

# Complete genome sequence of *Terriglobus saanensis* type strain SP1PR4<sup>T</sup>, an *Acidobacteria* from tundra soil

Suman R. Rawat<sup>1\*</sup>, Minna K. Männistö<sup>2</sup>, Valentin Starovoytov<sup>3</sup>, Lynne Goodwin<sup>5</sup>, Matt Nolan<sup>6</sup>, Lauren Hauser<sup>4</sup>, Miriam Land<sup>4</sup>, Karen Walston Davenport<sup>5</sup>, Tanja Woyke<sup>6</sup> and Max M. Häggblom<sup>1</sup>

<sup>1</sup> Department of Biochemistry and Microbiology, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, 08901-8520, USA

<sup>2</sup> Finnish Forest Research Institute, Rovaniemi, Finland

<sup>3</sup> Department of Cell Biology and Neuroscience, Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA.

<sup>4</sup> Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

<sup>5</sup> Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA

<sup>6</sup> DOE Joint Genome Institute, Walnut Creek, California, USA

\*Corresponding author: Suman R. Rawat (rawat@aesop.rutgers.edu)

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

---

*Terriglobus saanensis* SP1PR4<sup>T</sup> is a novel species of the genus *Terriglobus*. *T. saanensis* is of ecological interest because it is a representative of the phylum *Acidobacteria*, which are dominant members of bacterial soil microbiota in Arctic ecosystems. *T. saanensis* is a cold-adapted acidophile and a versatile heterotroph utilizing a suite of simple sugars and complex polysaccharides. The genome contained an abundance of genes assigned to metabolism and transport of carbohydrates including gene modules encoding for carbohydrate-active enzyme (CAZyme) family involved in breakdown, utilization and biosynthesis of diverse structural and storage polysaccharides. *T. saanensis* SP1PR4<sup>T</sup> represents the first member of genus *Terriglobus* with a completed genome sequence, consisting of a single replicon of 5,095,226 base pairs (bp), 54 RNA genes and 4,279 protein-coding genes. We infer that the physiology and metabolic potential of *T. saanensis* is adapted to allow for resilience to the nutrient-deficient conditions and fluctuating temperatures of Arctic tundra soils.

---

## Introduction

Strain SP1PR4<sup>T</sup> (= DSM 23119 = ATCC BAA-1853) is the type strain of *Terriglobus saanensis*. It is second of two validly ascribed species of the genus *Terriglobus*, with *T. roseus* first isolated from agricultural soils in 2007 [1]. *T. saanensis* SP1PR4<sup>T</sup> was isolated from Arctic tundra soil collected from a wind exposed site of Saana fjeld, north-western Finland (69°01'N, 20°50'E) [2,3]. The species name *saanensis* (sa.a.nen' sis. N.L. masc. adj. saanensis) pertains to Mount Saana in Finland.

*Acidobacteria* are found in diverse soil environments and are widely distributed in Arctic and boreal soils [4-8]. However, relatively little is still known about their metabolic potential and

ecological roles in these habitats. Despite a large collection of *Acidobacteria* 16S rRNA gene sequences in databases that represent diverse phylotypes from various habitats, few have been cultivated and described. *Acidobacteria* represent 26 phylogenetic subdivisions based on 16S rRNA gene phylogeny [9] of which subdivisions 1, 3, 4 and 6 are most commonly detected in soil environments [10]. The abundance of *Acidobacteria* has been found to correlate with soil pH [2,10,11] and carbon [1,12,13] with subdivision 1 *Acidobacteria* being most abundant in slightly acidic soils. The phylogenetic diversity, ubiquity and abundance of this group suggest that they play important ecological roles in soils.

Our previous studies on bacterial community profiling from Arctic alpine tundra soils of northern Finland have shown that *Acidobacteria* dominate in the acidic tundra heaths [2] and after multiple freeze-thaw cycles [6]. Using selective isolation techniques, including freezing soils at -20°C for 7 days, we have been able to isolate several slow growing and fastidious strains of *Acidobacteria*. On the basis of phylogenetic, phenotypic and chemotaxonomic data, including 16S rRNA, rpoB gene sequence similarity and DNA-DNA hybridization, strain SP1PR4<sup>T</sup> was classified as a novel species of the genus *Terriglobus* [3]. Here, we summarize the physiological features together with the complete genome sequence and annotation of *Terriglobus saanensis* SP1PR4<sup>T</sup>.

## Classification and features

Within the genus *Terriglobus*, two species are ascribed with validly published names, *T. saanensis* SP1PR4<sup>T</sup> [3] isolated from Arctic tundra soils and *T. roseus* KBS 63<sup>T</sup> (DSM 18391) isolated from agricultural soils (KBS-LTER site) [1]. Searching the NCBI non-redundant nucleotide database for homology to 16S rRNA gene sequence of *T. saanensis* SP1PR4<sup>T</sup> identified 10 cultured and 20 uncultured strains that were unclassified, with ≥97% 16S rRNA sequence identity. Phylogenetic tree based on 16S rRNA gene depicting the position of *T. saanensis* SP1PR4<sup>T</sup> relative to the other type strains within the family *Acidobacteriaceae* is shown in Figure 1. *T. saanensis* SP1PR4<sup>T</sup> is distinctly clustered into a separate branch with *T. roseus* KBS 63<sup>T</sup> (DQ660892) [1], as its closest described relative (97.1% 16S rRNA sequence identity). Strain SP1PR4<sup>T</sup> showed ~95% 16S rRNA gene identity to four strains in the genus *Granulicella* isolated from tundra soils, namely "*G. tundricola*" (95.9%), "*G. sapmiensis*" (95.8%), "*G. mallensis*" (95.5%) and "*G. arctica*" (94.9%) [3,15] (Figure 1).

Strain SP1PR4<sup>T</sup> grows at pH 4.5-7.5 with an optimum at 6.0 and at temperatures of +4 to +30°C with an optimum of +25°C on R2 medium [3]. On R2 agar, strain SP1PR4<sup>T</sup> forms small, circular, convex colonies with a diameter of approximately 1 mm. The pigment varies from light beige to light pink depending on the age of the culture. Cells of

strains SP1PR4<sup>T</sup> are Gram-negative, non-spore-forming, non-motile aerobic rods with a length of 1.5– 3.0 μm and a diameter of 0.5–0.7 μm. The cell-wall structure in ultrathin sections of electron micrographs of cells of strain SP1PR4<sup>T</sup> demonstrates numerous outer-membrane vesicles (Table 1, Figure 2).

Strain SP1PR4<sup>T</sup> utilized carbon substrates for growth which include cellobiose, D-fructose, D-galactose, D-glucose, lactose, D-maltose, D-mannose, D-ribose, sucrose, D-trehalose, D-xylose, D-melezitose, D-raffinose and N-acetyl-D-glucosamine. Strain SP1PR4<sup>T</sup> hydrolyzed polysaccharides such as starch, pectin, laminarin and aesculin but not gelatin, cellulose, xylan, lichenan, sodium alginate, pullulan, chitosan or chitin. Enzyme activities of strain SP1PR4<sup>T</sup> include chitinase, catalase, acid and alkaline phosphatase, leucine arylamidase, naphthol-AS-B1-phosphohydrolase, α- and β-galactosidase, α- and β-glucosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase [3,15].

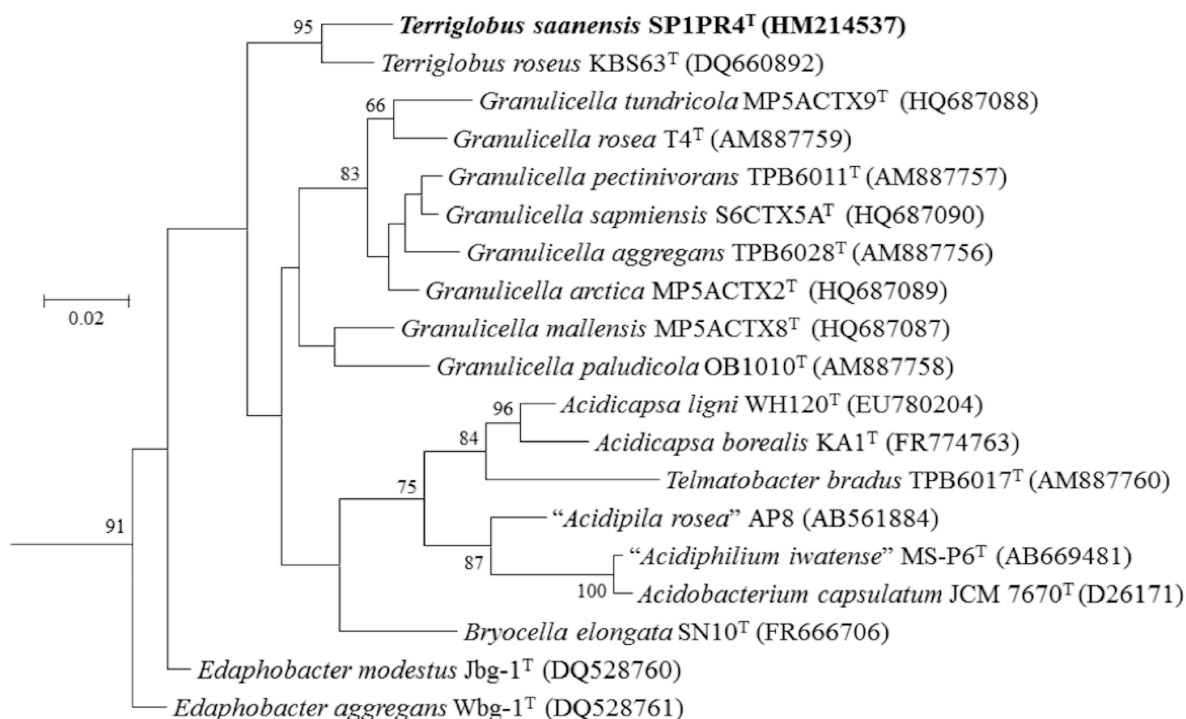
## Chemotaxonomy

The major cellular fatty acids in *T. saanensis* SP1PR4<sup>T</sup> are iso-C<sub>15:0</sub> (39.9%), C<sub>16:1 ω7c</sub> (28.4%), iso-C<sub>13:0</sub> (9.8%) and C<sub>16:0</sub> (9.8%). The cellular fatty acid compositions of strain SP1PR4<sup>T</sup> were relatively similar to that of *T. roseus* DSM 18391<sup>T</sup>, with higher relative abundance of iso-C<sub>13:0</sub> and a corresponding lower abundance of iso-C<sub>15:0</sub> in strain SP1PR4<sup>T</sup> [3].

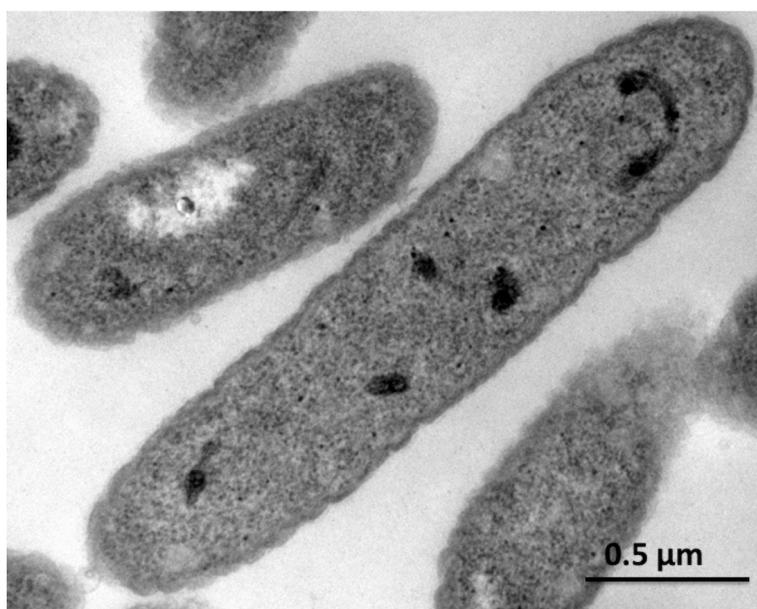
## Genome sequencing and annotation

### Genome project history

Strain SP1PR4<sup>T</sup> 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 August 6, 2010. The complete genome was made available on Jan 24, 2011. The genome project is deposited in the Genomes On-Line Database (GOLD) [25] and the complete genome sequence of strain SP1PR4<sup>T</sup> is deposited in GenBank. Table 2 presents the project information and its association with MIGS version 2.0 [16].



**Figure 1.** Phylogenetic tree highlighting the position of *T. saanensis* SP1PR4<sup>T</sup> relative to the other type strains within the family *Acidobacteriaceae*. The maximum likelihood tree was inferred from 1,359 aligned positions of the 16S rRNA gene sequences and derived using MEGA version 5 [14]. Bootstrap values (expressed as percentages of 1,000 replicates) of >50 are shown at branch points. Bar: 0.02 substitutions per nucleotide position. The strains (type strain=T) and their corresponding GenBank accession numbers are displayed in parentheses with strain *T. saanensis* SP1PR4<sup>T</sup> shown in bold. *Bryobacter aggregatus* MPL3 (AM162405) was used as outgroup. *T. saanensis* SP1PR4<sup>T</sup> and *T. roseus* KBS 63<sup>T</sup> (DSM 18391) genome sequences have been revealed.



**Figure 2.** Electron micrograph of cells of *T. saanensis* strain SP1PR4<sup>T</sup> (bar 0.5 μm).

**Table 1.** Classification and general features of *T. saanensis* SP1PR4<sup>T</sup> according to the MIGS recommendations [16].

MIGS ID	Property	Term	Evidence codes
		Domain <i>Bacteria</i>	TAS [17]
		Phylum <i>Acidobacteria</i>	TAS [18,19]
		Class <i>Acidobacteria</i>	TAS [20]
	Classification	Order <i>Acidobacteriales</i>	TAS [21,22]
		Family <i>Acidobacteriaceae</i>	TAS [18,23]
		Genus <i>Terriglobus</i>	TAS [1]
		Species <i>Terriglobus saanensis</i>	TAS [3]
		Type strain: SP1PR4 <sup>T</sup>	
	Gram stain	negative	TAS [3]
	Cell shape	rod	TAS [3]
	Motility	non-motile	TAS [3]
	Sporulation	non-spore forming	TAS [3]
	Temperature range	4–30°C	TAS [3]
	Optimum temperature	25°C	TAS [3]
	pH range	4.5-7.5	TAS [3]
	Optimum pH	6.0	TAS [3]
	Salinity	not reported	NAS
MIGS-22	Oxygen requirement	aerobe	TAS [3]
	Carbon source	cellobiose, D-fructose, D-galactose, D-glucose, lactose, D-maltose, D-mannose, D-ribose, sucrose, D-trehalose, D-xylose, D-melezitose, D-raffinose, starch, pectin, laminarin and aesculin	TAS [3]
MIGS-6	Habitat	terrestrial	TAS [3]
MIGS-15	Biotic relationship	free-living	TAS [3]
MIGS-14	Pathogenicity	non-pathogen	NAS
	Biosafety level	1	NAS
	Isolation	tundra soil	TAS [3]
MIGS-4	Geographic location	Saana fjeld, Arctic tundra, Finland	TAS [3]
MIGS-5	Sample collection time	2004-2005	TAS [3]
MIGS-4.1	Latitude	69°01'N,	TAS [3]
MIGS-4.2	Longitude	20°50'E	TAS [3]
MIGS-4.3	Depth	not reported	NAS
MIGS-4.4	Altitude	not reported	NAS

\*Evidence codes - IDA: Inferred from Direct Assay (first time in publication); 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 [24].

**Table 2.** Genome sequencing 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	Sequencing coverage	39× (454), 180× (Illumina)
MIGS 30	Assemblers	Newbler, Velvet, Phrap
MIGS 32	Gene calling method	Prodigal, GenePRIMP
	Locud Tag	AcIPR4
	INSDC / RefSeq ID	CP002467, NC_014963,
	GenBank Date of Release	October 7, 2011
	GOLD ID	Gc01604
	NCBI project ID	48971
MIGS 13	Source material identifier	ATCC BAA-1853, DSM 23119
	Project relevance	Environmental, Biogeochemical cycling of carbon, Biotechnological, GEBA

### Growth conditions and genomic DNA extraction

Strain SP1PR4<sup>T</sup> was cultivated in R2 medium as previously described [3]. 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.

### Genome sequencing and assembly

The finished genome of *T. saanensis* SP1PR4<sup>T</sup> (JGI ID 4088690) was generated at the DOE Joint Genome Institute (JGI) using a combination of Illumina [26] and 454 technologies [27]. For this genome, an Illumina GAii shotgun library which generated 23,685,130 reads totaling 916 Mb, a 454 Titanium standard library which generated 409,633 reads and a paired end 454 library with an average insert size of 10.8 kb which generated 180,451 reads totaling 157 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 [28]. The 454 Titanium standard data and the 454 paired end data were assembled together with Newbler, version 2.3. Illumina sequencing data was assembled with Velvet, version 0.7.63 [29]. We integrated the 454 Newbler consensus shreds, the Illumina Velvet consensus shreds and the read pairs in the 454 paired end library using parallel phrap, version SPS - 4.24 (High Performance Software, LLC). The software Consed [30,31] was used in the 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 mis-assemblies were corrected using gapResolution (Cliff Han, unpublished), Dupfinisher [32], 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 157 Mb of 454 data which provides an average 39× coverage and 916 Mb of Illumina data which provides an average 180× coverage of the genome.

### Genome annotation

Genes were identified using Prodigal [33] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [34]. 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) [35,36], 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 [37], RNAMMer [38], Rfam [39], TMHMM [40], and signalP [41]. Additional gene prediction analysis and functional annotation were performed within the Integrated Microbial Genomes Expert Review (IMG-ER) platform [42].

## Genome properties

The genome consists of one circular chromosome of 5,095,226 bp in size with a GC content of 57.3% and consists of 54 RNA genes (Figure 3, Table 3). Of the 4,333 predicted genes, 4,279 are protein-coding genes (CDSs) and 99 are pseudogenes. Of the total CDSs, 67% represent COG functional categories and 43% consist of signal peptides. The distribution of genes into COG functional categories is presented in Figure 3 and Table 4.

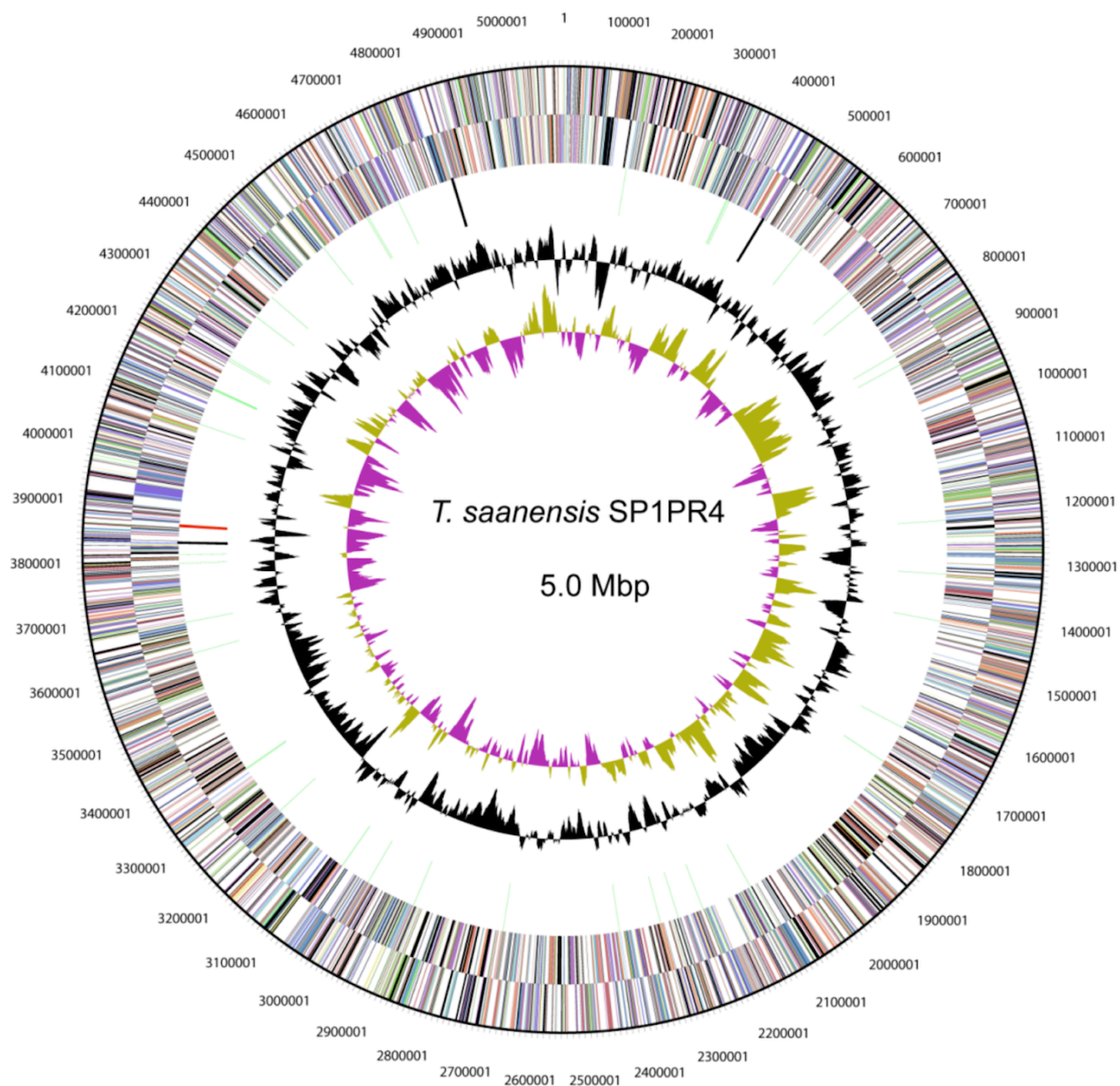
## Discussion

Genome analysis of *T. saanensis* identified a high abundance of genes assigned to COG functional categories for transport and metabolism carbohydrates (9.5%) and amino acids (7.6%), energy conversion (6.2%), cell envelope biogenesis (9.6%) and transcription (9.2%) [15]. This indicates that the *T. saanensis* genome encodes for functions involved in transport and utilization of nutrients, mainly carbohydrates and amino acids for energy production and cell biogenesis to maintain cell integrity in cold tundra soils. Further genome analysis revealed an abundance of gene modules for glycoside hydrolases, glycosyl transferases, polysaccharide lyases, carbohydrate esterases, and non-catalytic carbohydrate-binding modules within the carbohydrate-active enzymes (CAZy [43]) family involved in breakdown, utilization and biosynthesis of carbohydrates [15]. *T. saanensis* hydrolyzed complex carbon polymers, including pectin, laminarin, and starch,

and utilized sugars such as cellobiose, D-mannose, D-xylose, D-trehalose and laminarin. This parallels genome predictions for CDSs encoding for enzymes such as pectinases, chitinases, alginate lyases, trehalase and amylases. *T. saanensis* was unable to hydrolyze carboxymethyl cellulose (CMC) on plate assays and lacked CDSs encoding for cellulases involved in cellulose hydrolysis. However, the *T. saanensis* genome contained a BcsZ gene encoding for an endocellulase (GH8) as part of a bacterial cellulose synthesis (*bcs*) operon involved in cellulose biosynthesis in several species. This operon consists of clusters of genes in close proximity to the BcsZ gene which includes a cellulose synthase gene (*bcsAB*), a cellulose synthase operon protein (*bcsC*) and a cellulose synthase operon protein (*yhj*) [15]. In addition, the *T. saanensis* genome encoded for a large number of gene modules representing glycosyl transferases (GTs) involved in carbohydrate biosynthesis which include cellulose synthase (UDP-forming),  $\alpha$ -trehalose phosphate synthase [UDP-forming], starch glucosyl transferase, ceramide  $\beta$ -glucosyltransferase involved in biosynthesis of cellulose, trehalose, starch, hopanoid, and capsular/free exopolysaccharide (EPS) [15]. This suggests that *T. saanensis* is involved in hydrolysis of lignocellulosic soil organic matter, utilization of stored carbohydrates and biosynthesis of exopolysaccharides. Therefore, we surmise that *T. saanensis* may be central to carbon cycling processes in Arctic and boreal soil ecosystems.

**Table 3.** Genome statistics

Attribute	Value	% of Total
Genome size (bp)	5,095,226	100%
DNA coding (bp)	4,578,206	89.9%
DNA G+C (bp)	2,921,371	57.3%
Number of replicons	1	100%
Total genes	4,334	100%
RNA genes	54	1.3%
rRNA operons	1	-
Protein coding genes	4,180	98.8%
Pseudo genes	99	2.3%
Genes with function prediction	3,203	73.9%
Genes in paralog clusters	2,220	51.2%
Genes assigned to COGs	3,170	73.2%
Genes with Pfam domains	3,108	71.7%
Genes with signal peptides	1,867	43.1%
Genes with transmembrane helices	1,082	25%
CRISPR repeats	0	-



**Figure 3.** Graphical representation of circular map of the chromosome of *T. saanensis* strain SP1PR4<sup>T</sup> 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, GC skew.

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

Code	Value	%age	Description
J	163.0	4.6	Translation, ribosomal structure and biogenesis
A	2.0	0.1	RNA processing and modificatin
K	293.0	8.3	Transcription
L	142.0	4.0	Replication, recombination and repair
B	0.0	0.0	Chromatin structure and dynamics
D	24.0	0.7	Cell cycle control, Cell division, chromosome partitioning
Y	0.0	0.0	Nuclear structure
V	98.0	2.8	Defense mechanisms
T	174.0	4.9	Signal transduction mechanisms
M	307.0	8.7	Cell wall/membrane biogenesis
N	56.0	1.6	Cell motility
Z	2.0	0.1	Cytoskeleton
W	0.0	0.0	Extracellular structures
U	113.0	3.2	Intracellular trafficking and secretion
O	122.0	3.4	Posttranslational modification, protein turnover, chaperones
C	196.0	5.5	Energy production and conversion
G	303.0	8.6	Carbohydrate transport and metabolism
E	243.0	6.9	Amino acid transport and metabolism
F	69.0	2.0	Nucleotide transport and metabolism
H	134.0	3.8	Coenzyme transport and metabolism
I	116.0	3.3	Lipid transport and metabolism
P	134.0	3.8	Inorganic ion transport and metabolism
Q	85.0	2.4	Secondary metabolites biosynthesis, transport and catabolism
R	443.0	12.5	General function prediction only
S	323.0	9.1	Function unknown
-	1163.0	26.8	Not in COGs

## 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

1. Espinosa E, Marco-Noales E, Gomez D, Lucas-Elío P, Ordax M, Garcias-Bonet N, Duarte CM, Sanchez-Amat A. Taxonomic study of *Marinomonas* strains isolated from the seagrass *Posidonia oceanica*, with descriptions of *Marinomonas balearica* sp. nov. and *Marinomonas pollencensis* sp. nov. *Int J Syst Evol Microbiol* 2010; **60**:93-98. [PubMed](http://dx.doi.org/10.1099/ijs.0.008607-0) <http://dx.doi.org/10.1099/ijs.0.008607-0>
2. Lucas-Elío P, Marco-Noales E, Espinosa E, Ordax M, Lopez MM, Garcias-Bonet N, Marba N, Duarte CM, Sanchez-Amat A. *Marinomonas alcarazii* sp. nov., *M. rhizomae* sp. nov., *M. foliarum* sp. nov., *M. posidonica* sp. nov. and *M. aquiplantarum* sp. nov., isolated from the microbiota of the seagrass *Posidonia oceanica*. *Int J Syst Evol Microbiol* 2011; **61**:2191-2196. [PubMed](http://dx.doi.org/10.1099/ijs.0.027227-0) <http://dx.doi.org/10.1099/ijs.0.027227-0>



3. Arnaud-Haond S, Duarte CM, Díaz-Almela E, Marbà N, Sintes T, Serrão EA. Implications of extreme life span in clonal organisms: millenary clones in meadows of the threatened seagrass *Posidonia oceanica*. *PLoS ONE* 2012; **7**:e30454. [PubMed](#)  
<http://dx.doi.org/10.1371/journal.pone.0030454>
4. Celdran D, Espinosa E, Sanchez-Amat A, Atucha A. Effects of epibiotic bacteria on leaf growth and epiphytes of seagrass *Posidonia oceanica* *Mar Ecol Prog Ser* **456**: 21-27.  
<http://dx.doi.org/10.3354/meps09672>
5. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007; **24**:1596-1599. [PubMed](#)  
<http://dx.doi.org/10.1093/molbev/msm092>
6. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. I, Wilson G, and Wipat A: The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](#)  
<http://dx.doi.org/10.1038/nbt1360>
7. 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>
8. Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 1.
9. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. List no. 106. *Int J Syst Evol Microbiol* 2005; **55**:2235-2238.  
<http://dx.doi.org/10.1099/ijs.0.64108-0>
10. Garrity GM, Bell JA, Lilburn T. Class III. *Gammaproteobacteria* class. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 1.
11. Garrity GM, Bell JA, Lilburn T. Order VIII. *Oceanospirillales* ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 270.
12. Garrity GM, Bell JA, Lilburn T. Family I. *Oceanospirillaceae* fam. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 271.
13. Validation List no. 13. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. *Int J Syst Bacteriol* 1984; **34**:91-92.  
<http://dx.doi.org/10.1099/00207713-34-1-91>
14. van Landschoot A, De Ley J. Intra- and intergeneric similarities of the rRNA cistrons of *Alteromonas*, *Marinomonas*, gen. nov. and some other Gram-negative bacteria. *J Gen Microbiol* 1983; **129**:3057-3074.
15. Espinosa E, Marco-Noales E, Gómez D, Lucas-Elío P, Ordax M, Garcías-Bonet N, Duarte CM, Sanchez-Amat A. Taxonomic study of *Marinomonas* strains isolated from the seagrass *Posidonia oceanica*, with descriptions of *Marinomonas balearica* sp. nov. and *Marinomonas pollencensis* sp. nov. *Int J Syst Evol Microbiol* 2010; **60**:93-98. [PubMed](#)  
<http://dx.doi.org/10.1099/ijs.0.008607-0>
16. 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>
17. López-Serrano D, Solano F, Sanchez-Amat A. Identification of an operon involved in tyrosinase activity and melanin synthesis in *Marinomonas mediterranea*. *Gene* 2004; **342**:179-187. [PubMed](#)  
<http://dx.doi.org/10.1016/j.gene.2004.08.003>
18. Sanchez-Amat A, Lucas-Elio P, Fernández E, Garcia-Borrón JC, Solano F. Molecular cloning and functional characterization of a unique multipotent polyphenol oxidase from *Marinomonas mediterranea*. *Biochim Biophys Acta* 2001; **1547**:104-116. [PubMed](#)  
[http://dx.doi.org/10.1016/S0167-4838\(01\)00174-1](http://dx.doi.org/10.1016/S0167-4838(01)00174-1)
19. Curson ARJ, Todd JD, Sullivan MJ, Johnston AWB. Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nat Rev Microbiol* 2011; **9**:849-859. [PubMed](#)  
<http://dx.doi.org/10.1038/nrmicro2653>
20. Todd JD, Rogers R, Li YG, Wexler M, Bond PL, Sun L, Curson AR, Malin G, Steinke M, Johnston AWB. Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria. *Science*

- 2007; **315**:666-669. [PubMed](#)  
<http://dx.doi.org/10.1126/science.1135370>
21. Lucas-Elío P, Goodwin L, Woyke T, Pitluck S, Nolan M, Kyrpides N, Detter JC, Copeland A, Teshima H, Bruce D, *et al.* Complete genome sequence of the melanogenic marine bacterium *Marinomonas mediterranea* type strain (MMB-1<sup>T</sup>). *Stand Genomic Sci* 2012; **6**:63-73. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.2545743>
  22. Fernández E, Sanchez-Amat A, Solano F. Location and catalytic characteristics of a multipotent bacterial polyphenol oxidase. *Pigment Cell Res* 1999; **12**:331-339. [PubMed](#)  
<http://dx.doi.org/10.1111/j.1600-0749.1999.tb00767.x>
  23. Bennett S. Solexa Ltd. *Pharmacogenomics* 2004; **5**:433-438. [PubMed](#)  
<http://dx.doi.org/10.1517/14622416.5.4.433>
  24. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005; **437**:376-380. [PubMed](#)
  25. DOE Joint Genome Institute.  
<http://www.jgi.doe.gov>
  26. 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>
  27. 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](#)
  28. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 1998; **8**:186-194. [PubMed](#)
  29. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. *Genome Res* 1998; **8**:195-202. [PubMed](#)
  30. Han C, Chain P. Finishing repeat regions automatically with Dupfinisher. Arabnia, H. R. and Valafar, H. Proceeding of the 2006 International Conference on bioinformatics and computational biology, 141-146. 2006. CSREA Press.
  31. 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>
  32. 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>
  33. 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](#)
  34. 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>
  35. 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>
  36. 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>
  37. Bendtsen JD, Nielsen H, von Heinje 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>
  38. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 2009; **25**:2271-2278. [PubMed](#)  
<http://dx.doi.org/10.1093/bioinformatics/btp393>
  39. French-Constant RH, Dowling A, Waterfield NR. Insecticidal toxins from *Photorhabdus* bacteria and their potential use in agriculture. *Toxicon* 2007; **49**:436-451. [PubMed](#)  
<http://dx.doi.org/10.1016/j.toxicon.2006.11.019>
  40. Waterfield NR, Bowen DJ, Fetherston JD, Perry RD, French-Constant RH. The *tc* genes of *Photorhabdus*: a growing family. *Trends Microbiol* 2001; **9**:185-191. [PubMed](#)  
[http://dx.doi.org/10.1016/S0966-842X\(01\)01978-3](http://dx.doi.org/10.1016/S0966-842X(01)01978-3)
  41. Barth H, Aktories K. New insights into the mode of action of the actin ADP-ribosylating virulence factors *Salmonella enterica* SpvB and *Clostridium botulinum* C2 toxin. *Eur J Cell Biol* 2011; **90**:944-950. [PubMed](#)  
<http://dx.doi.org/10.1016/j.ejcb.2010.11.007>

42. Jackson AP, Thomas GH, Parkhill J, Thomson NR. Evolutionary diversification of an ancient gene family (*rhs*) through C-terminal displacement. *BMC Genomics* 2009; **10**:584. [PubMed](#)  
<http://dx.doi.org/10.1186/1471-2164-10-584>
43. Lin RJ, Capage M, Hill CW. A repetitive DNA sequence, *rhs*, responsible for duplications within the *Escherichia coli* K-12 chromosome. *J Mol Biol* 1984; **177**:1-18. [PubMed](#)  
[http://dx.doi.org/10.1016/0022-2836\(84\)90054-8](http://dx.doi.org/10.1016/0022-2836(84)90054-8)
44. Poole SJ, Diner EJ, Aoki SK, Braaten BA, t'Kint dR, Low DA, and Hayes CS. Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. *PLoS Genet* 2011; **7**:e1002217. [PubMed](#)  
<http://dx.doi.org/10.1371/journal.pgen.1002217>
45. Polz MF, Hunt DE, Preheim SP, Weinreich DM. Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philos Trans R Soc Lond B Biol Sci* 2006; **361**:2009-2021. [PubMed](#)  
<http://dx.doi.org/10.1098/rstb.2006.1928>
46. Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S, DeMaere MZ, Ting L, Ertan H, Johnson J, et al. The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci USA* 2009; **106**:15527-15533. [PubMed](#)  
<http://dx.doi.org/10.1073/pnas.0903507106>