

Complete genome sequence of *Granulicella mallensis* type strain MP5ACTX8^T, an acidobacterium from tundra soil

Suman R. Rawat¹, Minna K. Männistö², Valentin Starovoytov³, Lynne Goodwin⁴, Matt Nolan⁵, Loren J. Hauser⁶, Miriam Land⁶, Karen Walston Davenport⁴, Tanja Woyke⁵ and Max M. Häggblom^{1*}

¹ Department of Biochemistry and Microbiology, Rutgers, The State University of New Jersey, New Brunswick, New Jersey USA

² Finnish Forest Research Institute, Rovaniemi, Finland

³ Department of Cell Biology and Neuroscience, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA

⁴ Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA

⁵ DOE Joint Genome Institute, Walnut Creek, California, USA

⁶ Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

*Corresponding author: Max M Häggblom (haggblom@sebs.rutgers.edu)

Keywords: cold adapted, acidophile, tundra soil, *Acidobacteria*

Granulicella mallensis MP5ACTX8^T is a novel species of the genus *Granulicella* in subdivision 1 of *Acidobacteria*. *G. mallensis* is of ecological interest being a member of the dominant soil bacterial community active at low temperatures and nutrient limiting conditions in Arctic alpine tundra. *G. mallensis* is a cold-adapted acidophile and a versatile heterotroph that hydrolyzes a suite of sugars and complex polysaccharides. Genome analysis revealed metabolic versatility with genes involved in metabolism and transport of carbohydrates. These include gene modules encoding the carbohydrate-active enzyme (CAZyme) family involved in breakdown, utilization and biosynthesis of diverse structural and storage polysaccharides including plant based carbon polymers. The genome of *Granulicella mallensis* MP5ACTX8^T consists of a single replicon of 6,237,577 base pairs (bp) with 4,907 protein-coding genes and 53 RNA genes.

Introduction

Strain MP5ACTX8^T (= ATCC BAA-1857^T = DSM 23137^T), is the type strain of the species *Granulicella mallensis* [1]. The genus *Granulicella*, in subdivision 1 of *Acidobacteria*, was first described by Pankratov *et al.* in 2010 [2]. *Granulicella mallensis* (mal.len' sis. N. L. fem. adj. *mallensis*; pertaining to its isolation from soil of Malla Nature Reserve, Kilpisjärvi, Finland; 69°01'N, 20°50'E) was described along with other species of the genus *Granulicella* isolated from tundra soil [1] and is one of the two with sequenced genomes, out of eight validly described *Granulicella* species.

Acidobacteria is one of the most ubiquitous bacterial phyla found in diverse habitats and is abundant in most soil environments [3,4] including Arctic tundra soils [5,6]. *Acidobacteria* are phylogenetically and physiologically diverse [7]

represented by 26 phylogenetic subdivisions [8] of which only subdivisions 1, 3, 4, 8, and 10 are defined by taxonomically characterized representatives. To date, subdivision 1 is comprised of eight genera: *Acidobacterium* [9], *Terriglobus* [10,11], *Edaphobacter* [12], *Granulicella* [1,2], *Acidipila* [13], *Telmatobacter* [14], *Acidicapsa* [15] and *Bryocella* [16]. Subdivision 3, 4 and 10 include only one genus each, namely *Bryobacter* [17], *Blastocatella* [18] and *Thermotomaculum* [19], respectively, while subdivision 8 includes three genera; *Holophaga* [20], *Geothrix* [21] and *Acanthopleuribacter* [22]. Three species, 'Candidatus Koribacter versatilis' [23], 'Candidatus Solibacter usitatus' [23] and 'Candidatus Chloracidobacterium thermophilum' [24] have been described as 'Candidatus' taxa. *Acidobacteria* are relatively difficult to cultivate with slow growth

rates and typically require up to several weeks to develop visible colonies on solid media. Nevertheless, the phylogenetic diversity, ubiquity and abundance of this group suggest that they play important ecological roles in soils. The abundance of *Acidobacteria* has been found to correlate with soil pH [25,26] and carbon [27,28], with subdivision 1 *Acidobacteria* being most abundant in slightly acidic soils. Our previous studies have shown that *Acidobacteria* dominate in the acidic tundra heaths of northern Finland [25,29-31]. Using selective isolation techniques we have been able to isolate several slow growing and fastidious strains of *Acidobacteria* [1,11]. On the basis of phylogenetic, phenotypic and chemotaxonomic data, including 16S rRNA, rpoB gene sequence similarity and DNA-DNA hybridization, strain MP5ACTX8T was classified as a novel species of the genus *Granulicella* [1]. Here, we summarize the physiological features together with the complete genome sequence, annotation and data analysis of *Granulicella mallensis* MP5ACTX8T (Table 1).

Classification and features

Within the genus *Granulicella*, eight species are described with validly published names: *G. mallensis* MP5ACTX8T, *G. tundricola* MP5ACTX9T, *G. arctica* MP5ACTX2T and *G. sapmiensis* S6CTX5AT isolated from Arctic tundra soil [1] and *G. paludicola* OB1010T, *G. pectinivorans* TPB6011T, *G. rosea* TP01014T and *G. aggregans* TPB6028T isolated from sphagnum peat bogs [3]. Strain MP5ACTX8T showed 95.5 - 96.1% 16S rRNA gene sequence identity to tundra soil strains, *G. tundricola* MP5ACTX9T (95.5%), *G. sapmiensis* S6CTX5AT (96.2%) and *G. arctica* MP5ACTX2T (96.1%) and 94.6 - 97.4% to *G. rosea* TP01014T (94.6%), *G. aggregans* TPB6028T (96.0%), *G. pectinivorans* TPB6011T (96.1%), *G. paludicola* OB1010T (96.5%) and *G. paludicola* LCBR1 (97.4%). Phylogenetic analysis based on the 16S rRNA gene of taxonomically classified strains of family *Acidobacteriaceae* placed *G. paludicola* type strain OB1010 T as the closest taxonomically classified relative of *G. mallensis* MP5ACTX8T (Figure 1).

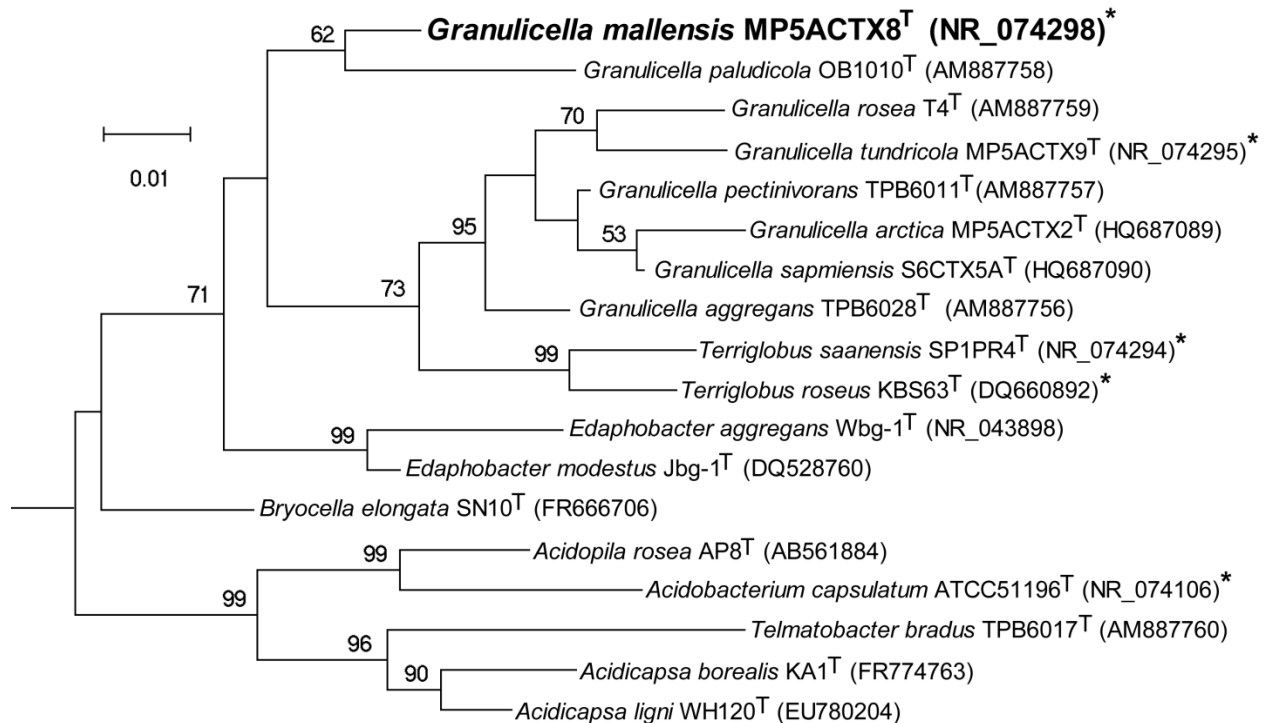


Figure 1. Phylogenetic tree highlighting the position of *G. mallensis* MP5ACTX8 (shown in bold) relative to the other type strains within SD1 *Acidobacteria*. The maximum likelihood tree was inferred from 1,361 aligned positions of the 16S rRNA gene sequences and derived based on the Tamura-Nei model using MEGA 5 [41]. Bootstrap values >50 (expressed as percentages of 1,000 replicates) are shown at branch points. Bar: 0.02 substitutions per nucleotide position. The corresponding GenBank accession numbers are displayed in parentheses. Strains whose genomes have been sequenced, are marked with an asterisk; *G. mallensis* MP5ACTX8 (CP003130), *G. tundricola* MP5ACTX9 (CP002480), *T. saanensis* SP1PR4 (CP002467), *T. roseus* KBS63 (CP003379) and *A. capsulatum* ATCC 51196 (CP001472). *Bryobacter aggregatus* MPL3 (AM162405) in SD3 *Acidobacteria* was used as an outgroup.

Table 1. Classification and general features of *G. mallensis* strain MP5ACTX8^T according to the MIGS recommendations [32]

MIGS ID	Property	Term	Evidence code ^a
		Domain <i>Bacteria</i>	TAS [33]
		Phylum <i>Acidobacteria</i>	TAS [34,35]
		Class <i>Acidobacteria</i>	TAS [36,37]
	Classification	Order <i>Acidobacteriales</i>	TAS [36,38]
		Family <i>Acidobacteriaceae</i>	TAS [34,39]
		Genus <i>Granulicella</i>	TAS [1,2]
		Species <i>Granulicella mallensis</i>	TAS [1]
		Type strain: MP5ACTX8 ^T (= ATCC BAA-1857 ^T = DSM 23137 ^T)	
	Gram stain	negative	TAS [1]
	Cell shape	rod	TAS [1]
	Motility	non-motile	TAS [1]
	Sporulation	not reported	NAS
	Temperature range	4–28 °C	TAS [1]
	Optimum temperature	24–27 °C	TAS [1]
	pH range	3.5–6.5	TAS [1]
	Optimum pH	5	TAS [1]
	Carbon source	D-glucose, maltose, D-fructose, D-galactose, lactose, lactulose, D-mannose, D-ribose, raffinose, sucrose, trehalose, cellobiose, D-xylose, glucuronate	TAS [1]
MIGS-6	Habitat	terrestrial	TAS [1]
MIGS-6.3	Salinity	Growth with up to 1.5% NaCl	TAS [1]
MIGS-22	Oxygen requirement	aerobic	TAS [1]
MIGS-15	Biotic relationship	free-living	TAS [1]
MIGS-14	Pathogenicity	non-pathogenic	NAS
MIGS-4	Geographic location	Arctic-alpine tundra, Finland	TAS [1]
MIGS-5	Sample collection	2006	TAS [1]
MIGS-4.1	Latitude	69°01'N,	TAS [1]
MIGS-4.2	Longitude	20°50'E	
MIGS-4.4	Altitude	700 m	TAS [1]

^aEvidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [40].

Morphology and physiology

G. mallensis grows on R2 medium (Difco) at pH 3.5–6.5 (optimum pH 5) and at +4 to +28 °C (optimum 24–27 °C) [1]. On R2 agar, strain MP5ACTX8^T forms opaque white mucoid colonies with a diameter of approximately 1 mm. Cells are Gram-negative, non-motile, aerobic rods, approximately 0.5–0.7 μm wide and 0.6–1.3 μm long. Growth observed with up to 1.5% NaCl (w/v) (Table 1). The cell-wall structure in ultrathin sections of electron micrographs of cells of MP5ACTX8^T is shown in Figure 2.

G. mallensis utilizes D-glucose, maltose, cellobiose, D-fructose, D-galactose, lactose, lactulose, D-mannose, D-ribose, raffinose, sucrose, trehalose, D-xylose, N-acetyl-D-glucosamine, glucuronate, glutamate, melezitose and salicin, but does not utilize D-arabinose, acetate, formate, pyruvate, malate, mannitol, D- or L-alanine, D-glycine, L-leucine, L-ornithine, gluconic acid, aspartate, dulcitol, butyrate, caproate, valerate, lactate, oxalate, propionate, fumarate, adonitol, methanol, ethanol, succinate, D-sorbitol or myoinositol, when grown using VL55 mineral medium with 100 mg yeast extract l⁻¹. *G. mallensis* hydrolyzes aesculin, starch, pectin, laminarin and lichenan, but not gelatin, cellulose, xylan, sodium alginate, pullulan, chitosan or chitin on R2 medium. Strains show positive reaction for acid and alkaline phosphatases, leucine arylamidase, α-chymotrypsin, naphthol-AS-BI-phosphohydrolase, α- and β-galactosidases, α- and β-glucosidases, N-acetyl-β-glucosaminidase, β-glucuronidase, trypsin and valine arylamidase, but negative for α-fucosidase, α-mannosidase, esterase (C4 and C8), lipase (C14) and cystine arylamidase. Strain MP5ACTX8^T reduces nitrate to nitrite. Strain MP5ACTX8^T is resistant to the antibiotics erythromycin, chloramphenicol, neomycin, rifampicin, streptomycin, gentamicin, polymyxin B and penicillin, but susceptible to ampicillin, kanamycin, tetracycline, lincomycin, novobiocin and bacitracin [1].

Chemotaxonomy

The major cellular fatty acids in *G. mallensis* are iso-C_{15:0} (45.3%), C_{16:1ω7c} (28.7%), iso-C_{13:0} (8.3%) and C_{16:0} (8.9%). The cellular fatty acid compositions of strain MP5ACTX8^T were relatively similar to that of other *Granulicella* strains with fatty acids iso-C_{15:0} and C_{16:1ω7c} being most abundant in all

strains. Strain MP5ACTX8^T contains MK-8 as the major quinone.

Genome sequencing and annotation

Genome project history

G. mallensis strain MP5ACTX8^T was selected for sequencing in 2009 by the DOE Joint Genome Institute (JGI) community sequencing program. The Quality Draft (QD) assembly and annotation were completed on December 26, 2010. The complete genome was made available on Dec. 1, 2011. The genome project is deposited in the Genomes On-Line Database (GOLD) [42] and the complete genome sequence of strain MP5ACTX8^T is deposited in GenBank (CP003130). Table 2 presents the project information and its association with MIGS version 2.0 [32].

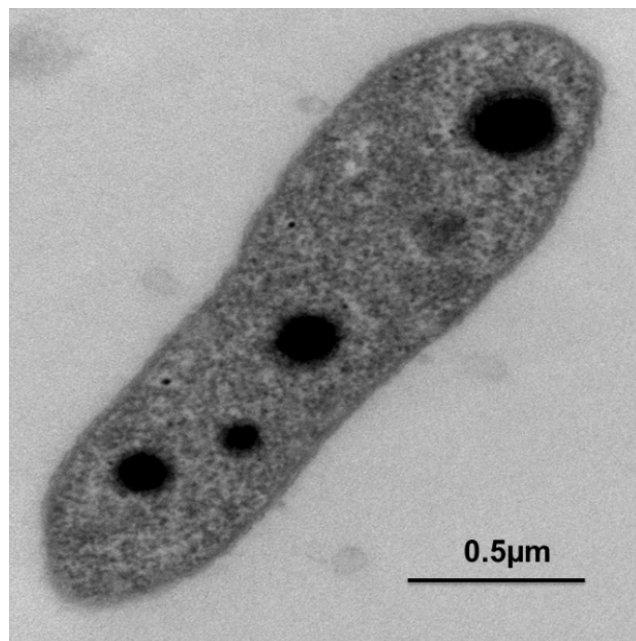


Figure 2. Electron micrograph of *G. mallensis* MP5ACTX8^T.

Growth conditions and genomic DNA extraction

G. mallensis MP5ACTX8^T was cultivated on R2 medium as previously described [1]. Genomic DNA (gDNA) of high sequencing quality was isolated using a modified CTAB method and evaluated according to the Quality Control (QC) guidelines provided by the DOE Joint Genome Institute [43].

Table 2. Project information.

MIGS ID	Property	Term
MIGS 31	Finishing quality	Finished
MIGS-28	Libraries used	Three libraries, an Illumina GAii shotgun library (GSGY), a 454 Titanium standard library (GSXT, GWTA) and a paired end 454 (GSFP) library
MIGS 29	Sequencing platforms	454 Titanium standard, 454 Paired End, Illumina
MIGS 31.2	Fold coverage	18.5× (454), 213× (Illumina)
MIGS 30	Assemblers	Newbler, VELVET, PHRAP
MIGS 32	Gene calling method	ProdigaL, GenePRIMP
	Locus Tag	AciX8
	Genbank ID	CP003130.1
	GenBank Date of Release	December 1, 2011
	GOLD ID	Gc02349
	BIOPROJECT	PRJNA49957, PRJNA47903
	Project relevance	Environmental, Biogeochemical cycling of Carbon, Biotechnological, GEBA

Genome sequencing and assembly

The finished genome of *G. mallensis* MP5ACTX8^T (JGI ID 4088692) was generated at the DOE Joint genome Institute (JGI) using a combination of Illumina [44] and 454 technologies [45]. For this genome, an Illumina GAii shotgun library which generated 59,701,420 reads totaling 4537.3 Mb, a 454 Titanium standard library which generated 136,708 reads and a paired end 454 library with an average insert size of 10.3 kb which generated 157,336 reads totaling 172.0 Mb of 454 data, were constructed and sequenced. All general aspects of library construction and sequencing performed at the JGI can be found at the JGI website [43]. The 454 Titanium standard data and the 454 paired end data were assembled with Newbler, version 2.3. Illumina sequencing data was assembled with Velvet, version 0.7.63 [46]. The 454 Newbler consensus shreds, the Illumina Velvet consensus shreds and the read pairs in the 454 paired end library were integrated using parallel phrap, version SPS - 4.24 (High Performance Software, LLC)

[47]. The software Consed [48] was used in the finishing process. The Phred/Phrap/Consed software package [49] was used for sequence assembly and quality assessment in the subsequent finishing process. Illumina data was used to correct potential base errors and increase consensus quality using the software Polisher developed at JGI (Alla Lapidus, unpublished). Possible misassemblies were corrected using gapResolution (Cliff Han, un-published), Dupfinisher [50] or sequencing cloned bridging PCR fragments with sub-cloning. Gaps between contigs were closed by editing in Consed, by PCR and by Bubble PCR (J-F Cheng, unpublished) primer walks. The final assembly is based on 74.2 Mb of 454 data which provides an average 18.5× coverage and 1318.5 Mb of Illumina data which provides an average 213× coverage of the genome.

Genome annotation

Genes were identified using Prodigal [51] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [52]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) non-redundant database, UniProt, TIGRFam, Pfam, PRIAM, KEGG, COGs [53,54], and InterPro. These data sources were combined to assert a product description for each predicted protein. Non-coding genes and miscellaneous features were predicted using tRNAscan-SE [55], RNAMMer [56], Rfam [57], TMHMM [58], and signalP [59]. Additional gene prediction

analysis and functional annotation were performed within the Integrated Microbial Genomes Expert Review (IMG-ER) platform [60].

Genome properties

The genome consists of one circular chromosome of 6,211,694 bp in size with a GC content of 57.8 mol% and consists of 53 RNA genes (Figure 3 and Table 3). Of the 4,960 predicted genes, 4,907 are protein-coding genes (CDSs) and 90 are pseudogenes. Of the total CDSs, 70.5% represent COG functional categories and 16% consist of signal peptides. The distribution of genes into COG functional categories is presented in Figure 3 and Table 4.

Table 3. Genome statistics

Attribute	Value	% of Total
Genome size (bp)	6,237,577	100
DNA coding region(bp)	5,499,388	88.2
DNA G+C content (bp)	3612173	57.9
DNA scaffolds	1	100
Total genes	4,960	100
Protein coding genes	4,907	98.9
RNA genes	53	1.3
Pseudo genes	90	1.8
Genes in internal clusters	2,679	54
Genes with function prediction	3,511	70.8
Genes assigned to COGs	3,496	70.5
Genes with Pfam domains	3,754	75.7
Genes with signal peptides	797	16.1
Genes with transmembrane helices	1,291	26.0
CRISPR repeats	0	-

The total is based on either the size of the genome in base pairs or the protein coding genes in the annotated genome.

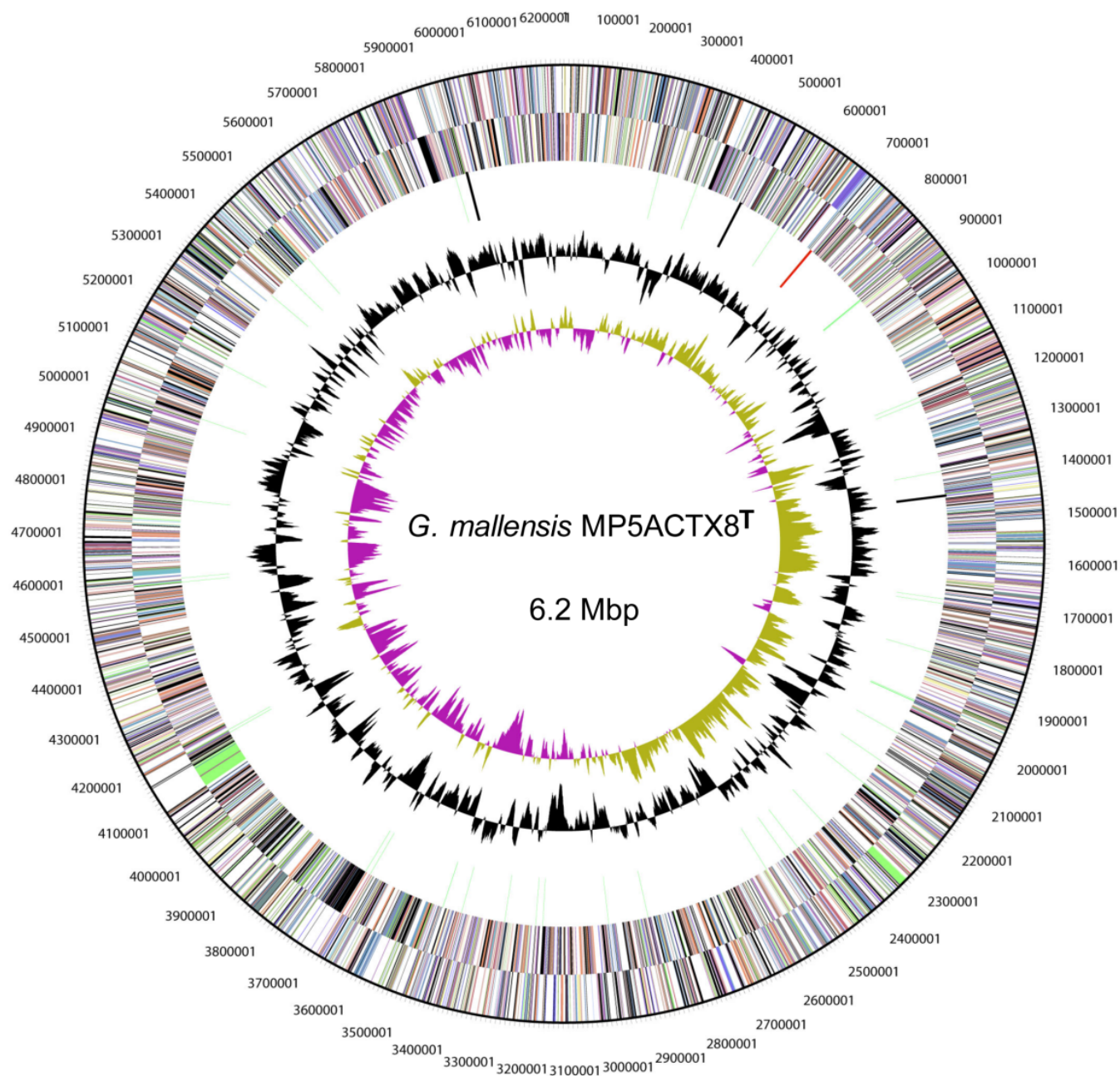


Figure 3. Circular representation of the chromosome of *G. mallensis* MP5ACTX8^T displaying relevant genome features. From outside to center; Genes on forward strand (color by COG categories), genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content and GC skew.

Table 4. Number of genes associated with general COG functional categories

Code	Value	%age	Description
J	167	4.32	Translation, ribosomal structure and biogenesis
A	2	0.05	RNA processing and modification
K	332	8.58	Transcription
L	156	4.03	Replication, recombination and repair
B	1	0.03	Chromatin structure and dynamics
D	27	0.7	Cell cycle control, Cell division, chromosome partitioning
Y	0.0	0.0	Nuclear structure
V	76	1.96	Defense mechanisms
T	139	3.59	Signal transduction mechanisms
M	322	8.32	Cell wall/membrane biogenesis
N	17	0.44	Cell motility
Z	0.0	0.0	Cytoskeleton
W	0.0	0.0	Extracellular structures
U	79	2.04	Intracellular trafficking and secretion
O	123	3.18	Posttranslational modification, protein turnover, chaperones
C	193	4.99	Energy production and conversion
G	355	9.18	Carbohydrate transport and metabolism
E	258	6.67	Amino acid transport and metabolism
F	76	1.96	Nucleotide transport and metabolism
H	155	4.01	Coenzyme transport and metabolism
I	164	4.24	Lipid transport and metabolism
P	157	4.06	Inorganic ion transport and metabolism
Q	125	3.23	Secondary metabolites biosynthesis, transport and catabolism
R	527	13.62	General function prediction only
S	418	10.8	Function unknown
-	1,464	29.52	Not in COGs

The total is based on the total number of protein coding genes in the genome.

Discussion

Granulicella mallensis type strain MP5ACTX8^T has the largest genome size of 6.2 Mbp. among the three tundra soil strains of subdivision 1 *Acidobacteria* [28]. Genome analysis of *Granulicella mallensis* identified a high abundance of genes assigned to COG functional categories for transport and metabolism of carbohydrates (9.1%) and amino acids (6.7%) and involved in cell envelope biogenesis (8.3%) and transcription (8.6%). Further genome analysis revealed an abundance of gene modules encoding for functional activities within the carbohydrate-active enzymes (CAZy) family [61] involved in breakdown, utilization and biosynthesis of carbohydrates. *G. mallensis* hydrolyzed complex carbon polymers, including CMC, pectin, lichenin, laminarin and starch, and utilized sugars such as cellobiose, D-mannose, D-xylose, D-trehalose. This parallels genome predictions for CDSs encoding

for enzymes such as cellulases, pectinases, alginate lyases, trehalase and amylases. In addition, the *G. mallensis* genome contained a cluster of genes in the neighborhood of the cellulose synthase gene (*bcsAB*) which included cellulase (*bcsZ*) (endoglucanase Y) of family GH8, cellulose synthase operon protein (*bcsC*) and a cellulose synthase operon protein (*yhjQ*) involved in cellulose biosynthesis. Detailed comparative genome analysis of *G. mallensis* MP5ACTX8^T with other *Acidobacteria* strains for which finished genomes were available is reported in Rawat et al. [28]. The data thus suggests that *G. mallensis* is involved in hydrolysis, the utilization of stored carbohydrates, and in the biosynthesis of exopolysaccharides from organic matter and plant based polymers in the soil. Therefore, we infer that strain *G. mallensis* may be central to carbon cycling processes in arctic and boreal soil ecosystems.

Acknowledgements

The work conducted by the US Department of Energy Joint Genome Institute is supported by the Office of Science of the US Department of Energy Under Contract

No. DE-AC02-05CH11231. This work was funded in part by the Academy of Finland and the New Jersey Agricultural Experiment Station.

References

- Männistö MK, Rawat S, Starovoytov V, Häggblom MM. *Granulicella arctica* sp. nov., *Granulicella mallensis* sp. nov., *Granulicella sapmiensis* sp. nov. and *Granulicella tundricola* sp. nov., novel *Acidobacteria* from tundra soil of Northern Finland. *Int J Syst Evol Microbiol* 2012; **62**:2097-2106. [PubMed](#)
<http://dx.doi.org/10.1099/ijs.0.031864-0>
- Pankratov TA, Dedysh SN. *Granulicella paludicola* gen. nov., sp. nov., *Granulicella pectinivorans* sp. nov., *Granulicella aggregans* sp. nov. and *Granulicella rosea* sp. nov., acidophilic, polymer degrading acidobacteria from Sphagnum peat bogs. *Int J Syst Evol Microbiol* 2010; **60**:2951-2959. [PubMed](#)
<http://dx.doi.org/10.1099/ijs.0.021824-0>
- Janssen PH. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol* 2006; **72**:1719-1728. [PubMed](#)
<http://dx.doi.org/10.1128/AEM.72.3.1719-1728.2006>
- Fierer N, Bradford MA, Jackson RB. Toward an ecological classification of soil bacteria. *Ecology* 2007; **88**:1354-1364. [PubMed](#)
<http://dx.doi.org/10.1890/05-1839>
- Campbell BJ, Polson SW, Hanson TE, Mack MC, Schuur EA. The effect of nutrient deposition on bacterial communities in Arctic tundra soil. *Environ Microbiol* 2010; **12**:1842-1854. [PubMed](#)
<http://dx.doi.org/10.1111/j.1462-2920.2010.02189.x>
- Chu H, Fierer N, Lauber CL, Caporaso JG, Knight R, Grogan P. Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environ Microbiol* 2010; **12**:2998-3006. [PubMed](#) <http://dx.doi.org/10.1111/j.1462-2920.2010.02277.x>
- Jones RT, Robeson MS, Lauber CL, Hamady M, Knight R, Fierer N. A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J* 2009; **3**:442-453. [PubMed](#)
<http://dx.doi.org/10.1038/ismej.2008.127>
- Barns SM, Cain EC, Sommerville L, Kuske CR. *Acidobacteria* phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum.

- Appl Environ Microbiol* 2007; **73**:3113-3116. [PubMed http://dx.doi.org/10.1128/AEM.02012-06](http://dx.doi.org/10.1128/AEM.02012-06)
9. Kishimoto N, Kosako Y, Tano T. *Acidobacterium capsulatum* gen. nov., sp. nov.: an acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. *Curr Microbiol* 1991; **22**:1-7. <http://dx.doi.org/10.1007/BF02106205>
 10. Eichorst SA, Breznak JA, Schmidt TM. Isolation and characterization of soil bacteria that define *Terriglobus* gen. nov., in the phylum *Acidobacteria*. *Appl Environ Microbiol* 2007; **73**:2708-2717. [PubMed http://dx.doi.org/10.1128/AEM.02140-06](http://dx.doi.org/10.1128/AEM.02140-06)
 11. Männistö MK, Rawat SR, Starovoytov V, Häggblom MM. *Terriglobus saanensis* sp. nov., an acidobacterium isolated from tundra soil. *Int J Syst Evol Microbiol* 2011; **61**:1823-1828. [PubMed http://dx.doi.org/10.1099/ijs.0.026005-0](http://dx.doi.org/10.1099/ijs.0.026005-0)
 12. Koch IH, Gich F, Dunfield PF, Overmann J. *Edaphobacter modestus* gen. nov., sp. nov., and *Edaphobacter aggregans* sp. nov., acidobacteria isolated from alpine and forest soils. *Int J Syst Evol Microbiol* 2008; **58**:1114-1122. [PubMed http://dx.doi.org/10.1099/ijs.0.65303-0](http://dx.doi.org/10.1099/ijs.0.65303-0)
 13. Okamura K, Kawai A, Yamada T, Hiraishi A. *Acidipila rosea* gen. nov., sp. nov., an acidophilic chemoorganotrophic bacterium belonging to the phylum *Acidobacteria*. *FEMS Microbiol Lett* 2011; **317**:138-142. [PubMed http://dx.doi.org/10.1111/j.1574-6968.2011.02224.x](http://dx.doi.org/10.1111/j.1574-6968.2011.02224.x)
 14. Pankratov TA, Kirsanova LA, Kaparullina EN, Kevbrin VV, Dedysh SN. *Telmatobacter bradus* gen. nov., sp. nov., a cellulolytic facultative anaerobe from subdivision 1 of the *Acidobacteria* and emended description of *Acidobacterium capsulatum* Kishimoto et al. *Int J Syst Evol Microbiol* 2012; **62**:430-437. [PubMed http://dx.doi.org/10.1099/ijs.0.029629-0](http://dx.doi.org/10.1099/ijs.0.029629-0)
 15. Kulichevskaya IS, Kostina LA, Valášková V, Rijpstra IC, Sinninghe Damsté JS, de Boer W, Dedysh SN. *Acidicapsa borealis* gen. nov., sp. nov. and *A. ligni* sp. nov., two novel subdivision 1 *Acidobacteria* from sphagnum peat and decaying wood. *Int J Syst Evol Microbiol* 2012; **62**:1512-1520. [PubMed http://dx.doi.org/10.1099/ijs.0.034819-0](http://dx.doi.org/10.1099/ijs.0.034819-0)
 16. Dedysh SN, Kulichevskaya IS, Serkebaeva YM, Mityaeva MA, Sorokin VV, Suzina NE, Rijpstra WI, Damsté JS. *Bryocella elongata* gen. nov., sp. nov., a novel member of Subdivision 1 of the *Acidobacteria* isolated from a methanotrophic enrichment culture, and emended description of *Edaphobacter aggregans* Koch et al. 2008. *Int J Syst Evol Microbiol* 2012; **62**:654-664. [PubMed http://dx.doi.org/10.1099/ijs.0.031898-0](http://dx.doi.org/10.1099/ijs.0.031898-0)
 17. Kulichevskaya IS, Suzina NE, Liesack W, Dedysh SN. *Bryobacter aggregatus* gen. nov., sp. nov., a peat-inhabiting, aerobic chemoorganotroph from subdivision 3 of the *Acidobacteria*. *Int J Syst Evol Microbiol* 2010; **60**:301-306. [PubMed http://dx.doi.org/10.1099/ijs.0.013250-0](http://dx.doi.org/10.1099/ijs.0.013250-0)
 18. Foessel BU, Rohde M, Overmann J. *Blastocatella fastidiosa* gen. nov., sp. nov., isolated from semi-arid savanna soil – The first described species of *Acidobacteria* subdivision 4. *Syst Appl Microbiol* 2013; **36**:82-89. [PubMed http://dx.doi.org/10.1016/j.syapm.2012.11.002](http://dx.doi.org/10.1016/j.syapm.2012.11.002)
 19. Izumi H, Nunoura T, Miyazaki M, Mino S, Toki T, Takai K, Sako Y, Sawabe T, Nakagawa S. *Thermotomaculum hydrothermale* gen. nov., sp. nov., a novel heterotrophic thermophile within the phylum *Acidobacteria* from a deep-sea hydrothermal vent chimney in the Southern Okinawa Trough. *Extremophiles* 2012; **16**:245-253. [PubMed http://dx.doi.org/10.1007/s00792-011-0425-9](http://dx.doi.org/10.1007/s00792-011-0425-9)
 20. Liesack W, Bak F, Kreft JU, Stackebrandt E. *Holophaga foetida* gen. nov., sp. nov., a new homoacetogenic bacterium degrading methoxylated aromatic compounds. *Arch Microbiol* 1994; **162**:85-90. [PubMed http://dx.doi.org/10.1007/BF00264378](http://dx.doi.org/10.1007/BF00264378)
 21. Coates JD, Ellis DJ, Gaw CV, Lovley DR. *Geothrix fermentans* gen. nov., sp. nov., a novel Fe(III)-reducing bacterium from a hydrocarbon contaminated aquifer. *Int J Syst Bacteriol* 1999; **49**:1615-1622. [PubMed http://dx.doi.org/10.1099/00207713-49-4-1615](http://dx.doi.org/10.1099/00207713-49-4-1615)
 22. Fukunaga Y, Kurahashi M, Yanagi K, Yokota A, Harayama S. *Acanthopleuribacter pedis* gen. nov., sp. nov., a marine bacterium isolated from a chiton, and description of *Acanthopleuribacteraceae* fam. nov., *Acanthopleuribacterales* ord. nov., *Holophagales* ord. nov. and *Holophagae* classis nov. in the phylum 'Acidobacteria'. *Int J Syst Evol Microbiol* 2008; **58**:2597-2601. [PubMed http://dx.doi.org/10.1099/ijs.0.65589-0](http://dx.doi.org/10.1099/ijs.0.65589-0)
 23. Ward NL, Challacombe JF, Jansen PH, Henrissat B, Coutinho PM, Wu M, Xie G, Haft DH, Sait M, Badger J, et al. Three genomes from the phylum *Acidobacteria* provide insight into the lifestyles of these microorganisms in soils. *Appl Environ*

- Microbiol* 2009; **75**:2046-2056. [PubMed](#)
<http://dx.doi.org/10.1128/AEM.02294-08>
24. Bryant DA, Amaya M, Garcia Costas AMG, Maresca JA, Chew AGM, Klatt CG, Bateson MM, Tallon LJ, Hostetler J, Nelson WC, Heidelberg JF, Ward DM. Candidatus *Chloracidobacterium thermophilum*: an aerobic phototrophic acidobacterium. *Science* 2007; **317**:523-526. [PubMed](#)
<http://dx.doi.org/10.1126/science.1143236>
 25. Männistö MK, Tiirola M, Häggblom MM. Microbial communities in Arctic fjelds of Finnish Lapland are stable but highly pH dependent. *FEMS Microbiol Ecol* 2007; **59**:452-465. [PubMed](#)
<http://dx.doi.org/10.1111/j.1574-6941.2006.00232.x>
 26. Sait M, Davis KE, Janssen PH. Effect of pH on isolation and distribution of members of subdivision 1 of the phylum *Acidobacteria* occurring in soil. *Appl Environ Microbiol* 2006; **72**:1852-1857. [PubMed](#)
<http://dx.doi.org/10.1128/AEM.72.3.1852-1857.2006>
 27. Eichorst SA, Kuske CR, Schmidt TM. Influence of plant polymers on the distribution and cultivation of bacteria in the phylum *Acidobacteria*. *Appl Environ Microbiol* 2011; **77**:586-596. [PubMed](#)
<http://dx.doi.org/10.1128/AEM.01080-10>
 28. Rawat SR, Männistö MK, Bromberg Y, Häggblom MM. Comparative genomic and physiological analysis provides insights into the role of *Acidobacteria* in organic carbon utilization in Arctic tundra soils. *FEMS Microbiol Ecol* 2012; **82**:341-355. [PubMed](#)
<http://dx.doi.org/10.1111/j.1574-6941.2012.01381.x>
 29. Rawat S, Männistö MK, Starovoytov V, Goodwin L, Nolan M, Hauser L, Land M, Davenport KW, Woyke T, Häggblom MM. Complete genome sequence of *Terriglobus saanensis* strain SP1PR4T, an *Acidobacteria* from tundra soil. *Stand Genomic Sci* 2012; **7**:59-69. [PubMed](#)
<http://dx.doi.org/10.4056/sigs.3036810>
 30. Männistö MK, Tiirola M, Häggblom MM. Effect of freeze-thaw cycles on bacterial communities of Arctic tundra soil. *Microb Ecol* 2009; **58**:621-631. [PubMed](#)
<http://dx.doi.org/10.1007/s00248-009-9516-x>
 31. Männistö MK, Kurhela E, Tiirola M, Häggblom MM. *Acidobacteria* dominate the active bacterial communities of sub-Arctic tundra with widely divergent winter-time snow accumulation and soil temperatures. *FEMS Microbiol Ecol* 2013; **84**:47-59. [PubMed](#)
<http://dx.doi.org/10.1111/1574-6941.12035>
 32. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](#)
<http://dx.doi.org/10.1038/nbt1360>
 33. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](#)
<http://dx.doi.org/10.1073/pnas.87.12.4576>
 34. Validation List No. 143. *Int J Syst Evol Microbiol* 2012; **62**:1-4.
<http://dx.doi.org/10.1099/ijs.0.039487-0>
 35. Thrash JC, Coates JD. Phylum XVII. *Acidobacteria* phyl. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 4, Springer, New York, 2011, p. 725.
 36. Judicial Commission of the International Committee on Systematics of Prokaryotes. The nomenclatural types of the orders *Acholeplasmatales*, *Halanaerobiales*, *Halobacteriales*, *Methanobacteriales*, *Methanococcales*, *Methanomicrobiales*, *Planctomycetales*, *Prochlorales*, *Sulfobobales*, *Thermococcales*, *Thermoproteales* and *Verrucomicrobiales* are the genera *Acholeplasma*, *Halanaerobium*, *Halobacterium*, *Methanobacterium*, *Methanococcus*, *Methanomicrobium*, *Planctomyces*, *Prochloron*, *Sulfobobus*, *Thermococcus*, *Thermoproteus* and *Verrucomicrobium*, respectively. Opinion 79. *Int J Syst Evol Microbiol* 2005; **55**:517-518. [PubMed](#)
<http://dx.doi.org/10.1099/ijs.0.63548-0>
 37. Cavalier-Smith T. The neomuran origin of archaeobacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int J Syst Evol Microbiol* 2002; **52**:7-76. [PubMed](#)
 38. Ludwig W, Euzéby J, Whitman WG. Draft taxonomic outline of the *Bacteroidetes*, *Planctomycetes*, *Chlamydiae*, *Spirochaetes*, *Fibrobacteres*, *Fusobacteria*, *Acidobacteria*, *Verrucomicrobia*, *Dictyoglomi*, and *Gemmatimonadetes*.
http://www.bergeys.org/outlines/Bergeys_Vol_4_Outline.pdf. Taxonomic Outline 2008.

39. Thrash JC, Coates JD. Family I. *Acidobacteriaceae* fam. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 4, Springer, New York, 2011, p. 728.
40. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29. [PubMed](#) <http://dx.doi.org/10.1038/75556>
41. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011; **28**:2731-2739. [PubMed](#) <http://dx.doi.org/10.1093/molbev/msr121>
42. Liolios K, Mavromatis K, Tavernarakis N, Kyrpides NC. The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 2007; **36**:D475-D479. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkm884>
43. DOE Joint Genome Institute. <http://www.jgi.doe.gov>.
44. Bennett S. Solexa Ltd. *Pharmacogenomics* 2004; **5**:433-438. [PubMed](#) <http://dx.doi.org/10.1517/14622416.5.4.433>
45. Margulies M, Egholm M, Altman WE. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005; **437**:376-380. [PubMed](#)
46. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 2008; **18**:821-829. [PubMed](#) <http://dx.doi.org/10.1101/gr.074492.107>
47. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 1998; **8**:175-185. [PubMed](#) <http://dx.doi.org/10.1101/gr.8.3.175>
48. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. *Genome Res* 1998; **8**:195-202. [PubMed](#) <http://dx.doi.org/10.1101/gr.8.3.195>
49. The Phred/Phrap/Consed software package. <http://www.phrap.com>
50. Han CS, Chain P. Finishing repeat regions automatically with Dupfinisher CSREA Press. In: Arabnia AR, Valafar H, editors. Proceedings of the 2006 international conference on bioinformatics & computational biology; 2006; June 26-29. CSREA Press. p 141-146.
51. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**:119. [PubMed](#) <http://dx.doi.org/10.1186/1471-2105-11-119>
52. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods* 2010; **7**:455-457. [PubMed](#) <http://dx.doi.org/10.1038/nmeth.1457>
53. Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. *Science* 1997; **278**:631-637. [PubMed](#) <http://dx.doi.org/10.1126/science.278.5338.631>
54. Clusters of Orthologous Groups. <http://www.ncbi.nlm.nih.gov/COG>.
55. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997; **25**:955-964. [PubMed](#)
56. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007; **35**:3100-3108. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkm160>
57. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. *Nucleic Acids Res* 2003; **31**:439-441. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkg006>
58. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001; **305**:567-580. [PubMed](#) <http://dx.doi.org/10.1006/jmbi.2000.4315>
59. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004; **340**:783-795. [PubMed](#) <http://dx.doi.org/10.1016/j.jmb.2004.05.028>
60. Markowitz VM, Mavromatis K, Ivanova N, Chen IM, Chu K, Kyrpides N. Expert Review of Functional Annotations for Microbial Genomes. *Bioinformatics* 2009; **25**:2271-2278. [PubMed](#) <http://dx.doi.org/10.1093/bioinformatics/btp393>
61. Carbohydrate-active enzymes. <http://www.cazy.org>