SUPPLEMENTARY METHODS

DNA extraction. DNA was extracted using a modified procedure of Sommerville et al.\textsuperscript{31}, as follows. Microcosm samples were centrifuged at 5,000 \( g \) for 15 min and mixed with freshly prepared lysozyme solution (10 mg ml\(^{-1}\) in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA), followed by incubation at 37\( ^{\circ} \)C for 1 h. SDS (2\% final concentration) and proteinase K (20 \( \mu \)g ml\(^{-1}\)) were then added, and samples were incubated at 37\( ^{\circ} \)C for an additional 5 h or overnight. 5 M NaCl solution was then added to the mixtures to a final concentration of 1.25 M. An equal volume of phenol–chloroform–isoamyl alcohol (25:24:1) was added and the mixture was incubated for 30 min at room temperature with horizontal shaking at 150 rpm, followed by centrifugation at 5,000 \( g \) for 15 min at 4\( ^{\circ} \)C. The aqueous phase was transferred to a clean centrifuge tube and treated again with an equal volume of phenol–chloroform–isoamyl alcohol (25:24:1), as above. An additional purification step, using an equal volume of chloroform, was conducted. The DNA was precipitated from the aqueous phase with 0.7 volume of isopropyl alcohol at room temperature, followed by centrifugation at 5,000 \( g \) for 15 min at 4\( ^{\circ} \)C. The pellet was washed with 70\% ethanol, dried, and re-suspended in 1 ml of TE buffer, pH 8.0. DNA concentration was measured spectrophotometrically.

Isopycnic centrifugation and DNA recovery. DNA extracted from the microcosms was prepared for CsCl-ethidium bromide density gradient ultracentrifugation as previously described\textsuperscript{29} and centrifuged at 265,000 \( g \) (Beckman VTi 65 rotor) for 16 h at 20\( ^{\circ} \)C. \textsuperscript{13}C-DNA fractions were visualized in UV (Fig. S1) and collected using 19-gauge needles. DNA was purified following standard procedures and used in a second CsCl-ethidium bromide density gradient ultracentrifugation, as described above.
**Array design.** The array design was based on the composite genomic sequence of *M. mobilis* (*Methylotenera* bin; Table 1). Probes for all identified potential genes were designed by Combimatrix Inc. using proprietary software. Probes were designed to have a melting temperature (Tm) of 72°C as calculated using the method of SantaLucia and Hicks\(^{36}\). Probes were chosen only if they fulfilled certain quality control metrics, as follows. They had to be 35–40 bp length, with a worst case probe hairpin Tm of < 40°C, there could be no single-base repeats greater than 6 and no two-base repeats greater than 4, GC percentage needed to be between 35 and 65%. The probe length criterion was relaxed to 30 bp for a total of 48 probes, which otherwise would have resulted in the corresponding genes not being represented on the array. The 12,951 gene sequences representing the *Methylotenera* composite genome were clustered using a version of a BLAST-based similarity algorithm. Clusters were made from the input sequences using a percent similarity minimum of 90%. We found a total of 7,195 singletons and 2,403 clusters. Multiple probes were designed for each of these. The specificity of a potential probe was determined by using a proprietary Combimatrix BLAST algorithm that uses the SantaLucia thermodynamic model to determine the Tm of each hit. This model takes gaps and mismatches into account. A hit was counted as a true hit if its Tm was within 12°C of the Tm of the probe itself. Probes were chosen first to be unique, to not hit any other singletons or clusters. One Unique probe was thus chosen for each singleton. Then, for each cluster, the minimal set of probes was chosen that hit all the members of that cluster. A total of 11,287 probes were chosen. 713 were replicated, bringing the total number of *in situ* synthesized probes to 12,000. In addition, the design included 545 manufacturer-designed quality control probes and 149 empty spots used for background correction.
DNA labeling and microarray hybridization. *M. mobilis* JLW8\textsuperscript{14} was cultivated as previously described\textsuperscript{14}. DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN, Valencia CA) and fragmented using an ultrasonic homogenizer Branson Sonifier 150 (for 5 seconds at setting 5), resulting in 300-500 base pair long DNA fragments. 5(3-aminoallyl)-d-UTP (Invitrogen, Carlsbad, CA) was incorporated into DNA using the Random Primed DNA Labeling Kit (Roche Applied Science, Indianapolis, IN USA), and amino modified DNA was labeled with the AlexaFluor 555 dye (Invitrogen, Carlsbad, CA) using the ARES DNA labeling kit (Invitrogen, Carlsbad, CA), in accordance with the manufacturer's instructions. Labeled DNA was purified using the QIAquick PCR purification kit (QIAGEN). Concentration of labeled DNA and efficiency of dye incorporation were analyzed using the NanoDrop ND-1000 instrument (NanoDrop Technologies, Wilmington, DE). DNA (25 µl) was hybridized to the microarray for 16 h at 58°C in a standard hybridization buffer (5xSSC, 20% formamide, 0.1% SDS, 0.01 mg Salmon DNA). Two replicate hybridizations were carried out. Arrays were scanned using the Axon GenePix 4000B microarray scanner (Molecular Devices Corporation, Sunnyvale, CA) at 5 µm resolution. Images were acquired using the Microarray Imager software (Combimatrix, https://webapps.combimatrix.com/customarray/submitandstatus.jsp). Poor quality spots were identified visually and flagged accordingly. Out of the 12,000 arrayed probes, 6,363 and 6,576 produced signal intensities above background (maximum intensity for control spots). More than 90% of the clustered probes (2,181 and 2,524 respectively) and more than 50% of the singleton probes (4,182 and 4,032, respectively) produced signals. These data suggest that the majority if not all genes in the genome of *M. mobilis* JLW8 had matching probes on the microarray.
Identification of Fae homologs. Peptide sequences of Fae and Fae homologs belonging to different phylogenetic groups (Fig. S6) were used as queries against the non-redundant database (NCBI) as well as against the database that is part of the JGI’s IMG/M system. Fae homologs only distantly related to the queries were identified in a number of microbes not capable of tetrahydromethanopterin (H₄MPT)-linked transformations, such as Yersinia species, Serratia species, Burkholderia species, and Arthrobacter. The genomes of these species then were queried with other peptides involved in H₄MPT-linked transformations in Betaproteobacteria, Archaea and Planctomycetes. The query peptides are listed in Table S2. No homologs for these genes were detected. Notably, a number of species of Burkholderia do possess complete sets of genes for H₄MPT-linked reactions, and these possess typical fae genes but no distant fae homologs that are present in Burkholderia species that lack other genes for H₄MPT-linked transformations.

Phylogenetic analysis. Fae and Fae homolog amino acid sequences (138 to 148) were aligned using the ClustalW program. For phylogenetic analyses, the Phylip package was used. Maximum likelihood, distance and parsimony methods were employed, 1000 bootstrap analyses were performed. Tree branching patterns were similar for the three analyses.
SUPPLEMENTARY REFERENCES


Supplementary Figure 1. Separation of $^{13}$C DNA from $^{12}$C DNA by isopycnic centrifugation.
1, methane; 2, methanol; 3, methylamine; 4, formaldehyde; 5, formate
### Supplementary Table 1. 16S rRNA genes identified in Lake Washington metagenomic datasets

<table>
<thead>
<tr>
<th>Gene identifier</th>
<th>Gene length (bp)</th>
<th>Contig length (bp)</th>
<th>Contig coverage (X)</th>
<th>Top hit (organism, % identity)</th>
<th>Coverage score#</th>
<th>Methylotroph representatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methane microcosm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006207859</td>
<td>293</td>
<td>4411</td>
<td>2.6</td>
<td><em>Methylobacter tundripaludum</em> (100)</td>
<td>2.6</td>
<td>Yes</td>
</tr>
<tr>
<td>2006207716</td>
<td>1527</td>
<td>2144</td>
<td>2.6</td>
<td><em>Methylobacter tundripaludum</em> (98)</td>
<td>2.6</td>
<td>Yes</td>
</tr>
<tr>
<td>2006208254</td>
<td>758</td>
<td>3852</td>
<td>1.7</td>
<td><em>Methylobacter tundripaludum</em> (95)</td>
<td>1.7</td>
<td>Yes</td>
</tr>
<tr>
<td>2006240650</td>
<td>787</td>
<td>800</td>
<td>-</td>
<td><em>Methylobacter tundripaludum</em> (97)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006207685</td>
<td>1395</td>
<td>2720</td>
<td>1.8</td>
<td><em>Methylotenera mobilis</em> (96)</td>
<td>1.8</td>
<td>Yes</td>
</tr>
<tr>
<td>2006207515</td>
<td>908</td>
<td>932</td>
<td>1.9</td>
<td><em>Methylotenera mobilis</em> (97)</td>
<td>1.9</td>
<td>Yes</td>
</tr>
<tr>
<td>2006249425</td>
<td>479</td>
<td>924</td>
<td>-</td>
<td><em>Methylotenera mobilis</em> (99)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006209782</td>
<td>834</td>
<td>1581</td>
<td>1.2</td>
<td>Uncultured <em>Verrucomicrobiales</em> (99)</td>
<td>1.2</td>
<td>Yes</td>
</tr>
<tr>
<td>2006245860</td>
<td>582</td>
<td>898</td>
<td>-</td>
<td>Uncultured <em>Verrucomicrobiales</em> (97)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006218740</td>
<td>630</td>
<td>922</td>
<td>-</td>
<td>Uncultured <em>Deltaproteobacteria</em> (94)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006248162</td>
<td>588</td>
<td>782</td>
<td>-</td>
<td>Uncultured <em>Comamonadaceae</em> (98)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006275380</td>
<td>288</td>
<td>679</td>
<td>-</td>
<td>Uncultured <em>Nitrospirae</em> (97)</td>
<td>0.5</td>
<td>No</td>
</tr>
</tbody>
</table>
### Methanol microcosm

<table>
<thead>
<tr>
<th>Date</th>
<th>Seqno</th>
<th>OTU</th>
<th>Abundance</th>
<th>Species</th>
<th>Identity</th>
<th>Cultured</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006292615</td>
<td>1319</td>
<td>5913</td>
<td>4.6</td>
<td><em>Methylobacterium mobilis</em> (98)</td>
<td>4.6</td>
<td>Yes</td>
</tr>
<tr>
<td>2006292610</td>
<td>1530</td>
<td>2576</td>
<td>1.7</td>
<td>Uncultured <em>Betaproteobacteria</em> (94)</td>
<td>1.7</td>
<td>Yes</td>
</tr>
<tr>
<td>2006291654</td>
<td>444</td>
<td>2124</td>
<td>1.4</td>
<td><em>Sphingomonas</em> (99)</td>
<td>1.4</td>
<td>Yes</td>
</tr>
<tr>
<td>2006291383</td>
<td>984</td>
<td>2015</td>
<td>1.4</td>
<td><em>Actinobacteria</em> (99)</td>
<td>1.4</td>
<td>Yes</td>
</tr>
<tr>
<td>2006289988</td>
<td>659</td>
<td>1790</td>
<td>1.6</td>
<td><em>Actinobacteria</em> (95)</td>
<td>1.6</td>
<td>Yes</td>
</tr>
<tr>
<td>2006289159</td>
<td>703</td>
<td>1403</td>
<td>1.4</td>
<td><em>Acinetobacter</em> (98)</td>
<td>1.4</td>
<td>No</td>
</tr>
<tr>
<td>2006293221</td>
<td>814</td>
<td>1418</td>
<td>1.4</td>
<td>Uncultured <em>Verrucomicrobiales</em> (99)</td>
<td>1.4</td>
<td>Yes</td>
</tr>
<tr>
<td>2006291811</td>
<td>454</td>
<td>1143</td>
<td>1.6</td>
<td>Uncultured <em>Verrucomicrobiales</em> (99)</td>
<td>1.6</td>
<td>Yes</td>
</tr>
<tr>
<td>2006289678</td>
<td>872</td>
<td>1368</td>
<td>2.2</td>
<td><em>Cyanobacteria</em> (98)</td>
<td>2.2</td>
<td>No</td>
</tr>
<tr>
<td>2006320920</td>
<td>829</td>
<td>831</td>
<td>-</td>
<td><em>Chloroflexi</em> (97)</td>
<td>0.5</td>
<td>No</td>
</tr>
<tr>
<td>2006330132</td>
<td>558</td>
<td>917</td>
<td>-</td>
<td><em>Acidobacteria</em> (98)</td>
<td>0.5</td>
<td>No</td>
</tr>
<tr>
<td>2006309983</td>
<td>512</td>
<td>905</td>
<td>-</td>
<td><em>Acidobacteria</em> (98)</td>
<td>0.5</td>
<td>No</td>
</tr>
</tbody>
</table>
### Methylamine microcosm

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>CFU</th>
<th>TSP</th>
<th>%</th>
<th>Genus</th>
<th>Species</th>
<th>%</th>
<th>Presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006368771</td>
<td>1413</td>
<td>6497</td>
<td>20.4</td>
<td><em>Methylotenera mobilis</em></td>
<td>(98)</td>
<td>20.4</td>
<td>Yes</td>
</tr>
<tr>
<td>2006377733</td>
<td>1175</td>
<td>5092</td>
<td>13.2</td>
<td><em>Methylotenera mobilis</em></td>
<td>(97)</td>
<td>13.2</td>
<td>Yes</td>
</tr>
<tr>
<td>2006367297</td>
<td>291</td>
<td>2202</td>
<td>1.8</td>
<td><em>Methylotenera mobilis</em></td>
<td>(100)</td>
<td>1.8</td>
<td>Yes</td>
</tr>
<tr>
<td>2006376104</td>
<td>1053</td>
<td>1061</td>
<td>1.7</td>
<td><em>Methylotenera mobilis</em></td>
<td>(99)</td>
<td>1.7</td>
<td>Yes</td>
</tr>
<tr>
<td>2006367454</td>
<td>1490</td>
<td>4538</td>
<td>2.2</td>
<td><em>Rhodoferax</em></td>
<td>(98)</td>
<td>2.2</td>
<td>No</td>
</tr>
<tr>
<td>2006377337</td>
<td>1018</td>
<td>1670</td>
<td>1.7</td>
<td><em>Burkholderiales</em></td>
<td>(97)</td>
<td>1.7</td>
<td>Yes</td>
</tr>
<tr>
<td>2006396024</td>
<td>835</td>
<td>835</td>
<td>-</td>
<td><em>Methylobacter tundripaludum</em></td>
<td>(99)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006392119</td>
<td>735</td>
<td>735</td>
<td>-</td>
<td><em>Methylomonas</em></td>
<td>(95)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006412713</td>
<td>266</td>
<td>707</td>
<td>-</td>
<td>Uncultured <em>Nitrospirae</em></td>
<td>(91)</td>
<td>0.5</td>
<td>No</td>
</tr>
<tr>
<td>2006388932</td>
<td>650</td>
<td>650</td>
<td>-</td>
<td>Uncultured <em>Nitrospirae</em></td>
<td>(98)</td>
<td>0.5</td>
<td>No</td>
</tr>
</tbody>
</table>
## Formaldehyde microcosm

<table>
<thead>
<tr>
<th>Date</th>
<th>FID</th>
<th>OID</th>
<th>Log10</th>
<th>Genus Species</th>
<th>Percent</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006425045</td>
<td>1319</td>
<td>1899</td>
<td>2.0</td>
<td><em>Methylobacter tundripaludum</em> (96)</td>
<td>1.9</td>
<td>Yes</td>
</tr>
<tr>
<td>2006512227</td>
<td>812</td>
<td>814</td>
<td>-</td>
<td><em>Methyloversatilis universalis</em> (98)</td>
<td>1.9</td>
<td>Yes</td>
</tr>
<tr>
<td>2006424484</td>
<td>983</td>
<td>1230</td>
<td>1.3</td>
<td>Uncultured <em>Verrucomicrobiales</em> (95)</td>
<td>1.3</td>
<td>Yes</td>
</tr>
<tr>
<td>2006437489</td>
<td>833</td>
<td>833</td>
<td>-</td>
<td>Uncultured <em>Verrucomicrobiales</em> (97)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006426163</td>
<td>766</td>
<td>875</td>
<td>2.0</td>
<td>Uncultured <em>Archaea</em> (99)</td>
<td>2.0</td>
<td>Yes</td>
</tr>
<tr>
<td>2006423173</td>
<td>1156</td>
<td>2183</td>
<td>1.8</td>
<td>Uncultured <em>Nitrospirae</em> (91)</td>
<td>1.8</td>
<td>No</td>
</tr>
<tr>
<td>2006425407</td>
<td>1500</td>
<td>2663</td>
<td>1.5</td>
<td><em>Nitrosomonas</em> (92)</td>
<td>1.5</td>
<td>No</td>
</tr>
<tr>
<td>2006507974</td>
<td>314</td>
<td>809</td>
<td>-</td>
<td><em>Nitrosomonas</em> (92)</td>
<td>0.5</td>
<td>No</td>
</tr>
<tr>
<td>2006423169</td>
<td>1098</td>
<td>1502</td>
<td>1.3</td>
<td><em>Acidobacteria</em> (92)</td>
<td>1.3</td>
<td>No</td>
</tr>
<tr>
<td>2006425656</td>
<td>897</td>
<td>1584</td>
<td>1.2</td>
<td><em>Aquaspirillum</em> (97)</td>
<td>1.2</td>
<td>No</td>
</tr>
<tr>
<td>2006426163</td>
<td>766</td>
<td>875</td>
<td>1.2</td>
<td>Uncultured <em>Planctomycete</em> (94)</td>
<td>1.2</td>
<td>No</td>
</tr>
<tr>
<td>2006430778</td>
<td>505</td>
<td>798</td>
<td>-</td>
<td>Uncultured <em>Gammaproteobacteria</em> (99)</td>
<td>0.5</td>
<td>Yes</td>
</tr>
<tr>
<td>2006506712</td>
<td>449</td>
<td>794</td>
<td>-</td>
<td><em>Chloroflexi</em> (98)</td>
<td>0.5</td>
<td>No</td>
</tr>
<tr>
<td>2006423039</td>
<td>1391</td>
<td>2875</td>
<td>1.7</td>
<td>Chloroplast (96)</td>
<td>1.7</td>
<td>*</td>
</tr>
</tbody>
</table>
Formate microcosm

<table>
<thead>
<tr>
<th>ID</th>
<th>NGSS</th>
<th>Id</th>
<th>-</th>
<th>Sequence</th>
<th>Coverage (%)</th>
<th>Score</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006530432</td>
<td>825</td>
<td>825</td>
<td>-</td>
<td>Ralstonia eutropha (99)</td>
<td>0.5</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>2006530476</td>
<td>248</td>
<td>870</td>
<td>-</td>
<td>Micromonospora (96)</td>
<td>0.5</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>2006538978</td>
<td>314</td>
<td>753</td>
<td>-</td>
<td>Uncultured Actinobacteria (99)</td>
<td>0.5</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>2006538694</td>
<td>314</td>
<td>846</td>
<td>-</td>
<td>Uncultured Deltaproteobacteria (99)</td>
<td>0.5</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>2006539641</td>
<td>574</td>
<td>615</td>
<td>-</td>
<td>Uncultured Deltaproteobacteria (96)</td>
<td>0.5</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

#Coverage score is sequence coverage for contigs. For singleton sequences, coverage was arbitrarily assumed at 0.5.

* Algae are known to consume methanol and oxidize it to formaldehyde and further to CO₂. Current knowledge on C₁ metabolism by algae and higher plants is extensively referenced in[37].
<table>
<thead>
<tr>
<th>Gene</th>
<th>Number of copies (coverage score)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>23 (11.5)</td>
</tr>
<tr>
<td>Fae</td>
<td>3 (1.5)</td>
</tr>
<tr>
<td>MtdB/MtdC</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Mch</td>
<td>7 (3.5)</td>
</tr>
<tr>
<td>FhcA</td>
<td>5 (2.5)</td>
</tr>
<tr>
<td>FhcB</td>
<td>4 (2.0)</td>
</tr>
<tr>
<td>FhcC</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>FhcD</td>
<td>2 (1.0)</td>
</tr>
<tr>
<td>MptG</td>
<td>7 (3.5)</td>
</tr>
<tr>
<td>Afp</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Orf5</td>
<td>4 (2.0)</td>
</tr>
<tr>
<td>Orf7</td>
<td>3 (1.5)</td>
</tr>
<tr>
<td>Orf9</td>
<td>8 (4.0)</td>
</tr>
<tr>
<td>Orf17</td>
<td>3 (6.0)</td>
</tr>
<tr>
<td>Orf19</td>
<td>3 (1.5)</td>
</tr>
<tr>
<td>Orf20</td>
<td>7 (3.5)</td>
</tr>
<tr>
<td>Orf21</td>
<td>6 (3.0)</td>
</tr>
<tr>
<td>Orf22</td>
<td>4 (2.0)</td>
</tr>
<tr>
<td>OrfY</td>
<td>3 (1.5)</td>
</tr>
<tr>
<td></td>
<td>12 (13.6)</td>
</tr>
<tr>
<td></td>
<td>12 (18.8)</td>
</tr>
<tr>
<td></td>
<td>10 (42.5)</td>
</tr>
<tr>
<td></td>
<td>18 (21.4)</td>
</tr>
<tr>
<td></td>
<td>13 (8.1)</td>
</tr>
<tr>
<td></td>
<td>7 (2.9)</td>
</tr>
<tr>
<td></td>
<td>27 (45.0)</td>
</tr>
<tr>
<td></td>
<td>11 (5.5)</td>
</tr>
<tr>
<td></td>
<td>12 (9.7)</td>
</tr>
<tr>
<td></td>
<td>6 (3.0)</td>
</tr>
<tr>
<td></td>
<td>9 (12.4)</td>
</tr>
<tr>
<td></td>
<td>4 (2.6)</td>
</tr>
<tr>
<td></td>
<td>2 (1.0)</td>
</tr>
<tr>
<td></td>
<td>7 (10.1)</td>
</tr>
<tr>
<td></td>
<td>3 (2.4)</td>
</tr>
<tr>
<td></td>
<td>8 (5.1)</td>
</tr>
<tr>
<td></td>
<td>16 (18.1)</td>
</tr>
<tr>
<td></td>
<td>6 (5.1)</td>
</tr>
<tr>
<td></td>
<td>6 (5.2)</td>
</tr>
<tr>
<td></td>
<td>5 (7.0)</td>
</tr>
<tr>
<td></td>
<td>3 (1.5)</td>
</tr>
<tr>
<td></td>
<td>14 (17.6)</td>
</tr>
<tr>
<td></td>
<td>4 (3.2)</td>
</tr>
<tr>
<td></td>
<td>3 (1.5)</td>
</tr>
<tr>
<td></td>
<td>7 (4.3)</td>
</tr>
<tr>
<td></td>
<td>11 (16.1)</td>
</tr>
<tr>
<td></td>
<td>4 (2.6)</td>
</tr>
<tr>
<td></td>
<td>2 (1.5)</td>
</tr>
<tr>
<td></td>
<td>10 (16.1)</td>
</tr>
<tr>
<td></td>
<td>4 (2.6)</td>
</tr>
<tr>
<td></td>
<td>2 (1.0)</td>
</tr>
<tr>
<td></td>
<td>10 (15.9)</td>
</tr>
<tr>
<td></td>
<td>3 (1.5)</td>
</tr>
<tr>
<td></td>
<td>18 (23.5)</td>
</tr>
<tr>
<td></td>
<td>5 (2.5)</td>
</tr>
<tr>
<td></td>
<td>7 (5.0)</td>
</tr>
<tr>
<td></td>
<td>7 (6.0)</td>
</tr>
<tr>
<td></td>
<td>18 (23.5)</td>
</tr>
<tr>
<td></td>
<td>5 (2.5)</td>
</tr>
<tr>
<td></td>
<td>8 (5.1)</td>
</tr>
<tr>
<td></td>
<td>3 (4.1)</td>
</tr>
<tr>
<td></td>
<td>10 (13.7)</td>
</tr>
<tr>
<td></td>
<td>2 (1.0)</td>
</tr>
<tr>
<td></td>
<td>5 (3.8)</td>
</tr>
<tr>
<td></td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>6 (10.5)</td>
</tr>
<tr>
<td></td>
<td>2 (1.0)</td>
</tr>
<tr>
<td></td>
<td>10 (19.1)</td>
</tr>
<tr>
<td></td>
<td>2 (1.7)</td>
</tr>
<tr>
<td></td>
<td>7 (20.8)</td>
</tr>
<tr>
<td></td>
<td>2 (1.0)</td>
</tr>
<tr>
<td></td>
<td>4 (2.0)</td>
</tr>
<tr>
<td></td>
<td>2 (1.0)</td>
</tr>
<tr>
<td></td>
<td>4 (6.6)</td>
</tr>
<tr>
<td></td>
<td>0 (0.0)</td>
</tr>
<tr>
<td></td>
<td>5 (3.8)</td>
</tr>
<tr>
<td></td>
<td>2 (1.0)</td>
</tr>
<tr>
<td></td>
<td>15 (18.9)</td>
</tr>
<tr>
<td></td>
<td>3 (2.4)</td>
</tr>
</tbody>
</table>

*Coverage score was calculated as in Table S1.
**Supplementary Table 3. Phylogenetic distribution of methylamine-utilizing strains isolated from Lake Washington sediment**

<table>
<thead>
<tr>
<th>Organism</th>
<th>Number of strains</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Hyphomicrobium</em> spp.</td>
<td>42</td>
</tr>
<tr>
<td><em>Arthrobacter</em> spp.</td>
<td>38</td>
</tr>
<tr>
<td><em>Methylopha capsulata</em></td>
<td>3</td>
</tr>
<tr>
<td><em>Xanthobacter</em> spp.</td>
<td>3</td>
</tr>
<tr>
<td><em>Paenibacillus amylolyticus</em></td>
<td>3</td>
</tr>
<tr>
<td><em>Labrys</em> spp.</td>
<td>2</td>
</tr>
<tr>
<td><em>Methylobacterium</em> spp.</td>
<td>2</td>
</tr>
<tr>
<td><em>Rhodobacter</em> sp.</td>
<td>1</td>
</tr>
<tr>
<td><em>Methylophilus</em> sp.</td>
<td>1</td>
</tr>
<tr>
<td><em>Ancylobacter aquaticus</em></td>
<td>1</td>
</tr>
<tr>
<td><em>Pseudomonas</em> sp.</td>
<td>1</td>
</tr>
<tr>
<td><em>Methylotenera mobilis</em></td>
<td>1</td>
</tr>
</tbody>
</table>

Enrichments were established by inoculating filter-sterilized Lake Washington water supplemented with 10 mM methylamine with 1 ml of sediment sludge. 10 ml of the original 100 ml culture were transferred twice into 90 ml of fresh medium, and the third transfer enrichment was diluted appropriately and plated onto solid 0.2X Hypho medium containing 10 mM methylamine, essentially as described\(^\text{14}\). 100 random colonies were selected for identification via sequencing the 16S rRNA gene fragment, as described\(^\text{15}\).
### Supplementary Table 4. Major metabolic pathways deduced from the composite genome of *Methylotherma mobilis*

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Protein/function</th>
<th>Number of genes in contigs/singletons</th>
<th>Major contigs</th>
<th>Coverage score*</th>
<th>Homolog</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methylamine oxidation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>mauF</em></td>
<td>TTQ biosynthesis</td>
<td>3/1</td>
<td>C3855, C5447, C5618</td>
<td>11.7</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauB</em></td>
<td>MADH large subunit</td>
<td>4/5</td>
<td>C6715, C5447, C1074, C3854</td>
<td>15.8</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauE</em></td>
<td>essential for small subunit maturation</td>
<td>6/1</td>
<td>C6715, C5447, C1445, C1074</td>
<td>19.5</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauD</em></td>
<td>essential for small subunit maturation</td>
<td>6/0</td>
<td>C6715, C5447, C1445, C1074</td>
<td>19.0</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauA</em></td>
<td>MADH small subunit</td>
<td>4/0</td>
<td>C1445, C2535, C6715, C5447</td>
<td>11.9</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauG</em></td>
<td>TTQ biosynthesis</td>
<td>6/1</td>
<td>C2498, C6715, C1445, C2367</td>
<td>14.6</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauL</em></td>
<td>unknown</td>
<td>7/0</td>
<td>C5777, C2498, C6715, C5550</td>
<td>16.1</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauM</em></td>
<td>unknown</td>
<td>5/2</td>
<td>C5777, C2498, C1444, C5550</td>
<td>14.9</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauN</em></td>
<td>unknown</td>
<td>6/3</td>
<td>C5777, C2269, C2498, C1444</td>
<td>17.0</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mauO</em></td>
<td>cytochrome, proposed electron acceptor from MADH</td>
<td>4/0</td>
<td>C5777, C2269, C2498, C1444</td>
<td>11.7</td>
<td>No</td>
</tr>
<tr>
<td><strong>H_{4}MPT-linked formaldehyde oxidation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>mptG</em></td>
<td>beta-ribofuranosylaminobenzene 5-phosphate synthase</td>
<td>3/5</td>
<td>C5192, C3441, C4480</td>
<td>12.9</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mtdB</em></td>
<td>methylene H_{4}MPT dehydrogenase</td>
<td>3/2</td>
<td>C5192, C3441, C1971</td>
<td>9.7</td>
<td>Yes</td>
</tr>
<tr>
<td><em>orfY</em></td>
<td>unknown</td>
<td>7/5</td>
<td>C5192, C4480, C2798, C4119</td>
<td>17.4</td>
<td>Yes</td>
</tr>
<tr>
<td><em>mch</em></td>
<td>methenyl H_{4}MPT cyclohydrolase</td>
<td>5/1</td>
<td>C5192, C6529, C5269, C6324</td>
<td>9.3</td>
<td>Yes</td>
</tr>
<tr>
<td><em>orf5</em></td>
<td>biosynthesis of H_{4}MPT</td>
<td>6/0</td>
<td>C5192, C6529, C2219, C1365</td>
<td>11.9</td>
<td>Yes</td>
</tr>
<tr>
<td><em>orf7</em></td>
<td>unknown</td>
<td>6/1</td>
<td>C5192, C6529, C2219, C6323</td>
<td>12.9</td>
<td>Yes</td>
</tr>
<tr>
<td><em>foxA</em></td>
<td>response regulator, DNA-binding</td>
<td>4/4</td>
<td>C5192, C2219, C6323, C518</td>
<td>13.2</td>
<td>Yes</td>
</tr>
<tr>
<td><em>foxB</em></td>
<td>signal transduction histidine kinase</td>
<td>3/4</td>
<td>C5192, C2219, C6323, C518</td>
<td>11.5</td>
<td>Yes</td>
</tr>
<tr>
<td><em>orf17</em></td>
<td>unknown</td>
<td>5/4</td>
<td>C1746, C2894, C1766, C6041</td>
<td>13.1</td>
<td>Yes</td>
</tr>
<tr>
<td><em>orf1</em></td>
<td>unknown</td>
<td>6/5</td>
<td>C1544, C1766, C1060, C6041</td>
<td>14.8</td>
<td>Yes</td>
</tr>
<tr>
<td><em>orf9</em></td>
<td>biosynthesis of H_{4}MPT</td>
<td>8/3</td>
<td>C3049, C1544, C1765, C1060</td>
<td>20.0</td>
<td>Yes</td>
</tr>
<tr>
<td><em>pabB</em></td>
<td>para-aminobenzoate synthase component I</td>
<td>9/3</td>
<td>C3049, C1544, C1765, C1060</td>
<td>21.3</td>
<td>Yes</td>
</tr>
<tr>
<td><em>orf21</em></td>
<td>biosynthesis of H_{4}MPT</td>
<td>7/3</td>
<td>C1059, C1544, C1764, C1765</td>
<td>20.1</td>
<td>Yes</td>
</tr>
<tr>
<td><em>fhcB</em></td>
<td>formyltransferase/hydrolase complex subunit B</td>
<td>3/3</td>
<td>C4455, C3501, C3795</td>
<td>7.7</td>
<td>Yes</td>
</tr>
<tr>
<td><em>fhcA</em></td>
<td>formyltransferase/hydrolase complex subunit A</td>
<td>7/6</td>
<td>C2703, C4455, C6865, C3501</td>
<td>14.8</td>
<td>Yes</td>
</tr>
<tr>
<td><em>fhcD</em></td>
<td>formyltransferase/hydrolase complex subunit D</td>
<td>10/5</td>
<td>C2703, C4455, C4456, C6865</td>
<td>21.3</td>
<td>Yes</td>
</tr>
<tr>
<td><em>fhcC</em></td>
<td>formyltransferase/hydrolase complex subunit C</td>
<td>7/1</td>
<td>C2703, C4456, C6865, C5719</td>
<td>14.6</td>
<td>Yes</td>
</tr>
<tr>
<td>Gene</td>
<td>Function</td>
<td>Chromosome Locations</td>
<td>Coverage</td>
<td>Presence</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td>----------------------</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>orf22</td>
<td>biosynthesis of H₄MPT</td>
<td>C7362, C2237, C3254</td>
<td>6.1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>orf19</td>
<td>biosynthesis of H₄MPT</td>
<td>C7362, C2237, C3254, C3325</td>
<td>10.5</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>orf20</td>
<td>biosynthesis of H₄MPT</td>
<td>C3650, C3762, C2237, C3254</td>
<td>17.6</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>afp</td>
<td>dihydroorotase</td>
<td>C3660, C3762, C2237, C3254</td>
<td>14.7</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fae</td>
<td>formaldehyde activating enzyme (phytolype 1)</td>
<td>C4584, C7418, C824, C3699</td>
<td>17.0</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fae2</td>
<td>formaldehyde activating enzyme (phytolype 2)</td>
<td>C6255, C5746, C978, C5988</td>
<td>15.9</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>fad</td>
<td>homolog of formaldehyde activating enzyme</td>
<td>C5074, C6223, C4592</td>
<td>8.6</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fdhC</td>
<td>formate dehydrogenase gamma subunit</td>
<td>C1316, C371, C346, C5753</td>
<td>19.0</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fdhB</td>
<td>formate dehydrogenase beta subunit</td>
<td>C1316, C371, C346, C5753</td>
<td>23.7</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fdhA</td>
<td>formate dehydrogenase alpha subunit</td>
<td>C2031, C1316, C346, C5753</td>
<td>34.1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fdhD</td>
<td>formate dehydrogenase accessory protein</td>
<td>C2031, C4481, C1498, C1081</td>
<td>20.1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fdhE</td>
<td>formate dehydrogenase delta subunit</td>
<td>C2031, C1498, C1315, C3858</td>
<td>14.9</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fdh4A</td>
<td>formate dehydrogenase 4</td>
<td>C7278, C5619, C1549, C5323</td>
<td>24.3</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>fdh4B</td>
<td>formate dehydrogenase 4-associated protein</td>
<td>C7278, C1549, C5323, C1523</td>
<td>13.5</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>hps</td>
<td>gexulosephosphate synthase</td>
<td>C1746, C5917, C2894, C1766</td>
<td>18.7</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>hpi</td>
<td>hexulosephosphate isomerase</td>
<td>C1746, C2894, C1766, C6041</td>
<td>9.3</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>tal</td>
<td>transaldolase</td>
<td>C5192, C1746, C5917, C2894</td>
<td>14.1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>pgl</td>
<td>glucose 6-phosphate isomerase</td>
<td>C1859, C6642, C2189, C6499</td>
<td>25.9</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>zwf</td>
<td>glucose 6-phosphate dehydrogenase</td>
<td>C32, C1656, C755, C2267</td>
<td>12.9</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>pgI</td>
<td>6-phosphogluconolactonase</td>
<td>C32, C1500, C1656, C755</td>
<td>14.7</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>gndB</td>
<td>6-phosphogluconate dehydrogenase (NADP)</td>
<td>C1965, C5143, C2943, C7502</td>
<td>20.0</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>edd</td>
<td>6-phosphogluconate dehydratase</td>
<td>C4951, C3028, C3520, C5290</td>
<td>14.1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>eda</td>
<td>2-keto 3-deoxy-6-phosphogluconate aldolase</td>
<td>C2750, C2309, C1915, C5861</td>
<td>17.2</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ppi</td>
<td>ribose 5-phosphate isomerase</td>
<td>C1186, C1572, C432, C3180</td>
<td>21.3</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ttk</td>
<td>transketolase</td>
<td>C1572, C432, C3180, C3325</td>
<td>17.1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>tpe</td>
<td>ribulose 5-phosphate 3-epimerase</td>
<td>C4986, C7173, C950, C6275</td>
<td>21.6</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

**Formate oxidation**

- **fdhC**: Formate dehydrogenase gamma subunit
- **fdhB**: Formate dehydrogenase beta subunit
- **fdhA**: Formate dehydrogenase alpha subunit
- **fdhD**: Formate dehydrogenase accessory protein
- **fdhE**: Formate dehydrogenase delta subunit
- **fdh4A**: Formate dehydrogenase 4
- **fdh4B**: Formate dehydrogenase 4-associated protein

**Ribulose monophosphate cycle for formaldehyde assimilation/oxidation**

- **hps**: Gexulosephosphate synthase
- **hpi**: Hexulosephosphate isomerase
- **tal**: Transaldolase
- **pgl**: Glucose 6-phosphate isomerase
- **zwf**: Glucose 6-phosphate dehydrogenase
- **pgI**: 6-Phosphogluconolactonase
- **gndB**: 6-Phosphogluconate dehydrogenase (NADP)
- **edd**: 6-Phosphogluconate dehydratase
- **eda**: 2-Keto 3-Deoxy-6-Phosphogluconate Aldolase
- **ppi**: Ribose 5-Phosphate Isomerase
- **ttk**: Transketolase
- **tpe**: Ribulose 5-Phosphate 3-Epimerase

**C3 interconversion reactions**

- **aceE**: E1 component, pyruvate dehydrogenase
- **aceF**: E2 component, pyruvate dehydrogenase
<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Coverage</th>
<th>Contigs</th>
<th>ID</th>
<th>Gene</th>
<th>Function</th>
<th>Coverage</th>
<th>Contigs</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>lpdA</em></td>
<td>E3 component, pyruvate dehydrogenase</td>
<td>9/3</td>
<td>C5882, C285, C2718, C4579</td>
<td>23.8</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>pyk</em></td>
<td>pyruvate kinase</td>
<td>7/1</td>
<td>C2212, C4282, C7289, C5497</td>
<td>15.3</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>pgk</em></td>
<td>phosphoglycerate kinase</td>
<td>7/4</td>
<td>C2212, C4282, C5497, C9</td>
<td>15.7</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>gpd</em></td>
<td>glyceraldehydes phosphate dehydrogenase</td>
<td>4/0</td>
<td>C2212, C236, C9, C2333</td>
<td>9.1</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>ppn</em></td>
<td>PEP synthase</td>
<td>6/10</td>
<td>C3441, C4600, C683, C1281</td>
<td>16.9</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>pgm</em></td>
<td>phosphoglyceromutase</td>
<td>5/1</td>
<td>C6678, C5987, C5079, C2721</td>
<td>10.3</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>eno</em></td>
<td>enolase</td>
<td>11/0</td>
<td>C6365, C5742, C4439, C4440</td>
<td>26.0</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tpi</em></td>
<td>triosephosphate isomerase</td>
<td>5/0</td>
<td>C2358, C6889, C764, C1801</td>
<td>11.2</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Citric acid and methylcitric acid cycles**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Coverage</th>
<th>Contigs</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>acnB</em></td>
<td>aconitate hydratase B</td>
<td>11/5</td>
<td>C3642, C2776, C2380, C927</td>
<td>23.9</td>
</tr>
<tr>
<td><em>fum</em></td>
<td>fumarate hydratase</td>
<td>7/6</td>
<td>C3642, C7122, C927, C3767</td>
<td>18.3</td>
</tr>
<tr>
<td><em>prpE</em></td>
<td>putative propionyl-CoA synthetase</td>
<td>8/0</td>
<td>C3642, C2834, C7122, C3767</td>
<td>19.4</td>
</tr>
<tr>
<td><em>sdhB</em></td>
<td>succinate dehydrogenase, Fe-S subunit</td>
<td>6/3</td>
<td>C3642, C2834, C7122, C3993</td>
<td>16.2</td>
</tr>
<tr>
<td><em>sdhA</em></td>
<td>succinate dehydrogenase, flavoprotein subunit</td>
<td>10/4</td>
<td>C3642, C2834, C7122, C3993</td>
<td>23.3</td>
</tr>
<tr>
<td><em>sdhD</em></td>
<td>succinate dehydrogenase, hydrophobic anchor subunit</td>
<td>5/3</td>
<td>C2834, C2947, C7082, C3643</td>
<td>11.7</td>
</tr>
<tr>
<td><em>sdhC</em></td>
<td>succinate dehydrogenase, cytochrome subunit</td>
<td>3/0</td>
<td>C2834, C2947, C7082</td>
<td>7.2</td>
</tr>
<tr>
<td><em>mdh</em></td>
<td>malate dehydrogenase</td>
<td>8/1</td>
<td>C2834, C2278, C2946, C2947</td>
<td>18.2</td>
</tr>
<tr>
<td><em>prpR</em></td>
<td>transcriptional regulator</td>
<td>5/1</td>
<td>C2834, C2278, C2946, C598</td>
<td>13.0</td>
</tr>
<tr>
<td><em>acnM</em></td>
<td>putative methyl cis-aconitate hydratase</td>
<td>9/3</td>
<td>C4991, C2278, C2946, C1354</td>
<td>18.9</td>
</tr>
<tr>
<td><em>prpB</em></td>
<td>methylisocitrate lyase</td>
<td>10/1</td>
<td>C4991, C1354, C2279, C2557</td>
<td>15.8</td>
</tr>
<tr>
<td><em>prpC</em></td>
<td>methylcitrate synthase</td>
<td>10/2</td>
<td>C4991, C2279, C2557, C5762</td>
<td>18.5</td>
</tr>
<tr>
<td><em>prpD</em></td>
<td>methylcitrate dehydratase</td>
<td>11/1</td>
<td>C4991, C2279, C2557, C5762</td>
<td>22.3</td>
</tr>
<tr>
<td><em>gltA</em></td>
<td>citrate synthase</td>
<td>4/3</td>
<td>C1115, C2089, C3834, C3682</td>
<td>9.5</td>
</tr>
<tr>
<td><em>idh</em></td>
<td>isocitrate dehydrogenase</td>
<td>7/1</td>
<td>C3539, C1504, C5866, C7016</td>
<td>14.1</td>
</tr>
<tr>
<td><em>sucC</em></td>
<td>succinyl-CoA synthase alpha subunit</td>
<td>8/3</td>
<td>C2556, C143, C5216, C2910</td>
<td>17.0</td>
</tr>
<tr>
<td><em>sucD</em></td>
<td>succinyl-CoA synthase beta subunit</td>
<td>6/2</td>
<td>C2556, C7272, C3326, C2910</td>
<td>13.0</td>
</tr>
</tbody>
</table>

TTQ, triptophan triptophylquinone; MMDH, methylamine dehydrogenase; H4MPT, tetrahydromethanopterin. Contiguous genes highlighted in the same color are clustered on the chromosome.

* Coverage score is calculated as a sum of average contig coverage (X) for each gene. For singleton reads, coverage was counted at 0.5X.
Supplementary Table 5. Presence of genes encoding major house keeping functions in the *Methylotenera* composite genome, compared to other betaproteobacterial methylotrophs

<table>
<thead>
<tr>
<th>COG</th>
<th>Function</th>
<th>Methylobacillus flagellatus (3.0 Mbp)</th>
<th>Methylibium petrolephilum (4.6 Mbp)</th>
<th>Methylophilales bacterium (1.3 Mbp)</th>
<th>M. mobilis composite (methylamine enrichment) (11.1 Mbp)</th>
<th>M. mobilis composite (combined assembly) (13.3 Mbp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0331</td>
<td>(acyl-carrier-protein) S-malonyltransferase</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Glycerol-3-phosphate O-acyltransferase (an alternative to acyl phosphate pathway)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG2937</td>
<td>Fatty acid/phospholipid biosynthesis enzyme (PisX protein in acyl phosphate pathway)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0416</td>
<td>Fatty acid/phospholipid biosynthesis enzyme (PisY protein in acyl phosphate pathway)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>COG0344</td>
<td>1-acyl-sn-glycerol-3-phosphate acyltransferase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>COG0204</td>
<td>Enzyme involved in the deoxyxylulose pathway of</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>COG0331</td>
<td>Deoxyxylulose-5-phosphate synthase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>COG0743</td>
<td>1-deoxy-D-xylulose 5-phosphate reductoisomerase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>COG1211</td>
<td>4-diphosphocytidyl-2-methyl-D-erythritol synthase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>COG1947</td>
<td>2C-methyl-D-erythritol 2,4-cyclodiphosphate</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>COG0245</td>
<td>synthase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>COG0821</td>
<td>Enzyme involved in the deoxyxylulose pathway of</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>
isoprenoid biosynthesis (hydroxymethylbutenyl diphosphate synthase)  
Penicillin tolerance protein (hydroxymethylbutenyl pyrophosphate reductase)  

**Nucleotide biosynthesis**

**Purine biosynthesis**

<table>
<thead>
<tr>
<th>COG</th>
<th>Name</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
<th>Value4</th>
<th>Value5</th>
<th>Value6</th>
<th>Value7</th>
<th>Value8</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0034</td>
<td>Glutamine phosphoribosylpyrophosphate amidotransferase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0151</td>
<td>Phosphoribosylamine-glycine ligase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0299</td>
<td>Formyltransferase PurN (an alternative to COG0027)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0027</td>
<td>Formyltransferase (GAR transformylase)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0046</td>
<td>Folate-dependent phosphoribosylglycinamide</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0047</td>
<td>Folate-dependent phosphoribosylglycinamide</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0150</td>
<td>Phosphoribosylaminoidazole (AIR) synthetase</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0026</td>
<td>Phosphoribosylaminoidazole synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0152</td>
<td>(SAICAR) synthase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0015</td>
<td>Adenylsuccinate lyase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0138</td>
<td>(only IMP cyclohydrodase domain in Aful)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0516</td>
<td>IMP dehydrogenase/GMP reductase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0518</td>
<td>GMP synthase - Glutamine amidotransferase domain</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0519</td>
<td>GMP synthase, PP-ATPase domain/subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0194</td>
<td>Guanylate kinase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Purine and pyrimidine biosynthesis**

<table>
<thead>
<tr>
<th>COG</th>
<th>Name</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
<th>Value4</th>
<th>Value5</th>
<th>Value6</th>
<th>Value7</th>
<th>Value8</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0105</td>
<td>Nucleoside diphosphate kinase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0208</td>
<td>Ribonucleotide reductase, beta subunit</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>7</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0209</td>
<td>Ribonucleotide reductase, alpha subunit</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>14</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0602</td>
<td>Organic radical activating enzymes (ribonucleoside-triphosphate reductase activase subunit, alternative to ribonucleoside-diphosphate reductase)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Oxygen-sensitive ribonucleoside-triphosphate reductase (alternative to ribonucleoside-diphosphate reductase)

| COG1328 | 0 | 2 | 0 | 1 | 1 |

**Pyrimidine biosynthesis**

- Carbamoylphosphate synthase large subunit (split gene in MJ)
  - COG0458
  - COG0505
  - COG0540
  - COG0418
  - COG0167
  - COG0461
  - COG0284
  - COG0528

- Carbamoylphosphate synthase small subunit
  - COG0504

- Aspartate carbamoyltransferase, catalytic chain
  - COG0418

- Dihydroorotase
  - COG0017

- Dihydroorotate dehydrogenase
  - COG0461

- Orotate phosphoribosyltransferase
  - COG0284

- Orotidine-5'-phosphate decarboxylase
  - COG0017

- Uridylate kinase
  - COG0017

- Thymidylate synthase
  - COG0017

- Thymidylate kinase
  - COG0017

- CTP synthase (UTP-ammonia lyase)
  - COG0017

**Coenzyme and cofactor biosynthesis**

**Coenzyme A biosynthesis**

- Ketopantoate hydroxymethyltransferase
  - COG0413

- Ketopantoate reductase
  - COG1893

- Panthothenate synthetase
  - COG0017

- Panthothenate kinase (one of the three alternative forms)
  - COG0017

- Putative transcriptional regulator, homolog of Bvg accessory factor (panthothenate kinase, one of the three alternative forms)
  - COG0017

- Panthothenate kinase, acetyl-CoA regulated (one of the three alternative forms)
  - COG0017

- Phosphopantothenoylcysteine synthetase/decarboxylase
  - COG0017

- Phosphopantetheine adenylyltransferase
  - COG0017

- Dephospho-CoA kinase
  - COG0017

**Riboflavin and FAD biosynthesis**

- 3,4-dihydroxy-2-butanone 4-phosphate synthase
  - COG0017

- GTP cyclohydrolase II
  - COG0017

- Uncharacterized conserved protein (archaeal GTP cyclohydrolase)
  - COG0017
<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>Count</th>
<th>DB</th>
<th>GL</th>
<th>GLA</th>
<th>LA</th>
<th>RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0117</td>
<td>Pyrimidine deaminase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>COG01985</td>
<td>Pyrimidine reductase, riboflavin biosynthesis</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>COG0307</td>
<td>Riboflavin synthase alpha chain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>COG0054</td>
<td>Riboflavin synthase beta-chain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>COG1731</td>
<td>Archaeal riboflavin synthase (alternative to riboflavin synthase)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>COG0196</td>
<td>FAD synthase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>COG1339</td>
<td>Transcriptional regulator of a riboflavin/FAD biosynthetic operon (archaeal riboflavin kinase)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**NAD biosynthesis**

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>Count</th>
<th>DB</th>
<th>GL</th>
<th>GLA</th>
<th>LA</th>
<th>RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0029</td>
<td>Aspartate oxidase</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>COG1712</td>
<td>Predicted dinucleotide-utilizing enzyme (aspartate dehydrogenase, alternative to aspartate oxidase)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>COG0379</td>
<td>Quinolinate synthase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>COG3483</td>
<td>Tryptophan 2,3-dioxygenase (vermilion) (alternative pathway of quinolinate biosynthesis)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>COG3844</td>
<td>Kynureninase (alternative pathway of quinolinate biosynthesis)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>COG0157</td>
<td>Nicotinate-nucleotide pyrophosphorylase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>COG0157</td>
<td>Nicotinic acid mononucleotide adenylyltransferase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>COG0017</td>
<td>NAD synthase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

**Molybdenum cofactor and molybdopterin guanine dinucleotide biosynthesis**

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>Count</th>
<th>DB</th>
<th>GL</th>
<th>GLA</th>
<th>LA</th>
<th>RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG2896</td>
<td>Molybdenum cofactor biosynthesis enzyme</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>COG0315</td>
<td>Molybdenum cofactor biosynthesis enzyme (GTP cyclohydrolase subunit MoaC)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>COG0314</td>
<td>Molybdopterin converting factor, large subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>COG0476</td>
<td>Dinucleotide-utilizing enzymes involved in molybdopterin and thiamine biosynthesis family 2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>COG0521</td>
<td>([molybdopterin synthase] sulfurylase)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>COG0303</td>
<td>Molybdopterin biosynthesis enzymes (molybdopterin adenylyltransferase)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>COG0746</td>
<td>Molybdopterin biosynthesis enzyme (molybdopterin molybdochelatase)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>COG0746</td>
<td>Molybdopterin-guanine dinucleotide biosynthesis</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>COG Code</td>
<td>Protein Name and Description</td>
<td>Annotation Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------------------------</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG2068</td>
<td>Uncharacterized MobA-related protein</td>
<td>1 1 0 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Heme biosynthesis**

- 7-keto-8-aminopelargonate synthetase and related enzymes
- Glutamyl-tRNA reductase
- Glutamate-1-semialdehyde aminotransferase
- Delta-aminolevulinic acid dehydratase
- Porphobilinogen deaminase
- Uroporphyrinogen-III synthase
- Uroporphyrinogen-III decarboxylase
- Coproporphyrinogen III oxidase
- Coproporphyrinogen III oxidase and related Fe-S oxidoreductases

**Thiamine diphosphate biosynthesis**

- Thiamine biosynthesis protein ThiC
- Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase
- Thiamine biosynthesis enzyme ThiH and related uncharacterized enzymes (alternative to glycine oxidase)
- Glycine/D-amino acid oxidases (deaminating)
- Rhodanese-related sulfuryltransferase (ThiL protein)
- Sulfur transfer protein involved in thiamine biosynthesis
- Dinucleotide-utilizing enzymes involved in thiamine biosynthesis
- Molybdopterin and thiamine biosynthesis family 2
- Uncharacterized enzyme of thiazole biosynthesis
- (thiazole phosphate synthase)
- Thiamine monophosphate synthase
- Thiamine monophosphate kinase

**Amino acid biosynthesis**
### Arginine biosynthesis

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1364</td>
<td>N-acetylglutamate synthase (N-acetylornithine aminotransferase)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>0548</td>
<td>Acetylglutamate kinase</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>0002</td>
<td>Acetylglutamate semialdehyde dehydrogenase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4992</td>
<td>Ornithine/acetylornithine aminotransferase</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>0078</td>
<td>Ornithine carbamoyltransferase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acetylornithine deacetylase/Succinyl-diaminopimelate desuccinylase and related deacylases</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>9</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>0624</td>
<td>Serine acetyltransferase</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0031</td>
<td>Cystathionine beta-lyases/cystathionine gamma-synthases</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>0032</td>
<td>Cysteine synthase</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>0201</td>
<td>Homoserine acetyltransferase</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>1897</td>
<td>Homoserine trans-succinylase</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0620</td>
<td>Methionine synthase II (cobalamin-independent)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Methionine synthase I (cobalamin-dependent),</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0646</td>
<td>methyltransferase domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>0111</td>
<td>Phosphoglycerate dehydrogenase and related dehydrogenases</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1932</td>
<td>Phosphoserine aminotransferase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>0560</td>
<td>Phosphoserine phosphatase</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>14</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

### Cysteine, methionine and serine biosynthesis

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0197</td>
<td>Homoserine trans-succinylase</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0620</td>
<td>Methionine synthase II (cobalamin-independent)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Methionine synthase I (cobalamin-dependent),</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0646</td>
<td>methyltransferase domain</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>0111</td>
<td>Phosphoglycerate dehydrogenase and related dehydrogenases</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1932</td>
<td>Phosphoserine aminotransferase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>0560</td>
<td>Phosphoserine phosphatase</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>14</td>
<td>18</td>
<td></td>
</tr>
</tbody>
</table>

### Biosynthesis of branched-chain amino acids and threonine

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
<th>Entry Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0119</td>
<td>Isopropylmalate/homocitrate/citramalate synthases</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>0065</td>
<td>3-isopropylmalate dehydratase large subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>0066</td>
<td>3-isopropylmalate dehydratase small subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>0473</td>
<td>Isocitrate/isopropylmalate dehydrogenase</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1171</td>
<td>Threonine dehydratase</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thiamine pyrophosphate-requiring enzymes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase,]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0028</td>
<td>phosphonopyruvate decarboxylase</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>COG</td>
<td>Description</td>
<td>Occurrences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0440</td>
<td>Acetolactate synthase, small (regulatory) subunit</td>
<td>1 1 1 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0059</td>
<td>Ketol-acid reductoisomerase</td>
<td>1 1 1 3 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0129</td>
<td>Dihydroxyacids dehydratase/phosphogluconate dehydratase</td>
<td>2 2 2 16 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0115</td>
<td>Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lyase</td>
<td>3 3 2 13 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0527</td>
<td>Aspartokinases</td>
<td>1 1 1 3 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0136</td>
<td>Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lyase</td>
<td>1 1 1 5 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0460</td>
<td>Homoserine dehydrogenase</td>
<td>2 1 1 7 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Lysine biosynthesis**

- Dihydrodipicolinate synthase/N-acetyleneuraminase lyase
  | Occurrences |
|-------------|-------------|
| COG0329     | 2 1 2 4 6   |
| COG0289     | 1 1 1 4 7   |
| COG2171     | 1 1 1 5 7   |
| COG4992     | 1 2 1 6 8   |

**Aromatic amino acid biosynthesis**

- 3-deoxy-D-arabino-heptulosonate 7-phosphate
  | Occurrences |
|-------------|-------------|
| COG2876     | 0 0 1 0 0   |
| COG1830     | 0 0 0 0 0   |
| COG0337     | 1 1 1 3 4   |
| COG1465     | 0 0 0 0 0   |
| COG0710     | 0 0 0 0 0   |
| COG0757     | 1 1 1 1 1   |
| COG0169     | 1 1 1 2 2   |
| COG0703     | 1 1 1 1 1   |
| COG1685     | 0 0 0 0 0   |
| COG   | Description                                                                 | Entries | Homologs |(
|-------|------------------------------------------------------------------------------|---------|----------|--
<p>| COG0128 | 5-enolpyruvylshikimate-3-phosphate synthase                              | 2       | 1        | 1    | 5    | 7    |
| COG0082 | Chorismate synthase                                                        | 1       | 1        | 1    | 4    | 8    |
| COG1605 | Chorismate mutase                                                          | 1       | 1        | 1    | 0    | 1    |
|        | Chorismate mutase (alternative to chorismate mutase)                      | 0       | 0        | 0    | 0    | 0    |
| COG4401 | ABC-type amino acid transport/signal transduction systems, periplasmic component/domain (periplasmic cyclohexadienyl dehydratase) | 1       | 5        | 1    | 5    | 5    |
| COG0834 | Prephenate dehydratase                                                      | 1       | 1        | 1    | 7    | 6    |
| COG0077 | Prephenate dehydrogenase                                                   | 2       | 1        | 1    | 6    | 7    |
| COG0287 | Aspartate/tyrosine/aromatic aminotransferase (alternative to aspartate/tyrosine/aromatic aminotransferase) | 4       | 5        | 2    | 18   | 22   |
| COG1448 | Anthranilate/para-aminobenzoate synthases                                 | 0       | 2        | 0    | 0    | 0    |
|        | Anthranilate/para-aminobenzoate synthases component I                      | 2       | 2        | 2    | 12   | 16   |
| COG0512 | Anthranilate/para-aminobenzoate synthases component II                     | 1       | 1        | 1    | 7    | 7    |
|        | Anthranilate phosphoribosyltransferase                                      | 1       | 2        | 1    | 11   | 11   |
|        | Phosphoribosylanthranilate isomerase                                       | 1       | 1        | 1    | 5    | 9    |
|        | Indole-3-glycerol phosphate synthase                                       | 1       | 1        | 1    | 9    | 7    |
| COG0159 | Tryptophan synthase alpha chain                                            | 1       | 1        | 1    | 8    | 11   |
|        | Tryptophan synthase beta chain                                             | 1       | 1        | 1    | 9    | 13   |
|        | <strong>Histidine biosynthesis</strong>                                                 |         |          |      |      |      |
| COG0040 | ATP phosphoribosyltransferase involved in histidine biosynthesis           | 1       | 1        | 1    | 6    | 7    |
| COG3705 | ATP phosphoribosyltransferase                                              | 1       | 1        | 1    | 5    | 7    |
| COG0140 | Phosphoribosyl-ATP pyrophosphohydrolase                                    | 1       | 1        | 1    | 1    | 1    |
| COG0139 | Phosphoribosyl-AMP cyclohydrolase                                          | 1       | 1        | 1    | 2    | 2    |
|        | Phosphoribosylformimino-5-aminoimidazole                                    |         |          |      |      |      |
| COG0106 | carboxamide ribonucleotide (ProFAR) isomerase                             | 1       | 1        | 1    | 3    | 3    |
| COG0107 | Imidazol glycercophosphate synthase                                        | 1       | 1        | 1    | 2    | 2    |
| COG0118 | Glutamine amidotransferase                                                 | 1       | 1        | 1    | 3    | 3    |
| COG0131 | Histidinol-phosphate dehydratase                                           | 1       | 1        | 1    | 3    | 5    |
|        | Histidinol-phosphate/aromatic aminotransferase                            | 3       | 2        | 2    | 20   | 23   |</p>
<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0241</td>
<td>Histidinol phosphatase and related phosphatases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Histidinol phosphatase and related hydrolases of the PHP family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1387</td>
<td>PHP family</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0141</td>
<td>Histidinol dehydrogenase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

**Polyamine biosynthesis**

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>7</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0421</td>
<td>Spermidine synthase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1586</td>
<td>S-adenosylmethionine decarboxylase</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**DNA replication and chromosome partitioning**

ATPase involved in DNA replication initiation

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0593</td>
<td>(DnaA protein)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), A subunit</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV), B subunit</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>COG1484</td>
<td>DNA replication protein</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG3935</td>
<td>Putative primosome component and related proteins</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG3611</td>
<td>Replication initiation/membrane attachment protein</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0050</td>
<td>Replicative DNA helicase</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

ATPase involved in DNA replication (DNA polymerase delta prime subunit)

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0470</td>
<td>DNA polymerase III, gamma/tau subunits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DNA polymerase sliding clamp subunit (PCNA homolog)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>COG0592</td>
<td>DNA polymerase III, alpha subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>COG0358</td>
<td>DNA primase (bacterial type)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>COG0587</td>
<td>DNA polymerase III, alpha subunit</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>COG0847</td>
<td>DNA polymerase III, epsilon subunit and related 3'-5' exonucleases</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>COG2927</td>
<td>DNA polymerase III, chi subunit</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>COG3050</td>
<td>DNA polymerase III, psi subunit</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0328</td>
<td>Ribonuclease HI</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>COG0164</td>
<td>Ribonuclease HII</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>COG0749</td>
<td>DNA polymerase I - 3'-5' exonuclease and polymerase domains</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>COG2916</td>
<td>DNA-binding protein H-NS</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0776</td>
<td>Bacterial nucleoid DNA-binding protein</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>COG2901</td>
<td>Factor for inversion stimulation Fis, transcriptional</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>COG Code</td>
<td>Description</td>
<td>Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG3096</td>
<td>Uncharacterized protein involved in chromosome partitioning (MukB protein, alternative to Smc complex)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG3006</td>
<td>Uncharacterized protein involved in chromosome partitioning (MukF protein, alternative to Smc complex)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG3095</td>
<td>Uncharacterized protein involved in chromosome partitioning (MukE protein, alternative to Smc complex)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1196</td>
<td>Chromosome segregation ATPases (Smc complex protein Smc)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1354</td>
<td>Uncharacterized conserved protein (Smc complex protein ScpA)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1386</td>
<td>Predicted transcriptional regulator containing the HTH domain (Smc complex protein ScpB)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1192</td>
<td>ATPases involved in chromosome partitioning (chromosome segregation protein ParA)</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1475</td>
<td>Predicted transcriptional regulators (chromosome segregation protein ParB)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transcription (basal factors)</strong></td>
<td>DNA-directed RNA polymerase, beta subunit/140 kD</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0085</td>
<td>DNA-directed RNA polymerase, beta' subunit/160 kD</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0086</td>
<td>kD subunit</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0202</td>
<td>DNA-directed RNA polymerase, alpha subunit/40 kD</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0568</td>
<td>DNA-directed RNA polymerase, sigma subunit (sigma70/sigma32)</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1758</td>
<td>DNA-directed RNA polymerase, subunit K/omega</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0195</td>
<td>Transcription elongation factor</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0782</td>
<td>Transcription elongation factor</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG250</td>
<td>Transcription antiterminator</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0781</td>
<td>Transcription termination factor</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1158</td>
<td>Transcription termination factor</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ribosomal proteins (large subunit)</strong></td>
<td>Ribosomal protein L11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

26
| COG0081 | Ribosomal protein L1 | 1 | 1 | 1 | 4 | 8 |
| COG0087 | Ribosomal protein L3 | 1 | 1 | 1 | 1 | 0 |
| COG0088 | Ribosomal protein L4 | 1 | 1 | 1 | 1 | 0 |
| COG0089 | Ribosomal protein L23 | 1 | 1 | 1 | 1 | 0 |
| COG0090 | Ribosomal protein L2 | 1 | 1 | 1 | 1 | 0 |
| COG0091 | Ribosomal protein L22 | 1 | 1 | 1 | 0 | 0 |
| COG0093 | Ribosomal protein L14 | 1 | 1 | 1 | 0 | 1 |
| COG0094 | Ribosomal protein L5 | 1 | 1 | 1 | 0 | 1 |
| COG0097 | Ribosomal protein L6P/L9E | 1 | 1 | 1 | 0 | 1 |
| COG0102 | Ribosomal protein L13 | 1 | 1 | 1 | 2 | 2 |
| COG0106 | Ribosomal protein L16/L10E | 1 | 1 | 1 | 2 | 2 |
| COG0197 | Ribosomal protein L24 | 1 | 0 | 1 | 0 | 1 |
| COG0198 | Ribosomal protein L24 | 1 | 0 | 1 | 0 | 1 |
| COG0200 | Ribosomal protein L15 | 1 | 1 | 1 | 0 | 1 |
| COG0203 | Ribosomal protein L17 | 1 | 1 | 1 | 3 | 4 |
| COG0211 | Ribosomal protein L27 | 1 | 1 | 1 | 4 | 5 |
| COG0222 | Ribosomal protein L7/L12 | 1 | 1 | 1 | 4 | 5 |
| COG0227 | Ribosomal protein L28 | 1 | 1 | 1 | 2 | 2 |
| COG0230 | Ribosomal protein L34 | 0 | 0 | 0 | 0 | 0 |
| COG0244 | Ribosomal protein L10 | 1 | 1 | 1 | 3 | 7 |
| COG0254 | Ribosomal protein L31 | 1 | 1 | 1 | 3 | 3 |
| COG0255 | Ribosomal protein L29 | 1 | 1 | 1 | 1 | 2 |
| COG0256 | Ribosomal protein L18 | 1 | 1 | 1 | 0 | 1 |
| COG0257 | Ribosomal protein L36 | 0 | 0 | 0 | 0 | 0 |
| COG0261 | Ribosomal protein L21 | 1 | 1 | 1 | 3 | 4 |
| COG0267 | Ribosomal protein L33 | 1 | 1 | 1 | 2 | 2 |
| COG0291 | Ribosomal protein L35 | 1 | 1 | 1 | 2 | 4 |
| COG0292 | Ribosomal protein L20 | 1 | 1 | 1 | 3 | 5 |
| COG0333 | Ribosomal protein L32 | 1 | 1 | 1 | 3 | 4 |
| COG0335 | Ribosomal protein L19 | 1 | 1 | 1 | 1 | 1 |
| COG0359 | Ribosomal protein L9 | 1 | 1 | 1 | 4 | 4 |
| COG1825 | Ribosomal protein L25 (general stress protein Ctc) | 1 | 1 | 1 | 5 | 5 |
| COG1841 | Ribosomal protein L30/L7E | 1 | 1 | 1 | 0 | 1 |

**Ribosomal proteins (small subunit)**

<p>| COG0048 | Ribosomal protein S12 | 1 | 1 | 1 | 0 | 0 |
| COG0049 | Ribosomal protein S7 | 1 | 1 | 0 | 1 | 1 |
| COG0051 | Ribosomal protein S10 | 1 | 1 | 1 | 1 | 0 |</p>
<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0052</td>
<td>Ribosomal protein S2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0092</td>
<td>Ribosomal protein S3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0096</td>
<td>Ribosomal protein S8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0098</td>
<td>Ribosomal protein S5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0099</td>
<td>Ribosomal protein S13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0100</td>
<td>Ribosomal protein S11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0103</td>
<td>Ribosomal protein S9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0184</td>
<td>Ribosomal protein S15P/S13E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0185</td>
<td>Ribosomal protein S19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0186</td>
<td>Ribosomal protein S17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0199</td>
<td>Ribosomal protein S14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0228</td>
<td>Ribosomal protein S16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0238</td>
<td>Ribosomal protein S18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0268</td>
<td>Ribosomal protein S20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0360</td>
<td>Ribosomal protein S6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0522</td>
<td>Ribosomal protein S4 and related proteins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0539</td>
<td>Ribosomal protein S1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0828</td>
<td>Ribosomal protein S21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Translation factors**

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0290</td>
<td>Translation initiation factor 3 (IF-3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0532</td>
<td>Translation initiation factor 2 (IF-2; GTPase)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0361</td>
<td>Translation initiation factor 1 (IF-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0264</td>
<td>Translation elongation factor Ts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0480</td>
<td>Translation elongation factors (GTPases)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0505</td>
<td>GTPases - translation elongation factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Translation elongation factor P (EF-P)/translation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0231</td>
<td>initiation factor 5A (eIF-5A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0216</td>
<td>Protein chain release factor A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG1186</td>
<td>Protein chain release factor B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0233</td>
<td>Ribosome recycling factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0193</td>
<td>Peptidyl-tRNA hydrolase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0242</td>
<td>N-formylmethionyl-tRNA deformylase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Aminoacyl-tRNA synthetases**

<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>21</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG0008</td>
<td>Glutamyl- and glutaminyl-tRNA synthetases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0013</td>
<td>Alanyl-tRNA synthetase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0016</td>
<td>Phenylalanyl-tRNA synthetase alpha subunit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG0017</td>
<td>Aspartyl/asparaginyl-tRNA synthetases</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0018</td>
<td>Arginyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>COG0060</td>
<td>Isoleucyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Asp-tRNAAsn/Glu-tRNAGln amidotransferase B subunit (PET112 homolog)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>COG0072</td>
<td>Phenylalanyl-tRNA synthetase beta subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>COG0124</td>
<td>Histidyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>COG0162</td>
<td>Tyrosyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>COG0172</td>
<td>Seryl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>COG0173</td>
<td>Aspartyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>COG0180</td>
<td>Tryptophanyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>COG0215</td>
<td>Cysteinyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>COG0423</td>
<td>Glycyl-tRNA synthetase (class II)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0441</td>
<td>Threonyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>COG0442</td>
<td>Prolyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>COG0495</td>
<td>Leucyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>COG0525</td>
<td>Valyl-tRNA synthetase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Asp-tRNAAsn/Glu-tRNAGln amidotransferase C subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>COG0721</td>
<td>Preprotein translocase subunit SecA (ATPase, RNA helicase)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>COG0751</td>
<td>Glycyl-tRNA synthetase, beta subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>COG0752</td>
<td>Glycyl-tRNA synthetase, alpha subunit</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>COG1190</td>
<td>Lysyl-tRNA synthetase (class II)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>COG1384</td>
<td>Lysyl-tRNA synthetase (class I)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Protein folding and secretion**

<p>| COG0653 | Preprotein translocase subunit SecA (ATPase, RNA helicase) | 1 | 1 | 1 | 10 | 14 |
| COG1952 | Preprotein translocase subunit SecB | 1 | 1 | 1 | 4 | 5 |
| COG0342 | Preprotein translocase subunit SecD | 1 | 1 | 1 | 4 | 6 |
| COG0690 | Preprotein translocase subunit SecE | 1 | 1 | 1 | 3 | 4 |
| COG0341 | Preprotein translocase subunit SecF | 1 | 1 | 1 | 3 | 4 |
| COG1314 | Preprotein translocase subunit SecG | 1 | 1 | 1 | 3 | 3 |
| COG0201 | Preprotein translocase subunit SecY | 1 | 1 | 1 | 1 | 1 |
| COG2443 | Preprotein translocase subunit Sss1 | 0 | 0 | 0 | 0 | 0 |
| COG1862 | Preprotein translocase subunit YajC | 1 | 1 | 1 | 5 | 6 |
| COG0706 | Preprotein translocase subunit YidC | 1 | 1 | 1 | 3 | 3 |
| COG0681 | Signal peptidase I | 1 | 1 | 1 | 4 | 4 |</p>
<table>
<thead>
<tr>
<th>COG</th>
<th>Description</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG1400</td>
<td>Signal recognition particle 19 kDa protein</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0541</td>
<td>Signal recognition particle GTPase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>COG0552</td>
<td>Signal recognition particle GTPase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>COG0459</td>
<td>Chaperonin GroEL (HSP60 family)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>COG0234</td>
<td>Co-chaperonin GroES (HSP10)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>COG0597</td>
<td>Lipoprotein signal peptidase</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>COG1076</td>
<td>DnaJ-domain-containing proteins 1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>COG1043</td>
<td>Molecular chaperone</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>COG0544</td>
<td>Fkbp-type peptidyl-prolyl cis-trans isomerase (trigger factor)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>COG0545</td>
<td>Fkbp-type peptidyl-prolyl cis-trans isomerase 1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>COG0546</td>
<td>Molecular chaperone GTPs (small heat shock protein)</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>COG0435</td>
<td>Molecular chaperone, HSP90 family</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>COG0071</td>
<td>Molecular chaperone (small heat shock protein)</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COG0576</td>
<td>Molecular chaperone GTPs (HSP60 family)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>COG0326</td>
<td>Molecular chaperone, HSP90 family</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>COG0760</td>
<td>Parvulin-like peptidyl-prolyl isomerase</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>COG0652</td>
<td>Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>13</td>
</tr>
</tbody>
</table>

### 339 COGs

<table>
<thead>
<tr>
<th>300 COGs in Methylobacillus flagellatus</th>
<th>299 COGs in Methylibium petroleiphilum</th>
<th>293 COGs in Methylophilales bacterium</th>
<th>280 COGs in Methylophilales (methylamine)</th>
<th>287 COGs in Methylophilales (combined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>93.3333333333</td>
<td>95.6666666666</td>
<td>97.95221843</td>
<td>95.6666666666</td>
<td>97.95221843</td>
</tr>
</tbody>
</table>

A number of genes encoding a number of COGs representing ribosomal proteins are missing in the *Methylophilales* genome. These genes are notoriously unclonable\(^3\).
Supplementary Figure 2. Central metabolic pathways reconstructed from the composite genome of *M. mobilis*. Enzyme description and statistics are shown in Supplementary Table 4.
Supplementary Figure 3. Phylogenetic diversity of *fae* genes detected in metagenomic datasets described in this work (only complete or nearly complete sequences were included in analysis). Red, methylamine microcosm, green, methane microcosm, blue, methanol microcosm, yellow, formaldehyde microcosm, purple, formate microcosm. Fae, formaldehyde activating enzyme. Fae2–4, homologs of Fae with no demonstrated function.
Supplementary Table 6. Indels containing more than 2 genes, mapped on the chromosome of *M. flagellatus* (Genbank accession CP000284)

<table>
<thead>
<tr>
<th>Coordinates (bp)</th>
<th>Number of genes</th>
<th>Predicted function(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,339-13,223</td>
<td>6</td>
<td>Transport</td>
</tr>
<tr>
<td>16, 108-22,554</td>
<td>9</td>
<td>Transport</td>
</tr>
<tr>
<td>55, 702-57,905</td>
<td>3</td>
<td>Cell shape</td>
</tr>
<tr>
<td>160,871-192,124</td>
<td>30</td>
<td>Transport</td>
</tr>
<tr>
<td>212,171-221,143</td>
<td>9</td>
<td>Dehydrogenase/azurin</td>
</tr>
<tr>
<td>310,255-315,299</td>
<td>3</td>
<td>Oxidoreductase</td>
</tr>
<tr>
<td>327,628-334,485</td>
<td>6</td>
<td>Transport</td>
</tr>
<tr>
<td>329,063-386,592</td>
<td>15</td>
<td>Transport/regulation/oxidoreductase</td>
</tr>
<tr>
<td>390,398-397,303</td>
<td>9</td>
<td>Superoxide dismutase</td>
</tr>
<tr>
<td>403,781-407,978</td>
<td>4</td>
<td>Regulation</td>
</tr>
<tr>
<td>411,416-418,023</td>
<td>8</td>
<td>Transport</td>
</tr>
<tr>
<td>430,704-480,568</td>
<td>43</td>
<td>Amine metabolism</td>
</tr>
<tr>
<td>493,123-506,838</td>
<td>10</td>
<td>Adhesion</td>
</tr>
<tr>
<td>Start Position</td>
<td>End Position</td>
<td>Number</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------</td>
<td>--------</td>
</tr>
<tr>
<td>548,425-571,527</td>
<td>21</td>
<td>Secretion</td>
</tr>
<tr>
<td>579,009-590,147</td>
<td>11</td>
<td>Azurin/ oxidoreductase</td>
</tr>
<tr>
<td>601,520-607,385</td>
<td>6</td>
<td>Multisubunit Na+/H+ antiporter</td>
</tr>
<tr>
<td>610,488-617,262</td>
<td>6</td>
<td>Fructose bisphosphatase, short chain dehydrogenase</td>
</tr>
<tr>
<td>624,682-638,803</td>
<td>7</td>
<td>CRISPR and CRISPR-associated proteins</td>
</tr>
<tr>
<td>657,412-668,991</td>
<td>12</td>
<td>bb-type cytochrome oxidase, oxidoreductase</td>
</tr>
<tr>
<td>687,077-704,922</td>
<td>16</td>
<td>Transport</td>
</tr>
<tr>
<td>731,388-746,254</td>
<td>17</td>
<td>Transport</td>
</tr>
<tr>
<td>770,311-786,617</td>
<td>10</td>
<td>Transport</td>
</tr>
<tr>
<td>802,677-807,100</td>
<td>5</td>
<td>Methyltransferase, glycosyltransferase</td>
</tr>
<tr>
<td>866,145-1,018,175</td>
<td>171</td>
<td>Prophage sequences surrounding an identical repeat of 133 Kb</td>
</tr>
<tr>
<td>1,029,922-1,034,827</td>
<td>6</td>
<td>Esterase, hydrolase</td>
</tr>
<tr>
<td>1,065,143-1,068,125</td>
<td>5</td>
<td>Multidrug resistance</td>
</tr>
<tr>
<td>1,076,530-1,089,872</td>
<td>10</td>
<td>Putative prophage</td>
</tr>
<tr>
<td>1,123,945-1,127,704</td>
<td>5</td>
<td>Transport</td>
</tr>
<tr>
<td>1,133,038-1,141,895</td>
<td>9</td>
<td>Cytochrome bd ubiquinol oxidase</td>
</tr>
<tr>
<td>Start Position</td>
<td>End Position</td>
<td>Length</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------</td>
<td>--------</td>
</tr>
<tr>
<td>1,176,779-1,186,415</td>
<td>8</td>
<td>Transport</td>
</tr>
<tr>
<td>1,193,775-1,297,144</td>
<td>88</td>
<td>Restriction/ modification</td>
</tr>
<tr>
<td>1,304,198-1,310,205</td>
<td>7</td>
<td>Regulation</td>
</tr>
<tr>
<td>1,345,124-1,397,790</td>
<td>51</td>
<td>Polysaccharide biosynthesis and transport</td>
</tr>
<tr>
<td>1,440,674-1,466,064</td>
<td>19</td>
<td>Polysaccharide biosynthesis</td>
</tr>
<tr>
<td>1,522,784-1,573,650</td>
<td>52</td>
<td>Amylase, cytochrome oxidase, oxidoreductase, putative prophage</td>
</tr>
<tr>
<td>1,594,568-1,603,261</td>
<td>8</td>
<td>Transport</td>
</tr>
<tr>
<td>1,642,034-1,652,046</td>
<td>9</td>
<td>Transport</td>
</tr>
<tr>
<td>1,656,940-1,662,626</td>
<td>4</td>
<td>Transport</td>
</tr>
<tr>
<td>1,666,030-1,670,536</td>
<td>3</td>
<td>Transport</td>
</tr>
<tr>
<td>1,712,573-1,724,799</td>
<td>13</td>
<td>Transport</td>
</tr>
<tr>
<td>1,839,792-1,843,884</td>
<td>4</td>
<td>Transport</td>
</tr>
<tr>
<td>1,852,729-1,860,211</td>
<td>9</td>
<td>Regulation</td>
</tr>
<tr>
<td>1,891,635-1,943,362</td>
<td>44</td>
<td>Urea metabolism</td>
</tr>
<tr>
<td>1,949,155-1,955,403</td>
<td>8</td>
<td>Regulation</td>
</tr>
<tr>
<td>Patent Range</td>
<td>Length</td>
<td>Function</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2,145,041-2,185,147</td>
<td>38</td>
<td>Polysaccharide biosynthesis, methanol dehydrogenase</td>
</tr>
<tr>
<td>2,319,107-2,331,500</td>
<td>14</td>
<td>Transport</td>
</tr>
<tr>
<td>2,338,624-2,350,096</td>
<td>11</td>
<td>Transport</td>
</tr>
<tr>
<td>2,371,334-2,376,055</td>
<td>3</td>
<td>Transport</td>
</tr>
<tr>
<td>2,387,080-2,390,286</td>
<td>4</td>
<td>Transport</td>
</tr>
<tr>
<td>2,451,215-2,456,811</td>
<td>5</td>
<td>Transport</td>
</tr>
<tr>
<td>2,497,672-2,507,795</td>
<td>8</td>
<td>Transport</td>
</tr>
<tr>
<td>2,518,800-2,531,513</td>
<td>11</td>
<td>Transport</td>
</tr>
<tr>
<td>2,539,472-2,583,043</td>
<td>23</td>
<td>Transport</td>
</tr>
<tr>
<td>2,587,455-2,596,334</td>
<td>9</td>
<td>Type II secretion</td>
</tr>
<tr>
<td>2,608,188-2,621,238</td>
<td>8</td>
<td>Transport</td>
</tr>
<tr>
<td>2,723,525-2,732,193</td>
<td>4</td>
<td>Regulation</td>
</tr>
<tr>
<td>2,735,722-2,739,760</td>
<td>4</td>
<td>Transport</td>
</tr>
<tr>
<td>2,752,465-2,769,855</td>
<td>16</td>
<td>Transport</td>
</tr>
<tr>
<td>2,786,888-2,798,698</td>
<td>13</td>
<td>Regulation</td>
</tr>
<tr>
<td>2,802,468-2,820,659</td>
<td>12</td>
<td>Polysaccharide degradation</td>
</tr>
<tr>
<td>Region</td>
<td>Length</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------</td>
<td>--------------------</td>
</tr>
<tr>
<td>2,826,963-2,840,605</td>
<td>10</td>
<td>Regulation</td>
</tr>
<tr>
<td>2,843,981-2,873,347</td>
<td>24</td>
<td>Hydrocarbon degradation</td>
</tr>
<tr>
<td>2,886,923-2,951,255</td>
<td>64</td>
<td>Putative prophage</td>
</tr>
</tbody>
</table>
Supplementary Table 7. Functional distribution of indels of more than two genes detected in the composite genome of *M. mobilis*

<table>
<thead>
<tr>
<th>Functional category</th>
<th>Number of indels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport</td>
<td>18</td>
</tr>
<tr>
<td>Enzyme</td>
<td>14</td>
</tr>
<tr>
<td>Regulation/signaling</td>
<td>10</td>
</tr>
<tr>
<td>Polysaccharide biosynthesis</td>
<td>6</td>
</tr>
<tr>
<td>Hypothetical proteins</td>
<td>6</td>
</tr>
<tr>
<td>Secretion (pili)</td>
<td>3</td>
</tr>
<tr>
<td>Lipid biosynthesis</td>
<td>2</td>
</tr>
<tr>
<td>Prophage</td>
<td>2</td>
</tr>
</tbody>
</table>
Supplementary Figure 4. Gene cluster encoding enzymes for the citric (incomplete) and methylcitric acid cycles in the composite genome of *M. mobilis*, compared to a cluster in *Methyphilales HTCC2181*\(^\text{38}\), and their schematic representation. Genes present in *M. flagellatus* are in blue and genes absent in *M. flagellatus* are in red.
<table>
<thead>
<tr>
<th>Electron transfer system</th>
<th><em>M. mobilis</em></th>
<th><em>M. flagellatus</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>NADH dehydrogenase (Complex I)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cytochrome oxidase (bb type)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Succinate dehydrogenase (Complex II)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>NADH-ubiquinol oxidoreductase (Rfn system)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Ubiquinol cytochrome c reductase (bc type)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cytochrome c oxidase (aa3 type)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cytochrome C5 oxidase (o type)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cytochrome c oxidase (cb type)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Nitric oxide reductase</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Na/H antiporter NADH quinone dehydrogenase</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cytochrome oxidase (cbb type)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cytochrome d ubiquinol oxidase</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cytochrome c oxidase (o type)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Cytochrome c oxidase</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Genes in question were identified in the *M. mobilis* and *M. flagellatus* genomes using word search against the annotated genomes in IMG/M. The annotations were verified by BLAST searches against the non-redundant (NCBI) and protein (SwissProt) databases. Reciprocal BLAST analyses were done between the genomes of *M. mobilis* and *M. flagellatus* and gene homologs (more than 30% amino acid identity) were identified.
**Supplementary Table 9.** Coverage of *Methylothenera mobilis* genomes in methane, methanol and formaldehyde microcosms, based on comparisons with the composite genome (12,719 protein queries).

<table>
<thead>
<tr>
<th>% identity (protein)</th>
<th>Number of proteins</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Methane</td>
<td>Methanol</td>
<td>Formaldehyde</td>
<td>Combined</td>
</tr>
<tr>
<td>90</td>
<td>1,245</td>
<td>1,612</td>
<td>706</td>
<td>3,563</td>
</tr>
<tr>
<td>80</td>
<td>2,450</td>
<td>3,116</td>
<td>1,638</td>
<td>7,204</td>
</tr>
<tr>
<td>70</td>
<td>4,260</td>
<td>4,804</td>
<td>3,170</td>
<td>12,234</td>
</tr>
<tr>
<td>60</td>
<td>7,143</td>
<td>7,105</td>
<td>5,854</td>
<td>20,102</td>
</tr>
<tr>
<td>50</td>
<td>11,738</td>
<td>10,809</td>
<td>10,216</td>
<td>32,763</td>
</tr>
</tbody>
</table>
**Supplementary Figure 5. Genomic structure of novel phages from lake Washington, the *M. flagellatus* prophage and bacteriophage PM2.** The circular genomes were linearized to align with the prophage. Different colors indicate different degrees of gene conservation (red, conserved in all; blue, conserved in Lake Washington phages and *M. flagellatus* prophage; turquoise, conserved in Lake Washington phages; light blue, conserved in two of the three phages; blank, unique)
Supplementary Figure 6. Gene cluster encoding pilus functions unique to *M. mobilis* from the methylvamine microcosm. Red, genes encoding pilus functions; yellow, regulatory genes; grey, genes encoding hypothetical proteins.
Supplementary Figure 7. Central metabolic pathways reconstructed from the composite genome of _M. tundripaludum_. Gene designations are as in Supplementary Fig. 2. _pmo_, particulate methane monooxygenase; _mxa_, methanol dehydrogenase functions; _pqq_, PQQ biosynthesis; _sucAB_, α-ketoglutarate dehydrogenase; _fba_, fructose bisphosphate aldolase.
### Supplementary Table 10. Phylum-specific binning statistics

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Total contigs</th>
<th>Total kb</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Methane dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>392</td>
<td>757</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>435</td>
<td>684</td>
</tr>
<tr>
<td>Methylophilaceae</td>
<td>85</td>
<td>139</td>
</tr>
<tr>
<td>Methylococcaceae</td>
<td>89</td>
<td>269</td>
</tr>
<tr>
<td><strong>Methanol dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>204</td>
<td>309</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>498</td>
<td>836</td>
</tr>
<tr>
<td>Methylophilaceae</td>
<td>316</td>
<td>56</td>
</tr>
<tr>
<td>Methylococcaceae</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td><strong>Methylamine dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>676</td>
<td>1,210</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>4337</td>
<td>11,580</td>
</tr>
<tr>
<td>Methylophilaceae</td>
<td>4079</td>
<td>11,160</td>
</tr>
<tr>
<td>Methylococcaceae</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Comamonadaceae</td>
<td>66</td>
<td>96</td>
</tr>
<tr>
<td><strong>Formaldehyde dataset</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>55</td>
<td>84</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>300</td>
<td>453</td>
</tr>
<tr>
<td>Family</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>Methylophilaceae</td>
<td>36</td>
<td>56</td>
</tr>
<tr>
<td>Rhodocyclaceae</td>
<td>64</td>
<td>105</td>
</tr>
</tbody>
</table>

**Combined assembly**

<table>
<thead>
<tr>
<th>Family</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betaproteobacteria</td>
<td>8,788</td>
<td>19,500</td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>1,600</td>
<td>2,800</td>
</tr>
<tr>
<td>Methylophilaceae</td>
<td>5,000</td>
<td>13,300</td>
</tr>
<tr>
<td>Methylococcaceae</td>
<td>118</td>
<td>340</td>
</tr>
<tr>
<td>Burkholderiaceae</td>
<td>787</td>
<td>1,570</td>
</tr>
<tr>
<td>Comamonadaceae</td>
<td>675</td>
<td>1,100</td>
</tr>
<tr>
<td>Rhodocyclaceae</td>
<td>475</td>
<td>882</td>
</tr>
</tbody>
</table>
Supplementary Table 11. MtaB translated from the Lake Washington metagenome, compared to homologs from *Verrucomicrobia*, methylotrophic *Clostridia* and methylotrophic Archaea

<table>
<thead>
<tr>
<th>Gene ID/ Microcosm/ Amino acids</th>
<th>Amino acid identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Opitutaceae strain TAV2</em></td>
</tr>
<tr>
<td>2006235727/ methane/ 108</td>
<td>45</td>
</tr>
<tr>
<td>2006268647/ methane/ 290</td>
<td>74</td>
</tr>
<tr>
<td>2006329978/ methanol/ 279</td>
<td>69</td>
</tr>
<tr>
<td>2006441833/ formaldehyde/ 245</td>
<td>50</td>
</tr>
<tr>
<td>2006478187/ formaldehyde/ 192</td>
<td>58</td>
</tr>
</tbody>
</table>