], AND VIDAL S[1]

(IN PREPARATION)

[1]DEPARTMENT OF CROP SCIENCES, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR.6, D-37077 GÖTTINGEN, GERMANY; [2]INSTITUTE OF MICROBIOLOGY AND GENETICS, GEORG-AUGUST-UNIVERSITY GÖTTINGEN, GRISEBACHSTR. 8, D-37077 GÖTTINGEN, GERMANY

Author contributions to the work:

Performed the experiments: FW, DK

Analyzed data: FW, BW

Wrote the publication: FW, RD, SV

Conceived and designed the experiments: FW, SV

Abstract

Most plant species are colonized by a diverse number of microorganisms including endophytic bacteria. Despite their importance for plant health and yield, the response of these bacteria to grassland management regimes is still largely unexplored. This study aimed at assessing the bacterial endophytic community structure in the agricultural important grass species *Lolium perenne* L., *Dactylis glomerata* L., and *Festuca rubra* L. with regard to different fertilizer and mowing treatments. For that purpose, above-ground plant material from the Grassland Management Experiment (GrassMan) in Germany was collected in September 2010 and 2011. To evaluate seasonal effects, additional samples were taken in April and July 2011. DNA was extracted from the plant material and subjected to 16S rRNA gene PCRs. The endophytic community structure was subsequently studied by Denaturing Gradient Gel Electrophoresis (DGGE). Management regimes did not impact the endophytic community structure in the grasses in the same manner. Fertilization and mowing frequency significantly altered the endophytic communities in *L. perenne* and *F. rubra* but not in *D. glomerata*. On the other hand, season significantly affected the community structure in all three grass species. Moreover, as community structures were subjected to temporal variations, the recorded impact of management regimes differed between the two investigated years.

Introduction

Almost all plant species are colonized by a high number of microorganisms including endophytic bacteria (Senthilkumar *et al.*, 2011). Endophytic bacteria are defined as bacteria that can be extracted from within plants or isolated from surface-disinfested plant tissue, and that have no visibly harmful effects on the plant (Hallmann *et al.*, 1997). They are found in a wide range of plants (Sturz *et al.*, 2000).

Many biotic factors including plant species, plant age, plant tissue, or the presence of phytopathogenic fungi, as well as abiotic factors such as soil conditions, temperature, or crop rotation influence the bacterial endophytic community (e.g., Hallmann *et al.*, 1997; Fuentes-Ramírez *et al.*, 1999; Sessitsch *et al.*, 2002; Seghers *et al.*, 2004; Hardoim *et al.*, 2012). Moreover, plant species vary in their biochemical composition, which may affect the endophytic bacterial community (Hallmann & Berg, 2006). As endophytic bacteria rely on the nutritional supply offered by the plant, any factor influencing the nutritional or physiological status of the plant may consequently have an impact on the endophytic community (Hallmann *et al.*, 1997; Fuentes-Ramírez *et al.*, 1999).

Several endophytic bacteria have been reported to promote plant growth, plant yield, and the overall plant health by a number of mechanisms. These include the production of phytohormones and antibiotics (Bacon & Hinton, 2006; Compant *et al.*, 2010) as well as enhanced nutrient availability and nitrogen fixation (Stoltzfus *et al.*, 1997; Rosenblueth & Martinez-Romero, 2006). Furthermore, plants infected with endophytic bacteria have a higher resistance to plant pathogens (e.g., Hallmann *et al.*, 1998; Hallmann, 2001; Krechel *et al.*, 2002; Siddiqui & Shaukat, 2003; Compant *et al.*, 2005) and environmental stresses (Sturz & Nowak, 2000; Bacon & Hinton, 2006; Bacon & Hinton, 2011).

Although their important role in agricultural cropping systems is frequently appreciated (e.g., Hallmann *et al.*, 1997; Kobayashi & Palumbo, 2000; Bacon & Hinton, 2006; Maksimov *et al.*, 2011; Senthilkumar *et al.*, 2011), the diversity of interactions between endophytic bacteria, plant species, and management regimes is not fully understood. Previous studies on the impact of different management regimes, such as fertilizer application, have mainly focused on root endophytic bacteria (Tan *et al.*, 2003; Seghers *et al.*, 2004; Kuklinsky-Sobral *et al.*, 2005), and nitrogen-fixing (diazotrophic) bacteria (Fuentes-Ramírez *et al.*, 1999; Sturz *et al.*, 2000; Tan *et al.*, 2003; Doty *et al.*, 2009; Prakamhang *et al.*, 2009).

The aim of this study was to investigate the influence of combined fertilizer applications and mowing regimes as well as the effect of season on the overall diversity of bacterial endophytes in three abundant and important agricultural grass species (*Dactylis glomerata* L., *Festuca rubra* L., and *Lolium perenne* L.). We hypothesized (1) that the overall endophytic community structure is different between the three examined grass species as the grasses differ in their physiological state. We further hypothesized (2) that the overall bacterial endophytic community structure of the investigated grasses is influenced by fertilizer application and different mowing frequencies as these management regimes affect the host plants and, thus, indirectly

the endophytes in the grasses. Moreover, we hypothesized (3) that the endophytic community in the grass species is influenced by season as the physiological state of the plant is altered with season.

For this purpose, above-ground plant material was taken from the Grassland Management Experiment (GrassMan), a long-term experimental field on a semi-natural, moderately species-rich grassland site. The aim of this experiment was to investigate the effects of fertilizer application, mowing frequencies, and sward composition on diversity and ecosystem functioning. For this purpose, ten samples per grass species and plot were collected in both September 2010 and 2011. To investigate the influence of season on the endophytic communities, 10 samples per grass species were collected from fertilized plots in April and July 2011. DNA was extracted from the plant material and subjected to 16S RNA gene PCR. Obtained PCR products were subsequently studied by DGGE analysis. In addition to the culture-independent approach, non-specialized endophytes were isolated from the grass species and classified by 16S rRNA gene analysis.

Materials and Methods

Study site

The Grassland Management Experiment (GrassMan) is a long-term field experiment with different management intensity treatments. It was established in spring 2008 at a semi-natural, moderately species-rich grassland site in the Solling Mountains in Lower Saxony, central Germany (51°44'53" N, 9°32'43" E, 490 m a.s.l.). At least since the late 19th century, this grassland site has been traditionally used as pasture or for hay making (Geological Map of Prussia 1910 (based on the topographic inventory of 1896), topographic maps of Sievershausen and Neuhaus/Solling 1924, 1956 and 1974). The pasture has been improved by annual fertilization (80 kg N ha^{-1} yr^{-1}), liming, and overseeding with high value forage species (farm records Relliehausen since 1966). The moderate fertilization stopped two years before the first experiments started. The vegetation consists of a nutrient poor, moderately wet Lolio-Cynosuretum (Petersen *et al.*, 2012). The mean annual temperature is 6.9°C and the mean annual precipitation is 1028 mm (Deutscher Wetterdienst 1960-1990, Station Silberborn-Holzminden, 440 m a.s.l.). During the study period, mean temperature and precipitation were 11.42°C and 93.6 mm in September 2010, 11.26°C and 41.75 mm in April 2011, 14.48°C and 110.85 mm in July 2011, and 14.75°C and 54.75 mm in September 2011, respectively. The dominating soil type of the experimental area has been determined as a shallow (40-60 cm), stony Haplic Cambisol (Keuter *et al.*, 2013) with a pH_{KCl} ranging from 4.18 to 5.47.

Experimental design

The three-factorial design of this study included two mowing frequencies (once per year in July vs. three times per year in May, July, and September) and two fertilizer treatments (no vs. NPK fertilizer application). All plots were cut to a height of 7 cm with a Haldrup® harvester. The N fertilizer was applied as calcium ammonium nitrate N27 in two equal doses (180 kg N ha^{-1} yr^{-1}) in April and end of May. In addition, 30 kg P ha^{-1} yr^{-1} plus 105 kg K ha^{-1} yr^{-1} as Thomaskali® (8% P_2O_5, 15% K_2O, 20% CaO) were also applied at the end of May. A third parameter manipulated was the sward composition (monocot-reduced, dicot-reduced, species-rich). This was achieved by selective herbicide application which either reduced dicot (Mecoprop-P and Fluroxypyr/ Triclopyr; 3 l ha^{-1} each) or monocot species diversity (Clethodim; 0.5 l ha^{-1}). One third of the plots was left untreated as control (species-rich). The application of herbicides took place on 31st July 2008 resulting in significant changes in species richness and in functional group abundances (Petersen *et al.*, 2012). Each treatment was replicated six times, resulting in 72 plots of 15 x 15 m size arranged in a Latin rectangle.

Sampling

Above-ground plant material was collected on 19th September 2010 and on 12th September 2011 (shortly before the third annual mowing application) from dicot-reduced plots. To investigate seasonal effects on the bacterial endophytic community structure in the three investigated grass species, samples from the intensively managed (fertilized, thrice mown), dicot-reduced plots were additionally collected on 12th April 2011 (prior to fertilizer application or mowing) and on 18th July 2011 (after fertilizer application and shortly before the second annual mowing application). Ten plants per grass species and plot were randomly selected for sampling, with one exception: due to the low number of *L. perenne* in the plots mown once a year in September 2010,

above-ground plant material was collected only from two non-fertilized and from three fertilized plots.

Collected plants did not show obvious disease symptoms, such as leaf spots, chlorosis, or other types of pathogen-induced lesions. Following cutting of above-ground plant material with sterilized scissors, the collected plant samples were immediately cooled down (below 4°C), transported to the laboratory, and kept frozen at -80°C until further use. Plant material derived from the same plot and plant species was pooled prior to DNA extraction.

Surface sterilization of plants

Surface-sterilization of plant tissues was performed according to Schulz *et al.* (1993), with slight modifications. Plant material was immersed in 37% formaldehyde for 3 min and rinsed two times with autoclaved and sterile-filtered water. To remove DNA, samples were rinsed with DNA-Exitus (Applichem, Darmstadt, Germany) for 30 s and subsequently washed three times in autoclaved and sterile-filtered distilled water. The surface-sterilized plant material was triturated with an autoclaved mortar and pestle or directly used for an isolation experiment. Until DNA extraction, the powdered samples were stored at -80°C.

Isolation of non-specialized endophytes

For the isolation experiment, surface-sterilized plant material from 9 plots (at least 2 of each treatment) was cut into several pieces of approximately 5 to 15 mm length. Ten to 15 plant fragments were placed on malt extract agar (MEA), Luria-Bertani-Agar (LB), and potato dextrose agar (PDA) plates. Moreover, at least 10 plant fragments were incubated in 1 mL NaCl-solution (1% (w/v). The tubes were extensively shaken for 10 s and then incubated for 20 to 30 min. Prior to shaking, five to six glass beads (3 mm) were added to increase the extraction efficiency. 400 µl of the resulting solution were pipetted onto an agar plate. The plates were incubated in the dark at 25°C for at least two weeks. Colonies were further cultivated in liquid culture (LB media). After one day growing at 25°C, DNA was extracted using the peqGold Plant DNA Mini Kit (Peqlab, Erlangen, Germany) were subjected to PCR-based amplification targeting the bacterial 16S rRNA gene.

Amplification of the 16S rRNA genes of isolated endophytic strains

PCR amplification of bacterial 16S rRNA genes was performed with the primers 8F 5'-AGAGTTTGATCMTGGC-3 (Muyzer *et al.*, 1995) and 1114R 5'-GGGTTGCGCTCGTTRC-3' (Wilmotte *et al.*, 1993). The PCR reaction mixture (25 µl) contained 2.5 µl of 10-fold Mg-free Taq polymerase buffer (Fermentas), 200 µM of each of the four desoxynucleoside triphosphates, 2 mM $MgCl_2$, 0.4 µM of each primer, 5% DMSO, 0.5 U of *Taq* DNA polymerase (Fermentas), and approximately 10 ng of the DNA sample as template. Negative controls were performed by using the reaction mixture without template. The following thermal cycling scheme was used: initial denaturation at 95°C for 2 min and 25 cycles of: 1 min at 95°C, 1 min at 55°C and 1.5 min at 72°C. The final extension was carried out at 72°C for 5 min. The resulting PCR products were checked for appropriate size and then purified using the peqGOLD Gel Extraction Kit (Peqlab) as recommended by the manufacturer. Sequences of the purified PCR products were determined by Sanger sequencing at the Göttingen Genomics Laboratory.

Extraction of total community DNA

Total microbial community DNA was extracted employing the peqGOLD Plant DNA Mini Kit (Peqlab) according to the manufacturer's instructions with two modifications. Glass beads were used in the first step to grind plant material. Furthermore, 10 µl Proteinase K (20 mg mL^{-1}) were added to improve initial cell lysis. DNA was eluted in 30 µl DEPC water.

Amplification of the 16S rRNA genes for DGGE analysis

For DGGE analysis, a nested PCR approach was applied. In the first PCR, the primers 799f (AACMGGATTAGATACCCKG) and 1492R (GCYTACCTTGTTACGACTT) were used to suppress co-amplification of plant chloroplast 16S rRNA gene DNA (Chelius & Triplett, 2001). PCR amplification with this primer pair resulted in two PCR products: a mitochondrial product with approximately 1.1 kbp and a bacterial product of approximately 735 bp.

The PCR reaction mixture (25 µl) for amplification of the target gene contained 2.5 µl of 10-fold Mg-free *Taq* polymerase buffer (Fermentas, St. Leon-Rot, Germany), 200 µM of each of the four desoxynucleoside triphosphates, 1.75 mM $MgCl_2$, 0.4 µM of each primer, 5% DMSO, 1.5 U

of *Taq* DNA polymerase (Fermentas), and approximately 25 ng of the DNA sample as template. Negative controls were performed by using the reaction mixture without template. Three independent PCR reactions were performed per sample and obtained PCR products were pooled in equal amounts. The following thermal cycling scheme was used: initial denaturation at 95°C for 5 min and thirty cycles of: 1 min at 94°C, 1 min at 53°C and 1 min at 72°C. The final extension was carried out at 72°C for 8 min. The resulting PCR amplicons were electrophoretically separated and bands specific for bacteria were excised from the gel. DNA was subsequently purified using the peqGOLD Gel Extraction Kit (Peqlab) according to manufacturer's instructions.

Purified products were subjected to nested PCR with the primer pair F968-GC (5'-AACGCGAAGAACCTTAC-3') and R1401 (5'-CGGTGTGTACAAGACCC-3') (Nübel *et al.*, 1996). To prevent complete denaturation of the fragment, a GC-rich sequence (5'-CGCCCGCCGCGCCCCGCGCCCGTCCCGCCG CCCCCGCCCG-3') was attached at the 5'- end of the primer F968-GC (Muyzer *et al.*, 1993). The same PCR reaction mixture as for the first PCR was used for nested PCR with one modification: only 1 U of *Taq* DNA polymerase (Fermentas) was added to the mixture. The thermal cycling scheme of the nested PCR was as follows: initial denaturation at 94°C for 5 min, 11 cycles of: 1 min at 94°C, 1 min at 60°C (minus 1°C per cycle) and 2 min at 72°C, followed by 17 cycles of: 1 min at 94°C, 1 min at 53°C and 2 min at 72°C. The final extension was carried out at 72°C for 10 min. The resulting PCR products were checked for appropriate size by agarose gel electrophoresis. Three independent PCR reactions were performed per sample and obtained PCR products were pooled in equal amounts.

Denaturing Gradient Gel Electrophoresis (DGGE)

To investigate the bacterial endophytic diversity, the products derived from 16S rRNA gene PCRs were studied by DGGE analysis. DGGEs were carried out by using a PhorU2 machine (Ingeny, Goes, the Netherlands) with a double gradient. The first gradient ranged from 55 to 68% denaturant with a second gradient of 6.2 to 9% acrylamide. The acrylamide gradient was applied to enhance band sharpness and resolution (Cremonesi *et al.*, 1997). The denaturant (100%) contained 7 M urea and 40% formamide. Approximately 100 ng of the PCR product were loaded. The DGGE run was performed in 1xTris-acetate-EDTA buffer (40 mM Tris, 20 mM sodium acetate, 1 mM Na_2EDTA [pH 7.4]) at 60°C. Following electrophoresis for 16 h at 100 V, the gels were stained for 60 min with SYBRGold (Invitrogen, Darmstadt, Germany) and subsequently photographed on a UV transillumination table. To compare the reproducibility of the statistical analysis of the DGGE profiles, at least two independent DGGE runs were performed.

DGGE data analysis

Analysis of DGGE profiles was carried out using the software package GELCOMPAR II, version 5.1 (Applied Math, Ghent, Belgium). Cluster analyses (UPGMA) based on Jaccard correlation indices considering band presence and absence were performed to evaluate the percentage of similarity shared among the samples from the different treatments and sampling dates. Due to the low plant number obtained for *L. perenne* in September 2010, these data were excluded from the cluster analysis. To further evaluate the impact of management regimes and sampling time, the results of the DGGE were analysed in R employing the vegan package (version 3.0.1). For this purpose, similarity matrices exported from GelCompare were converted into dissimilarity objects and subsequently analysed by *Permutational Multivariate Analysis of Variance Using Distance Matrices* (*adonis*) [http://cran.r-project.org/web/packages/vegan/vegan.pdf].

Identification of abundant bacterial community members by DGGE

To identify the most abundant members of the bacterial endophytic community, several dominant bands were excised from DGGE gels, re-amplified, and sequenced. Excised bands were incubated in 30 µl sterile TE buffer (pH 8) overnight at 4°C. One µl of the resulting solution was subjected to PCR reaction to re-amplify the 16S rRNA gene fragment. The PCR was performed as described for the nested PCR reaction with one exception: the forward primer F968 did not carry the GC clamp. The resulting PCR products were checked for appropriate size and purified using the peqGOLD Gel Extraction Kit (Peqlab) as recommended by the manufacturer. The Göttingen Genomics Laboratory determined the sequences of the purified PCR products by Sanger sequencing.

Further Analysis of 16S rRNA gene sequences

All obtained 16S rRNA gene sequences were further analyzed employing the QIIME software package (version 1.6) (Caporaso *et al.*, 2010) and other tools. The Uchime algorithm implemented in Usearch (version 6.0.152) was initially applied in reference mode to identify and remove putative chimeric sequences using the most recent SILVA database (SSURef 115 NR) (Quast *et al.*, 2013) as reference dataset. Afterwards sequences were clustered into operational taxonomic units (OTUs) at 99% genetic similarity by BLAST alignment against the above-mentioned SILVA database using the pick_otus.py script (QIIME). The phylogenetic composition was determined by classifying the sequences with respect to the silva taxonomy of their closest match.

Nucleotide sequence accession numbers

Nucleotide sequences of the isolated strains and sequenced DGGE bands were deposited in GenBank under accession numbers KF699892 to KF699947 and KF699948 to KF700039, respectively.

Results and Discussion

Community structure differs with grass species and analysis approach

To assess endophytic community structures in the three grass species, DNA was extracted from plant material and subjected to 16S rRNA gene PCRs. Obtained PCR products were studied by DGGE analysis. DGGE fingerprints revealed patterns with 10 to 20 bands for each sample (Figs. S1-3). Prominent bands were excised and sequenced. Analysis of the obtained sequences revealed that bacterial diversity on class level was lowest and highest in *L. perenne* and *D. glomerata* (Fig. 1), respectively. *Gammaproteobacteria* were the most dominant bacterial phylum in both *D. glomerata* and *F. rubra*. This is in agreement with other studies (Chelius & Triplett, 2001; Sun *et al.*, 2008; Gottel *et al.*, 2011). Endophytic bacteria in *L. perenne* were dominated by *Betaproteobacteria.* The second most dominant groups were *Bacilli* (*D. glomerata*), *Betaproteobacteria* (*F. rubra*), or *Gammaproteobacteria* (*L. perenne*). Within the *Gammaproteobacteria*, we identified *Pseudomonas* as the most common genus (Table S1). One interesting species identified was *Herbaspirillum seropediacae* which is known as a nitrogen-fixing endophyte in sorghum, maize, sugarcane, and other plants (Baldani *et al.*, 1986; Olivares *et al.*, 1996).

We further examined how similar/dissimilar the endophytic communities are between the three investigated grass species. The number of calculated operational taxonomic units (OTUs) shared between the species was lower than the number exclusively found in one species (Fig. 2) which may refer to the different physiological states of the grass species investigated. Whereas 10 of the 29 identified OTUs of *D. glomerata* were also detected in *F. rubra* and *L. perenne*, the latter species shared 7 OTUs. Only 5 OTUs were found being present in all three grass species: one uncultured bacterium of the *Comamonadaceae*, *Staphylococcus aureus*, *S. epidermidis*, *Janthinobacterium lividum*, and *Pseudomonas balearica* (Table S1). The recorded findings support our first hypothesis that the grass species differ in their endophyte community structure. This is in accordance with a study of McInroy and Kloepper (1995A) who found differences in the

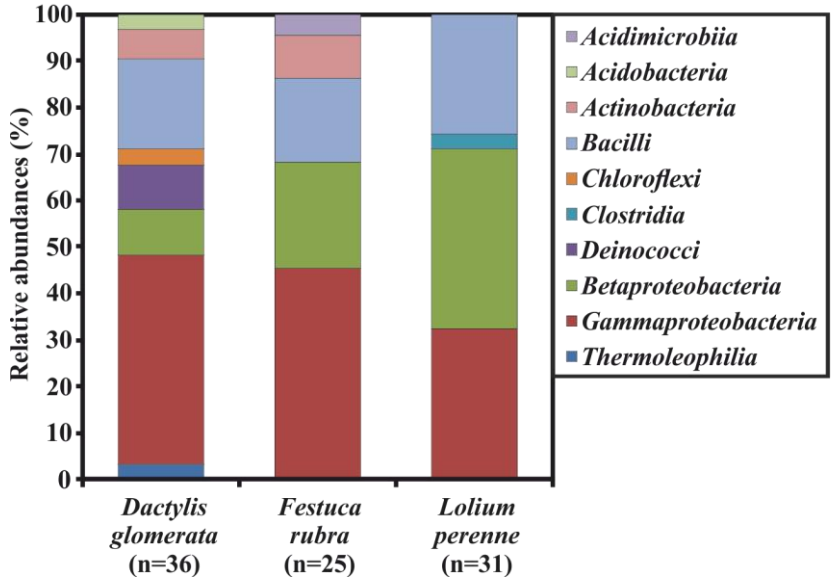


Fig. 1. Composition of the endophytic bacterial communities in the three grass species as revealed by sequencing of prominent DGGE bands. The number below the species name refers to the number of 16S rRNA genes sequences used in the analysis.

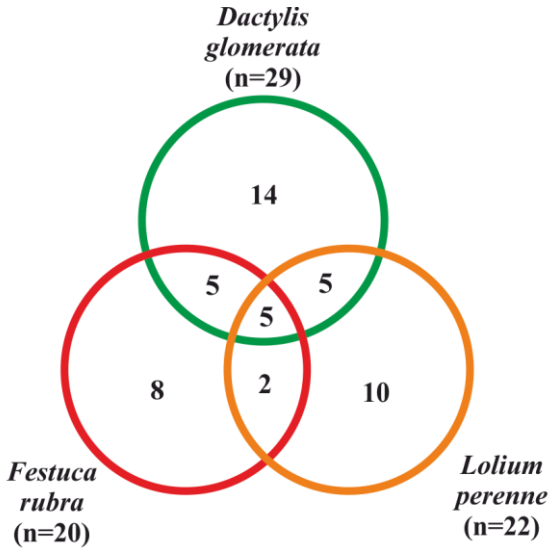


Fig. 2. Number of shared operational taxonomic units (OTUs) at 1% genetic distance. The number below the species name refers to the number of OTUs used in the analysis.

bacterial endophyte population in field-grown sweet corn and cotton grown side by side. They suggested that internal plant niches are colonized by a wide variety of bacteria. According to Hallmann (2001), the differences in bacterial endophytic community structures between different plant species growing next to each other can only be explained by plant species-specific selection mechanisms. Different plant species vary in their biochemical composition, which may affect bacterial endophyte community (Hallmann & Berg, 2006).

Moreover, the spectrum of indigenous endophytic bacteria in roots is not only affected by niche specialization, but also by differences in colonization pathway (Hallmann & Berg, 2006). According to Hallmann and Berg (2006), the soil and the rhizosphere are the main sources of endophytic colonizers. Many bacteria in these environments are able to penetrate and colonize root tissues (Quadt-Hallmann *et al.*, 1997; Reinhold-Hurek & Hurek, 1998). Plant wounding either by abiotic (e.g., tillage, extreme temperature fluctuations) or by biotic factors (e.g., fungi, plant-parasitic nematodes, insects) can also result in microbes entering the plant tissue (reviewed in Siddiqui & Shaukat, 2003). Other possible sources for endophytic bacteria include the anthosphere, the seeds, and the phyllosphere (Hallmann *et al.*, 1997; Hallmann, 2001; Compant *et al.*, 2010).

We also tried to assess the endophytic community structure by isolating strains from the three grass species. The most dominant groups isolated from the grasses were members of the *Bacilli* and *Gammaproteobacteria*, with *Pseudomonas* and *Bacillus* being the most abundant genera (Table S1). This is in accordance with other studies (as reviewed in Hallmann & Berg, 2006). However, a comparison of OTUs calculated for the 16S rRNA gene datasets obtained from the culturing-dependent and from the culturing-independent approach exhibited no overlap of the endophytic communities (Table S1). Consequently, the isolated strains do not necessarily represent the dominant endophytes in the three grasses which is supported by other studies (e.g., Chelius & Triplett, 2001; Garbeva *et al.*, 2001; Araujo *et al.*, 2002; Conn & Franco, 2004). For example, *Araujo et al.* (2002) showed that some endophytic bacteria in citrus plants were only observed by DGGE and not by the culture-dependent approach. In a study with potato plants, several non-culturable or so far uncultured endophytic organisms were detected. According to Chelius and Triplett (2001), the culturable component of the bacterial community reflected a community composition different from that of the clone library. Thus, only the community structures assessed by the metagenomic approach were further examined for their response to different management regimes and season.

Fertilizer application and mowing regimes differently shape bacterial endophytic community composition in D. glomerata, L. perenne, and F. rubra

In order to validate our second hypothesis that different fertilizer application and mowing regimes alter the bacterial endophytic communities, we compared DGGE band patterns with respect to the different management practises. UPGMA dendrograms of endophytic bacterial communities in *D. glomerata*, *L. perenne*, *and F. rubra* revealed differences with regard to fertilizer treatments and mowing frequencies (Figs. 3-5). Plants of *D. glomerata* sampled in September 2010 (Fig. 3A) and 2011 (Fig. 3B) did not cluster with respect to the applied management regimes. Furthermore, a significant influence of fertilizer application or mowing frequency was not recorded (Table 1). In contrast to *D. glomerata*, cluster analysis for *F. rubra* revealed a strong impact of the fertilizer treatment on bacterial endophytic community in September 2010 (Fig. 4A), but to a lesser extend in

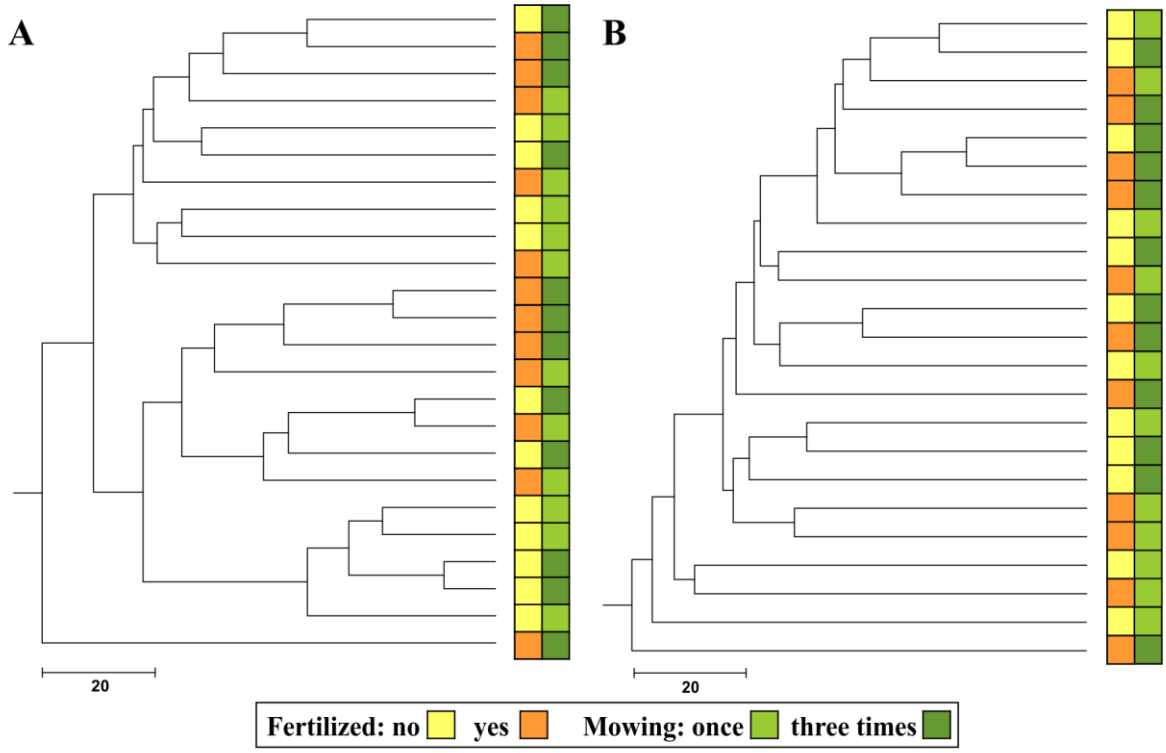


Fig. 3. UPGMA dendrogram generated by cluster analysis of DGGE fingerprints on the influence of different fertilization and mowing regimes on bacterial endophyte communities in above-ground plant parts of *D. glomerata*. Plant samples were taken in September 2010 (A) and 2011 (B). The dendrogram was constructed using the Jaccard correlation coefficient. The scale shows similarity values.

2011 (Fig. 4B). Furthermore, fertilizer application affected the community structure of bacterial endophytes in plants of *L. perenne* in September 2011 (Fig. 5). Such clear patterns were not recorded for the mowing regime. These results are in concordance with the statistical evaluation: fertilization and the interaction of fertilizer application and mowing frequency but not of mowing itself significantly influenced the structure of the endophytic community in *F. rubra* in September 2010 and in *L. perenne* in September 2011 (Table 1).

It is well-known that different management practices have an impact on bacterial endophytic communities, but most previous research has focused on root endophytes (Tan *et al.*, 2003; Seghers *et al.*, 2004; Kuklinsky-Sobral *et al.*, 2005) or on nitrogen-fixing (diazotrophic) endophytes (Fuentes-Ramírez *et al.*; 1999, Sturz *et al.*, 2000; Tan *et al.*, 2003; Doty *et al.*, 2009; Prakamhang *et*

Table 1. Statistical evaluation of the influence of management regimes and season towards the bacterial endophyte community in *D. glomerata*, *F. rubra*, and *L. perenne*. Abbreviation: Fert.:Mow. = the interaction of fertilization and mowing.

Species		***Management regimes***			***Time***	
		Fertilization	***Mowing***	***Fert.:Mow.***	***Season***	***Year***
D. glomerata	2010	-	-	-		
	2011	-	-	-		
	-				***	**
F. rubra	2010	**	-	***		
	2011	-	-	-		
	-				***	***
L. perenne	2010	NA	NA	NA		
	2011	*	-	**		
	-				***	***

not significant (-); significant with $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***)

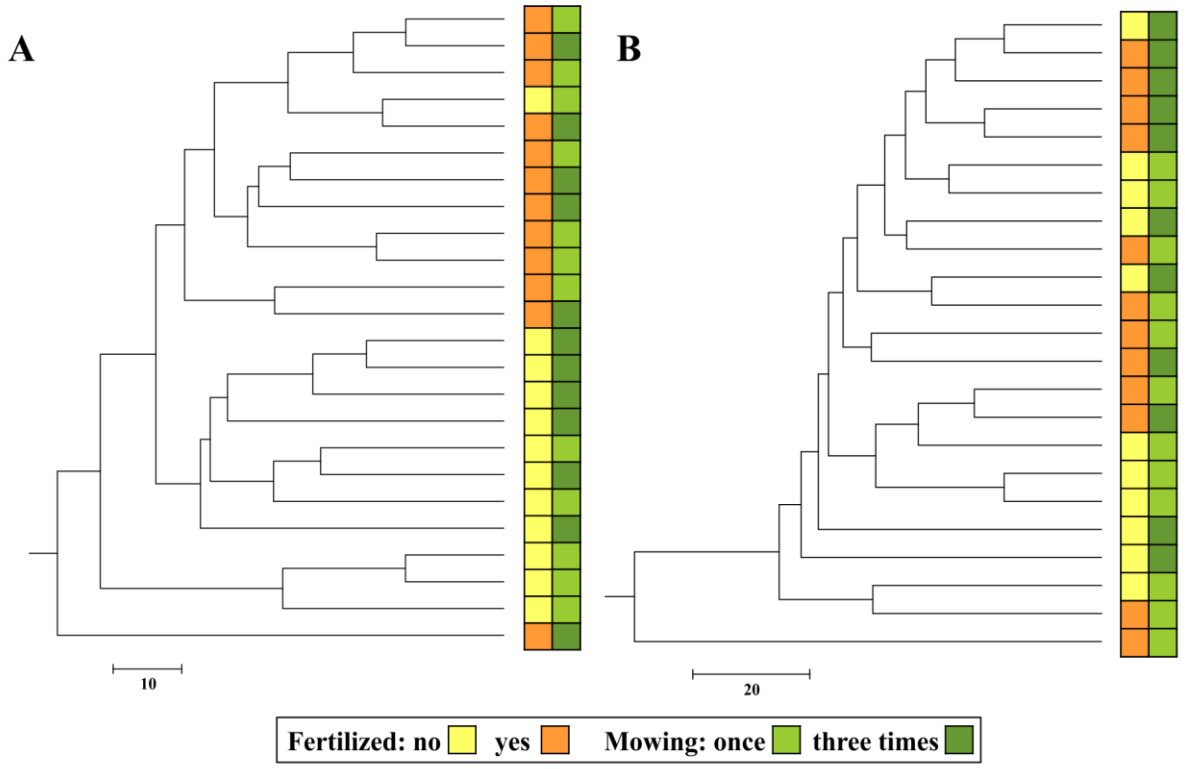


Fig. 4. UPGMA dendrogram generated by cluster analysis of DGGE fingerprints on the influence of different fertilization and mowing regimes on bacterial endophyte communities in above-ground plant parts of *F. rubra*. Plant samples were taken in September 2010 (A) and 2011 (B). For details see Fig. 3.

al., 2009). For example, endophytic populations in cotton roots are affected by application of nitrogen-containing chitin as an organic amendment (Hallmann *et al.*, 1999). Moreover, a higher diazotrophic bacterial diversity in the roots of rice cultivated in unfertilized and previously uncultivated soil than in paddy soil amended with nitrogen fertilizer were recorded by Prakamhang *et al*. (2009). According to Tan *et al*. (2003), a rapid change of both the population and the activity of nitrogen-fixing bacteria in rice roots were observed within 15 days after N-fertilization. Although these studies investigated the endophytic community in cotton and rice roots, they are in accordance with the results of the present study. Plant samples of *D. glomerata* in both years investigated and plant samples of *F. rubra* taken in September 2011 showed no significant impact of any management regime. This result is concordant with a study of Seghers *et al.* (2004), which showed that mineral fertilizer as well as herbicide application exhibited no impact on bacterial endophytic community structure in maize kernels. The recorded findings partly support our initial hypothesis as some but not all investigated plant samples were affected by the applied management regimes.

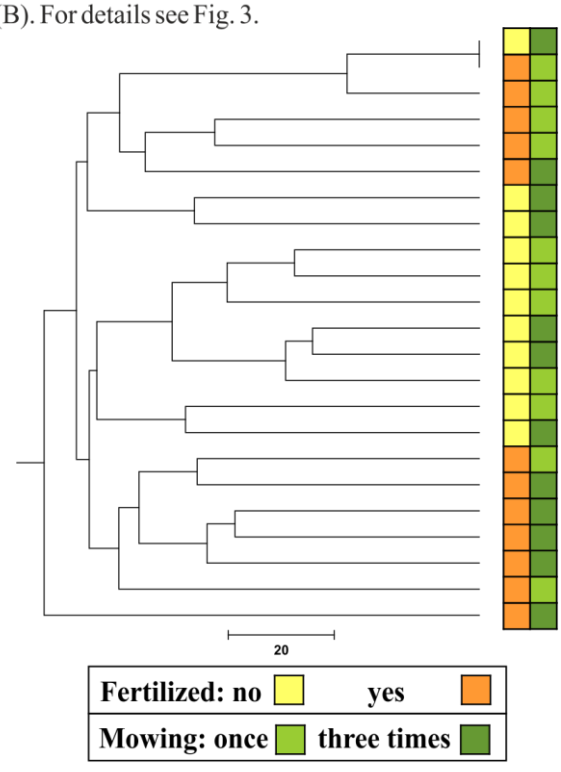


Fig. 5. UPGMA dendrogram generated by cluster analysis of DGGE fingerprints on the influence of different fertilization and mowing regimes on bacterial endophyte communities in above-ground plant parts of *L. perenne*. Plant samples were taken in September 2011. For details see Fig. 3.

Moreover, as the recorded effects on endophytic communities were different between the three grass species examined in this study, it is most likely that also the grasses are affected differently by management regimes which is in concordance with our second hypothesis.This was supported by an experiment in the Fraser Valley of British Columbia (Parish *et al.*, 1990). In five consecutive years, the authors investigated the effects of two different fertilizer levels (non-fertilized, fertilized) and four frequencies of mowing on the botanical composition of a pasture. At the end of the study, only *D. glomerata* was found in all treatments, while the abundance of *Lolium* spp. declined considerably. Mowing and fertilizer application every 3 weeks had a significant impact on the abundance of all investigated species. Furthermore, there was a significant fertilization - mowing interaction effect on all species except *Festuca* sp. The authors suggested that the plants differ in their growth rates and tolerance to shading and fertilizer application.

Additionally, the grass species investigated in this study differ in their indicator values such as tolerance against mowing or grazing (Dierschke & Briemle, 2002). Both *D. glomerata* and *L. perenne* have a higher tolerance against mowing compared to *F. rubra*. In contrast, *L. perenne* shows a higher indicator value for nitrogen than the other two grass species. As mentioned earlier, plants vary in their biochemical composition which might explain differences in the bacterial endophytic community (Hallmann & Berg, 2006). Hallmann *et al.* (1999) suggested that changes in plant physiology may result in the development of distinct bacterial endophytic communities. Moreover, endophytic bacteria rely on the nutritional supply offered by their host plant. As a consequence, changes in the nutritional or physiological status of the host plant may have an influence on the plant's endophytic community (Hallmann *et al.*, 1997; Fuentes-Ramírez *et al.*, 1999).

Seasonal impact on the abundance of bacterial endophytic community in the three grass species D. glomerata, F. rubra, and L. perenne

To verify our third hypothesis that the season has an effect on the bacterial endophytic community structure, we compared DGGE band patterns obtained from plant samples collected in September 2010 and April, July, and September 2011 (Figs. 6 and S4). Band patterns of *F. rubra* samples taken during the same season clustered together indicating a more similar community composition at the same season (Fig. 6B). Four of the six July samples cluster together with samples taken in September 2011. The other two samples showed higher similarities to samples taken in April 2011 and September 2010. This may indicate that the bacterial community composition in *F. rubra* followed a within year pattern. Plant species that propagate vegetatively are able to transmit their endophytes to the next generation so that no infection is required (Rosenblueth & Martinez-Romero, 2006). *Festuca rubra* is propagated mainly by rhizomes. Therefore, this propagation pattern might explain our findings that endophytic communities were quite similar in July and September 2011 in this grass species.

Moreover, three of the six samples taken in September 2010 and all samples from April 2011 formed a coherent cluster suggesting that they harbor a similar endophytic community. This distinct cluster pattern might be explained by seed transfer although this mechanism was not specifically tested in our study. It is known for some perennial plant species that several bacterial endophytes are seed-borne. These species are transferred from one plant generation to the next through the seeds of many plant species such as tobacco (Mastretta *et al.*, 2009), rice (Hardoim *et al.*, 2012), or Norway spruce (Cankar *et al.*, 2005). In a study of endophytic bacteria in switchgrass, some bacterial species were found in plants that originated from seeds sampled a year earlier (Gagne-Bourgue *et al.*, 2013). The authors regarded this as evidence for a vertical transmission to the next generation within this host plant.

Cluster analysis of the bacterial endophytic community in *L. perenne* revealed a clear separation of groups based on sampling year and season (Fig. 6C). Samples taken in 2011 formed a coherent cluster and exhibit a higher similarity to each other compared to samples taken in 2010. This finding suggests that the community structures in this grass species were different between both investigated years. Interestingly, samples collected in April 2011 and September 2011 were more similar to each other compared to samples taken in July 2011. This indicates that endophytic communities in *L. perenne* followed a seasonal pattern and that endophytic communities respond to changing climatic conditions.

DGGE band patters derived from *D. glomerata* samples revealed that samples taken in April 2011 clustered together, suggesting that they harbor a homogenous community composition (Fig. 6A). In accordance with *L. perenne*, three of the six samples taken in September 2011 were more similar to samples taken in April 2011. The other three samples of September 2011 were related to samples taken in July 2011. Such a pattern was already reported for *F. rubra*. Furthermore, samples taken in September 2010 were more similar to some of the samples taken in September 2011. These data suggested that the bacterial endophytic community in *D. glomerata* was less variable over consecutive years as, for example, the community in *L. perenne*. This might be explained by the higher endophytic diversity (number of OTUs) observed in *D. glomerata* compared to *L. perenne*; the smaller the community size the stronger the impact of seasonal fluctuations of single species on community structure.

The different seasonal patterns recorded for the three grass species confirm our first hypothesis that the overall endophytic community structure is different between the three examined grass species. Moreover, statistical analysis supported our third hypothesis that the season has an effect on the bacterial endophytic community in the three grasses as both season and year significantly influenced the composition of these communities (Table 1). This result is consistent with other studies. According to McInroy and Kloepper (1995B), the bacterial endophytic population in sweet corn and cotton fluctuated seasonally. The season also influenced the bacterial endophytic community in elm (Mocali *et al.*, 2003) and in soybean (Kuklinsky-Sobral *et al.*, 2004). However, only cultivable endophytes were investigated in these studies.

During the year, plants undergo physiological changes that probably increase nutrient availability and thus bacterial diversity in the roots (Hallmann & Berg, 2006). This might also play a role for endophytic bacteria in the above-ground plant tissues and could explain the high similarity of *F. rubra* and *D. glomerata* samples from September and July 2011 compared to samples from April 2011. Tan *et al.* (2003) showed that environmental conditions strongly influenced the diazotrophic endophytic community structure in rice roots. Several factors, such as temperature or precipitation, have a direct effect on the plant physiology and thus an indirect impact on the colonization and the survival of bacteria in the endosphere (Hallmann *et al.*, 1997; Hardoim *et al.*, 2012). This might explain the fact that the endophytic community structure in *L. perenne* in spring and autumn showed a higher similarity compared to the community in summer due to higher precipitation in summer.

In conclusion, our results demonstrate that different management regimes affect certain bacterial endophyte communities in grass species. However, this influence varies between the applied management regimes as the effect of the fertilizer application is clearer visible compared to the impact of different mowing frequencies. In addition, the influence of the management regimes can alter with time as seasonal changes also have an impact on the endophytic community composition. Interestingly, the effect of different management regimes and season is dependent on the host species as differences between the three investigated grass species were recorded. So far, the majority of the studies examined the effect of only one management regime in one single year, or focused on culturable endophytes or one functional group only. This study provides first insights into structural changes of endophyte communities in three agricultural important grass species as response to combined fertilizer application and mowing regimes as well as season. More studies targeting the influence of management regimes in combination with the impact of season and plant species are required to unravel the diversity of interactions between endophytic bacteria, plant species and management regimes.

Acknowledgments

We thank the technical staff of the Department of Crop Sciences and the Department of Genomic and Applied Microbiology at the University of Göttingen for help with experimental maintenance. In addition, we are grateful to Birgit Pfeiffer for providing assistance with the DGGE. This study was funded by the Ministry of Science and Culture of Lower Saxony and the 'Niedersächsisches Vorab' as part of the Cluster of Excellence 'Functional Biodiversity Research'.

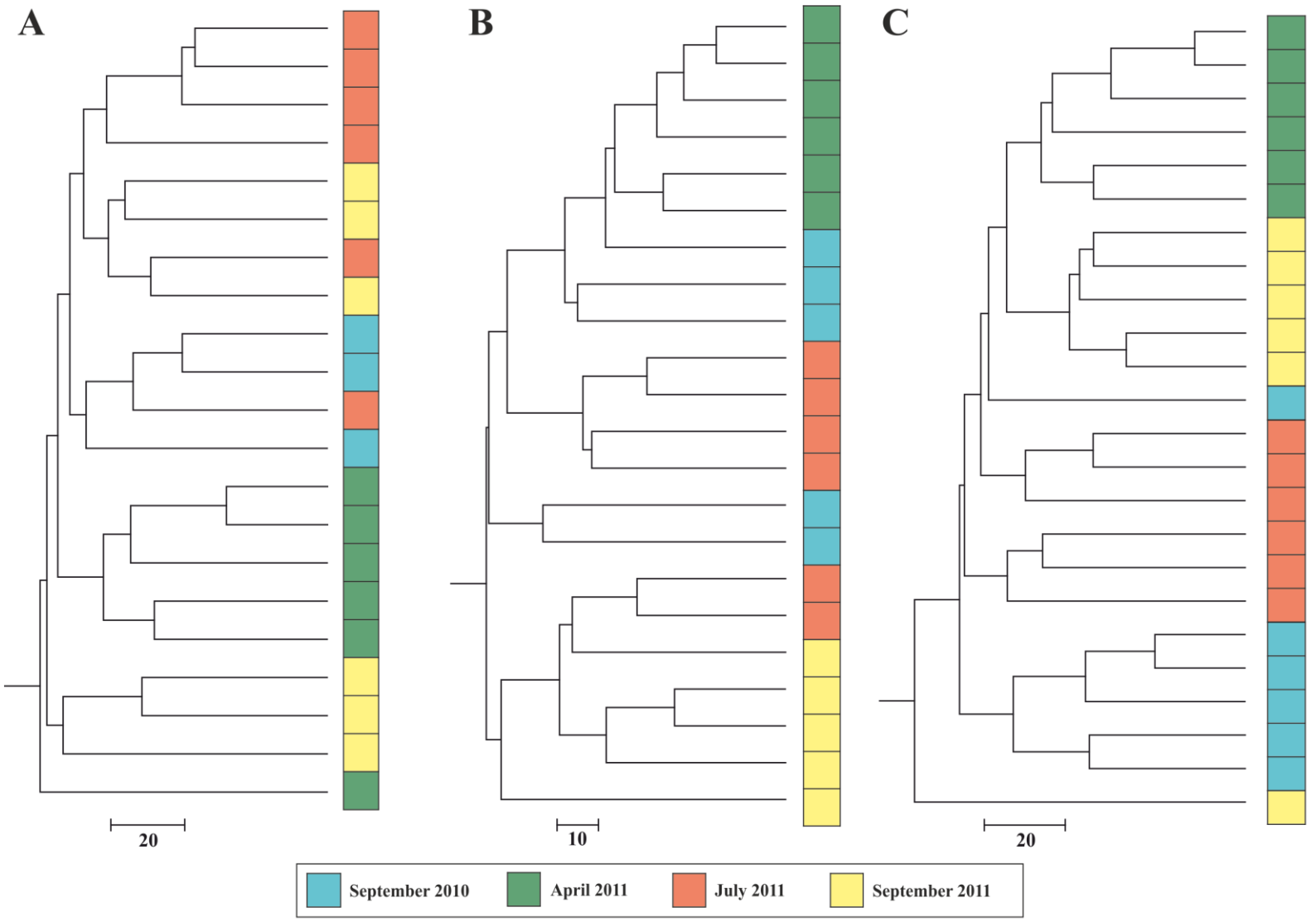


Fig. 6. UPGMA dendrogram generated by cluster analysis of DGGE fingerprints on the seasonal effect on the bacterial endophytic community composition in above-ground plant parts of *D. glomerata* (A), *F. rubra* (B), and *L. perenne* (C). For details see Fig. 3

References

Araujo WL, Marcon J, Maccheroni W, Jr., Van Elsas JD, Van Vuurde JW & Azevedo JL (2002) Diversity of endophytic bacterial populations and their interaction with *Xylella fastidiosa* in citrus plants. *Appl Environ Microbiol* **68**: 4906-4914.

Bacon C & Hinton D (2006) Bacterial endophytes: The endophytic niche, its occupants, and its utility. *Plant-Associated Bacteria* (Gnanamanickam S, ed), pp. 155-194. Springer Netherlands.

Bacon C & Hinton D (2011) *Bacillus mojavensis*: Its Endophytic Nature, the Surfactins, and Their Role in the Plant Response to Infection by Fusarium verticillioides. *Bacteria in Agrobiology: Plant Growth Responses* (Maheshwari DK, ed), pp. 21-39. Springer Berlin Heidelberg, Germany.

Baldani J, Baldani V, Seldin L & Döbereiner J (1986) Characterization of *Herbaspirillum seropedicae* gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium. *Int J Syst Bacteriol* **36**: 86-93.

Cankar K, Kraigher H, Ravnikar M & Rupnik M (2005) Bacterial endophytes from seeds of Norway spruce (*Picea abies* L. Karst). *FEMS Microbiol Lett* **244**, 341-345.

Caporaso JG, Kuczynski J, Stombaugh J, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335-336.

Chelius MK & Triplett EW (2001) The Diversity of Archaea and Bacteria in Association with the Roots of *Zea mays* L. *Microbial Ecol* **41**: 252-263.Compant S, Clément C & Sessitsch A (2010) Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biol and Bioch* **2**: 669-678.

Compant S, Duffy B, Nowak J, Clement C & Barka EA (2005) Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Appl Environ Microbiol* **71**: 4951-4959.

Conn VM & Franco CM (2004) Analysis of the endophytic actinobacterial population in the roots of wheat (*Triticum aestivum* L.) by terminal restriction fragment length polymorphism and sequencing of 16S rRNA clones. *Appl Environ Microbiol* **70**: 1787-1794.

Cremonesi L, Firpo S, Ferrari M, Righetti PG & Gelfi C (1997) Double-gradient DGGE for optimized detection of DNA point mutations. *Biotechniques* **22**: 326-330.

Dierschke H & Briemle G (2002) *Kulturgrasland*. Eugen Ulmer, Stuttgart, Germany.

Doty S, Oakley B, Xin G, Kang J, Singleton G, Khan Z, Vajzovic A & Staley J (2009) Diazotrophic endophytes of native black cottonwood and willow. *Symbiosis* **47**: 23-33.

Fuentes-Ramírez LE , Caballero-Mellado J , Sepúlveda J & Martínez-Romero E (1999) Colonization of sugarcane by *Acetobacter diazotrophicus* is inhibited by high N-fertilization. *FEMS Microbiol Ecol* **29**: 117-128.

Gagne-Bourgue F, Aliferis KA, Seguin P, Rani M, Samson R & Jabaji S (2013) Isolation and characterization of indigenous endophytic bacteria associated with leaves of switchgrass (*Panicum virgatum* L.) cultivars. *J Appl Microbiol* **114**: 836-853.

Garbeva P, Overbeek LS, Vuurde JW & Elsas JD (2001) Analysis of Endophytic Bacterial Communities of Potato by Plating and Denaturing Gradient Gel Electrophoresis (DGGE) of 16S rDNA Based PCR Fragments. *Microb Ecol* **41**: 369-383.

Gottel NR, Castro HF, Kerley M, *et al.* (2011) Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. *Appl Environ Microbiol* **77**: 5934-5944.

Hallmann J (2006) Plant Interactions with Endophytic Bacteria. *Biotic Interactions in Plant-pathogen Associations* (Jeger MJ & Spence NJ, eds), pp. 87-119. CABI, Wallingford.

Hallmann J & Berg G (2006) Spectrum and Population Dynamics of Bacterial Root Endophytes. *Microbial Root Endophytes, Vol. 9* (Schulz BE, Boyle CC & Sieber T, eds.), pp. 15-31. Springer Berlin Heidelberg, Germany.

Hallmann J , Rodríguez-Kábana R & Kloepper JW (1999) Chitin-mediated changes in bacterial communities of the soil, rhizosphere and within roots of cotton in relation to nematode control. *Soil Biol and Bioch* **31**: 551-560.

Hallmann J, Quadt-Hallmann A, Mahaffee WF & Kloepper JW (1997) Bacterial endophytes in agricultural crops. *Can J Microbiol* **43**: 895-914.

Hallmann J, Quadt-Hallmann A, Rodríguez-Kábana R & Kloepper JW (1998) Interactions between *Meloidogyne incognita* and endophytic bacteria in cotton and cucumber. *Soil Biol and Bioch* **30**: 925-937.

Hardoim PR, Hardoim CCP, van Overbeek LS & van Elsas JD (2012) Dynamics of Seed-Borne Rice Endophytes on Early Plant Growth Stages. *PLoS ONE* **7**: e30438.

Keuter A, Hoeft I, Veldkamp E & Corre M (2013) Nitrogen response efficiency of a managed and phytodiverse temperate grassland. *Plant Soil* **364**: 193-206.

Kobayashi DY & Palumbo JD (2000) Bacterial endophytes and their effects on plants and uses in agriculture. *Microbial endophytes* (Bacon CW,White JF, eds), pp. 199-236. *Marcel Dekker, New York*.

Krechel A, Faupel A, Hallmann J, Ulrich A & Berg G (2002) Potato-associated bacteria and their antagonistic potential towards plant-pathogenic fungi and the plant-parasitic nematode *Meloidogyne incognita* (Kofoid & White) Chitwood. *Can J Microbiol* **48**: 772-786.

Kuklinsky-Sobral J, Araújo W, Mendes R, Pizzirani-Kleiner A & Azevedo J (2005) Isolation and characterization of endophytic bacteria from soybean (*Glycine max*) grown in soil treated with glyphosate herbicide. *Plant Soil* **273**: 91-99.

Kuklinsky-Sobral J, Araujo WL, Mendes R, Geraldi IO, Pizzirani-Kleiner AA & Azevedo JL (2004) Isolation and characterization of soybean-associated bacteria and their potential for plant growth promotion. *Environ Microbiol* **6**: 1244-1251.

Maksimov IV, Abizgil'dina RR & Pusenkova LI (2011) Plant growth promoting rhizobacteria as alternative to chemical crop protectors from pathogens. *Appl Biochem Micro+* **47**: 333-345.

Mastretta C, Taghavi S, van der Lelie D, Mengoni A, Galardi F, Gonnelli C, Barac T, Boulet J, Weyens N & Vangronsveld J (2009) Endophytic bacteria from seeds of *Nicotiana tabacum* can reduce cadmium phytotoxicity. *Int J Phytoremediat* **11**: 251-267.

McInroy J & Kloepper J (1995A) Survey of indigenous bacterial endophytes from cotton and sweet corn. *Plant Soil* **173**: 337-342.

McInroy JA & Kloepper JW (1995B) Population dynamics of endophytic bacteria in field-grown sweet corn and cotton. *Can J Microbiol* **41**: 895-901.

Mocali S, Bertelli E, Di Cello F, Mengoni A, Sfalanga A, Viliani F, Caciotti A, Tegli S, Surico G & Fani R (2003) Fluctuation of bacteria isolated from elm tissues during different seasons and from different plant organs. *Res Microbiol* **154**: 105-114.

Muyzer G, de Waal EC & Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695-700.

Muyzer G, Teske A, Wirsen CO & Jannasch HW (1995) Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Arch Microbiol* **164**: 165-172.

Nübel U, Engelen B, Felske A, Snaidr J, Wieshuber A, Amann RI, Ludwig W & Backhaus H (1996) Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *J Bacteriol* **178**: 5636-5643.

Olivares F, Baldani VD, Reis V, Baldani J & Döbereiner J (1996) Occurrence of the endophytic diazotrophs *Herbaspirillum* spp. in roots, stems, and leaves, predominantly of Gramineae. *Biol Fert Soils* **21**: 197-200.

Parish R, Turkington R & Klein E (1990) The influence of mowing, fertilization, and plant removal on the botanical composition of an artificial sward. *Can J Bot* **68**: 1080-1085.

Petersen U, Wrage N, Köhler L, Leuschner C & Isselstein J (2012) Manipulating the species composition of permanent grasslands - A new approach to biodiversity experiments. *Basic Appl Ecoly* **13**: 1-9.

Prakamhang J, Minamisawa K, Teamtaisong K, Boonkerd N & Teaumroong N (2009) The communities of endophytic diazotrophic bacteria in cultivated rice (*Oryza sativa* L.). *Appl Soil Ecol* **42**: 141-149.

Quadt-Hallmann A, Kloepper JW & Benhamou N (1997) Bacterial endophytes in cotton: mechanisms of entering the plant. *Can J Microbioly* **43**: 577-582.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J & Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590-D596.

Reinhold-Hurek B & Hurek T (1998) Life in grasses: diazotrophic endophytes. *Trends Microbiol* **6**: 139-144.

Rosenblueth M & Martinez-Romero E (2006) Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact* **19**: 827-837.

Schulz B, Wanke U, Draeger S & Aust HJ (1993) Endophytes from herbaceous plants and shrubs: effectiveness of surface sterilization methods. *Mycol Res* **97**: 1447-1450.

Seghers D, Wittebolle L, Top EM, Verstraete W & Siciliano SD (2004) Impact of agricultural practices on the *Zea mays* L. endophytic community. *Appl Environ Microbiol* **70**: 1475-1482.

Senthilkumar M, Anandham R, Madhaiyan M, Venkateswaran V & Sa T (2011) Endophytic Bacteria: Perspectives and Applications in Agricultural Crop Production. *Bacteria in Agrobiology: Crop Ecosystems* (Maheshwari DK, ed), pp. 61-96. Springer Berlin Heidelberg, Germany.

Sessitsch A, Reiter B, Pfeifer U & Wilhelm E (2002) Cultivation-independent population analysis of bacterial endophytes in three potato varieties based on eubacterial and Actinomycetes-specific PCR of 16S rRNA genes. *FEMS Microbiol Ecol* **39**: 23-32.

Siddiqui IA & Shaukat SS (2003) Endophytic bacteria. Prospects and opportunities for the biological control of plant-parasitic nematodes. *Nematologia Mediterranea* **31**: 110-120.

Stoltzfus JR, So R, Malarvithi PP, Ladha JK & de Bruijn FJ (1997) Isolation of endophytic bacteria from rice and assessment of their potential for supplying rice with biologically fixed nitrogen. *Plant Soil* **194**: 25-36.

Sturz AV & Nowak J (2000) Endophytic communities of rhizobacteria and the strategies required to create yield enhancing associations with crops. *Appl Soil Ecol* **15**: 183-190.

Sturz AV, Christie BR & Nowak J (2000) Bacterial Endophytes: Potential Role in Developing Sustainable Systems of Crop Production. *CRC Cr Rev Plant Sci* **19**: 1-30.

Sun L, Qiu F, Zhang X, Dai X, Dong X & Song W (2008) Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. *Microb Ecol* **55**: 415-424.

Tan Z, Hurek T & Reinhold-Hurek B (2003) Effect of N-fertilization, plant genotype and environmental conditions on nifH gene pools in roots of rice. *Environ microbiol* **5**: 1009-1015.

Wilmotte A, Van der Auwera G & De Wachter R (1993) Structure of the 16 S ribosomal RNA of the thermophilic cyanobacterium *Chlorogloeopsis* HTF (*'Mastigocladus laminosus* HTF') strain PCC7518, and phylogenetic analysis. *FEBS Lett* **317**: 96-100.

SUPPORTING INFORMATION FOR STUDY 9

CONTENTS:

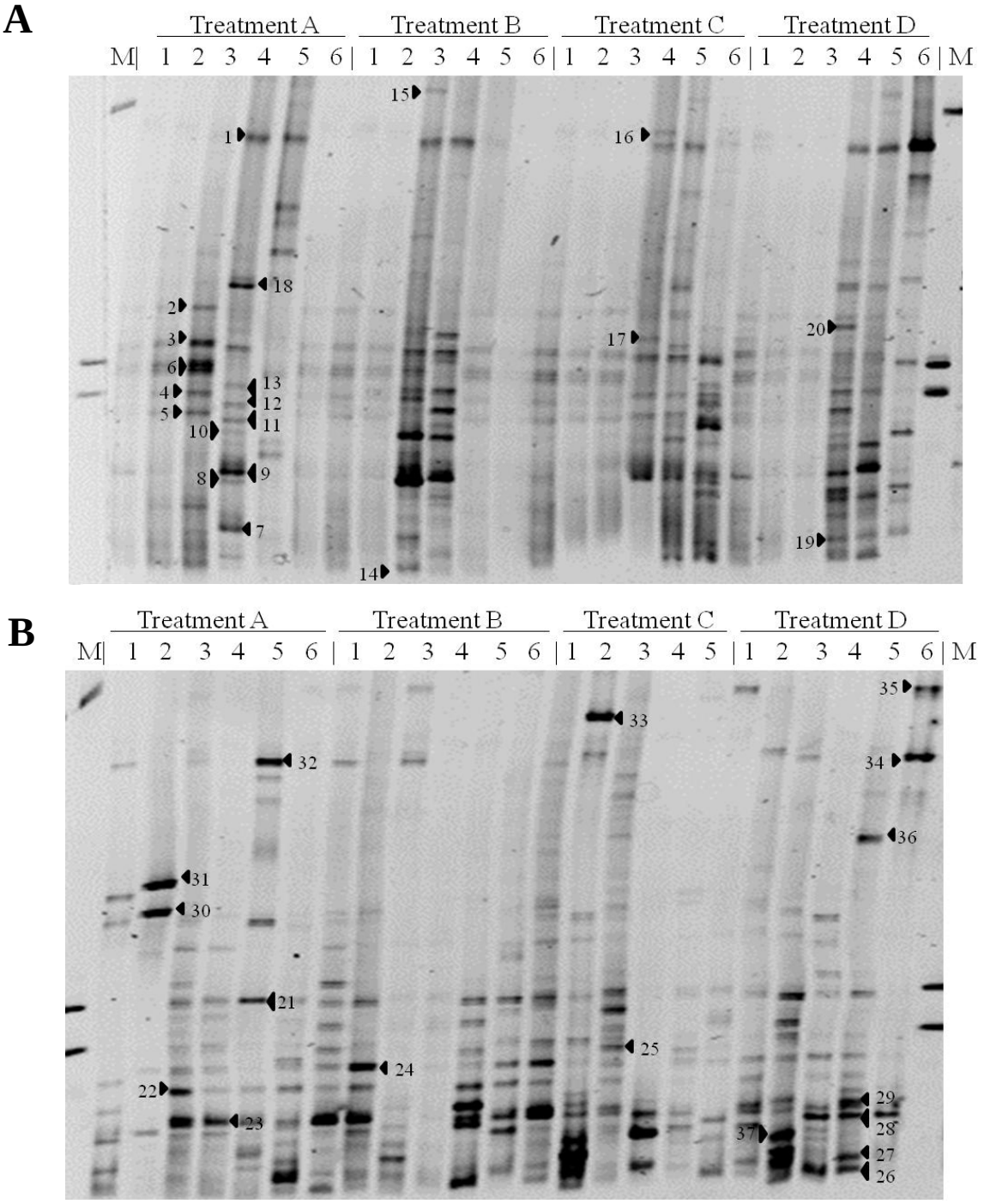


Fig. S1. 16S-DGGE profile showing the influence of different fertilization and mowing regimes on bacterial endophyte communities in above-ground plant parts of *D. glomerata*. Plant samples were taken in September 2010 (A) and 2011 (B). Independent replicates are indicated with numbers from 1 to 6. Excised bands are labelled with numbers.

Treatment A: 1 x mowing/ year, no NPK; treatment B: 3 x mowing/ year, no NPK; treatment C: 1 x mowing/ year, NPK; treatment D: 3 x mowing/ year, NPK. M: GeneRuler *1 kb* DNA *Ladder* (Fermentas, St. Leon-Rot, Germany).

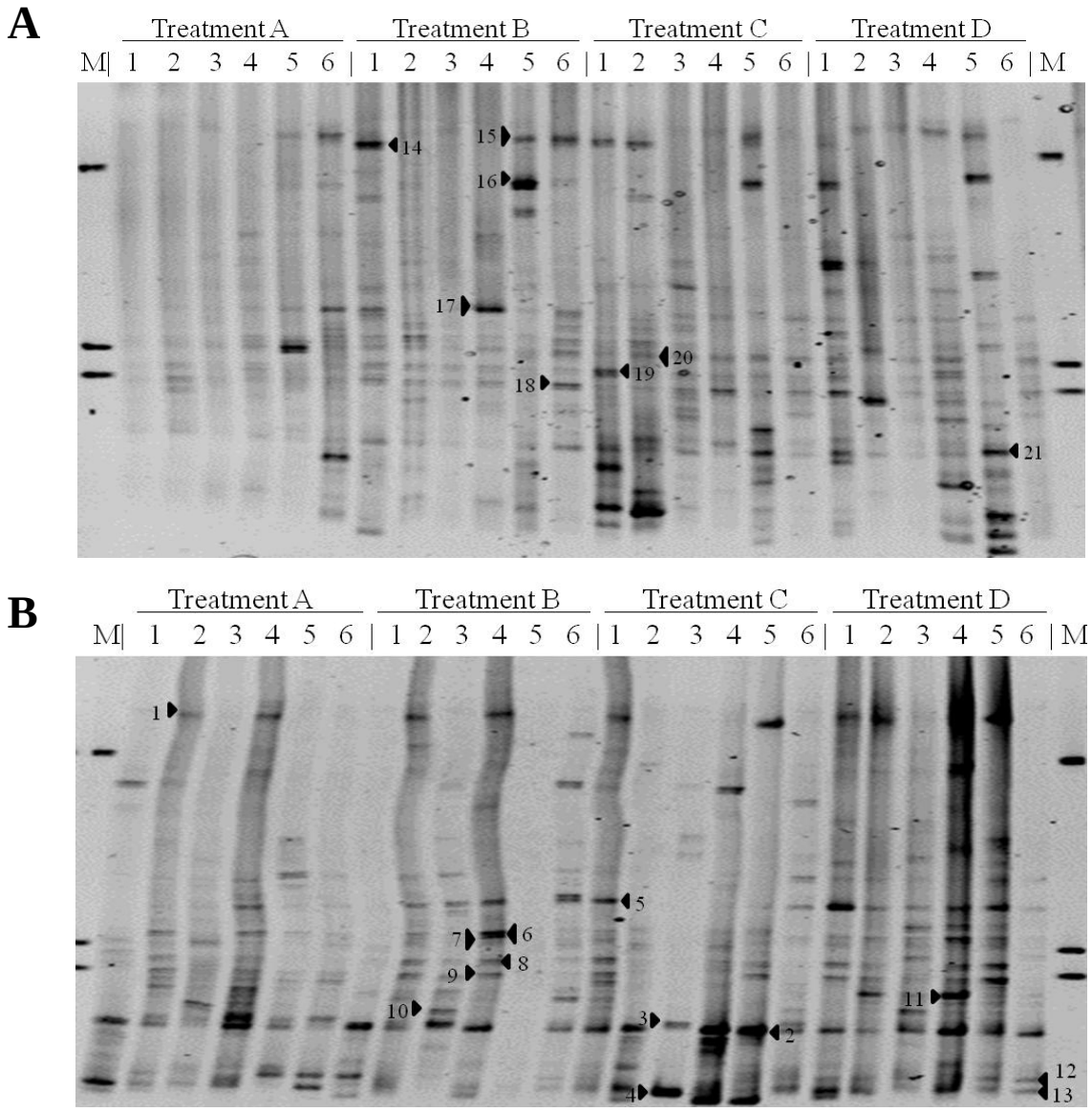
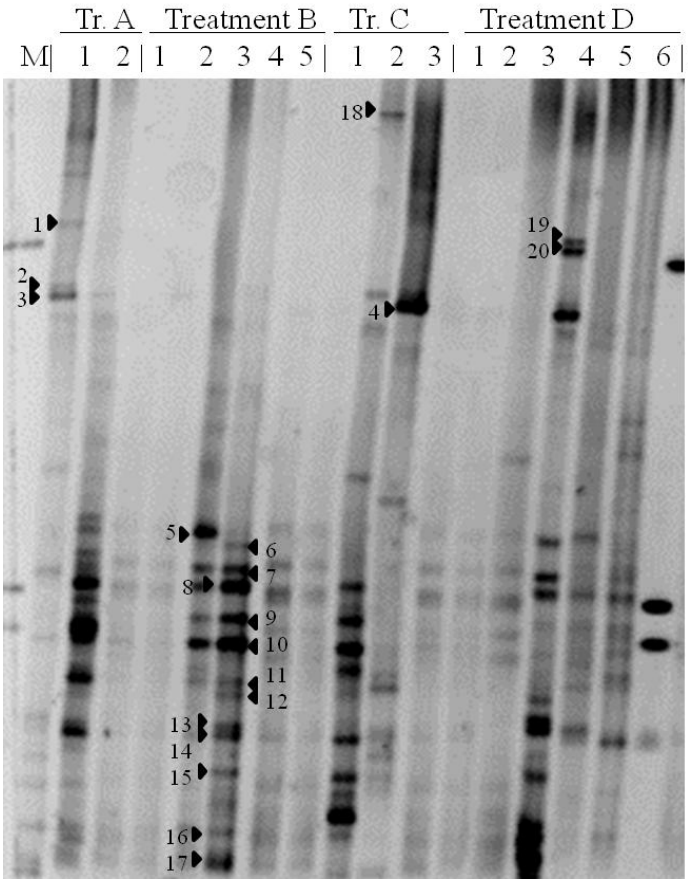


Fig. S2. 16S-DGGE profile showing the influence of different fertilization and mowing regimes on bacterial endophyte communities in above-ground plant parts of *F. rubra*. Plant samples were taken in September 2010 (A) and 2011 (B). Independent replicates are indicated with numbers from 1 to 6. Excised bands are labelled with numbers, respectively.

Treatment A: 1 x mowing/ year, no NPK; treatment B: 3 x mowing/ year, no NPK; treatment C: 1 x mowing/ year, NPK; treatment D: 3 x mowing/ year, NPK. M: GeneRuler *1 kb* DNA *Ladder* (Fermentas, St. Leon-Rot, Germany).

A


Tr. A
Treatment B
Tr. C
Treatment D
M| 1 2 | 1 2 3 4 5 | 1 2 3 | 1 2 3 4 5 6 | M
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

B

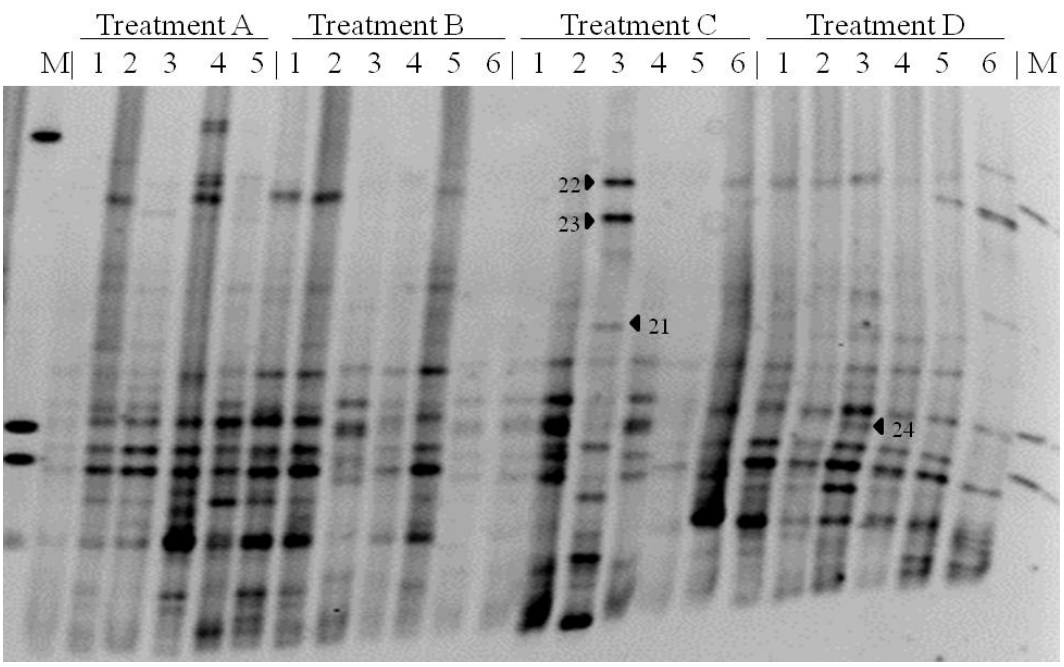
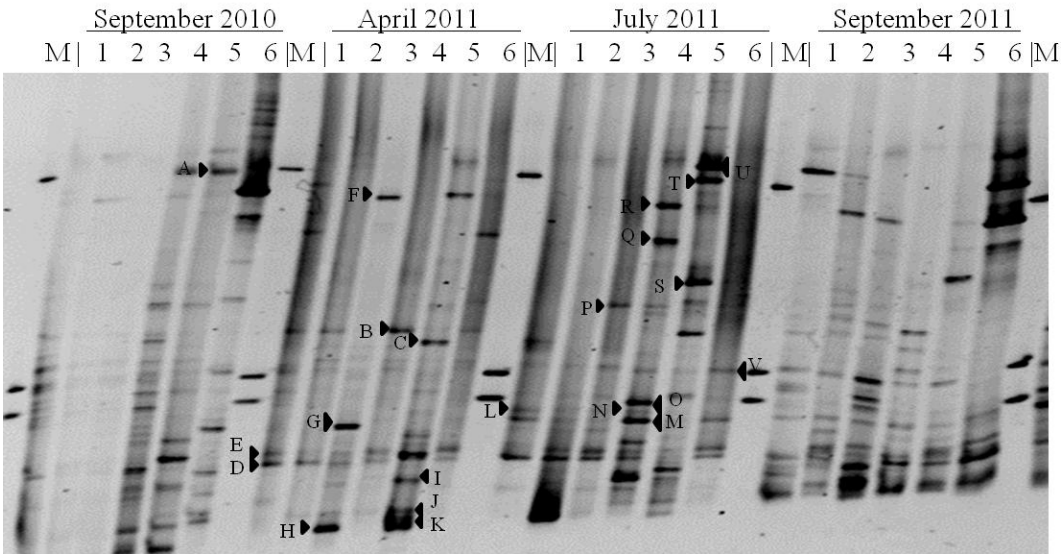

Treatment A
Treatment B
Treatment C
Treatment D
M| 1 2 3 4 5 | 1 2 3 4 5 6 | 1 2 3 4 5 6 | 1 2 3 4 5 6 | M
21
22
23
24

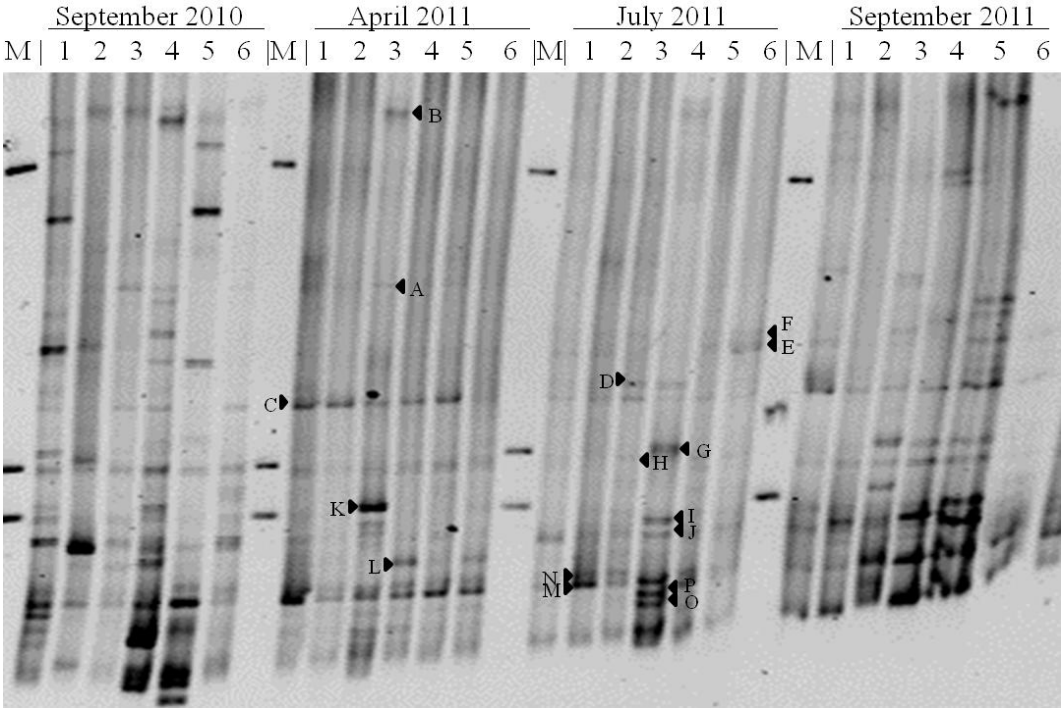
Fig. S3. 16S-DGGE profile showing the influence of different fertilization and mowing regimes on bacterial endophyte communities in above-ground plant parts of *L. perenne*. Plant samples were taken in September 2010 (A) and 2011 (B). Independent replicates are indicated with numbers from 1 to 6. Excised bands are labelled with numbers.

Treatment A: 1 x mowing/ year, no NPK; treatment B: 3 x mowing/ year, no NPK; treatment C: 1 x mowing/ year, NPK; treatment D: 3 x mowing/ year, NPK. M: GeneRuler *1 kb* DNA *Ladder* (Fermentas, St. Leon-Rot, Germany).

A


September 2010
April 2011
July 2011
September 2011
M | 1 2 3 4 5 6 |M | 1 2 3 4 5 6 |M| 1 2 3 4 5 6 | M| 1 2 3 4 5 6 |M
A
F
B
C
G
E
D
H
I
J
K
L
T
U
R
Q
S
P
V
N
O
M

B


September 2010
April 2011
July 2011
September 2011
M | 1 2 3 4 5 6 |M | 1 2 3 4 5 6 |M| 1 2 3 4 5 6 | M | 1 2 3 4 5 6
B
A
C
K
L
D
F
E
H
G
I
J
N
M
P
O

C

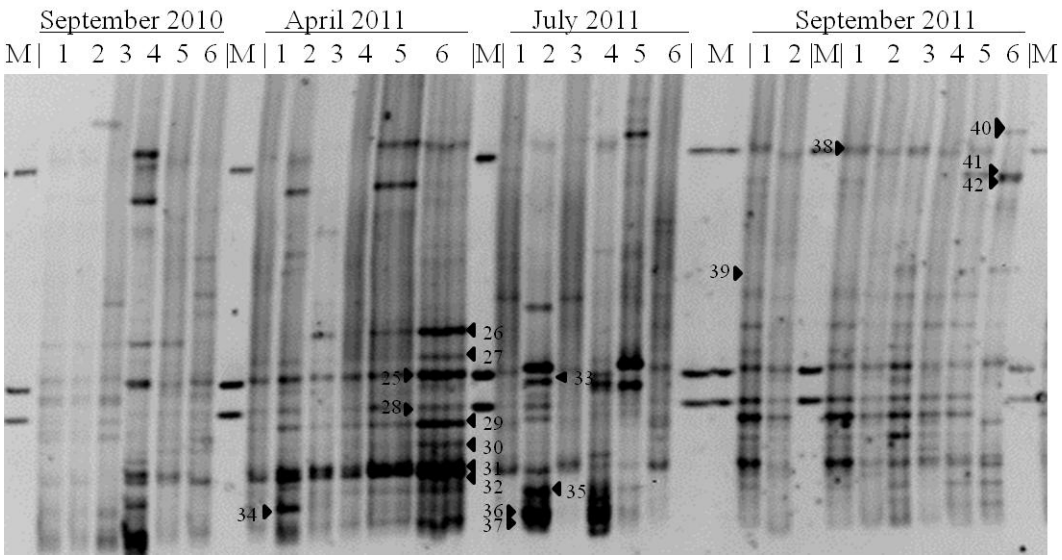


Fig. S4. 16S-DGGE profile showing the seasonal effect on bacterial endophyte communities in above-ground plant parts of *D. glomerata* (A), *F. rubra* (B), and *L. perenne* (C). Independent replicates are indicated with numbers from 1 to 6. M: GeneRuler *1 kb* DNA *Ladder* (Fermentas, St. Leon-Rot, Germany). Excised bands are labelled with numbers and letters, respectively. Samples were taken on fertilized plots three times a year in September 2010 as well as in April, July, and September 2011.

Table S1. Overview about all OTUs obtained by analysis of 16S rRNA sequences derived from the isolation and the DGGE analysis.

ID	*Sequences affiliated*				*Closest hit in the SILVA database*		
	Dactylis	*Festuca*	*Lolium*	*Isolates*	*acession*	*e value*	*SILVA taxonomy*
1	1	0	0	0	HQ598842	0	*Bacteria; Acidobacteria; Acidobacteria; Subgroup 3; Family IncertaeSedis; Bryobacter; uncultured Acidobacteria bacterium*
2	0	0	1	0	JF176919	0	*Bacteria; Actinobacteria; Acidimicrobiia; Acidimicrobiales; Acidimicrobiaceae; CL500-29 marine group; uncultured bacterium*
3	0	0	0	1	KC236620	0	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Curtobacterium; Curtobacterium sp. 4136*
4	0	0	1	0	Y17233	0	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Microbacterium; Microbacteriumkeratanolyticum*
5	1	0	1	0	KC169799	0	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Microbacterium; Microbacterium sp. CC-AMFLN-3*
6	1	0	0	0	Y17240	1E-168	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Microbacterium; Microbacteriumtrichothecenolyticum*
7	0	0	0	1	JX133202	0	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; Plantibacter; Plantibactercousiniae*
8	0	0	0	1	JQ071511	0	*Bacteria; Actinobacteria; Actinobacteria; Micrococcales; Micrococcaceae; Micrococcus; Micrococcus yunnanensis*
9	1	0	0	0	AB672179	0	*Bacteria; Actinobacteria; Thermoleophilia; Gaiellales; uncultured; uncultured bacterium*
10	1	0	0	0	JX091739	0	*Bacteria; Chloroflexi; S085; uncultured bacterium*
11	1	0	0	0	AM696939	2E-158	*Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales; Trueperaceae; Truepera; uncultured bacterium*
12	1	0	0	0	AB374378	0	*Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales; Trueperaceae; Truepera; uncultured endolithic bacterium*
13	1	0	0	0	KC120646	0	*Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae; Thermus; uncultured bacterium*
14	0	0	0	3	KC441733	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus licheniformis*
15	0	0	0	1	KC434960	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus safensis*
16	0	0	0	1	KC434960	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus safensis*
17	0	0	0	2	KC310814	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus sp. A8(2013)*
18	0	0	0	1	FN395277	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus sp. FR-W2C1*
19	0	0	0	3	KC441785	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Bacillus subtilis*
20	0	0	0	11	JX436372	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; Firmicutes bacterium Man17*
21	0	0	2	0	HE974809	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Bacillaceae; Bacillus; uncultured Bacillus sp.*
22	0	0	0	2	EU282459	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Family XII IncertaeSedis; Exiguobacterium; Exiguobacterium sp. TC38-2b*
23	0	0	0	1	AB363733	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Paenibacillaceae; Paenibacillus; Paenibacilluslautus*
24	0	0	0	10	JX897938	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Paenibacillaceae; Paenibacillus; Paenibacillusxylanexedens*
25	0	0	0	1	FM173819	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Paenibacillaceae; Paenibacillus; Pseudomonas sp. CL4.14*
26	0	0	0	1	JX990163	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Planococcaceae; Lysinibacillus; Bacillales bacterium Cul_0304*
27	0	0	0	1	JX898015	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Planococcaceae; Lysinibacillus; Bacillus sp. FBst09*
28	0	0	0	1	JN208189	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Planococcaceae; Lysinibacillus; Lysinibacillus sp. DT3*
29	0	0	0	1	JX996174	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Planococcaceae; Solibacillus; Solibacillussilvestris*
30	0	0	0	1	JX996174	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Planococcaceae; Solibacillus; Solibacillussilvestris*
31	3	5	1	0	X70648	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Staphylococcaceae; Staphylococcus; Staphylococcus aureus*
32	1	2	1	0	L37605	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Staphylococcaceae; Staphylococcus; Staphylococcus epidermidis*
33	1	1	0	0	KC153285	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Staphylococcaceae; Staphylococcus; Staphylococcus sp. G2-10*
34	0	0	0	1	KC012992	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Staphylococcaceae; Staphylococcus; Staphylococcus sp. JP44SK55*
35	1	0	0	0	HQ792508	0	*Bacteria; Firmicutes; Bacilli; Bacillales; Staphylococcaceae; Staphylococcus; uncultured organism*
36	0	1	0	0	JQ901473	0	*Bacteria; Firmicutes; Clostridia; Clostridiales; Family XI IncertaeSedis; Peptoniphilus; uncultured bacterium*
37	0	0	0	1	KC003398	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Ralstonia; unidentified marine bacterioplankton*
38	0	2	0	0	JN713899	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; Tepidimonas; Tepidimonas sp. AT-A2*
39	1	2	1	0	JX271982	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; uncultured; uncultured bacterium*
40	1	0	0	0	HQ222272	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; uncultured; Variovorax sp. enrichment culture clone Van40*
41	0	1	0	0	KC286834.	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; Variovorax; uncultured bacterium*
42	0	1	0	0	X74914	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Duganella; Zoogloearamigera*
43	0	1	0	0	Y10146	2E-129	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Herbaspirillum; Herbaspirillumseropedicae*

Table S1 *continued.*

ID	*Sequences affiliated*				*Closest hit in the SILVA database*		
	Dactylis	*Festuca*	*Lolium*	*Isolates*	*acession*	*e value*	*SILVA taxonomy*
44	1	2	1	0	Y08846	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Janthinobacterium; Janthinobacteriumlividum*
45	0	1	2	0	JN024091	0	*Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Massilia; uncultured bacterium*
46	0	0	1	0	JQ278953	0	*Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; uncultured; uncultured beta proteobacterium*
47	0	2	0	0	KC331513	0	*Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales; Rhodocyclaceae; Azospira; uncultured bacterium*
48	1	0	1	0	Z96082	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Pantoea; Pantoeaagglomerans*
49	0	1	0	0	HQ801751	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus; uncultured organism*
50	0	0	0	1	JX849037	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; Enhydrobacter; Moraxella osloensis*
51	0	1	0	0	JX849037	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; Enhydrobacter; Moraxella osloensis*
52	0	0	0	2	HQ178997	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; bacterium OC25(2011)*
53	0	0	1	0	X99541	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas anguilliseptica*
54	1	1	1	0	AF054936	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas balearica*
55	1	1	0	0	KC342251	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas chlororaphis*
56	1	0	0	0	Z76673	7E-114	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas chlororaphis*
57	1	0	1	0	FJ976601	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas putida*
58	0	0	1	0	KC310832	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas sp. C2(2013)*
59	1	1	0	0	KC310832	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas sp. C2(2013)*
60	0	0	1	0	JX899644	6E-174	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas sp. REm-amp_189*
61	0	1	1	0	U65012	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas stutzeri*
62	1	0	1	0	Z76669	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas syringae*
63	1	1	0	0	DQ469202	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; uncultured bacterium*
64	0	0	0	5	HM261524	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; uncultured bacterium*
65	1	0	1	0	GQ262820	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; uncultured; uncultured bacterium*
66	1	0	0	0	JN023904	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; uncultured; uncultured bacterium*
67	1	0	0	0	EF018613	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; uncultured; uncultured proteobacterium*
68	1	1	0	0	JN872548	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Luteibacter; Xanthomonadaceae bacterium SAP40_3*
69	1	0	0	0	FJ164060	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Rhodanobacter; gamma proteobacterium CH23i*
70	0	1	0	0	FJ380140	2E-170	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Rhodanobacter; uncultured bacterium*
71	0	0	1	0	JF180263	2E-151	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Rhodanobacter; uncultured bacterium*
72	0	0	0	2	JN897284	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Stenotrophomonas; Pseudomonas poae*
73	0	1	0	0	JX205209	0	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Stenotrophomonas; Pseudomonas sp. MLB-42*
74	1	0	0	0	X95923	2E-155	*Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Stenotrophomonas; Stenotrophomonasmaltophilia*

CHAPTER D

GENERAL DISCUSSION

Bacteria and *Archaea* are the most abundant organisms on Earth (Oren, 2004). They have been found in every environment that has been investigated so far. Moreover, these prokaryotes are responsible for various key processes in many globally important biogeochemical cycles (Lengeler *et al.*, 1999, Madigan *et al.*, 2008). As they influence types and rates of nutrient cycling, *Bacteria* and *Archaea* play a crucial role in ecosystem structuring (Azam & Malfatti, 2007). However, the knowledge about their ecological contribution is still limited as most prokaryote species cannot be cultivated and characterized (Oren, 2004). The development of culture-independent molecular approaches has greatly advanced the understanding of microbial diversity including bacterial and archaeal communities in marine ecosystems (e.g., Giebel *et al.*, 2009, Giebel *et al.*, 2011, Sperling *et al.*, 2012). However, most studies so far used DNA-based techniques which are not able to distinguish between active and inactive community members. Recent developments, expanding culture-independent approaches to RNA and proteins, were applied to overcome this limitation and allowed conclusion on active players, their ecological contributions, and putative key functions (e.g., Teeling *et al.*, 2012, Gifford *et al.*, 2013, Ottesen *et al.*, 2013).

In this thesis, the diversity and ecology of archaeal and bacterial communities in the pelagic realm of the North Sea was studied using different culture-independent approaches. In particular, the abundance and ecological role of the *Roseobacter* clade in the North Sea was investigated. In addition, the response of the prokaryotic communities to changing environmental conditions was examined. Contrary to most previous studies, mainly metatranscriptomic approaches were applied and thus the active community players were studied.

1 MARINE MICROBES

1.1 MICROBIAL ECOLOGY OF MARINE ECOSYSTEMS

The majority of studies presented in this thesis investigated microbial diversity in the pelagic realm of the North Sea (studies 1-4). The examined bacterial communities were dominated by different marine groups affiliated to the *Bacteroidetes* as well as *Alpha-* and *Gammaproteobacteria* including different subclusters of *Roseobacter* clade and members of the SAR116 clade. Moreover, *Cyanobacteria* were found in minor abundance with *Synechococcus* as the dominant genus.

During the past decades, marine communities in the North Sea have been frequently studied (Schmidt *et al.*, 1991, Eilers *et al.*, 2000, Zubkov *et al.*, 2001, Brakstad & Lødeng, 2005, Alderkamp *et al.*, 2006, Rink *et al.*, 2011). These studies showed that distinct lineages of *Alpha-* and *Gammaproteobacteria* as well as *Flavobacteria* constitute the majority of these communities (Alderkamp *et al.*, 2006, Rink *et al.*, 2011, Teeling *et al.*, 2012, Sintes *et al.*, 2013). The *Alphaproteobacteria* mainly composed of the SAR11 clade and the *Roseobacter* clade with the *Roseobacter* clade affiliated (RCA) cluster (Selje *et al.*, 2004, Giebel *et al.*, 2011, Sperling *et al.*, 2012). Although being concordant with the results presented in this thesis, most previous studies used DNA-based approaches and thus analyzed the total biodiversity but not the active fraction. Moreover, these surveys mainly used traditional techniques such as Sanger sequencing-based analysis of 16S rRNA gene libraries to examine bacterial community structures. Consequently, they failed to investigate the full extent of bacterial diversity due the size and complexity of marine bacterial communities.

In addition to bacteria, the archaeal diversity in the North Sea was studied in this thesis. The investigated archaeal communities were dominated by *Halobacteria* and minor proportions of *Thermoplasmata* as well as members of the Marine Group I. Recent metagenomic studies provided evidence for ammonium-oxidizing archaea capable of nitrification (Cavicchioli *et al.*, 2007). Some marine crenarchaeal lineages are thought to be important nitrifiers in the pelagic realm (Auguet *et al.*, 2009). These studies indicated that *Archaea* are important players in the global nitrogen cycle. However, detailed comparative

studies to understand archaeal community patterns and ecology as well as environmental drivers that shape these communities are missing (Auguet *et al.*, 2009).

1.2 ECOLOGY OF THE ROSEOBACTER CLADE

In the fifth study, an abundant *Roseobacter* clade member in the North Sea was investigated to elucidate its ecological role and to identify putative key traits contributing to the success of this clade in marine ecosystems. The *Roseobacter* clade is an abundant marine group constituting up to 25% of the total prokaryote community in pelagic realm (Brinkhoff *et al.*, 2008). The success of this group is usually referred to the use of a multitude of organic compounds, the production of secondary metabolites, and other metabolic pathways such as aerobic anoxygenic photosynthesis or carbon monoxide oxidation as most of the sequenced genomes contain genes for at least one of these features (Buchan *et al.*, 2005). However, only one genome of an abundant *Roseobacter* subcluster, the genome of *Planktomarina temperata* RCA23, is currently available (Giebel *et al.*, 2013). Consequently, our knowledge about the ecology of the *Roseobacter* clade mainly derives from genome analyses of isolates, which do not represent abundant subclusters (Fig. 8).

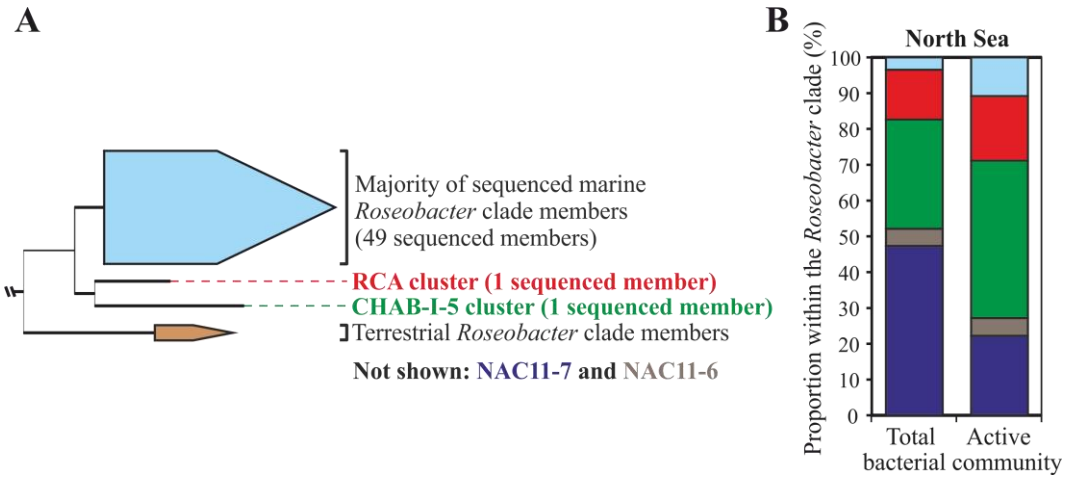


Fig. 8. Gene content tree of sequenced *Roseobacter* clade members (as from July 2013) (A) and the average composition of this clade based on the results from studies 2 and 3 (B). The colors in (B) refer to the colors in (A).

Despite the high 16S rRNA gene similarity within the *Roseobacter* clade, inferring physiology from 16S phylogeny is inherently biased (Jaspers & Overmann, 2004). One example of such a bias is the genus *Octadecabacter* which is abundant in polar regions. Both type strains *O. arcticus* 238 and *O. antarcticus* 307 share more than 99% identity on 16S level (Vollmers *et al.*, 2013). Although sharing some features unusual for other *Roseobacter* clade members, they also show remarkable differences. Unique genes of each strain comprise 20% and 23% of the genome in *O. antarcticus* and *O. arcticus*, respectively, indicating a high potential for individual adaptations (Vollmers *et al.*, 2013). A similar result was found for two other closely related *Roseobacter* clade members, *Roseobacter litoralis* and *Roseobacter denitrificans* (Kalhoefer *et al.*, 2011). The comparative analysis of both genomes showed that 1122 of the 4537 (24.7%) genes predicted for *R. litoralis* were not present in the genome of *R. denitrificans (Kalhoefer et al., 2011)*. This includes several potential heavy metal resistance genes. Moreover, some genomic regions were exclusively detected in these two strains but not in other *Roseobacter* clade members.

To investigate the gene expression of *Planktomarina temperata* in situ, environmental mRNA was sequenced and the retrieved sequencing information were mapped on the genome of this species (see study 5). Almost 86% of the

entire genome was transcribed. Active features included the photosynthetic operon, reflecting the biosynthesis of bacteriochlorophyll *a* (Bchl *a*), which is an indicator for energy harvesting via aerobic anoxygenic photosynthesis. Moreover, both CO dehydrogenase systems were fully transcribed, suggesting an active carboxidovoric metabolism. In addition, transcripts for the assimilatory sulfate reduction pathway were identified. The observed presence of transcripts of the cleavage and demethylase pathways reflected a versatile metabolism of dimethylsulfoniopropionate, a major organosulfur compound in marine ecosystems.

Mapping of environmental mRNA sequences on available genomes or genetic data stored in public databases is a common approach to study gene expression and active contribution of single marine groups to ecosystem functioning (Frias-Lopez, 2008, McCarren *et al.*, 2010, Hollibaugh *et al.*, 2011, Gifford *et al.*, 2013, Ottesen *et al.*, 2013). An early study analyzed gene expression in the open ocean and found active genes associated with key metabolic pathways including genes involved in photosynthesis, carbon fixation, and nitrogen acquisition (Frias-Lopez, 2008). Ottesen *et al.* (2013) studied transcriptome profiles of eukaryotic (*Ostreococcus*) and bacterial (*Synechococcus*) photosynthetic picoplankton as well as proteorhodopsin-containing heterotrophs, including *Pelagibacter*, SAR86-cluster *Gammaproteobacteria*, and marine *Euryarchaea*. The investigated photosynthetic picoplankton exhibited strong circadian rhythms over thousands of gene transcripts which was remarkably consistent with observations from pure cultures experiments. Investigated heterotrophs did not follow such a daily routine.

Although novel insights into gene expression patterns and functions of an abundant *Roseobacter* clade member were gained, our knowledge about the active ecological contribution of the *Roseobacter* clade is still very limited. Further research is needed addressing this issue.

1.3 MICROBIAL RESPONSE TO A CHANGING ENVIRONMENT

Four studies presented in this thesis analyzed changes of microbial community structures as a response to changing environments conditions (studies 1-4). The structure and function of microbial communities in marine ecosystems is largely driven by environmental conditions. Many biotic factors including phytoplankton blooms (West *et al.*, 2008, Teeling *et al.*, 2012) as well as abiotic factors such as salinity, temperature, and depth (Thompson *et al.*, 2011, Campbell & Kirchman, 2013) influence the structure and function of marine microbial communities.

The first two studies presented in this thesis investigated the impact of a phytoplankton bloom on the active archaeal and bacterial community structure in the southern North Sea employing different metatranscriptomic approaches (studies 1 & 2). Both prokaryotic communities responded to the presence of the bloom or environmental parameters correlated with bloom presence. The number of studies examining the impact of a phytoplankton bloom on microbial community structures is limited especially for *Archaea*. Herfort *et al.* (2007) studied archaeal communities in the southwestern North Sea via denaturing gradient gel electrophoresis and showed a positive correlation between the abundance of *Euryarchaeota* and chlorophyll concentrations, whereas the abundance of *Crenarchaeota* was negatively correlated. Teeling *et al.* (2012) investigated the impact of a phytoplankton bloom bacterial communities near Helgoland. They demonstrated that the structure of these communities followed the development stages of the examined bloom. In a study from the Southern Ocean, community structures were different in and outside of a natural iron-fertilized algal bloom (West *et al.*, 2008).

In the third and fourth study, bacterial communities along a latitudinal gradient ranging from the German Bight to the coast of Norway were investigated. Community structures significantly changed along the examined latitudinal gradient. The abundance of certain marine groups such as the *Roseobacter* CHAB-I-5 and NAC11-7 clusters either decreased or increased with rising latitude. This is supported by other studies. For example, high numbers of *Proteobacteria* in coastal bacterioplankton communities along a latitudinal

gradient in Latin America were observed by Thompson *et al.* (2011). Another study showed that the SAR11 clade as well as the genus *Prochlorococcus* were the dominant phylogenetic groups in oceanic surface waters on a latitudinal gradient ranging from South Africa to the English Channel (Schattenhofer *et al.*, 2009). According to Lefort & Gasol (2013), marine surface bacterioplankton groups varied along gradients of salinity, chlorophyll, and temperature. The authors suggested that both different niche preferences and a coherent response to environmental factors were responsible for the observed patterns. Another study examined total and active bacterial communities along a salinity gradient in Delaware Bay by pyrotag sequencing of 16S rRNA and found significant structural changes in surface water communities along the investigated gradient (Campbell & Kirchman, 2013). Fuhrman *et al.* (2008) investigated community patterns of marine planktonic bacteria worldwide by amplified ribosomal intergenic spacer analysis (ARISA). Geographic patterns similar to those of higher organisms were recorded. The authors concluded that the kinetics of metabolisms strongly influenced bacterial diversity and global distribution.

Although being intensively studied over the past years, our knowledge about prokaryotic communities in marine ecosystems and their response to changing environmental conditions is still limited. However, it is of fundamental importance to understand how they react as the reaction can have global implications.

2 FURTHER STUDIES ON MICROBIAL ECOLOGY

2.1 MICROBIAL DIVERSITY AND ECOLOGY OF TWO HYPERTHERMAL SPRINGS

In one of the studies presented in this thesis, the prokaryotic community composition and ecology of two hyperthermal springs (AI: 51°C, pH3; AIV: 92°C, pH 8) on the Azores was investigated (study 7). Bacterial and archaeal communities were dominated by a limited number of genera and lineages. This is not surprising as hyperthermal springs are considered to be low-diversity habitats due to the high physicochemical constraints and the high degree of geographical isolation (Papke *et al.*, 2003).

Several studies revealed that the abundance and diversity of diverse microbial communities in terrestrial hot springs is influenced by chemical compositions, pH, temperature, dissolved hydrogen sulfide levels, and biogeography (Ward & Castenholz, 2002, Whitaker *et al.*, 2003, Purcell *et al.*, 2007). It is known that temperature is one major driver of the prokaryotic diversity in hot springs. For example, Hiraishi et al. (1999) showed that the microbial diversity of sulfide-containing neutral pH hot spring mats in Japan decreased as the temperature of the environment increased. Skirnisdottir et al. (Skirnisdottir *et al.*, 2000) found the lowest bacterial diversity at the highest temperature (80°C) and at the highest sulfide concentration when comparing hot springs in Iceland with low and high concentration of sulfide. These findings were also supported by a study about the bacterial diversity of microbial mats in two alkaline hot springs from Yellowstone National Park Park (Miller *et al.*, 2009). On the other hand, no statistical correlation between prokaryotic diversity and temperature was observed in almost pH neutral hot springs in Central and Central-Eastern Tibet (Huang *et al.*, 2011) or in microbial mats and streamers in alkaline geothermal springs from Tibet (Yim *et al.*, 2006).

Interestingly, the prokaryotic diversity on species level was lower in the acidic spring AI compared to the hot, slightly alkaline spring AIV. This result is supported by one of our previous studies of two hot springs on the Kamchatka peninsula (Wemheuer *et al.*, 2013). We showed that the prokaryotic community in the thermophilic (81°C, pH 7.2-7.4) spring was more diverse than in the

thermoacidophilic (70°C, pH 3.5-4) spring. The pH value was the major driver of the prokaryotic diversity in the investigated hot springs.

Geothermal springs are windows in the plant's past. Therefore, studies about these springs provide a unique inside in the early days of our planet and might help us to understand the evolution of life on Earth. Additionally, the low diversity found in these springs makes them ideal systems to study principles of community structure and function

2.2 DIVERSITY, GENOMIC POTENTIAL, AND FUNCTION IN A PHOTOBIOREACTOR

The community structure of algae-associated bacteria in a photobioreactor was analyzed in one of the presented studies (study 8). Furthermore, the microbial potential and function was evaluated. The bacterial diversity within the bioreactor was low with approximately 30 algae-associated bacterial species. Moreover, the analysis suggested that the catabolic and metabolic potential of the microbes living in the photobioreactor biofilm is highly flexible and diverse.

Microalgae comprise a large and phylogenetically very heterogeneous group of microorganisms. They carry out oxygenic photosynthesis and thereby convert CO_2 to significant amounts of cellular lipids as potential source of biofuels (Abomohra *et al.*, Chisti, 2007, Gouveia & Oliveira, 2009, Williams & Laurens, 2010, Wiley *et al.*, 2011). Since photobioreactors allow the continuous cultivation of microalgae under relatively controlled conditions, they have been in the focus of many research projects often linked to algal fuel production (Chisti, 2007, Ugwu *et al.*, 2008, Scott *et al.*, 2010, Williams & Laurens, 2010).

In this study, the metagenome sequences suggest that the bacteria associated within the photobioreactor were mainly heterotrophs metabolizing a wide range of carbon and energy sources provided by the algae. In addition, the relatively high number of predicted and functional esterase and lipase genes indicated that the algae-associated bacteria might be a major sink for lipids and fatty acids released by the microalgae. Heterotrophic organisms in algal cultures are generally considered as contaminants. Moreover, they are harmful for algal

growth in microalgal biomass production systems (Belay, 1997, Huntley & Redalje, 2007). According to Joint et al. (2002), bacteria and microalgae in a photobioreactor compete for the available nutrients such as N and P. Furthermore, bacteria may produce metabolites that are inhibitory to microalgal growth. On the other hand, certain bacteria may have positive impacts on growth of microalgae (Watanabe *et al.*, 2005, Park *et al.*, 2008), for example by increasing the solubility of nutrients and trace elements for algae (Keshtacher-Liebso *et al.*, 1995) or by supplying vitamin B12 (Croft *et al.*, 2005). One example for such a close mutualistic relationship is *Dinoroseobacter shibae*, another member of the *Roseobacter* clade. It is associated with dinoflagellates and supplies its algal host with vitamin B12 (Wagner-Döbler *et al.*, 2010).

Photobioreactors can be used to produce algal fuel which is a promising alternative to fossil fuel and thus a putative solution for the global energy crisis. Moreover, photobioreactors can be used for the reduction of anthropogenic carbon dioxide. However, there is still a knowledge gap on the physiology and metabolism of complex microbial communities associated with eukaryotic microalgae. Therefore, a better understanding of the complex biology and interactions of microalgae with the environment is crucial to bypass known problems of photobioreactors and to improve this technology.

2.3 ENDOPHYTIC COMMUNITIES IN THREE GRASS SPECIES

The last study presented in this thesis evaluated the impact of different management regimes and seasons on the endophytic bacterial community composition in three agricultural important grass species, *L. perenne*, *F. rubra*, and *D. glomerata* (study 9). Community structures were significantly influenced by the applied management regimes. This is in accordance with other studies. For example, a higher diazotrophic bacterial diversity in the roots of rice cultivated in unfertilized and previously uncultivated soil than in paddy soil amended with nitrogen fertilizer were recorded by Prakamhang *et al*. (Prakamhang *et al.*, 2009).

Moreover, differences in the community structures of the analyzed grass species and their response to the examined management regimes were recorded.

According to Hallmann & Berg (Hallmann & Berg, 2006), plants vary in their biochemical composition which might explain differences in the bacterial endophytic community. Hallmann *et al.* (Hallmann *et al.*, 1999) suggested that changes in plant physiology may result in the development of distinct bacterial endophytic communities. Several factors including temperature or precipitation have a direct effect on the plant physiology and thus an indirect impact on the colonization and the survival of bacteria in the endosphere (Hallmann *et al.*, 1997, Tan *et al.*, 2003, Hardoim *et al.*, 2012). During the year, plants undergo physiological changes that probably increase nutrient availability and thus bacterial diversity (Hallmann & Berg, 2006). This might explain the effect of season on the endophytic community composition recorded in this study.

As endophytes can promote plant growth, yield, and health, understanding what drives the composition of these microbes is important for agricultural production systems. However, the complex interaction between plant and endophyte is still not fully understood.

3 CONCLUDING REMARK

The majority of studies presented in this thesis investigated the diversity and dynamics of prokaryotic communities in diverse ecosystems using a network of different culture-independent approaches. Our knowledge about the dynamics of these communities and their contribution to ecosystem functioning is still limited. Therefore, the results presented in this thesis might pave the way for a better understanding of these dynamics and the ecological role of prokaryotic communities. Moreover, the presented results also support the conclusion that only a combination of different, culture-dependent and culture-independent as well as DNA- and RNA-based ,approaches is able to draw a more comprehensive and realistic picture of a prokaryotic community.

CHAPTER E

GENERAL REFERENCES

Abomohra AE-F, Wagner M, El-Sheekh M & Hanelt D Lipid and total fatty acid productivity in photoautotrophic fresh water microalgae: screening studies towards biodiesel production. *Journal of Applied Phycology* 1-6.

Alderkamp AC, Sintes E & Herndl GJ (2006) Abundance and activity of major groups of prokaryotic plankton in the coastal North Sea during spring and summer. *Aquat Microb Ecol* **45**: 237-246.

Atlas RM & Bartha R (1993) Microbial Ecology: Fundamentals and Applications. Benjamin Cummings, Redwood City, CA.

Auguet J-C, Barberan A & Casamayor EO (2009) Global ecological patterns in uncultured Archaea. *ISME J* **4**: 182-190.

Azam F & Malfatti F (2007) Microbial structuring of marine ecosystems. *Nat Rev Microbiol* **5**: 782-791.

Bacon C & Hinton D (2006) Bacterial endophytes: The endophytic niche, its occupants, and its utility. *Plant-Associated Bacteria*,(Gnanamanickam S, ed.) pp. 155-194. Springer Netherlands.

Bacon C & Hinton D (2011) Bacillus mojavensis: Its Endophytic Nature, the Surfactins, and Their Role in the Plant Response to Infection by Fusarium verticillioides. *Bacteria in Agrobiology: Plant Growth Responses*,(Maheshwari DK, ed.) p.^pp. 21-39. Springer Berlin Heidelberg.

Belay A (1997) Mass culture of Spirulina outdoors: the Earthrise Farms experience. *Spirulina platensis (Arthrospira): Physiology, cell-biology and biotechnology London, Taylor and Francis* 131-158.

Bolhuis H & Stal LJ (2011) Analysis of bacterial and archaeal diversity in coastal microbial mats using massive parallel 16S rRNA gene tag sequencing. *The ISME journal* **5**: 1701-1712.

Brakstad OG & Lødeng AG (2005) Microbial diversity during biodegradation of crude oil in seawater from the North Sea. *Microb Ecol* **49**: 94-103.

Brinkhoff T, Giebel HA & Simon M (2008) Diversity, ecology, and genomics of the Roseobacter clade: a short overview. *Arch Microbiol* **189**: 531-539.

Brinkmeyer R, Knittel K, Jürgens J, Weyland H, Amann R & Helmke E (2003) Diversity and structure of bacterial communities in Arctic versus Antarctic pack ice. *Appl Environ Microbiol* **69**: 6610-6619.

Britschgi TB & Giovannoni SJ (1991) Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. *Appl Environ Microbiol* **57**: 1707-1713.

Buchan A, González JM & Moran MA (2005) Overview of the marine roseobacter lineage. *Appl Environ Microbiol* **71**: 5665-5677.

Buchan A, Gonzalez JM & Moran MA (2005) Overview of the marine roseobacter lineage. *Applied and environmental microbiology* **71**: 5665-5677.

Campbell BJ & Kirchman DL (2013) Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. *ISME J* **7**: 210-220.

Cavicchioli R, DeMaere MZ & Thomas T (2007) Metagenomic studies reveal the critical and wide-ranging ecological importance of uncultivated archaea: the role of ammonia oxidizers. *BioEssays* **29**: 11-14.

Chisti Y (2007) Biodiesel from microalgae. *Biotechnology advances* **25**: 294-306.

Compant S, Clément C & Sessitsch A (2010) Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biology and Biochemistry* **42**: 669-678.

Compant S, Duffy B, Nowak J, Clement C & Barka EA (2005) Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Appl Environ Microbiol* **71**: 4951-4959.

Croft MT, Lawrence AD, Raux-Deery E, Warren MJ & Smith AG (2005) Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* **438**: 90-93.

Curtis TP, Sloan WT & Scannell JW (2002) Estimating prokaryotic diversity and its limits. *Proc Natl Acad Sci U S A* **99**: 10494-10499.

Doty S, Oakley B, Xin G, Kang J, Singleton G, Khan Z, Vajzovic A & Staley J (2009) Diazotrophic endophytes of native black cottonwood and willow. *Symbiosis* **47**: 23-33.

Eilers H, Pernthaler J, Glockner FO & Amann R (2000) Culturability and In situ abundance of pelagic bacteria from the North Sea. *Applied and environmental microbiology* **66**: 3044-3051.

Elend C, Schmeisser C, Leggewie C, Babiak P, Carballeira JD, Steele HL, Reymond JL, Jaeger KE & Streit WR (2006) Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl Environ Microbiol* **72**: 3637-3645.

Frias-Lopez J (2008) Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA* **105**: 3805-3810.

Fuentes-Ramírez LE, Caballero-Mellado J, Sepúlveda J & Martínez-Romero E (1999) Colonization of sugarcane by Acetobacter diazotrophicus is inhibited by high N-fertilization. *FEMS Microbiology Ecology* **29**: 117-128.

Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL & Brown JH (2008) A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences* **105**: 7774-7778.

Giebel H-A, Kalhoefer D, Gahl-Janssen R, *et al.* (2013) Planktomarina temperata gen. nov., sp. nov., belonging to the globally distributed RCA cluster of the marine Roseobacter clade, isolated from the German Wadden Sea. *International Journal of Systematic and Evolutionary Microbiology* **63**: 4207-4217.

Giebel HA, Brinkhoff T, Zwisler W, Selje N & Simon M (2009) Distribution of Roseobacter RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environ Microbiol* **11**: 2164-2178.

Giebel HA, Kalhoefer D, Lemke A, Thole S, Gahl-Janssen R, Simon M & Brinkhoff T (2011) Distribution of Roseobacter RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J* **5**: 8-19.

Gifford SM, Sharma S, Booth M & Moran MA (2013) Expression patterns reveal niche diversification in a marine microbial assemblage. *ISME J* **7**: 281-298.

Giovannoni SJ & Stingl U (2005) Molecular diversity and ecology of microbial plankton. *Nature* **437**: 343-348.

Giovannoni SJ, Britschgi TB, Moyer CL & Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60-63.

Gouveia L & Oliveira AC (2009) Microalgae as a raw material for biofuels production. *Journal of industrial microbiology & biotechnology* **36**: 269-274.

Hallmann J & Berg G (2006) Spectrum and Population Dynamics of Bacterial Root Endophytes. *Microbial Root Endophytes*, Vol. 9 (Schulz BE, Boyle CC & Sieber T, eds.), pp. 15-31. Springer Berlin Heidelberg.

Hallmann J, Rodríguez-Kábana R & Kloepper JW (1999) Chitin-mediated changes in bacterial communities of the soil, rhizosphere and within roots of cotton in relation to nematode control. *Soil Biology and Biochemistry* **31**: 551-560.

Hallmann J, Quadt-Hallmann A, Mahaffee WF & Kloepper JW (1997) Bacterial endophytes in agricultural crops. *Canadian Journal of Microbiology* **43**: 895-914.

Hallmann J, Quadt-Hallmann A, Rodríguez-Kábana R & Kloepper JW (1998) Interactions between Meloidogyne incognita and endophytic bacteria in cotton and cucumber. *Soil Biology and Biochemistry* **30**: 925-937.

Hardoim PR, Hardoim CCP, van Overbeek LS & van Elsas JD (2012) Dynamics of Seed-Borne Rice Endophytes on Early Plant Growth Stages. *PLoS ONE* **7**: e30438.

Heidelberg KB, Gilbert JA & Joint I (2010) Marine genomics: at the interface of marine microbial ecology and biodiscovery. *Microb Biotechnol* **3**: 531-543.

Herfort L, Schouten S, Abbas B, Veldhuis MJW, Coolen MJL, Wuchter C, Boon JP, Herndl GJ & Sinninghe Damsté JS (2007) Variations in spatial and temporal distribution of Archaea in the North Sea in relation to environmental variables. *FEMS Microbiology Ecology* **62**: 242-257.

Hiraishi A, Umezawa T, Yamamoto H, Kato K & Maki Y (1999) Changes in quinone profiles of hot spring microbial mats with a thermal gradient. *Appl Environ Microbiol* **65**: 198-205.

Hollibaugh JT, Gifford S, Sharma S, Bano N & Moran MA (2011) Metatranscriptomic analysis of ammonia-oxidizing organisms in an estuarine bacterioplankton assemblage. *ISME J* **5**: 866-878.

Huang Q, Dong C, Dong R, *et al.* (2011) Archaeal and bacterial diversity in hot springs on the Tibetan Plateau, China. *Extremophiles* **15**: 549-563.

Hugenholtz P, Pitulle C, Hershberger KL & Pace NR (1998) Novel Division Level Bacterial Diversity in a Yellowstone Hot Spring. *Journal of Bacteriology* **180**: 366-376.

Huntley ME & Redalje DG (2007) CO2 mitigation and renewable oil from photosynthetic microbes: a new appraisal. *Mitigation and adaptation strategies for global change* **12**: 573-608.

Jaspers E & Overmann J (2004) Ecological Significance of Microdiversity: Identical 16S rRNA Gene Sequences Can Be Found in Bacteria with Highly Divergent Genomes and Ecophysiologies. *Applied and Environmental Microbiology* **70**: 4831-4839.

Joint I, Henriksen P, Fonnes GA, Bourne D, Thingstad TF & Riemann B (2002) Competition for inorganic nutrients between phytoplankton and bacterioplankton in nutrient manipulated mesocosms. *Aquatic Microbial Ecology* **29**: 145-159.

Kalhoefer D, Thole S, Voget S, Lehmann R, Liesegang H, Wollher A, Daniel R, Simon M & Brinkhoff T (2011) Comparative genome analysis and genome-guided physiological analysis of Roseobacter litoralis. *BMC Genomics* **12**: 324.

Keshtacher-Liebso E, Hadar Y & Chen Y (1995) Oligotrophic Bacteria Enhance Algal Growth under Iron-Deficient Conditions. *Applied and environmental microbiology* **61**: 2439-2441.

Kobayashi D & Palumbo J (2000) Bacterial endophytes and their effects on plants and uses in agriculture. *Microbial endophytes Marcel Dekker, New York* 199-236.

Krechel A, Faupel A, Hallmann J, Ulrich A & Berg G (2002) Potato-associated bacteria and their antagonistic potential towards plant-pathogenic fungi and the plant-parasitic nematode Meloidogyne incognita (Kofoid & White) Chitwood. *Canadian Journal of Microbiology* **48**: 772-786.

Kuklinsky-Sobral J, Araújo W, Mendes R, Pizzirani-Kleiner A & Azevedo J (2005) Isolation and characterization of endophytic bacteria from soybean (Glycine max) grown in soil treated with glyphosate herbicide. *Plant and Soil* **273**: 91-99.

Kvist T, Ahring BK & Westermann P (2007) Archaeal diversity in Icelandic hot springs. *FEMS Microbiology Ecology* **59**: 71-80.

Lee OO, Wang Y, Yang J, Lafi FF, Al-Suwailem A & Qian P-Y (2011) Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. *ISME J* **5**: 650-664.

Maksimov IV, Abizgil'dina RR & Pusenkova LI (2011) Plant growth promoting rhizobacteria as alternative to chemical crop protectors from pathogens (review). *Applied Biochemistry and Microbiology* **47**: 333-345.

McCarren J, Becker JW, Repeta DJ, Shi Y, Young CR, Malmstrom RR, Chisholm SW & DeLong EF (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proceedings of the National Academy of Sciences* **107**: 16420-16427.

Meyer-Dombard DR, Shock EL & Amend JP (2005) Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology* **3**: 211-227.

Miller SR, Strong AL, Jones KL & Ungerer MC (2009) Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl Environ Microbiol* **75**: 4565-4572.

Olsen GJ, Woese CR & Overbeek R (1994) The winds of (evolutionary) change: breathing new life into microbiology. *J Bacteriol* **176**: 1-6.

Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA & DeLong EF (2013) Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proceedings of the National Academy of Sciences*.

Papke RT, Ramsing NB, Bateson MM & Ward DM (2003) Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* **5**: 650-659.

Park Y, Je K-W, Lee K, Jung S-E & Choi T-J (2008) Growth promotion of Chlorella ellipsoidea by co-inoculation with Brevundimonas sp. isolated from the microalga. *Hydrobiologia* **598**: 219-228.

Prakamhang J, Minamisawa K, Teamtaisong K, Boonkerd N & Teaumroong N (2009) The communities of endophytic diazotrophic bacteria in cultivated rice (Oryza sativa L.). *Applied Soil Ecology* **42**: 141-149.

Purcell D, Sompong U, Yim LC, Barraclough TG, Peerapornpisal Y & Pointing SB (2007) The effects of temperature, pH and sulphide on the community structure of hyperthermophilic streamers in hot springs of northern Thailand. *FEMS Microbiol Ecol* **60**: 456-466.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J & Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**: D590-D596.

Rink B, Seeberger S, Martens T, Duerselen C-D, Simon M & Brinkhoff T (2011) Regional patterns of bacterial community composition and biogeochemical properties in the southern North Sea. *Aquatic Microbial Ecology* **63**: 207-222.

Rosenblueth M & Martinez-Romero E (2006) Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact* **19**: 827-837.

Rusch DB, Halpern AL, Sutton G, *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology* **5**: e77.

Sahm K, John P, Nacke H, Wemheuer B, Grote R, Daniel R & Antranikian G (2013) High abundance of heterotrophic prokaryotes in hydrothermal springs of the Azores as revealed by a network of 16S rRNA gene-based methods. *Extremophiles* **17**: 649-662.

Schattenhofer M, Fuchs BM, Amann R, Zubkov MV, Tarran GA & Pernthaler J (2009) Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ Microbiol* **11**: 2078-2093.

Schmidt TM, DeLong EF & Pace NR (1991) Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J Bacteriol* **173**: 4371-4378.

Scott SA, Davey MP, Dennis JS, Horst I, Howe CJ, Lea-Smith DJ & Smith AG (2010) Biodiesel from algae: challenges and prospects. *Current Opinion in Biotechnology* **21**: 277-286.

Seghers D, Wittebolle L, Top EM, Verstraete W & Siciliano SD (2004) Impact of agricultural practices on the Zea mays L. endophytic community. *Appl Environ Microbiol* **70**: 1475-1482.

Selje N, Simon M & Brinkhoff T (2004) A newly discovered Roseobacter cluster in temperate and polar oceans. *Nature* **427**: 445-448.

Senthilkumar M, Anandham R, Madhaiyan M, Venkateswaran V & Sa T (2011) Endophytic Bacteria: Perspectives and Applications in Agricultural Crop Production. *Bacteria in Agrobiology: Crop Ecosystems*,(Maheshwari DK, ed.) p.^pp. 61-96. Springer Berlin Heidelberg.

Sessitsch A, Reiter B, Pfeifer U & Wilhelm E (2002) Cultivation-independent population analysis of bacterial endophytes in three potato varieties based on eubacterial and Actinomycetes-specific PCR of 16S rRNA genes. *FEMS Microbiol Ecol*, Vol. 39 p.^pp. 23-32. England.

Siddiqui I & Shaukat S (2003) Endophytic bacteria. Prospects and opportunities for the biological control of plant parasitic nematodes. *Nematologia Mediterranea* **31**.

Simon C & Daniel R (2011) Metagenomic analyses: past and future trends. *Appl Environ Microbiol* **77**: 1153-1161.

Sintes E, Witte H, Stodderegger K, Steiner P & Herndl GJ (2013) Temporal dynamics in the free-living bacterial community composition in the coastal North Sea. *FEMS Microbiology Ecology* **83**: 413-424.

Skirnisdottir S, Hreggvidsson GO, Hjorleifsdottir S, Marteinsson VT, Petursdottir SK, Holst O & Kristjansson JK (2000) Influence of sulfide and temperature on species composition and community structure of hot spring microbial mats. *Appl Environ Microbiol* **66**: 2835-2841.

Sperling M, Giebel HA, Rink B, Grayek S, Staneva J, Stanev E & Simon M (2012) Differential effects of hydrographic and biogeochemical properties on the SAR11 clade and Roseobacter RCA cluster in the North Sea. *Aquatic Microbial Ecology* **67**: 25-34.

Stoltzfus JR, So R, Malarvithi PP, Ladha JK & de Bruijn FJ (1997) Isolation of endophytic bacteria from rice and assessment of their potential for supplying rice with biologically fixed nitrogen. *Plant and Soil* **194**: 25-36.

Sturz AV & Nowak J (2000) Endophytic communities of rhizobacteria and the strategies required to create yield enhancing associations with crops. *Applied Soil Ecology* **15**: 183-190.

Sturz AV, Christie BR & Nowak J (2000) Bacterial Endophytes: Potential Role in Developing Sustainable Systems of Crop Production. *Critical Reviews in Plant Sciences* **19**: 1-30.

Tan Z, Hurek T & Reinhold-Hurek B (2003) Effect of N-fertilization, plant genotype and environmental conditions on nifH gene pools in roots of rice. *Environmental microbiology* **5**: 1009-1015.

Teeling H, Fuchs BM, Becher D, *et al.* (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608-611.

Thompson F, Bruce T, Gonzalez A, *et al.* (2011) Coastal bacterioplankton community diversity along a latitudinal gradient in Latin America by means of V6 tag pyrosequencing. *Archives of Microbiology* **193**: 105-114.

Ugwu C, Aoyagi H & Uchiyama H (2008) Photobioreactors for mass cultivation of algae. *Bioresource Technology* **99**: 4021-4028.

Venter JC, Remington K, Heidelberg JF, *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66-74.

Vila-Costa M, Gasol JM, Sharma S & Moran MA (2012) Community analysis of high- and low-nucleic acid-containing bacteria in NW Mediterranean coastal waters using 16S rDNA pyrosequencing. *Environmental Microbiology* **14**: 1390-1402.

Vollmers J, Voget S, Dietrich S, Gollnow K, Smits M, Meyer K, Brinkhoff T, Simon M & Daniel R (2013) Poles Apart: Arctic and Antarctic *Octadecabacter* strains Share High Genome Plasticity and a New Type of Xanthorhodopsin. *PLoS ONE* **8**: e63422.

Wagner-Döbler I, Ballhausen B, Berger M, *et al.* (2010) The complete genome sequence of the algal symbiont Dinoroseobacter shibae: a hitchhiker's guide to life in the sea. *ISME J* **4**: 61-77.

Ward D & Castenholz R (2002) Cyanobacteria in Geothermal Habitats. *The Ecology of Cyanobacteria*,(Whitton B & Potts M, eds.), p.^pp. 37-59. Springer Netherlands.

Waschkowitz T, Rockstroh S & Daniel R (2009) Isolation and characterization of metalloproteases with a novel domain structure by construction and screening of metagenomic libraries. *Appl Environ Microbiol* **75**: 2506-2516.

Watanabe K, Takihana N, Aoyagi H, Hanada S, Watanabe Y, Ohmura N, Saiki H & Tanaka H (2005) Symbiotic association in Chlorella culture. *FEMS microbiology ecology* **51**: 187-196.

Wemheuer B, Wemheuer F & Daniel R (2012) RNA-Based Assessment of Diversity and Composition of Active Archaeal Communities in the German Bight. *Archaea*,vol. **2012**, Article ID 695826.

Wemheuer B, Taube R, Akyol P, Wemheuer F & Daniel R (2013) Microbial Diversity and Biochemical Potential Encoded by Thermal Spring Metagenomes Derived from the Kamchatka Peninsula. *Archaea*, vol **2013**, Article ID 136714.

Wemheuer B, Güllert S, Billerbeck S, Giebel H-A, Voget S, Simon M & Daniel R (2013) Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiology Ecology* (accepted)

West NJ, Obernosterer I, Zemb O & Lebaron P (2008) Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. *Environ Microbiol* **10**: 738-756.

Whitaker RJ, Grogan DW & Taylor JW (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **301**: 976-978.

Whitman WB, Coleman DC & Wiebe WJ (1998) Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* **95**: 6578-6583.

Wiley PE, Campbell JE & McKuin B (2011) Production of biodiesel and biogas from algae: a review of process train options. *Water Environment Research* **83**: 326-338.

Williams PJlB & Laurens LM (2010) Microalgae as biodiesel & biomass feedstocks: Review & analysis of the biochemistry, energetics & economics. *Energy & Environmental Science* **3**: 554-590.

Yim LC, Hongmei J, Aitchison JC & Pointing SB (2006) Highly diverse community structure in a remote central Tibetan geothermal spring does not display monotonic variation to thermal stress. *FEMS Microbiol Ecol* **57**: 80-91.

Zubkov MV, Fuchs BM, Archer SD, Kiene RP, Amann R & Burkill PH (2001) Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ Microbiol* **3**: 304-311.

APPENDIX

OTHER PUBLICATIONS

JOURNAL ARTICLES (PEER-REVIEWED)

- **Wemheuer B**, Taube R, Akyol P, Wemheuer F, Daniel R. Microbial diversity and biochemical potential encoded by thermal spring metagenomes derived from the Kamchatka peninsula. Archaea, vol. 2013: Article ID 136714, 2013. doi:10.1155/2013/136714.

- Mientus M, Brady S, Angelov A, Zimmermann P, **Wemheuer B,** Schuldes J, Daniel R, Liebl W (2013) Thermostable xylanase and β-glucanase derived from the metagenome of the Avachinsky Crater in Kamchatka (Russia). Current Biotechnology 2: 284-293.

TALKS AND POSTER PRESENTATIONS

TALKS AT CONFERENCES

- **Wemheuer B** (2013): Metagenomics and Metatranscriptomics at the Göttingen Genomics Laboratory (invited speaker). German Conference on Bioinformatics (GCB), Göttingen (10.-13.09.2013).

- Wemheuer F, **Wemheuer B**, Kretzschmar D, Daniel R, Vidal S (2013): Multitrophic interaction between microorganisms, plant and herbivores: Does fertilizing, mowing or herbivory on plants alter the microbial community diversity in the rhizosphere? Annual Conference of the Association for General and Applied Microbiology (VAAM), Bremen (10.-13.03.2013).

POSTERS AT CONFERENCES (SELECTION)

- **Wemheuer B**, Güllert S, Meier D, and Daniel R (2013). Diversity, ecology, and physiology of the Roseobacter clade and other marine microbes in the North Sea. Annual Conference of the Association for General and Applied Microbiology (VAAM), Bremen (10.-13.03.2013).

- **Wemheuer B**, Güllert S, Meier D, and Daniel R (2012). Metatranscriptomics of the Roseobacter clade. "International Symposium on Microbial Ecology", Copenhagen, 19.-24.08.2012.

- **Wemheuer B** and Daniel R, (2010): Characterization of metagenomes derived from volcanic sites of Kamchatka. Annual Conference of the Association for General and Applied Microbiology (VAAM), Hannover (28.-31.03.2010).

ACKNOWLEDGEMENTS

It would not have been possible to write this thesis without the help and support of many kind people around me.

First of all, I would like to thank Prof. Dr. Rolf Daniel for giving me the opportunity to work on this interesting, interdisciplinary topic as well as for the continuous support and supervision during my thesis. Many thanks for letting me work independently on my thesis but providing valuable advice when needed, and for the critically and constructive reading of my manuscripts.

I am very grateful to Prof. Dr. Stefanie Pöggeler for accepting to co-examine this thesis and always having a friendly ear.

This study is part of the TRR51'Ecology, Physiology and Molecular Biology of the *Roseobacter* clade'. I want to thank the Deutsche Forschungsgesellschaft (DFG) for funding. I am very grateful to Prof. Dr. Meinhard Simon, the speaker of the TRR, for helpful comments and feedbacks on my manuscripts and the cookies during the various research cruises. I also appreciated the cooperation with so many people in the TRR: Sara Billerbeck, Marco Dogs, Bert Engelen, Helge-Ansgar Giebel, Saranya Kanukollu, Judith Lucas, and all the other people. It was a great time!

I was not able to conduct all this work by myself and thus I want to thank: Simon Güllert, Peggy Klempert, and Dimitri Meier who wrote their master theses under my supervision and contributed significantly to this work, Pinar Akyol and Robert Taube who wrote their Bachelor theses under my supervision, and Sonja Voget and John Vollmers, the other '*Roseobacters'* in our group.

I would also like to thank my colleagues of the Department of Genomic and Applied Microbiology and the Göttingen Genomics Laboratory for the pleasant atmosphere. I want to thank: Mechthild Bömeke, Silja Brady, Sarah Herzog, Heiko Nacke, Birgit Pfeiffer, Dominik Schneider, Tanja Waschkowitz, and all the

other people. A special thanks to Andrea Thürmer, Frauke-Dorothee Meyer, and Kathleen Gollnow for all the sequencing stuff. I would also like to thank some former members of our group: Jenny Hüpeden, Boris Nowka, Steffi Rockstroh, Juliane Schnabel, Marie Schnapp, Carola Schröder, Carola Simon, and Jörn Voss.

Finally, I am very grateful to my wife Franziska for the love and understanding during all the time, the critical reading of my manuscripts, and the statistical advice. I am also grateful to my son Erik who kept me successfully away from sleeping and to my parents for their patience (direct quote: “Wann wirst du endlich fertig?!!”).

CURRICULUM VITAE

Personal information

Name	Bernd Wemheuer
Nationality	German
Date of birth	21 April 1984
Place of Birth	Northeim, Germany
Marital status	Married with one child

Education

04/2010 – 01/2014	PhD candidate in the TRR51: "Ecology, Physiology and Molecular Biology of the *Roseobacter* clade" Genomic and Applied Microbiology and Göttingen Genomics Laboratory, University of Göttingen
10/2003 – 09/2009	Study of Biology at the University of Göttingen Degree obtained: Dipl.-Biol. Univ Main subject: Microbiology Minor subjects: Human Genetics and Physical Chemistry Diploma thesis: „Characterization of microbial metagenomes derived from extreme habitats of the Kamchatka peninsula"
09/ 1990 – 06/2003	High school graduation (*Allgemeine Hochschulreife*) Ernst-Moritz-Arndt Gymnasium, Herzberg am Harz

Work Experience

10/2009 – 03/2010	Research assistant Genomic and Applied Microbiology, University of Göttingen
09/2010	Student assistant Genomic and Applied Microbiology, University of Göttingen

Other Experiences

03/2012 – 04/2012	Research trip in the Southern Ocean on board of the RV Polarstern (AntXXVIII-4)
07/2011	Research trip in the North Sea on board of the RV Heincke (HE361)
05/2010	Research trip in the North Sea on board of the RV Heincke (HE327)
09/2009	Research trip to South Africa University of the Free State, Bloemfontein, South Africa