Genome sequence of strain HIMB624, a cultured representative from the OM43 clade of marine Betaproteobacteria

Megan Huggett
Darren Hayakawa
Michael Rappe

Follow this and additional works at: https://ro.ecu.edu.au/ecuworks2012

Part of the Genetics and Genomics Commons

10.4056/sigs.2305090
Huggett, M. J., Hayakawa, D., & Rappe, M. (2012). Genome sequence of strain HIMB624, a cultured representative from the OM43 clade of marine Betaproteobacteria. Standards in Genomic Sciences, 6(1), 1-15.
This Journal Article is posted at Research Online.
https://ro.ecu.edu.au/ecuworks2012/685
Genome sequence of strain HIMB624, a cultured representative from the OM43 clade of marine Betaproteobacteria

Megan J. Huggett1, Darin H. Hayakawa1,2 and Michael S. Rappé1*

1 Hawaii Institute of Marine Biology, SOEST, University of Hawaii, Kaneohe, Hawaii, USA
2 Department of Microbiology, University of Hawaii, Honolulu, Hawaii, USA

*Corresponding author: Michael S. Rappé (rappe@hawaii.edu)

Keywords: methylotroph, methanol, OM43, marine, bacterioplankton, Methylophilaceae

Introduction

Strain HIMB624 was isolated from surface seawater of Kaneohe Bay, a subtropical bay on the northeastern shore of Oahu, Hawaii, via dilution to extinction culturing methods [1,2]. This strain is of interest because it belongs to a globally ubiquitous clade of aquatic bacterioplankton known as OM43, within the obligately methylotrophic family Methylophilaceae of the class Betaproteobacteria. The OM43 lineage was first described in 1997 from a 16S rRNA gene survey of coastal bacterioplankton from the Atlantic coast of the United States [3], and the first published report describing the isolation of OM43 strains via modified extinction to dilution culturing methods was reported in 2002 [1]. Recently, the genome sequence of a member of the OM43 lineage was reported for a strain isolated from the Pacific coast of the United States (HTCC2181) [4]. Here we present a preliminary set of features for strain HIMB624 (Table 1), together with a description of the genomic sequencing and annotation, as well as a preliminary comparative analysis with the genome of strain HTCC2181.

Classification and features

Strain HIMB624 was isolated from seawater collected off of the coast of Hawaii, USA, in the subtropical North Pacific Ocean by a high throughput, dilution-to-extinction approach [1,2]. The strain was re-grown in seawater that was sterilized by tangential flow filtration and by autoclaving. Attempts to cultivate cells on solidified seawater media or artificial seawater media (liquid or solidified) failed. However, amendment of sterile seawater with either methanol or formaldehyde increased the maximum cell density from ca. 1×10^6 cells ml^{-1} to ca. 1×10^7 cells ml^{-1}. Phylogenetic analyses based on 16S rRNA gene sequence comparisons revealed strain HIMB624 to be closely related to a large number of environmental gene clones obtained predominantly from seawater. Alignment of the HIMB624 16S rRNA gene sequence with the Silva release 104 reference database containing only high quality, aligned 16S rRNA sequences with a minimum length of 1,200 bases for Bacteria released in October 2010 (n=512,037 entries) [13], revealed 350 entries that belong to the same phylogenetic lineage within the Betaproteobacteria. Of these, only the entries from HTCC2181, HIMB624 and one other strain...
strain HIMB624 (AB022337) originated from cultivated isolates and all entries in the lineage were derived either from seawater, freshwater, or the marine environment. In phylogenetic analyses with taxonomically described members of the Betaproteobacteria, strains HIMB624 and HTCC2181 formed a monophyletic lineage within the family Methylophilaceae (Figure 1; 96.5% sequence similarity). The 16S rRNA gene of strain HIMB624 was most similar to the type strains of Methylophilus luteus strain Mim (94.4%) and Methylophilus flavus strain Ship (94.3%), both isolated from plants [18]; Methylophilus methylotrophus strain NCIMB 10515 (93.7%), isolated from activated sludge [19]; Methylophilus mobilis strain J6W8 (93.7%), isolated from freshwater sediment [20]; Methylobacillus flagellatus strain KT (93.5%) isolated from sewage [21]; Methylovorus mays strain C isolated from maize phyllosphere (92.5%) [22]; and Methylobacillus pratensis strain F31 (91.8%), isolated from meadow grass [23].

Table 1. Classification and general features of strain HIMB624 according to the MIGS recommendations [5].

| MIGS ID | Property                  | Term                           | Evidence code |
|---------|---------------------------|--------------------------------|---------------|
|         | Domain                     | Bacteria                       | TAS [6]       |
|         | Phylum                     | Proteobacteria                 | TAS [7]       |
|         | Class                      | Betaproteobacteria             | TAS [8,9]     |
|         | Order                      | Methylophilales                | TAS [8,10]    |
| Current classification | Family                    | Methylophilaceae               | TAS [8,11]    |
|         | Genus not assigned         |                                | IDA           |
|         | Species not assigned       |                                | IDA           |
|         | Strain HIMB624             |                                | IDA           |
|         | Gram stain                 | negative                       | IDA           |
|         | Cell shape                 | rod-shaped                     | IDA           |
|         | Motility                   | unknown                        |               |
|         | Sporulation                | non-sporulating                | NAS           |
|         | Temperature range           | mesophilic                     | IDA           |
|         | Optimum temperature        | 27-29°C                       | IDA           |
|         | Carbon source              | methanol, formaldehyde        | IDA           |
|         | Energy source              | chemoorganoheterotrophic      | IDA           |
|         | Habitat                    | sea water                      | IDA           |
| MIGS-6  | Salinity                   | ~35.0 %                       | NAS           |
| MIGS-6.3| Oxygen                     | aerobic                        | IDA           |
| MIGS-22 | Biotic relationship        | free-living                    | IDA           |
| MIGS-15 | Pathogenicity              | none                           | NAS           |
| MIGS-4  | Geographic location        | Kaneohe Bay, Hawaii, subtropical Pacific Ocean | IDA |
| MIGS-5  | Sample collection time     | 15 March 2004                  | IDA           |
| MIGS-4.1| Latitude                   | 21.44                          | IDA           |
| MIGS-4.2| Longitude                  | -157.78                        | IDA           |
| MIGS-4.3| Depth                      | ~1 m                           | IDA           |

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 [12]. If the evidence code is IDA, then the property was directly observed by one of the authors or an expert mentioned in the acknowledgements.
Figure 1. Phylogenetic tree based comparisons between 16S rRNA gene sequences from strain HIMB624, strain HTCC2181, type strains of related species within the family Methylophilaceae, and more distantly related Betaproteobacteria. Several Gammaproteobacteria and Alphaproteobacteria strains were used as outgroups. Sequence selection and alignment improvements were carried out using the ‘All-Species Living Tree’ project database [14] and the ARB software package [15]. The tree was inferred from 1,223 alignment positions using the RAxML maximum likelihood method [16]. Bootstrap support values, determined by RAxML [17], are displayed above branches if larger than 60% from 1000 replicates. The scale bar indicates substitutions per site.

In actively growing cultures, cells of strain HIMB624 are long, thin slightly curved rods between 0.1-0.3 μm wide and 0.6-1.8 μm long (Figure 2). Cells in stationary phase are spherical and approximately 0.2 μm in diameter. Strain HIMB624 can replicate in sterile unamended seawater, reaching cell densities of approximately 1×10⁶ cells ml⁻¹. However, in the presence of either methanol or formaldehyde, HIMB624 can achieve a significantly higher growth rate and cellular abundance, similar to the phylogenetically related strain HTCC2181 [4].

Chemotaxonomy

The fatty acid profile of strain HIMB624 was dominated by anteiso-C₁₇:₁, C₁₄:₀ and C₁₆:₀. This is similar to known obligate and restricted facultative methylotrophs within the Betaproteobacteria, which are typically dominated by anteiso-C₁₇:₁ and C₁₆:₀ [20]. All of the fatty acids detected in strain HIMB624 are either found in closely related strains or in strains isolated from marine environments. C₁₃:₀-2:OH was detected in HIMB624 but not in HTCC2181, and C₁₅:₁ iso G was only found in strain HTCC2181.

Genome sequencing and annotation

Genome project history

Strain HIMB624 was selected for whole genome sequencing because of its phylogenetic affiliation with a lineage (OM43) of coastal marine bacterioplankton that is common in 16S rRNA gene surveys of coastal and estuarine systems [24], but is underrepresented in culture collections [1,4]. In addition, a sister lineage is common in freshwater systems [24]. The respective genome project is deposited in the Genomes OnLine Database (GOLD) as project Gi02451, and in GenBank under the accession number ABXG00000000. A summary of the main project is given in Table 2.

Growth conditions and DNA isolation

Strain HIMB624 was grown at 27°C in 100 L of coastal Hawaii seawater sterilized by tangential flow filtration and autoclaving. Cells from liquid culture were collected on a 0.1 μm pore-sized polyethersulfone membrane filter, and DNA was isolated from the microbial biomass using a standard phenol/chloroform/isooamyl alcohol extraction protocol. A total of 74 μg of DNA was obtained.
strain HIMB624

Figure 2. Scanning electron micrograph of strain HIMB624 during exponential phase of growth. Scale bar corresponds to 0.5 μm.

| MIGS ID | Property                          | Term                                      |
|---------|-----------------------------------|-------------------------------------------|
| MIGS-31 | Finishing quality                 | Final draft                               |
| MIGS-28 | Libraries used                    | Sanger (one each of 1-4 and 10-12 kbp inserts) |
| MIGS-29 | Sequencing platforms              | ABI 3730XL                                |
| MIGS-31.2| Sequencing coverage              | 19.78×                                    |
| MIGS-30 | Assemblers                        | Celera Assembler30                         |
| MIGS-32 | Gene calling method               | Glimmer                                   |
|         | INSDC ID                          | 13602                                     |
|         | Genbank Date of Release           | 17 March 2008                              |
|         | GOLD ID                           | Gi02451                                   |
|         | NCBI taxon ID                     | 314607                                    |
|         | Database: IMG                     | 2503283018                                |
| MIGS-13 | Source material identifier        | HIMB624                                   |
|         | Project relevance                 | environmental                             |

**Genome sequencing and assembly**
The genome of strain HIMB624 was sequenced by the J. Craig Venter Institute (Rockville, MD) as part of the Gordon and Betty Moore Foundation Marine Microbial Genome Sequencing Project. Two genomic libraries of insert sizes of 1-4 and 10-12 kb were constructed [25]. Clones were sequenced from both ends on ABI 3730XL DNA sequencers (Applied Biosystems, Carlsbad, CA) at the JCVI Joint Technology Center to provide paired-end reads. A total of 27,957 reads with average read length of 943 bp were assembled using the Celera Assembler30, resulting in four contigs of 1,272; 146,687; 709,553 and 474,927 bp in length. Sequencing provided 19.78× coverage of the genome.

**Genome annotation**
The whole genome sequence was automatically annotated using the genome annotation pipeline in the Integrated Microbial Genomes Expert Review (IMG-ER) system [26]. Genes were identified using Glimmer [27]. The predicted CDSs
were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. The tRNAscanSE tool [28] was used to find tRNA genes, whereas ribosomal RNAs were found by using the tool RNAmmer [29]. Other non-coding RNAs were identified by searching the genome for the Rfam profiles using INFERNAL (v0.81) [30]. Additional gene prediction analysis and manual functional annotation was performed within IMG-ER.

**Genome properties**

The genome is 1,333,209 bp long and comprises four contigs in a single scaffold, with an overall GC content of 35.37% (Table 3 and Figure 3). Of the 1,420 genes predicted, 1,381 were protein-coding genes and 39 were RNAs. The majority (83.59%) of the protein coding genes was assigned with a putative function, while the remaining genes were annotated as hypothetical proteins. The distribution of genes into COGS functional categories is presented in Table 4.

**Insights from the Genome**

Of 1,381 protein encoding genes in the genome of HIMB624, 1,135 are shared with HTCC2181, representing 82-84% of the two genomes (Figure 4). Pathways for the synthesis of all twenty amino acids are present in both strains, as well as for the synthesis of all major vitamins except B12. The family *Methylophilaceae* consists of obligate methylotrophs and, while HIMB624 and HTCC2181 lack genes coding for either the large (mxaF) or small (mxaI) subunit of a confirmed methanol dehydrogenase, both organisms appear to have genes coding for a related analog of mxaF, known as xoxF. Methanol dehydrogenase activity of this paralog has been questioned for some time (see [4] and references therein), but current evidence suggests that the xoxF genes in these organisms code for a large subunit having methanol dehydrogenase activity [4]. The xoxF gene in HIMB624 is 87.4% similar in protein sequence to the xoxF gene in HTCC2181. Strains HTCC2181 and HIMB624 also have many of the other subunits required to form a methanol dehydrogenase holoenzyme including mxaA,C,D,E,G,J,K,R,L and S, and operons *pqBCD* or *pqBCDEFG*. Neither strain possesses genes coding for the E1 subunit (*sucA*, EC:1.2.4.2) of the α-ketoglutarate dehydrogenase complex, though they do appear to possess the E2 subunit (*sucB*, EC: 2.3.1.61). Both subunits are required to complete the tricarboxylic acid (TCA) cycle, and the absence of the E1 subunit suggests that these strains are obligate methylotrophs.

| Attribute                        | Value  | % of total |
|----------------------------------|--------|------------|
| Genome size (bp)                 | 1,333,209 | 100.00    |
| DNA coding region (bp)           | 1,284,895 | 96.38     |
| DNA G+C content (bp)             | 471,303  | 35.37     |
| Total genes                      | 1,420   | 100.00    |
| RNA genes                        | 39      | 2.75      |
| Protein-coding genes             | 1,381   | 97.25     |
| Genes with function prediction   | 1,187   | 83.59     |
| Genes assigned to COGs           | 1,174   | 82.68     |
| Genes assigned to Pfam domains   | 1,195   | 824.15    |
| Genes with signal peptides       | 229     | 16.13     |
| Genes with transmembrane helices | 315     | 22.18     |

a) The total is based on either the size of the genome in base pairs or the total number of protein coding genes in the annotated genome.
The genomes of HIMB624 and HTCC2181 were compared to two closely related species within the family *Methylophilaceae* whose whole genomes are publicly available: *Methylothera mobilis* (NC_012968) and *Methylovorus glucosotrophus* SIP3-4 (NC_012969, NC_012970, NC_012972). For this comparison only, the four strains were automatically annotated using the RAST annotation server [31] and protein sequences were compared using the sequence based analysis tool in order to identify all shared and unique gene combinations (Figure 4). In addition to a single large chromosome, *Methylovorus glucosotrophus* SIP3-4 has 2 plasmids, while the remaining three genomes are all single chromosomes only. Strain HIMB624 contains one gene for a Type 4 fimbrial assembly/ATPase PilB that shares 43.44% protein identity with a gene located on one of the plasmids of *Methylovorus glucosotrophus* SIP3-4, and strain HTCC2181 contains a single DNA methylase gene that shares 31.1% protein identity with the same plasmid. Other than these, all genes located on the plasmids are
exclusive to *Methylovorus glucosotrophus* SIP3-4, and the large majority of the genes on the plasmids are hypothetical proteins. The genomes of *Methylotenera mobilis* and *Methylovorus glucosotrophus* SIP3-4 share over 100 genes associated with motility (twitching, flagella related, pili), along with 13 genes for chemotaxis and 13 genes for secretion that are absent from the genomes of HIMB624 and HTCC2181, while the two smaller genomes have a higher percentage of their genomes (9.13% and 9.19%) dedicated to amino acid transport and metabolism than *Methylovorus glucosotrophus* SIP3-4 (6.76%) and *Methylotenera mobilis* (5.81%); and a higher percentage of translation, ribosomal structure and biogenesis genes (11.08% and 11.47%) than *Methylovorus glucosotrophus* SIP3-4 (6.12%) and *Methylotenera mobilis* (7.16%). Due to the small size of the two OM43 lineage genomes, the higher percentages result in a similar total number of genes between all genomes in these categories, at approximately 120 genes for amino acid transport and metabolism and approximately 140 genes for translation, ribosomal structure and biogenesis. The general distribution of genes in all other predicted COG categories are comparable between the four strains, resulting in smaller numbers of total genes in each COG category for the two members of the OM43 lineage due to their comparatively smaller genome sizes.

Table 4. Number of genes associated with the 25 general COG functional categories

| Code | Value | %age* | Description |
|------|-------|-------|-------------|
| J    | 142   | 11.1  | Translation |
| A    | 1     | 0.1   | RNA processing and modification |
| K    | 47    | 3.7   | Transcription |
| L    | 67    | 5.2   | Replication, recombination and repair |
| B    | 1     | 0.1   | Chromatin structure and dynamics |
| D    | 22    | 1.7   | Cell cycle control, mitosis and meiosis |
| Y    | 0     | 0     | Nuclear structure |
| V    | 8     | 0.6   | Defense mechanisms |
| T    | 27    | 2.1   | Signal transduction mechanisms |
| M    | 110   | 8.6   | Cell wall/membrane biogenesis |
| N    | 13    | 1.0   | Cell motility |
| Z    | 0     | 0     | Cytoskeleton |
| W    | 0     | 0     | Extracellular structures |
| U    | 40    | 3.1   | Intracellular trafficking and secretion |
| O    | 78    | 6.1   | Posttranslational modification, protein turnover, chaperones |
| C    | 91    | 7.1   | Energy production and conversion |
| G    | 63    | 4.9   | Carbohydrate transport and metabolism |
| E    | 117   | 9.1   | Amino acid transport and metabolism |
| F    | 44    | 3.4   | Nucleotide transport and metabolism |
| H    | 92    | 7.2   | Coenzyme transport and metabolism |
| I    | 40    | 3.1   | Lipid transport and metabolism |
| P    | 45    | 3.5   | Inorganic ion transport and metabolism |
| Q    | 16    | 1.3   | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 123   | 9.6   | General function prediction only |
| S    | 95    | 7.4   | Function unknown |
| -    | 246   | 17.3  | Not in COGs |

a) The total is based on the total number of protein coding genes in the annotated genome.
Figure 4. Proportional Venn diagram depicting the shared and unique gene fractions between HIMB624, HTCC2181, and two closely related strains from within the family Methylophilaceae, Methylotenera mobilis and Methylovorus glucosotrophus SIP3-4.

Acknowledgments
The authors thank Cornelia Schmidt for her work in isolating strain HIMB624, and the Gordon and Betty Moore Foundation, which funded sequencing of this genome through its Marine Microbial Sequencing Project. The authors also thank the J. Craig Venter Institute for performing the sequencing and assembly, Tina Carvalho for assistance with electron microscopy, and Steve Giovannoni and H. James Tripp for useful discussion. We also thank Steve Giovannoni for providing access to annotation and analysis tools through the Oregon State University Center for Genome Research and Biocomputing. This research was supported by National Science Foundation Grant DEB-0207085, and NSF Science and Technology Center Award EF-0424599. This is SOEST contribution 8496 and HIMB contribution 1467.

References
1. Connon SA, Giovannoni SJ. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. Appl Environ Microbiol 2002; 68:3878-3885. PubMed
   http://dx.doi.org/10.1128/AEM.68.8.3878-3885.2002
2. Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature 2002; 418:630-633. PubMed
   http://dx.doi.org/10.1038/nature00917
3. Rappé MS, Kemp PF, Giovannoni SJ. Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. Limnol Oceanogr 1997; 42:811-826. PubMed
   http://dx.doi.org/10.4319/lo.1997.42.5.0811
4. Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, Cho JC, Oh HM, Kitner JB, Vergin KL, Rappé MS. The small genome of an abundant coastal ocean methylotroph. Environ Microbiol 2008; 10:1771-1782. PubMed
5. 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

6. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains , , and Eucarya. Proc Natl Acad Sci USA 1990; 87:4576-4579. PubMed http://dx.doi.org/10.1073/pnas.87.12.4576

7. Garrity GM, Bell JA, Lilburn T. Phylum XIV. phyl. nov. In: DJ Brenner, NR Krieg, JT Staley, GM Garrity (eds), Bergey’s Manual of Systematic Bacteriology, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 1.

8. Validation List No. 107. List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol 2006; 56:1-6. PubMed http://dx.doi.org/10.1099/ijs.0.64188-0

9. Garrity GM, Bell JA, Lilburn T. Class II. class. nov. In: DJ Brenner, NR Krieg, JT Staley, GM Garrity (eds), Bergey’s Manual of Systematic Bacteriology, Second Edition, Volume 2, Part C, Springer, New York, 2005, p. 575.

10. Garrity GM, Bell JA, Lilburn T. Order III. ord. nov. In: DJ Brenner, NR Krieg, JT Staley, GM Garrity (eds), Bergey’s Manual of Systematic Bacteriology, Second Edition, Volume 2, Part C, Springer, New York, 2005, p. 770.

11. Garrity GM, Bell JA, Lilburn T. Family I. fam. nov. In: DJ Brenner, NR Krieg, JT Staley, GM Garrity (eds), Bergey’s Manual of Systematic Bacteriology, Second Edition, Volume 2, Part C, Springer, New York, 2005, p. 770.

12. 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. Nat Genet 2000; 25:25-29. PubMed http://dx.doi.org/10.1038/75556

13. Pruess E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 2007; 35:7188-7196. PubMed http://dx.doi.org/10.1093/nar/gkm864

14. Yarza P, Richter M, Peplies J, Euzéby J, Amann R, Schleifer KH, Ludwig W, Glöckner FO, Rosselló-Móra R. The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. Syst Appl Microbiol 2008; 31:241-250. PubMed http://dx.doi.org/10.1016/j.syapm.2008.07.001

15. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, et al. ARB: a software environment for sequence data. Nucleic Acids Res 2004; 32:1363-1371. PubMed http://dx.doi.org/10.1093/nar/gkh293

16. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 2006; 22:2688-2690. PubMed http://dx.doi.org/10.1093/bioinformatics/btl446

17. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol 2008; 57:758-771. PubMed http://dx.doi.org/10.1080/10635150802429642

18. Gogleva AA, Kaparullina EN, Doronina NV, Trotsenko YA, sp. nov. and sp. nov., aerobic, methylotrophic bacteria associated with plants. Int J Syst Evol Microbiol 2010; 60:2623-2628. PubMed http://dx.doi.org/10.1099/ijs.0.019455-0

19. Jenkins O, Byrom D, Jones D. a new genus of methanol-utilizing bacteria. Int J Syst Evol Microbiol 1987; 37:446-448.

20. Kalyuzhnaya MG, Bowerman S. Lara, J.C., Lidstron, M.E. and Chistoserdova. gen. nov., sp. nov., an obligately methylamine-utilizing bacterium within the family . Int J Syst Evol Microbiol 2006; 56:2819-2823. PubMed http://dx.doi.org/10.1099/ijs.0.64191-0

21. Govorukhina NI, Kletsova LV, Tsygankov YD, Trotsenko YA, Netrusov AI. Characteristics of a new obligate methylotroph. Microbiology 1987; 56:849-854.

22. Doronina NV, Kudinova LV, Trotsenko YA. sp. nov.: A new species of aerobic, obligately methylotrophic bacteria associated with plants. [English translation of Mikrobiologiya]. Microbiology 2000; 69:599-603. http://dx.doi.org/10.1007/BF02756815

23. Doronina NV, Trotsenko YA, Kolganova TV, Tourova TP, Salkinoja-Salonen MS. sp. nov., a novel non-pigmented, aerobic, obligately methylotrophic bacterium isolated from meadow grass. Int J Syst Evol Microbiol 2004; 54:1453-
24. Rappé MS, Vergin K, Giovannoni SJ. Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems. *FEMS Microbiol Ecol* 2000; 33:219-232. PubMed http://dx.doi.org/10.1016/S0168-6496(00)00064-7

25. Goldberg SMD, Johnson J, Busam D, Feldblyum T, Ferriera S, Friedman R, Halpern A, Khouri H, Kravitz SA, Lauro FM, et al. A Sanger/pyrosequencing hybrid approach for the generation of high-quality draft assemblies of marine microbial genomes. *Proc Natl Acad Sci USA* 2006; 103:11240-11245. PubMed http://dx.doi.org/10.1073/pnas.0604351103

26. 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

27. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 1999; 27:4636-4641. PubMed http://dx.doi.org/10.1093/nar/27.23.4636

28. 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 http://dx.doi.org/10.1093/nar/25.5.955

29. Lagesen K, Hallin P, Rognes T, Ussery DW. RNAmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007; 35:3100-3108. PubMed http://dx.doi.org/10.1093/nar/gkm160

30. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res* 2004; 33:D121-D124. PubMed http://dx.doi.org/10.1093/nar/gkm081

31. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 2008; 9:75. PubMed http://dx.doi.org/10.1186/1471-2164-9-75