Title: Diversity of *Rhodopirellula* and related planctomycetes in a North Sea coastal sediment employing *carB* as molecular marker

Authors: Marina Zure[1], Colin B. Munn[2] and Jens Harder[1]*

[1]Dept. of Microbiology, Max Planck Institute for Marine Microbiology, D-28359 Bremen, Germany, and [2]School of Marine Sciences and Engineering, University of Plymouth, Plymouth PL4 8AA, United Kingdom

*corresponding author:
Jens Harder
Max Planck Institute for Marine Microbiology
Celsiusstrasse 1
28359 Bremen, Germany
Email: jharder@mpi-bremen.de
Phone: ++49 421 2028 750
Fax: ++49 421 2028 590

Running title: *carB* diversity detects *Rhodopirellula* species

Abstract

Rhodopirellula is an abundant marine member of the bacterial phylum *Planctomycetes*. Cultivation studies revealed the presence of several closely related *Rhodopirellula* species in European coastal sediments. Because the 16S rRNA gene does not provide the desired taxonomic resolution to differentiate *Rhodopirellula* species, we performed a comparison of the genomes of nine *Rhodopirellula* strains and six related planctomycetes and identified *carB*, coding for the large subunit of carbamoylphosphate synthetase, as a suitable molecular marker. In this study we investigated the diversity of *Rhodopirellula* in coastal intertidal surface sediments of Sylt island, North Sea, using the 16S rRNA and *carB* genes as molecular markers. The *carB* clone and pyrosequencing libraries revealed the presence of 12 species of *Rhodopirellula* and of 66 species in closely related undescribed genera, a diversity that was not detected with a 16S rRNA gene library. This study demonstrates that the

carB gene is a powerful molecular marker for detecting *Rhodopirellula* species in the environment and may be used for the taxonomic evaluation of new strains.

Keywords: *Rhodopirellula*, 16S rRNA gene, *carB* gene, diversity, molecular marker

1. Introduction

The genus *Rhodopirellula* belongs to the widespread bacterial phylum *Planctomycetes* in the *Planctomycetes-Verrucomicrobia-Chlamydiae* (PVC) super-phylum (Wagner and Horn, 2006). The planctomycetes possess phenotypic characteristics unusual for the domain *Bacteria*, including reproduction by budding, and an intracellular membrane-bounded compartmentalization which extent varies between species (Fuerst and Sagulenko, 2011; Speth, van Teeseling and Jetten, 2012; Santarella-Mellwig *et al.*, 2013). Recently, cyro transmission electron microscopy and cyro electron tomography studies provided evidence for a peptidoglycan layer (van Teeseling *et al.* 2015, Jeske *et al.* 2015). *Planctomycetes* are abundant and highly diverse and have been proposed to contribute to the global carbon cycle by turnover of complex carbohydrates in marine sediments and marine snow (Glöckner *et al.*, 2003). Planctomycetes include free-living as well as attached-living organisms. Members have been associated with phytoplankton blooms (Morris, Longnecker and Giovannoni, 2006; Pizzetti *et al.*, 2011) marine snow particles (DeLong, Franks and Alldredge, 1993; Gade *et al.*, 2004; Fuchsman *et al.*, 2012) and in association with several eukaryotic organisms like prawns and sponges (Fuerst *et al.*, 1997; Pimentel-Elardo *et al.*, 2003; Izumi *et al.*, 2013). Several *Planctomycetes* belonging to new genera or new *Rhodopirellula* species were isolated from the microbial community on macroalgae (Bengtsson and Øvreås, 2010, 2010; Lage and Bondoso, 2011) Winkelmann and Harder (2009) isolated 70 strains from European seas, which affiliated according to 16S rRNA gene analysis with the strain *R. baltica* SH1^{T}. In a multilocus sequence analysis (MLSA) and in combination with DNA–DNA hybridization experiments, those isolates were arranged into 13 operational taxonomic units suggesting a high diversity of *Rhodopirellula*, which could not be deduced using the 16S rRNA gene solely as a marker gene (Winkelmann *et al.*, 2010).

The 16S rRNA gene is a powerful marker for classification of microorganisms and the taxonomic resolution of this gene enables the differentiation of genera, but not of closely-related species in a genus (Yarza *et al.*, 2014). Therefore, the bacterial species

definition can never be based solely on sequence similarity of 16S rRNA genes. Two organisms with almost identical 16S rRNA gene sequences can still be recognized as two different species based on DNA-DNA hybridisation (DDH) (Fox, Wisotzkey and Jurtshuk, 1992). Stackebrandt and Ebers (2006) recommended that above a value between 98.7 and 99 % in the 16S rRNA gene sequence similarity DDH would be necessary. With the advent of rapid genome sequencing, DDH is being replaced by determination of the average nucleotide identity (ANI) of shared genes between two strains (Richter and Rosselló-Móra, 2009). Konstantinidis and Tiedje (2005) showed that ANI values of 94–95 % correspond to 70% similarity in DNA-DNA hybridization experiments, the current standard of the species definition. Besides the 16S rRNA gene, other molecular markers can be used to resolve bacterial phylogenetic relationships. Recently, the *rpoB* gene, coding for the beta subunit of the RNA polymerase, has been used for the taxonomic affiliation of *Planctomycetes* strains (Bondoso, Harder and Lage, 2013).

The aim of this study was to investigate the diversity of *Rhodopirellula* in sandy intertidal sediments of Sylt island, Germany, using as molecular markers 16S rRNA and *carB* genes, the latter coding for the large subunit of carbamoylphosphate synthetase. The *carB* gene was selected based on a comparison of genomes of nine *Rhodopirellula* strains (Glöckner *et al.*, 2003; Klindworth *et al.*, 2014; Richter *et al.*, 2014a, 2014b; Richter-Heitmann *et al.*, 2014; Wegner *et al.*, 2014) and six related planctomycetes following the requirements for candidate genes used in a species prediction (Stackebrandt *et al.*, 2002).

2. Materials and methods

2.1 Sampling and DNA extraction

Samples from the upper 2 cm of sandy coarse sediment were obtained from two locations, Hausstrand beach (55.01518 N, 8.43814 E) and Weststrand beach (55,03840 N 8,38490 E) on the coast of Sylt island, Germany. Samples were collected into 15 ml falcon tubes and frozen at -80°C till further processing. Genomic DNA was extracted from 500 mg of sediment using FastDNA® Spin Kit for Soil (MP Biomedicals, USA) according to manufacturer instructions. The quantity and quality of extracted DNA was determined with a NanoDrop 100 Spectrophotometer (Thermo

Fisher Scientific, USA) followed by agarose gel electrophoresis and ethidium bromide (EtBr) staining.

2.2 PCR amplification of 16S rRNA gene

The general bacterial 16S rRNA gene forward primer 8-27F (Juretschko *et al.*, 1998) and the planctomycete-specific reverse primer PLA886 (Neef *et al.*, 1998) were used to amplify partial 16S rRNA gene in a 30µl of PCR mixture containing 10-15 ng of genomic DNA, 3.3 µM of each primer and 15 µl of GoTaq® Hot Start Colorless Master Mix (Promega, USA). The template DNA was denatured for 4 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 65°C, 3 min at 72°C, and a final extension of 10 min at 72 °C.

2.3 Design and PCR amplification of *carB* gene

Alignments of functional genes that were present in single copy in nine *Rhodopirellula* and six other planctomycetes genomes (*Blastopirellula marina*, *Gemmata obscuriglobus*, *Planctomyces maris*, *Planctomyces limnophilus*, *Planctomyces staleyi* and *Kuenenia stuttgartiensis*) were used to identify suitable marker genes. Criteria for marker gene regions included a unique PFAM model, an annotated function, two conserved regions flanking 500-800 bases of high variability, and primer sites conserved in all nine *Rhodopirellula* genome and very different in the other planctomycetal genomes. Genes were selected after manual alignment inspection of 373 candidate genes and primers were designed using the conservation plots. For the *carB* gene, a forward degenerated primer at the position 2095-2114 (5´-GCHCGBAACATGGAMGAAGC-3´) and a reverse degenerated primer at the postion 2827-2808 (5´-CVGCGAKTTGGCTYTTKGCR-3´) were highly specific for *Rhodopirellula* strains and used to generate 693bp long *carB* amplicons. The PCR mixture contained 10-15 ng of genomic DNA, 3.3 µM of each *carB* primer, and a 2x PCR master mix (Promega, USA) in 25 µl volume. The template DNA was denatured for 4 min at 94°C, followed by 40 cycles of 1 min at 94°C, 1 min at 60°C, 3 min at 72°C, and a final extension of 10 min at 72 °C.

2.4 Gene library construction and sequencing of 16S rRNA gene and *carB* amplicons

16S rRNA gene and *carB* PCR amplicons were purified on Sephadex™ G-50 Superfine columns (Amersham Biosciences AB, Uppsala, Sweden) and ligated into

the vector pCR4 applying the TOPO® TA cloning (Invitrogen™, USA). Inserts were sequenced using the BigDye Terminator v3.1 Cyle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA) on an ABI PRISM 3130xl Genetic Analyzer.

2.5 454 pyrosequencing of *carB* amplicon

In addition to the clone library, a *car*B amplicon from a Hausstrand sample was analyzed by 454 pyrosequencing. PCR amplification was performed in two steps. First PCR amplicons were obtained according to the aforementioned protocol with 20 cycles of amplification. The PCR products were purified with the QIAquick ® PCR purification kit (Qiagen, Hilden, Germany) and 1 µl of purified PCR amplicon was used for a second PCR amplification with fusion primers according to the aforementioned protocol for 20 cycles. Fusion primers contained in front of the *carB* primers a linker required for 454 sequencing; forward fusion primer (5´-GATGGCCATTACGGCC - GCHCGBAACATGGAMGAAGC-3´), reverse fusion primer (5´- GGTGGCCGAGGCGGCCACACGT - CVGCGAKTTGGCTYTTKGCR-3´). The amplicon was purified and sequenced on a 454 GS FLX sequencer at the Max Planck Genome Centre, Cologne.

2.6 Analysis of 16S rRNA gene and *carB* clone library sequences

16S rRNA gene and *carB* sequences were analyzed with Sequencing Analysis 5.2 (Applied Biosystems, Carlsbad, CA, USA) and truncated within Sequencer 4.6 (Gene Codes, Michigan, USA). The 16S rRNA gene sequences were aligned in ARB software (Ludwig *et al.*, 2004) using the SINA_aligner (Pruesse, Peplies and Glöckner, 2012). Manual refinement was conducted by comparison with their closest relatives. The *carB* sequences were first translated into protein sequences and visually inspected for the presence of stop codons. Translated sequences were aligned with MAFFT (Katoh, Asimenos and Toh, 2009) and the nucleotide sequence alignment was generated according to the protein alignment. Distance matrices of 16S rRNA gene and *carB* sequences were calculated using the neighbor-joining method as implemented in ARB. OTU clustering was performed based on those matrices using the software Mothur v 1.29.1 (Schloss *et al.*, 2009). Representative OTU sequences were used to calculate phylogenetic maximum likelihood (ML) trees in ARB using RAxML 7.0.4 (Stamatakis, 2006) with 50% minimal similarity filter.

2.7 Analysis of *carB* pyrosequencing reads

Pyrosequencing reads were first processed in Mothur. The analysis included removal of primer sequences, quality control to remove sequences with ambiguous nucleotides, denoising and removal of chimeric sequences using the UCHIME algorithm (Edgar *et al.*, 2011). Sequences were translated, inspected for stop codons and aligned with MAFFT. The corresponding nucleotide sequence alignment was generated according to the protein alignment. A distance matrix was calculated using the neighbor-joining method as implemented in ARB and OTU clustering was performed based on that matrix using the software Mothur. OTU representatives of 454 reads were aligned together with the OTU representative sequences of the *carB* clone libraries and the alignment confidence scores were assessed using the GUIDANCE algorithm (Penn *et al.*, 2010b) implemented in the GUIDANCE web server (Penn *et al.*, 2010a). Sequences below the confidence score of 0.826 were removed from the analysis. The 454 reads were then placed in the reference ML *carB* tree using the evolutionary placement algorithm (EPA) (Berger, Krompass and Stamatakis, 2011) implemented in RAxML 8 (Stamatakis, 2014).

2.8 Accession numbers

The 16S rRNA gene and *carB* nucleotide sequences obtained in this study were deposited in the Genbank Acc (Accession numbers will be provided in the later stage)

3. Results

3.1. Analysis of 16S rRNA gene clone library

The clone libraries from both sample sites yielded 158 sequences, of which 119 affiliated with *Planctomycetes*. The 119 planctomycetal sequences were aligned covering the 16S rRNA gene positions 57-906 (numbering related to *E. coli* sequence) and grouped into OTUs using either a genus-related threshold (95%), resulting in 63 OTUs, or a species-related threshold (97%), resulting in 73 OTUs. The phylogenetic analysis placed 7 OTUs at the 97% threshold closer to the genus *Rhodopirellula* than to its next validly described relative *Blastopirellula*, but none of the clone sequences affiliated closely with the species in the genus *Rhodopirellula* (Fig 1). OTUs with the closest affiliation to *Rhodopirellula* formed a group with the strain SM50, which

represents a novel genus based on a gene identity of 93.7% for the complete 16S rRNA gene (Winkelmann *et al.*, 2010) and is tentatively named `*Rhodopilula apulia*`.

3.2. Correlation between *carB* gene and ANI

To resolve intra and interspecies relationship and to define the species threshold for the partial *carB* gene (693bp) used in this study, the pairwise similarity values of the *carB* gene were plotted against the ANI values between each pair of genomes (Fig.2). The ANIs were determined by using the *in silico* DNA-DNA hybridization method of the JSpecies software (Richter and Rosselló-Móra 2009). Strains used for the calculation and *carB* identities of strains to *Rhodopirellula baltica* SH1^T are shown in Table 1. Genome analyses of four strains of *R. baltica* and two strains of the related species `*Rhodopirellula europaea*` established an intraspecies identity of the amplicon region of at least 97.4%. The pairwise interspecies average nucleotide identity (ANI) for the two species was 88% for the common genes in the genomes and less than 93.1% for the *carB* amplicon. Thus, a threshold of 95% was well suited to separate closely related *Rhodopirellula* species. *B. marina* and *P. staleyi* had higher sequence similarity (81.1%) than some strains within *Rhodopirellula* e.g. `*R. sallentina*` SM41 and `*R. maiorica*` SM1 (77.7%), so a genus threshold could not be established solely on the basis of *carB* partial gene sequences.

The *carB* sequences of the 454 pyrosequences covered an alignment of 204 base in length. In contrast to the 693 base alignment of Sanger sequences, this *carB* region was slightly higher conserved, resulting in a species border threshold of 97%.

3.3 Analysis of *carB* sequences

Two *carB* clone libraries gave 233 sequences which clustered into 48 OTUs on a 95% similarity threshold. The representative sequences of the OTUs and the *carB* sequences of *Planctomycetes* stains were used for the phylogenetic analysis (Fig 3). Strains and corresponding *carB* gene accession numbers are listed in Table 1. The tree revealed that the majority of the OTUs were more closely related to *Rhodopirellula* than to *Blastopirellula*. We detected one OTU that affiliated with *R. baltica* SH1^T with 99.6% similarity and one OTU that clustered with `*R. maiorica*` SM1 with 96.7% similarity.

A deeper insight into the species diversity was obtained by 454 pyrosequencing of a *carB* amplicon. After preprocessing, the dataset included 7763 reads with a length of

204 bp. The *carB* sequences of the genomes had indicated for the amplicon size of 204 bp a higher species border threshold of 97%. Applying this threshold, the reads clustered in 290 OTUs, of which 157 OTUs were singletons. Chao1, a conservative richness estimator, predicted the presence of 650 OTUs (548-792). After removing singletons and sequences with alignment confidence score below 0.826, 95 OTUs remained and were added to an already constructed ML tree of *carB* clone and strain sequences (Fig 3). Considering the relatedness to the *Rhodopirellula* strain sequences, *carB* sequences obtained in this study clustered in 4 groups.

Group B contained cultured *Rhodopirellula* strains, clone library sequences of *R. baltica* SH1^{T} and `*R. maiorica*` SM1 as well as pyrosequencing reads of *R. baltica*, *R. rosea*, `*R. europaea*`, *Rhodopirellula* sp. CS14 and ´*R. islandica*´ K833. In absolute numbers, group B included 154 reads and 13 clone sequences. In addition to known strains, phylogenetic analysis of the group B suggested a presence of six so far uncultured *Rhodopirellula* species. Pairwise identities within group B were 73-100%, between members of the group to ´*R. apulia*´ SM50 (group C) 72-75% and to *B. marina* and *P. staleyi* (group D) 67-73% and 67-73%, respectively. Group C comprised ´*R. apulia*´ SM50 together with six OTU representatives.

Group A represents a taxon related to the genus *Rhodopirellula*, but so far no strain has been brought into culture. In this group, phylogenetic analysis of 28 clones with pairwise identities of 74-91% and 43 pyrosequencing OTUs suggest the presence of 60 novel species. Group A had 73-80% similarity to cultured *Rhodopirellula* strains, 72-77% to ´*R. apulia*´ SM50 and 67-73% to *B. marina* and *P. staleyi*.

Group D comprises sequences far away from *Rhodopirellula*. It includes OTUs with a relationship of equal distance to *Blastopirellula*, *Pirellula* and *Rhodopirellula*, with similarities of 64-73%. Group D formed in the tree (Fig. 3) a monophyletic branch with *Blastopirellula* and *Pirellula*.

4. Discussion

The diversity of *Rhodopirellula* in coastal sediments of Sylt was characterized by 16S rRNA gene libraries, and by libraries and 454 sequencing of *carB* gene, with primers developed to target specifically the genus *Rhodopirellula*. Sylt was chosen for the diversity study as *Planctomycetes* were reported to be very abundant in the upper layers of sandy intertidal sediments on the island, accounting for between 3 and 19 % of all cells (Musat *et al.*, 2006).

For a planctomycete-enriched 16S rRNA gene study, the PLA886 probe developed for *in situ* hybridization (Neef *et al.*, 1998) was used as a *Planctomycetes*-specific reverse primer for the generation of 16S rRNA gene libraries (Pynaert *et al.*, 2003). This probe covered 91.6% of *Planctomycetaceae* and 91.5% of *Rhodopirellula* sequences of the SILVA database (Quast *et al.*, 2013) and was selected because earlier *in situ* hybridization studies on the coastal sand under study were performed with this probe (Musat *et al.*, 2006). As the *Planctomycetes* specific forward primers Pla40F and Pla46F had low percentage matches within the target group (Pollet, Tadonléké and Humbert, 2011), we used the general bacterial primer 8-27F as forward primer (Juretschko *et al.*, 1998). The libraries yielded 119 *Planctomycetes* sequences of which 14 clustered more closely to *Rhodopirellula* then to other validly described *Planctomycetes*. Although ~12% of the sequences clustered closely to *Rhodopirellula*, we did not detect any of the cultured *Rhodopirellula* strains. Similar sequences were also found in water samples taken at Helgoland in the German Bight (Pizzetti *et al.*, 2011). In that study, the *Rhodopirellula* related group accounted also for ~12% of the planctomycetal diversity. In a 16S rRNA V6 region 454 pyrosequencing study of Sylt sediments (project ICM_FIS_Bv6, vamps.mbl.edu, Huse *et al.*, 2014), 4566 planctomycetal reads contained 86 reads affiliating to *Rhodopirellula baltica* SH1^{T} and 75 reads affiliating to strain SM50 (more than 95% identity over 60 bp). The V6 regions of SM50 and *R. baltica* had an identity of just below 95%, of *B. marina* and *R. baltica* had an identity of 83%. This deep sequencing study also indicated that cells of *Rhodopirellula* present a small fraction of all planctomycetes in Sylt sediments.

To get a better insight into the genus *Rhodopirellula*, we wanted to go beyond the resolution of the 16S rRNA gene and therefore studied the diversity of *carB* gene. Functional genes are less conserved than 16S rRNA genes and more suitable for a characterization at a higher taxonomic level. However, the dissimilarities between sequences of species within one genus and of species of closely related genera are too similar, so a reliable prediction of genus borders is not possible (Kim *et al.*, 2014). The availability of nine genomes of *Rhodopirellula* strains allowed the correlation of the *carB* gene similarity with the ANI which showed that threshold value for the *carB* gene of 95% is well suited to separate closely related species, but the genus threshold could not be established. In these kinds of investigations, the threshold for species borders should be clearly outside the technical uncertainty of the sequencing

technology. For example, the *rpoB* gene sequence has a higher degree of conservation with a species border threshold of 98.2% (Bondoso, Harder and Lage, 2013), clearly above the ANI of shared genes between species making it less suitable for a diversity study.

The *carB* study revealed the presence of sequences with high similarity with *R. baltica* SH1^{T}, `*R. maiorica*` SM1, *R. rosea*, `*R. europaea*`, ´*R. islandica*´ K833, *Rhodopirellula* sp. CS14 and six so far uncultured *Rhodopirellula* species within the group B, plus 66 novel species closely related to *Rhodopirellula* within the groups A and C. We also detected 50 new species of *Planctomycetes* in the group D which are not closely related to *Rhodopirellula*.

This study demonstrated a high diversity of *Rhodopirellula* in North Sea sediments from Sylt. The detection of sequences related to strains isolated from the Baltic Sea (*R. baltica*), Iceland (`*R. islandica*` K833), as well as from the Belgium coast and the Mediterranean Sea (`*R. europaea*`) raises the question of the biogeography of these species. The dispersal of organisms with water currents may add to the diversity in Sylt sediments. This place is considered to be in contact with the water masses of the coastal current that originate from the North Atlantic Drift entering the North Sea at Scotland and from Atlantic waters entering through the British Channel (Otto *et al.*, 1990).

In summary, the high taxonomic resolution of the *carB* gene amplicon, together with a calibration of thresholds derived from a set of reference genomes in the genus of interest, provided a deep insight into the microdiversity of a genus, here *Rhodopirellula*, in the environment. The *carB* gene has more discriminatory power than the 16S rRNA gene when analyzing closely related *Rhodopirellula* species and it is suitable to discriminate strains on the intraspecies level and may be used for the taxonomic evaluation of the new isolates.

5. Funding

This work was supported by the Max Planck Society and through a MARES Grant. MARES is a Joint Doctorate programme selected under Erasmus Mundus coordinated by Ghent University (FPA 2011-0016). See www.mares-eu.org for extra information.

6. Acknowledgements

We thank Christina Probian, Tim Richter-Heitmann, Bhabananda Biswas and Samudrahalli Murthy for technical assistance.

7. References

Abraham W-R, Strömpl C, Meyer H *et al.* Phylogeny and polyphasic taxonomy of *Caulobacter* species. Proposal of *Maricaulis* gen. nov. with *Maricaulis maris* (Poindexter) comb. nov. as the type species, and emended description of the genera *Brevundimonas* and *Caulobacter*. *Int J Syst Bacteriol* 1999;**49**:1053–73.

Bauld J, Staley JT. *Planctomyces maris* sp. nov., nom. rev. *Int J Syst Bacteriol* 1980;**30**:657.Bengtsson MM, Øvreås L. Planctomycetes dominate biofilms on surfaces of the kelp *Laminaria hyperborea*. *BMC Microbiol* 2010;**10**:261.

Berger SA, Krompass D, Stamatakis A. Performance, Accuracy, and web server for evolutionary placement of short sequence reads under Maximum Likelihood. *Syst Biol* 2011;**60**:291–302.

Bondoso J, Albuquerque L, Lobo-da-Cunha A *et al. Rhodopirellula lusitana* sp. nov. and *Rhodopirellula rubra* sp. nov., isolated from the surface of macroalgae. *Syst Appl Microbiol* 2014;**37**:157–64.

Bondoso J, Harder J, Lage OM. rpoB gene as a novel molecular marker to infer phylogeny in *Planctomycetales*. *Antonie Van Leeuwenhoek* 2013;**104**:477–88.

DeLong EF, Franks DG, Alldredge AL. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol Oceanogr* 1993;**38**:924–34.

Edgar RC, Haas BJ, Clemente JC *et al.* UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;**27**:2194–200.

Fox GE, Wisotzkey JD, Jurtshuk P. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J Syst Bacteriol* 1992;**42**:166–70.

Fuchsman CA, Staley JT, Oakley BB *et al.* Free-living and aggregate-associated *Planctomycetes* in the Black Sea. *FEMS Microbiol Ecol* 2012;**80**:402–16.

Fuerst JA, Gwilliam HG, Lindsay M *et al.* Isolation and molecular identification of planctomycete bacteria from postlarvae of the giant tiger prawn, Penaeus monodon. *Appl Environ Microbiol* 1997;**63**:254–62.

Fuerst JA, Sagulenko E. Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. *Nat Rev Microbiol* 2011;**9**:403–13.

Gade D, Schlesner H, Glöckner FO *et al.* Identification of planctomycetes with order-, genus-, and strain-specific 16S rRNA-targeted probes. *Microb Ecol* 2004;**47**:243–51.

Giovannoni SJ, Schabtach E, Castenholtz RW. *Isosphaera pallida*, gen. and comb. nov., a gliding, budding eubacterium from hot springs. *Arch Microbiol* 1987;**147**:276-284.

Glöckner FO, Kube M, Bauer M *et al.* Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc Natl Acad Sci* 2003;**100**:8298–303.

Hirsch P, Müller M. *Planctomyces limnophilus* sp. nov., a stalked and budding bacterium from freshwater. *Syst Appl Microbiol* 1986;**6**:276-280.

Huse SM, Mark Welch DB, Voorhis A *et al.* VAMPS: a website for visualization and analysis of microbial population structures. BMC Bioinformatics 2014;**15**: 41.

Izumi H, Sagulenko E, Webb RI *et al.* Isolation and diversity of planctomycetes from the sponge *Niphates* sp., seawater, and sediment of Moreton Bay, Australia. *Antonie Van Leeuwenhoek* 2013;**104**:533–46.

Jeske O, Schüler M, Schumann P *et al.* Planctomycetes do possess a peptidoglycan cell wall. *Nature Comm* 2015;**6**:7116

Juretschko S, Timmermann G, Schmid M *et al.* Combined molecular and conventional analyses of nitrifying bacterium diversity in activated sludge: *Nitrosococcus mobilis* and *Nitrospira*-like bacteria as dominant populations. *Appl Environ Microbiol* 1998;**64**:3042–51.

Katoh K, Asimenos G, Toh H. Multiple alignment of DNA sequences with MAFFT. *Bioinformatics for DNA Sequence Analysis*. Springer, 2009, 39–64.

Kim M, Oh H-S-. S, Park S-C-. C *et al.* Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;**64**:346–51.

Klindworth A, Richter M, Richter-Heitmann T *et al.* Permanent draft genome of *Rhodopirellula rubra* SWK7. *Mar Genomics* 2014;**13**:11–2.

Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* 2005;**102**:2567–72.

Kulichevskaya IS, Ivanova AO, Baulina OI *et al. Singulisphaera acidiphila* gen. nov., sp. nov., a non-filamentous, *Isosphaera*-like planctomycete from acidic northern wetlands. *Int J Syst Evol Microbiol* 2008;**58**:1186–93.

Lage OM, Bondoso J. Planctomycetes diversity associated with macroalgae. *FEMS Microbiol Ecol* 2011;**78**:366–75.

Ludwig W, Strunk O, Westram R *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;**32**:1363–71.

Morris RM, Longnecker K, Giovannoni SJ. *Pirellula* and OM43 are among the dominant lineages identified in an Oregon coast diatom bloom. *Environ Microbiol* 2006;**8**:1361–70.

Musat N, Werner U, Knittel K *et al.* Microbial community structure of sandy intertidal sediments in the North Sea, Sylt-Rømø Basin, Wadden Sea. *Syst Appl Microbiol* 2006;**29**:333–48.

Neef A, Amann R, Schlesner H *et al.* Monitoring a widespread bacterial group: in situ detection of planctomycetes with 16S rRNA-targeted probes. *Microbiology* 1998;**144**:3257–66.

Otto L, Zimmerman JTF, Furnes GK *et al.* Review of the physical oceanography of the North Sea. *Neth J Sea Res* 1990;**26**:161–238.

Penn O, Privman E, Ashkenazy H *et al.* GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Res* 2010a;**38**:W23–8.

Penn O, Privman E, Landan G *et al.* An alignment confidence score capturing robustness to guide tree uncertainty. *Mol Biol Evol* 2010b;**27**:1759–67.

Pimentel-Elardo S, Wehrl M, Friedrich AB *et al.* Isolation of planctomycetes from Aplysina sponges. *Aquat Microb Ecol* 2003;**33**:239–45.

Pizzetti I, Fuchs BM, Gerdts G *et al.* Temporal variability of coastal *Planctomycetes* clades at Kabeltonne station, North Sea. *Appl Environ Microbiol* 2011;**77**:5009–17.

Pollet T, Tadonléké RD, Humbert J-F-. F. Comparison of primer sets for the study of *Planctomycetes* communities in lentic freshwater ecosystems. *Environ Microbiol Rep* 2011;**3**:254–61.

Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;**28**:1823–9.

Pynaert K, Smets BF, Wyffels S *et al.* Characterization of an autotrophic nitrogen-removing biofilm from a highly loaded lab-scale rotating biological contactor. *Appl Environ Microbiol* 2003;**69**:3626–35.

Quast C, Pruesse E, Yilmaz P *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:D590–6.

Richter-Heitmann T, Richter M, Klindworth A *et al.* Permanent draft genomes of the two *Rhodopirellula europaea* strains 6C and SH398. *Mar Genomics* 2014;**13**:15–6.

Richter M, Richter-Heitmann T, Klindworth A *et al.* Permanent draft genomes of the *Rhodopirellula maiorica* strain SM1. *Mar Genomics* 2014a;**13**:19–20.

Richter M, Richter-Heitmann T, Klindworth A *et al.* Permanent draft genomes of the three *Rhodopirellula baltica* strains SH28, SWK14 and WH47. *Mar Genomics* 2014b;**13**:13–4.

Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 2009;**106**:19126–31.

Roh SW, Lee H-W-. W, Yim KJ *et al. Rhodopirellula rosea* sp. nov., a novel bacterium isolated from an ark clam Scapharca broughtonii. *J Microbiol* 2013;**51**:301–4.

Santarella-Mellwig R, Pruggnaller S, Roos N *et al.* Three-dimensional reconstruction of bacteria with a complex endomembrane system. *PLoS Biol* 2013;**11**:e1001565.

Schlesner H. *Planctomyces brasiliensis* sp. nov., a halotolerant bacterium from a salt pit. *Syst Appl Microbiol* 1989;**12**:159–61.

Schlesner H, Hirsch P. Rejection of the genus name *Pirella* for pear-shaped budding bacteria and proposal to create the genus *Pirellula* gen. nov. *Int J Syst Bacteriol* 1987;**37**:441–441.

Schlesner H, Rensmann C, Tindall BJ *et al.* Taxonomic heterogeneity within the *Planctomycetales* as derived by DNA–DNA hybridization, description of *Rhodopirellula baltica* gen. nov., sp. nov., transfer of *Pirellula marina* to the genus *Blastopirellula* gen. nov. as *Blastopirellula marina* comb. nov. and emended description of the genus *Pirellula*. *Int J Syst Evol Microbiol* 2004;**54**:1567–80.

Schloss PD, Westcott SL, Ryabin T *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009;**75**:7537–41.

Schmid M, Twachtmann U, Klein M *et al.* Molecular evidence for genus level diversity of bacteria capable of catalyzing anaerobic ammonium oxidation. *Syst Appl Microbiol* 2000;**23**:93–106.

Speth DR, van Teeseling MC, Jetten MS. Genomic analysis indicates the presence of an asymmetric bilayer outer membrane in *Planctomycetes* and *Verrucomicrobia*. *Front Microbiol* 2012;**3**.

Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006;**33**:152.

Stackebrandt E, Frederiksen W, Garrity GM *et al.* Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* 2002;**52**:1043–7.

Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006;**22**:2688–90.

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;**30**:1312–3.

Takai K, Kobayashi H, Nealson KH *et al. Deferribacter desulfuricans* sp. nov., a novel sulfur-, nitrate- and arsenate-reducing thermophile isolated from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* 2003;**53**:839–46.

van Teeseling MCF, Mesman RJ, Kuru E *et al.* Anammox planctomycetes have a peptidoglycan cell wall. *Nature Comm* 2015;**6**:6878

Wagner M, Horn M. The *Planctomycetes*, *Verrucomicrobia*, *Chlamydiae* and sister phyla comprise a superphylum with biotechnological and medical relevance. *Curr Opin Biotechnol* 2006;**17**:241–9.

Wegner C-E-. E, Richter M, Richter-Heitmann T *et al.* Permanent draft genome of *Rhodopirellula sallentina* SM41. *Mar Genomics* 2014;**13**:17–8.

Winkelmann N, Harder J. An improved isolation method for attached-living *Planctomycetes* of the genus *Rhodopirellula*. *J Microbiol Methods* 2009;**77**:276–84.

Winkelmann N, Jaekel U, Meyer C *et al.* Determination of the diversity of *Rhodopirellula* isolates from European seas by multilocus sequence analysis. *Appl Environ Microbiol* 2010;**76**:776–85.

Yarza P, Yilmaz P, Pruesse E *et al.* Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;**12**:635–45.