

The taxonomic and phenotypic diversity of *Methylobacterium* species.

2022, March

Ola ALESSA

Graduate School of Environmental and Life Science

(Doctor's Course)

OKAYAMA UNIVERSITY

Table of Contents

Abstract.....	4
➤ 1. Introduction.....	8
1.1. Methylophilic bacteria and plants.....	8
1.2. Methylophilic genes.....	9
1.3. Taxonomic status of the PPFMs	11
1.4. Identification, classification of the PPFMs in the old and new whole-genome studies era	12
1.5. Species with special characteristics in genus <i>Methylobacterium</i>	13
➤ 2. The methodology.....	15
2.1. DNA extraction, genome sequencing, annotation and the 16S rRNA gene phylogeny	15
2.2. Whole genome-based analysis and pan-genome analysis.....	17
2.3. Functional annotation	17
2.4. Genomics of methylophilic, phyllosphere adaptation, and plant colonization	18
2.5. Phylogenetic tree construction of PQQ-dependent alcohol dehydrogenase homologues	19
2.6. Phenotypic characteristics.....	19
2.7. Antifungal activity	21
2.8. Indole acetic acid (IAA) production through L-tryptophan pathway	21
2.9. Elicitation of rice cell.....	22
2.10. Data availability.....	22
➤ 3. Results and discussion	23
3.1. General consideration	23
3.2. Genome assembling, annotation, and quality check	23
3.3. 16S rRNA gene-based phylogeny.....	24
3.4. Average nucleotide identity, digital DNA-DNA hybridization	25
3.5. Whole genome-genome phylogeny	26
3.6. Pan-genome analysis.....	26
3.7. Pathway and module analysis.....	27
3.8. Methylophilic genes cluster	28
3.9. MDH and alcohol dehydrogenase (ADH) homologues	31
3.10. Phenotyping of PPFMs	34
3.11. Secondary metabolites clusters.....	38
3.12. Plant growth-promoting genes.....	38
3.13. Indole acetic acid (IAA) production through the L-tryptophan pathway	39
3.14. Antifungal activity	40
3.16. Elicitation of rice cells	40
➤ 4. Discussion	41
4.1. General discussion	41

4.2. Taxonomic consequences: synonymous species, new combinations and emendation of the genus ..	44
➤ 5. Conclusion.....	47
➤ 6. Acknowledgements.....	48
➤ 7. Figure captions	49
➤ 8. Tables captions.....	52
➤ 9. References	54

Abstract

The pink pigmented facultative methylotrophs (PPFM), a major bacterial group found in the plant phyllosphere, comprises two genera *Methylobacterium* and recently established *Methylobacterium* (Green 2018). PPFMs colonize plants to establish a symbiotic relationship. They utilize methanol emitted from the plants through the function of two paralogous methanol dehydrogenases (MDHs); MxaFI, Ca^{2+} -dependent MDH and XoxF, lanthanide (Ln^{3+})-dependent MDH. In return, they produce phytohormones (auxin and cytokinin) and fix nitrogen, that contribute to plant growth and increase plant tolerance to biotic and abiotic stresses. The two genera contain 63 species that have been isolated from different habitats mostly of plants. It is unknown what kind of genetic trait(s) determine their habitat. Although most of the PPFMs are pink-pigmented and able to utilize methanol, some species are non-pigmented and the others show no methylotrophic ability. These things leave big questions what truly defines the PPFMs. Because sequencing of bacterial genomes is now more affordable recently, it may be possible to answer to these questions by comparative genomics. In the present study, a comprehensive comparative genomics approach was employed to reveal their phylogenetic relationship and to explain their genotypic differences that confer their different phenotypes.

I newly sequenced the genomes of 29 relevant type strains to complete a dataset for almost all validly published type strains in the genera. All assemblies including those downloaded from the database were filtered based on >90% completeness and <5% contamination. In total, PPFM type strains and additional 16 type strains from other genera in the *Methylobacteriaceae* family were used in the analyses. Average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) analyses suggested that *Mb. phyllosphaerae* and *Mb. oryzae* are later heterotypic synonyms of *M. fujisawaense*, and that *Mr. populi* is a later heterotypic synonym of *Mr. thiocyanatum*. Thus, the completion of genome information resulted in finding of previously erroneously assigned species.

The phylogeny of this bacterial group was studied and constructed with three different approaches; Genome Blast Distance Phylogeny (GBDP), the alignment of 711 core genes, and the presence and

absence matrix of the genes in the pan genome (55128 genes). The three trees were consistent each other, but not with the previously proposed 16S rRNA gene phylogeny (Green 2018). The 16S rRNA phylogeny divided the PPFMs into three clades A, B, and C. Clade A represented the “core” *Methylobacterium* species, and clade B represented *Methylobacterium* species. The main difference between the whole genome phylogeny and the 16S rRNA phylogeny was that the two genera were not separated clearly, and that *Mb. organophilum* NBRC 15689^T, the type species of *Methylobacterium* that was newly sequenced in this study, clustered with *Methylobacterium* group. These phylogenetic trees also showed that clade A is separated into 6 sub-clades, one of which I named as B2 due to its close phylogeny with clade B. Finally, Clade C members were clearly separated from the other two genera, and their genome size and GC% were higher than the others. One more interesting characteristic in clade C members is a lack of methylophily, which is detailed below.

By investigating the enriched KEGG pathways and modules for the genomes, some pathways were uniquely found in the PPFMs compared with other species in the *Methylobacteriaceae* family. These pathways included nitrate reduction and assimilation modules, and anoxygenic photosystem II. Between the PPFMs clades, the members of clades B and C formed distinct clusters. But some species from sub-clades A1, A2, and A3 clustered differently, which may mean that they have acquired or lost some modules keeping their core genes less modified, during evolution and adaptation to their habitats. The sub-clades A4 and A5 maintained a close phylogeny throughout the analysis.

Methylophily genes were another unique feature of the PPFMs, and *xoxF* was conserved in most of the genomes of PPFMs. Genes related to Ln^{3+} uptake and transport (*lut*) genes were also found in all the species but not in the other genera. Thus, the conservation of *xox* and methanol utilization ability within these bacterial groups is the hallmark to differentiate from the others. Other important genes related to C1 utilization such as formaldehyde and formate oxidation, H₄F pathway, folate synthesis, serine cycle, and ethylmalonyl-CoA pathway were all well conserved in the PPFMs genomes. Methylamine utilization gene clusters which were previously suggested as a main feature in clade B species was also found in many species of the PPFMs. Two clusters of N-methylglutamate (NMG) and methylamine dehydrogenase

(MaDH) were identified as pathways for methylamine utilization.

The genomes of PPFMs encode 2~7 (average 3.7) genes for PQQ-dependent alcohol dehydrogenases that include MxaF and XoxF, and their phylogeny showed distinct correlation with their genomic phylogeny. A set of 254 PQQ-ADH and MDH orthologs proteins were found in the pan-genome, and they were divided into eight major groups. The most abundant group was XoxF5-type MDH. It was also found that the ADHs and MDHs from clades A and B were close, but those from clade C were phylogenetically distant.

The strains were subjected to growth phenotyping. The data showed that some species cannot utilize methanol due to the lack of MxaF. But most of such species still harbour XoxF that enables growth on methanol in the presence of La^{3+} and is especially necessary for the species which were previously regarded as non-methylotrophic bacteria. Methylamine utilization ability was thought to be a unique feature of clade B members (*Methylobacterium* genus) previously, but I found that 11 strains from clade A were also able to grow on methylamine. Interestingly, they included the proposed sub-clade B2. A variation regarding assimilation of different carbon sources as a source of energy was noticed, while sub-clade A1, A2, and A3 strains were negative or weak, sub-clade A4 and A5 strains were stronger. Most strains were urease positive, and few strains from clade A4 and A5 were able to reduce nitrate.

All clade B members were able to grow on C1 compounds but some of them showed a weaker growth on complex mediums or complex carbon sources. This might indicate that these strains are stricter methylotrophs compared with the other PPFMs strains. Clade C members showed strong growth on different carbon sources and complex mediums, but most of them needed Ln^{3+} supplementation to grow on methanol.

Most strains produced indole acetic acid (IAA), of which concentration ranged within 8-29 $\mu\text{g/ml}$. The strains (especially *Mb. gnaphalii* DSM 24027^T, *Mb. oxalidis* DSM 24028^T, *Mb. brachythecii* DSM 24105^T, and *Mb. dankookense* DSM 22415^T) were the highest producers, making them an ideal plant growth promoter. I also found that all PPFM strains did not induce any immune response in rice culture cells and

were able to inhibit the growth of the pathogenic fungi *Fusarium oxysporum* f. sp. *lycopersici* *in vitro*, which can be an indirect means to promote plant growth. The genome mining data showed that chitinase, cellulase, pectinase, and several secondary metabolites clusters such as terpene, siderophores, non-ribosomal peptide synthases, and polyketide synthase were conserved across PPFM strains. These functions may contribute to the antifungal activity.

Overall, this study provides the first inclusive insight into the phylogeny, genotypes, and phenotypes of the PPFMs through completion of their genome data. It also approaches many factors related to their symbiotic interaction with the plant and key genes related to their adaptation to phyllosphere. It also divides the PPFMs into more homogenous groups, and embosses their unique features. In addition, it also showed that the genomes of these bacteria encoded a number of ADH and MDH genes that should be further biochemically characterized. Most importantly the *xoxF* gene clusters and Ln^{3+} transport genes were conserved in almost all species. The conservation of the cluster clearly emphasizes the importance of methanol utilization ability for these bacterial species to grow and flourish where methanol is abundant.

➤ 1. Introduction

1.1. Methylotrophic bacteria and plants

The ability to utilize single carbons with no carbon-carbon bond as a source of energy is a unique feature found in a group of organisms known as the methylotrophs. This definition includes a wide range of organisms from all forms of life such as yeast, fungi, archaea, and bacteria. Methylotrophic bacteria are divided based on the oxidation reactants (methane or methanol) in two main categories; methanotrophs and methylotrophs. Methanotrophs are a unique subset of methylotrophic bacteria capable of starting the C1 oxidation reaction from methane, a single carbon gas. This exclusive feature is due to the presence of methane monooxygenase (MMO), the key enzyme for methane oxidation to methanol. MMO type varies depending on the methanotroph species or type, and it can be either soluble (sMMO) or particulate (pMMO). Methylotrophs can be also divided into obligate methylotrophs, or facultative methylotrophs based on the flexibility of utilizing other carbon sources than C1 compounds. Obligate methylotrophs possess no ability to grow on compounds with carbon-carbon bonds, and grow only on single carbons or C1 compounds. Facultative methylotrophs are more flexible and have a broader range of compound to grow on (Anthony, 1983). They possess different carbon assimilation pathways, while obligate methylotrophs possess only the ribulose monophosphate (RuMP) pathway, the facultative methylotrophs possess either RuMP, serine cycle, or Calvin Benson cycle (Lidstrom & Stirling, 1990)

In nature, both of methanotrophs and methylotrophs are highly important for global cycle of methane, a potent greenhouse gas. Methane is produced by methanogenic archaea, and recently the plant cells were identified as a source of methane too (Bruhn et al., 2012). Methanol is either released by the oxidation of methane or released from plant pectin, a major component of plant cell wall, through the function of pectin methyl esters (MacDonald & Fall, 1993; Nemecek-Marshall et al., 1995). Methanol and other carbons are released from plant stomata to its surroundings, making the phyllosphere a rich environment for methylotrophic bacteria.

One of the most dominant methylotrophs on the plant phyllosphere is *Methylobacterium* species. They

represent 10-20% of bacterial community on the plant phyllosphere and found to be associated with a variety of plant hosts (Delmotte et al., 2009; Knief et al., 2008). Based on their pink color and their ability to utilize methanol and multi-carbon substances, they are called the pink-pigmented facultative methylotrophs (PPFMs). The PPFMs are ubiquitous in nature and can be found either living freely in water, soil, and air (Gallego et al., 2005; Kim et al., 2019; Park et al., 2018), or associated with plant rhizosphere (Grossi et al., 2020) and phyllosphere (Munusamy Madhaiyan et al., 2011; Wellner et al., 2013). When associated with the plant the PPFMs colonize the plant as epiphytes or endophytes (Jorge et al., 2019; Verginer et al., 2010), and in some cases they reside in intracellular space of plant meristematic cells (Pirttilä et al., 2000) (or root nodules (Jourand et al., 2004), to establish a symbiotic relationship. It is believed that the PPFMs are persistent plant colonizer, transmitted from one plant to its progeny through the seed (Mizuno et al., 2013). The ability to utilize methanol is considered as extra advantage for PPFM species to compete on the plant phyllosphere (Sy et al., 2005). As a plant symbiont the PPFMs are capable of producing phytohormones (auxin and cytokinin) and secondary metabolites (Ivanova et al., 2001; Lidstrom & Chistoserdova, 2002), and fix nitrogen (Jourand et al., 2004), that contribute to plant growth (Abanda-Nkpwatt et al., 2006). They also generate fruit flavour (Verginer et al., 2010) and contribute to plant tolerance to biotic and abiotic stresses (Ardanov et al., 2012; Jorge et al., 2019). *Methylobacterium* species are also found as symbionts of microalga and mosses (Krug et al., 2020; Tani et al., 2012).

1.2. Methylotrophy genes

Methylotrophy genes refer to countless number of genes that are involved in the single carbon assimilation pathways. Because methanol is the most abundant in the PPFMs environment, it is the focus for much research. The first step of methanol assimilation pathway involves the oxidation of methanol to formaldehyde, the enzyme responsible for this step is methanol dehydrogenase (MDH) (Figure 1). *Methylobacterium* species possess two types of methanol dehydrogenase; a Ca^{2+} -dependent methanol dehydrogenase, and an Ln^{3+} -dependent methanol dehydrogenase. The Ca^{2+} -dependent MDH is a heterotetrametric protein containing pyrroloquinoline quinone (PQQ) and was believed to be the sole indispensable enzyme for methanol oxidation in *Methylobacterium extorquens* AM1. This enzyme is

encoded by a large cluster known as *mx*a cluster or *mox* cluster previously (Anderson et al., 1990; Anthony, 1983; Nunn & Lidstrom, 1986). At least 10 genes from the *mx*a cluster were found to be necessary for the strain AM1 to grow on methanol (Nunn & Lidstrom, 1986). These genes include *mx*a*F* and *mx*a*I* that encode the large subunit (66-kDa) and small (8.5-kDa) subunit, respectively. The gene *mx*a*G* encodes a cytochrome c protein necessary for the electron transfer from the cofactor of MxaFI, pyrroloquinoline quinone (PQQ). The genes *mx*a*LKCA* are known to be essential for Ca²⁺ insertion into MxaF (Nunn et al., 1989).

Another region related to C1 utilization in the strain AM1 genome was identified (Chistoserdova & Lidstrom, 1997). The region contained an *mx*a*F* homologue gene with a 50% amino acid sequence identity, and homologous genes for *mx*a*G* and *mx*a*J* (solute binding protein). Although the putative protein encoded by the *mx*a*F* homologue seemed to resemble MxaFI, its function had been unknown and was called *xoxF*. The trials to reveal the enzymatic function of XoxF had been unsuccessful although it had been shown to be important for MxaF expression (Skovran et al., 2011). It was a serendipity to find that lanthanide metals promoted the growth of *Methylobacterium* species, the cells exhibited higher MDH activity, the purified MDH was XoxF, and the purified enzyme fraction contained the lanthanide (Hibi et al., 2011). Despite that lanthanide had not been believed to be biologically important, now we know lanthanide-dependent MDH, XoxF, is widely distributed in the genomes of Gram negative methylotrophs (Keltjens et al., 2014).

The availability of the full genome sequence of the strain AM1 allowed the discovery of another MDH-homologue gene, *xoxF2* that share 50% and 90% amino acid sequence identity with *mx*a*F* and *xoxF*, respectively (Vuilleumier et al., 2009). Both *xoxF1* (former *xoxF*) and *xoxF2* genes were found to be active in the presence of a range of rare earth elements (Vu et al., 2016). In addition, a protein encoded by *exaF* in strain AM1 was shown to be lanthanide-dependent ethanol dehydrogenase that is also active on methanol and formaldehyde. The presence of a number of Ln³⁺-dependent enzymes, and the fact that lanthanide is rare earth element (REE) pose a huge question: what biological function do these enzymes have on the plant phyllosphere, and whether they could be active in nature. This mystery was recently

uncovered by the presence of several REE element detected in the phyllosphere (Ochsner et al., 2019), and the fact that proteomic studies of bacterial communities colonizing plant phyllosphere have shown that both types of MDHs from *Methylobacterium* species were highly enriched (Delmotte et al., 2009).

Methylated amines are another important C1 carbon source. Gram negative bacteria, including some species of the PPFMs, utilize these compounds either by methylamine dehydrogenase (MaDH) or N-methylglutamate (NMG) pathway. Unlike the former pathway that is direct (one step) oxidation reaction of methylamine to formaldehyde, the latter is indirect and involves many steps. The genes related to MaDH activity, are clustered in the strain AM1 genome, and 11 genes of this cluster (*mau*) are needed for this reaction (Chistoserdov et al., 1994). The NMG pathway requires three enzymes; methylglutamate synthase, N-methylglutamate dehydrogenase, and gamma-glutamylmethylamide synthetase, and these enzymes are encoded by 8 genes in the NMG pathway cluster (Latypova et al., 2010). The NMG pathway is not only abundant in methylotrophic bacteria (Kröber et al., 2021) but can also be found in non-methylotrophic bacteria, which rely on it for their nitrogen source (Chen et al., 2010). Both gene sets were found in the strain AM1 genome, but in the presence on methylamine the MaDH pathway was the one expressed (Chistoserdov et al., 1994). Reports have shown that other closely related strains of *Mr. extorquens* were found to possess only the NMG pathway (Gruffaz et al., 2014; Nayak & Marx, 2014).

1.3. Taxonomic status of the PPFMs

Taxonomically the PPFMs belong to the class of alpha-protobacteria and consists of 63 validated species, within the *Methylobacteriaceae* family. The first PPFM strain described in the literature was isolated from earthworm contents and named as '*Bacillus extorquens*' (Bassalik 1913). The genus *Methylobacterium* was established with *Methylobacterium organophilum* as the only species within it (Patt et al., 1976). Since then, the number of new species have been increasing (<https://lpsn.dsmz.de/genus/methylobacterium>; Hedlund et al., 2021; Maeng et al., 2021). Various attempts have been made to approach PPFMs on a taxonomic level (Green & Bousfield, 1982; Hood et al., 1987; Tani et al., 2015; Tsuji et al., 1990). Green and Ardley recently divided *Methylobacterium*

species into three clades (A, B, and C) based on phenotypic characteristics, multi-locus sequence analysis (MLSA), and 16S rRNA gene sequence analysis. Clade A represents the core *Methylobacterium*. They proposed a new genus *Methylobacterium* for clade B species, which includes the most well-studied model *Methylobacterium extorquens* strain AM1. The members in clade C were considered as species *incertae sedis*: most of the species within it cannot utilize methanol, and some species have a unique ability such as nodulation in legumes (Jourand et al., 2004). A recent large study on entire alpha-proteobacterial genomes, however, suggested re-consideration of *Methylobacterium* as *Methylobacterium* until a greater number of genome data of the species become available (Hördt et al., 2020).

1.4. Identification, classification of the PPFMs in the old and new whole-genome studies era

Identification, classification, and species delineation of prokaryotes require various analyses on phenotypic features like morphological and biochemical differences, genomic DNA homology (DNA-DNA hybridization, DDH), genome GC% content (Johnson, 1973; McCarthy & Bolton, 1963; Schildkraut et al., 1961), or the 16S rRNA gene polymorphism (Weisburg et al., 1991). Although the methodologies for these analyses have been long used, experimental outcomes may be unstable. For example, DDH requires special experimental setups and depends highly on the quality and purity of genomic DNA. Fatty acid methyl ester (FAME) analysis is affected much by culture condition. 16S rRNA gene sequencing is widely used as gold standard but sequencing of the PCR product masks the intragenomic heterogeneity of the gene copies. On the other hand, recent availability of genome data paved the way for more robust high-throughput analyses on bacterial taxonomy. Sophisticated methods and tools including average nucleotide identity (ANI) and *in silico* (digital) DNA-DNA hybridization (dDDH) with a species cut-off 94-95% and 70%, respectively, are now taking over the old methods (Auch et al., 2010; Jain et al., 2018; Meier-Kolthoff et al., 2013; Meier-Kolthoff & Göker, 2019). Many annotation tools and pathway/module analysis pipelines for bacterial genomes have also been developed to characterize the genotype of bacterial species. Nowadays, the cost of whole-genome shotgun sequencing has become more affordable and enabled us to compare large sets of genome data, to understand their function and evolution history

through the core- and pan-genome analysis (Chaudhari et al., 2016; Tettelin et al., 2005).

To date, no comprehensive comparative genomic analysis has been done for *Methylobacterium* and *Methylobacterium* species due to the unavailability of genome assemblies of all validly published species. In this study, I sequenced the genomes of un-sequenced type strains and those of low quality. Sixty-two genome assemblies of type strains were used for comparative genomic analyses. The genome data set was used to assess their phylogeny based on the genomes but not on a single gene like 16S rRNA gene. Using genome mining I also tried to highlight important phenotypic characteristics related to their methylotrophy and plant symbiosis, which are experimentally verified. This study covered all available PPFM genome information and analysed the genome content in detail accompanied with phenotyping data, and I provide a comprehensive understanding on the physiology and diversity of PPFMs.

1.5. Species with special characteristics in genus *Methylobacterium*

Although most of *Methylobacterium* species are pink in colour and capable of utilizing methanol, it is worth mentioning that there are some species with different characteristic and unique phenotypes. The most interesting ones will be listed as follows.

1.5.1. *Methylobacterium nodulans*

This species contains strains which are nonpigmented, and capable of fixing nitrogen. These strains were isolated from the legume plant (*Crotalaria podocarpa*) root nodules and found to have an ability to utilize C1 compounds such as methanol and formaldehyde (Jourand et al., 2004). Phylogenetically, *Mb. nodulans* belong to clade C2, the furthest cluster in the genus *Methylobacterium* (Green & Ardley, 2018). The type strain of this species is ORS 2060^T, it is white in color and its genome size is the biggest among other *Methylobacterium* species, estimated around 8 Mb, and consist of one chromosome and seven plasmids (Table 1).

1.5.2. *Methylobacterium dichloromethanicum* subsp. *dichloromethanicum*

The type strain of this species is designated as DM4^T. DM4^T is a pink pigmented strain capable of growing

on dichloromethane (DCM) as the sole carbon for energy (Doronina et al., 2000). DCM or methylene chloride is a toxic solvent highly found in the industrial emissions. Interestingly DM4 can oxidize DCM into formaldehyde, which would be utilized through the C1 assimilation pathway. The strain DM4 possesses a DCM dehalogenase encoded by *dcmA* which convert DCM to formaldehyde and hydrochloric acid (Kayser et al., 2002). Phylogenetically strain DM4 is in clade B and highly similar to *Methylobacterium extorquens* AM1 (Green and Ardley, 2018), but recently was reclassified into *Methylobacterium dichloromethanicum* (Hördt et al., 2020).

1.5.3. *Methylobacterium dichloromethanicum* subsp. *chloromethanicum*

The type strain is *Methylobacterium dichloromethanicum* subsp. *chloromethanicum* CM4 and it is very close to *Methylobacterium dichloromethanicum* subsp. *dichloromethanicum* DM4^T. This strain can utilize chloromethane, but not dichloromethane as the source of carbon and energy (Doronina et al., 1996). Phylogenetically the strain CM4 is in clade B and highly similar to *Methylobacterium extorquens* AM1 (Green and Ardley, 2018), but recently was reclassified into *Methylobacterium dichloromethanicum* (Hördt et al., 2020).

1.5.4. *Methylobacterium jeotgali*

The type strain is S2R03-9^T, it is non-pigmented, strain. The strain S2R03-9^T was isolated from Korean fermented food called jeotgal and can grow on methanol as the source of energy and carbon. Phylogenetically this strain belongs to clade A (Green and Ardley, 2018).

➤ 2. The methodology

2.1. DNA extraction, genome sequencing, annotation and the 16S rRNA gene phylogeny

Completing the PPFMs genomic data

All publicly available genome assemblies of type strains of *Methylobacterium* and *Methylobacterium* species (hereafter abbreviated as *Mb.* and *Mr.*, respectively) were downloaded from the databases (NCBI and JGI) (Table 1). During the preparation of this research I found only 39 type strains assemblies available in the databases. Six of these assemblies were excluded because they were found to be contaminated or incomplete. Therefore, I sequenced 29 type strains; 23 type strains which were not yet available and six strains that have available assemblies but with low quality (*Mb. dankookense* SW08-7^T (GCA_902141855.1), *Mb. frigidaeris* GIER25-16^T (GCA_002759055.1), *Mb. mesophilicum* SR1.6/6^T (GCA_000364445.2), *Mb. crusticola* MIMD6^T (GCA_003574465.1), *Mb. soli* YIM 48816^T (GCA_008806385.1), and *Mr. thiocyanatum* JCM10893^T (GCA_001310875.1)). The quality of the available assembly of *Mb. organophilum* (GCA_003096615.1) was sufficiently high, but the source strain DSM 760^T is believed to have been contaminated such that the available assembly did not accurately represent its taxonomic status (Kato et al., 2005; Volpiano et al., 2021). Therefore, the strain *Mb. organophilum* NBRC 15689^T was newly sequenced in this study.

All type strains were collected from different culture centres. I collected the cells grown on Reasoner's 2A (R2A, Becton Dickinson) agar plates for one week, washed, suspended in 3 ml of Tris-HCl buffer, and the cell suspension was adjusted to OD₆₀₀=1. All reagents were prepared as suggested in the JGI bacterial DNA extraction CTAB protocol (William et al., 2012). DNA quality was checked by electrophoresis with 1% agarose gel and the concentration was checked with a spectrophotometer (NANODROP2000C, Thermofisher). RNase treatment was done as recommended in the protocol but with a lower RNase concentration, 10 mg/ml. In some cases, I noticed that the RNase treatment degraded

the DNA. Therefore, RNase solution was boiled for 30 minutes to inactivate any (contaminated) DNase activity. The samples were treated with 1/100 amount of RNase of the original protocol. By this treatment the DNA samples were in a good quality to proceed for next steps.

The 16S rRNA gene was amplified from the pure DNA to confirm that there is no contamination before sending for genome sequencing.

The 16S rRNA gene size is about 1500 bp and was amplified and sequenced in two steps. The first step aims to amplify the whole gene sequence using a forward primer Eu8f (AGAGTTTGATCCTGGCTCAG) and a reverse primer Eu149r (GGCTACCTTGTTACGACTT). The total PCR reaction volume was 25 µl and it contained the following reagents: DMSO 1.25 µl, KOD FXNeo buffer 12.5 µl, dNTP 5 µl, KOD FXNeo 0.25 µl, DNA 1 µl (DNA concentration 100 ng/µl), and water 3 µl. The PCR conditions were one cycle of 96°C for 2 min, 30 cycles of 98°C for 10 s, 52°C for 30 s, and 68°C for 1 min, and final cycle; 68°C for 1 min. The PCR products were purified and used as a template DNA for sequencing. The PCR reaction volume was 20 µl and contained the following reagents: super dyeTM (ADS) 1 µl, 5x sequencing buffer 4 µl, DNA template 1 µl, primer 1 µl, and water 13 µl. In this step either Eu8f or Eu803r (CATCGTTTACGGCGTGGAC) primer was used. The PCR conditions were: one cycle of 96°C for 2 min, 30 cycle of 96°C for 10 s, 50°C for 5 s, and 60°C for 4 min. A single stranded DNA around 700 bp of the gene sequence were amplified and sent for sequencing.

I checked the 16S rRNA sequence identity on the EZBioCloud database (<https://www.ezbiocloud.net>). The genome DNA samples were purified with ethanol, diluted with TE buffer, checked one more time on 1 % agarose gel and sent for sequencing.

The genome sequencing was done using Illumina MiSeq or MGI DNBseq. The reads were checked for quality using fastq QC tool available in galaxy (Afgan et al., 2018; Cuccuru et al., 2014; Seemann, 2014). I confirmed that the adapter sequence was removed from all reads. I did *de novo* assembling using CLC Genomics Workbench 20.0.4 or SPAdes (<https://usegalaxy.org>).

Assemblies' quality checking and annotation

All the assemblies were subjected to a final quality check. I added 16 assemblies from other species within the family *Methylobacteriaceae* (15 genomes from the genus *Microvirga*, and one genome from the genus *Enterovirga*) were added to the analysis as an out-group genome set. The reason of this addition was to see the unique features present or absent only in the PPFMs within the *Methylobacteriaceae* family. In total, 78 genomes were subjected to the analysis; 62 assemblies from the PPFMs (60 type strains and two non-type strains) and 16 from the other genera in the family *Methylobacteriaceae*. I confirmed the identity of each assembly by extracting the contigs encoding 16S rRNA genes using the ContEst16S tool (Lee et al., 2017) and check their identity with EZBioCloud database. I checked for contamination and completeness using CheckM (v1.0.18) (Parks et al., 2015), and set the threshold of >90% for completeness and <5% for contamination. All assemblies were within the range. For consistent annotation all assemblies were re-annotated using Prokka (version 1.14.5) available on the Galaxy web platform (Afgan et al., 2018; Cuccuru et al., 2014; Seemann, 2014).

2.2. Whole genome-based analysis and pan-genome analysis

In silico DDH values were calculated using TYGS (Type Strain Genome Server) platform (<https://tygs.dsmz.de/>) (Meier-Kolthoff & Göker, 2019), and ANI was calculated using FastANI on the Galaxy web platform (Afgan et al., 2018; Jain et al., 2018). Bacteria pan genome analysis (BPGA) was done using the amino acid sequence data under the default parameter settings (50% identity, Chaudhari et al., 2016).

2.3. Functional annotation

KEGG and COG functional clustering analysis within the BPGA program was done to find the distribution of PPFMs core and unique genes. For pathway/module analysis, metabolic and physiological potential evaluator GenomappleTM pipeline version 2.3.2 was used. Only modules that occurred in at least four but fewer than 65 genomes among 78 genomes were taken into consideration; 90 modules met this criterion.

The data into a presence and absence binary matrix and visualized with R heatmap.3 package.

Gene clusters for secondary metabolites synthesis were mined using antiSMASH pipeline version 6 (Blin et al., 2019; Medema et al., 2011). I used the whole assembly's data file and run the analysis under the default settings.

2.4. Genomics of methylotrophy, phyllosphere adaptation, and plant colonization

I collected all reference genes for C1 assimilation, namely the metabolic and regulatory genes for methanol, methylamine, and dichloromethane. The genes responsible for utilization of methanol and methylamine were extracted from the reference genome of the model organism *Mr. extorquens* AM1 (GCA_000083545.1) (Vuilleumier et al., 2009). The genes of N-methylglutamate pathway (NMG) for methylamine utilization were extracted from *Mb. extorquens* PA1 genome sequence (GCA_000018845.1). Dichloromethane utilization genes were from *Mr. dichloromethanicum* subsp. *dichloromethanicum* DM4 (GCA_000083545.1). The genes involved in adaptation to the phyllosphere and plant growth promotion (UV damage and oxidative stress protection, biosynthesis of cytokinin, aminocyclopropane-1-carboxylic acid (ACC) deaminase, and chitinase) were from *Mb. oryzae* CBMB20^T (GCA_000757795.1, Kwak et al., 2014) and *Mb. aquaticum* MA-22A (GCA_001548015.1, Tani et al., 2015). Enzymes related to plant colonization (pectate lyase, cellulase, and endoglucanase) were from *Mr. extroquens* AM1 and *Mb. aquaticum* MA-22A. The genes for indole acetic acid production through the L-tryptophan pathway (monoamine oxidase, L-tryptophan decarboxylase, and aldehyde dehydrogenase) were from *Mb. nodulans* ORS 2060^T (GCA_000022085.1).

The amino acid sequences of the genes of interest were used as queries for the BlastP analysis (50% identity threshold) against genome data of the strains. The duplicate hits were removed, and from each genome only a single hit for each gene with the highest identity value was kept. The resultant gene presence/absence table was used to examine the conservation of the important genes. Additional check was done for *xoxF* to confirm that it exists in its complete cluster, since the genomes contain multiple

copies of *xoxF* homologues. Reference sequence source accession number and annotation ID are summarized in Table 2.

2.5. Phylogenetic tree construction of PQQ-dependent alcohol dehydrogenase homologues

As mentioned above, one genome may contain multiple copies of PQQ-dependent alcohol dehydrogenases (ADHs) homologues. To examine the ADH types encoded in each genome, the amino acid sequences of various ADHs (49 sequences) from different sources were manually picked up from the tree presented by Keltjens et al., 2014. The reference sequences were used as queries for BlastP-search. The homologues of >50% identity and <0.001 e-value were taken. After duplicate removal the Blast hit of the 254 sequences were aligned with MAFFT together with the reference sequences, and a maximum likelihood phylogenetic tree was constructed with a Fasttree web tool on the Galaxy web platform (Afgan et al., 2018). MDH and ADH reference sequence sources, accession numbers, and types are summarized in Table 3.

2.6. Phenotypic characteristics

2.6.1. General notes

For phenotypic analysis, among all 63 validly published *Methylobacterium* and *Methylobacterium* type strains, I did not use *Mr. populi* BJ001^T (a patented strain), *Mr. pseudosasa* NBRC105205^T (contaminated culture), *Mb. symbioticum* SB0023/3^T, *Mb. ajmalii* IF7SW-B2^T, and *Mb. radiodurans* 17Sr1-43^T (which were published recently), but added *Mr. dichloromethanicum* subsp. *chloromethanicum* CM4, *Mr. dichloromethanicum* subsp. *dichloromethanicum* DM4, *Mr. extorquens* AM1 (ATCC 14718), and *Mr. lusitanum* (NCIMB 13779) which is a later synonym of *Mr. rhodesianum* DSM 5687^T.

2.6.2. Phenotyping assay on solid mediums

All the available strains were subjected to the following growth tests: growth on solidified nutrient broth (NB, EIKEN chemical), R2A agar (DifcoTM), and Potato-Dextrose agar (PDA, DifcoTM); utilization of methanol, methylamine, dichloromethane, glycerol, ethanol, glucose, fructose, and sucrose, using a mineral medium (MM) solidified with 1.5% agar (Table 4); salt tolerance (1% NaCl using 0.5% glycerol as a carbon source); ampicillin resistance (25 and 50 mg/l on R2A), and the effect of 30 μ M LaCl₃ supplementation for methanol and glycerol growths. A colony of each strain grown on R2A agar for one week was picked and streaked on a test plate. Plates were incubated at 28°C for 10 days. The carbon source concentrations were set to 0.5% (w/v for solid substances or w/w for liquid substances), except for dichloromethane. For dichloromethane test, I inoculated the strains on solidified MM medium prepared in 12-well plates, and a small beaker containing 2 ml dichloromethane was placed in a glass desiccator (22.6 L). The cultivation was done in the dark. The growth was evaluated as either positive or negative after 10 days of incubation at 28°C. The oligotrophic growth on MM agar plates with no added carbon source was regarded as a negative control.

2.6.3. Growth assay on liquid medium

Fresh cells grown on R2A plates for one week were resuspended in an 0.9% NaCl saline solution. A cell suspension of OD₆₀₀=1 was prepared, 4 μ l of which was added to 200 μ l of the liquid culture medium. The culture was kept on at 28°C for 7-14 days through rotary shaking (300 rpm). Growth (OD₆₀₀) was monitored two times a day. Minimum Mineral medium (MM) supplemented with Methanol (MMM) and MMM medium supplemented with 30 μ M LaCl₃ were used to test the ability to utilize methanol with and without the presence of Ln³⁺.

2.6.4. API 20NE strip (bioMérieux) tests

The API 20NE strips were used according to the manufacturer's instructions. In short, fresh cells grown on R2A plates were collected and suspended in saline solution 0.85%. The cell concentration was adjusted

to OD₆₀₀=0.1 or equivalent to 0.5 McFarland. For the first 8 tests only (enzyme activity), bacteria suspended in the saline solution was added. For the remaining 14 assimilation tests bacterial saline suspension (200 µl, OD₆₀₀=0.1) was added to 7 ml of the API medium, mixed well, and 270 µl of the suspension was inoculated into the remaining tubes. The strips were incubated at 28°C and monitored for 7 days. Positive growth for the first three tests (reduction of nitrates, production of indole, and fermentation of glucose) was recorded after 24 hours of inoculation. The 17 other tests were checked on the second, third, fifth, and the seventh day, but only the final day result was recorded.

2.7. Antifungal activity

Growth inhibition assay against *Fusarium oxysporum* f. sp. *lycopersici* strain (9859-1) race 2 (a kind gift from Mr. Matsusaki, Aichi Agricultural Research Center) was done as follows. I used Potato dextrose agar medium plated for the co-culture assay. I tested the PPFMs strains growth on the PDA medium before the experiment to know which strains will be able to grow the PDA medium, and I used only the strains which grew well.

I sub cultured the fungal strain on PDA plates and the PPFMs on R2A medium. After one week I collected the bacterial cells and suspended them in 0.9% NaCl (OD₆₀₀=1.0). I spread 10 µl of the suspension with a loop on one of the halves of the PDA plates. An agar plug (5 × 5 mm) containing *Fusarium* mycelium was placed in the middle of the plate. After 7-10 days at 25°C, plates were pictured. The inhibition percentage was measured by comparing the fungal colony radius in the inoculated half of the plate (ir) with that in the non-inoculated half (cr) as follows:

$$\text{Inhibition \%} = 100 \times (\text{cr} - \text{ir}) / \text{cr}$$

2.8. Indole acetic acid (IAA) production through L-tryptophan pathway

Bacterial cells grown on R2A plates were suspended in 0.9% NaCl (OD₆₀₀=1.0), and 10 µl of the suspension was added to 1 ml of the King's B medium composed of 20 g/l protease peptone, 1.15 g/l K₂HPO₄, and 1.5 g/l MgSO₄ supplemented with 0.51 g/l L-tryptophan, 0.5% (w/v) glucose, and 0.5%

(v/v) methanol prepared in 12-well cell culture plates. Cultivation was done at 28°C for 7 days through rotary shaking (300 rpm). One ml of the culture was centrifuged ($9,100 \times g$) for 5 min at room temperature, and 100 μ l of the supernatant was added to the reagent R1 (12 g/l FeCl_3 and 7.9 M H_2SO_4). The absorbance at 530 nm was measured after 30 min. Authentic IAA (1-100 mg/l) was used as standards. *Pseudomonas syringae* strain DC3000 (ATCC BAA-871) was used as a positive control. The experiment was done in technical triplicates.

2.9. Elicitation of rice cell

The PPFM strains were tested for defense elicitation in rice cells as described previously (Shinya et al., 2016). In short, 40 mg fresh rice cells grown in N6 liquid medium for 4 days at 25°C were suspended in 995 μ l N6 medium in 24-well plates. Bacterial suspension (5 μ l, $\text{OD}_{600}=1$) was then added. *P. syringae* DC3000, chitin oligomer ((GlcNAc)₈, final concentration of 10 nM), and *Serratia marcescens* subsp. *marcescens* isolates 1-7 and 4-24 that have shown a positive elicitation response previously (Wari et al., 2019), were used as positive controls. The plates were incubated at 25°C with shaking (300 rpm) and pictured at 24 and 40 h post-inoculation. The experiment was repeated two times. Responses were recorded as positive if the rice cells changed their color, which is indicative for accumulation of defense metabolites.

2.10. Data availability

The genome sequence data are deposited in the DNA data bank of Japan (DDBJ) under BioProject number PRJDB11873. The Prokka-annotation files are accompanied with the publication (Alessa et al., 2021).

➤ 3. Results and discussion

3.1. General consideration

In this study, I aim to complete the genome information of this taxonomically heterogeneous group of methylotrophic bacteria and compare their genotypes and phenotypes. I collected the strains from culture collections as well as genome assembly data from the databases. At the time I started the project, there were 63 *Methylobacterium* and *Methylobacterium* species validly published (<https://lpsn.dsmz.de>). Among them, *Mr. pseudosasa* strain NBRC 105205^T I obtained was found to be contaminated, and its genome assembly was unavailable from public data repositories, so I excluded this strain from the genomic analysis. Within the set of remaining 62 species, I used the cultures of non-type strains of *Mb. trifolii* strain TA88+4-73 (DSM 23632) and *Mb. thuringiense* strain C61+2-34 (DSM 23674), both of which are closely related to their correspondent type strains *Mb. trifolii* strain TA73^T and *Mb. thuringiense* strain C34^T, respectively. Other strains were all type strains of the species. I PCR-amplified 16S rRNA gene fragments of all strains to confirm their identity before starting genome sequencing and phenotypic analysis.

3.2. Genome assembling, annotation, and quality check

Among 62 *Methylobacterium* and *Methylobacterium* species type strains used in this study, 33 genome assemblies of the type strains were downloaded from GenBank (Table 1). In total 29 strains were sequenced; 23 of which are reported for the first time in this study; the remaining six had publicly available assemblies, but these were of low quality (Table 1). The quality information of the newly sequenced assemblies is summarized in Table 5. For consistent CDS analysis, the genome assemblies were all re-annotated with Prokka on the Galaxy web platform (Afgan et al., 2018), and quality checked by CheckM (Parks et al., 2015). The genome size of the species ranged 4.4~8.8 Mb with CDSs ranging 4171~8622, and the GC% ranging 65.9~72.7%. The average completeness was 98.8% and average contamination was 0.4%. The 16S rRNA gene sequences extracted from the assemblies by ContEst16S were compared with

their PCR-based sequences available in the EZBioCloud database. I found problems in five downloaded assemblies as follows: The 16S rRNA contig from *Mb. terricola* 17Sr1-39^T was short (1251 bp), but it shows a high enough identity (99.5%). The 16S rRNA contig from *Mb. planium* YIM 132548^T was also short (1,208 bp). One of the two 16S rRNA contigs of *Mb. oryzihabitans* TER-1^T showed high homology to *Pseudomonas asplenii* ATCC 23835, as an apparent contamination. Three assemblies ; *Mb. ajmalii* IF7SW-B2^T, *Mb. segetis* 17J42-1^T and *Mr. salsuginis* CGMCC 1.6474^T did not contain 16S rRNA gene contigs. I sequenced the PCR-amplified 16S rRNA gene fragments from the strains and confirmed their identity to those of the database. Thus, I regarded that overall, the genome assemblies were generally of good quality for comparative genomics, and that I could use the cultures for phenotypic analysis.

As reported previously (Kato et al., 2005), *Mb. dichloromethanicum* (Doronina et al., 2015) and *Mb. chloromethanicum* (McDonald et al., 2001) are later heterotypic synonyms of *Mb. extorquens*. Hördt et al., (2020) recently suggested the reunion of the genera *Methylobacterium* and *Methylobacterium*. Due to the high dDDH values (70~79%), they also proposed classification of *Mb. chloromethanicum* as *Mb. dichloromethanicum* subsp. *chloromethanicum*, subsp. nov. and *Mr. extorquens* as *Mb. dichloromethanicum* subsp. *extorquens*, subsp. nov. In a recent announcement for validation of new names (Oren & Garrity, 2020), the name *Mb. dichloromethanicum* subsp. *chloromethanicum* has been validated, but *Mb. dichloromethanicum* subsp. *extorquens*, subsp. nov. had not. Instead, the name *Mb. dichloromethanicum* subsp. *dichloromethanicum* was also created automatically according to a taxonomical rule in the same validation. However, the fact that the name *Mb. extorquens* (Urakami & Komagata, 1984) had a priority, was ignored. Therefore, I propose classifying the names as *Mb. extorquens* subsp. *dichloromethanicum* subsp. nov., *Mb. extorquens* subsp. *chloromethanicum* subsp. nov., and *Mb. extorquens* subsp. *extorquens* subsp. nov.

3.3. 16S rRNA gene-based phylogeny

The phylogenetic tree based on 16S rRNA gene sequences extracted from the assemblies using TYGS web tool, is shown in Figure 2. This tree divided the species into four clades. Three of them, A, B, and C

were described previously (Green & Ardley, 2018). Clade A (37 species) is the largest clade containing the type species of *Methylobacterium*, *Mb. organophilum*, and comprising the genus *Methylobacterium*. Clade B consists of 8 *Methylorubrum* species containing *Mr. extorquens* as the type species of the genus *Methylorubrum*. Clade C was divided into clade C1 of 11 species and C2 of two species (*Mb. isbiliense* DSM 17168^T and *Mb. nodulans* ORS 2060^T), and *Mb. oryzihabitans* TER-1^T. The out group *Microvirga brassicacearum* clustered alone and close to clade C2. As described above and also pointed out by Hördt et al., (2020), because 16S rRNA gene sequences extracted from genome assemblies had problems or were missing, I do not use the 16S rRNA tree to infer their phylogenetic relationship but rather use the whole genome information.

3.4. Average nucleotide identity, digital DNA-DNA hybridization

Most of the pairwise FastANI values ranged between 78.8 and 92.9%, with an average of 81.7%. Most of the pairwise dDDH values ranged between 21 and 45.8% with an average of 23.5% (Tables 6 and 7). A clear linear correlation was seen between ANI and dDDH values (Figure 3). Some combinations showed high ANI and dDDH values of more than 95% and 70%, respectively. *Mb. oryzae* CBMB20^T, *Mb. fujisawaense* DSM 5686^T, and *Mb. phyllosphaerae* CBMB27^T were genomically close (98.8% ANI, 90.3% dDDH, and 0.03 ~ 0.47% GC% difference). Another combination was found between *Mr. thiocyanatum* JCM 10893^T (newly sequenced in this study) and *Mr. populi* BJ001^T, with relatively high ANI values of 97.8%, high dDDH% of 83%, and a low GC% difference of 0.1%. Thus, as already pointed (Green & Ardley, 2018; Hördt et al., 2020), *Mb. Phyllosphaerae* and *Mb. oryzae* are later heterotypic synonyms of *Mb. fujisawaense*. I also propose that *Mr. populi* BJ001^T is a later synonym of *Mr. thiocyanatum* JCM 10893^T. These combinations of the species showed close relationships even in the following analyses. FastANI showed that there were some clusters comprised of relatively highly similar species, and they are arranged in 8 homogenous clusters. These clusters are named as A1, A2, A3, A4, A5, B, B2 and C1. (Figure 4).

3.5. Whole genome-genome phylogeny

A whole-genome phylogenetic tree was created based on Genome Blast Distance Phylogeny (GBDP), including *Microvirga brassicacearum* CDVBN77^T as an outgroup. Clade A of *Methylobacterium* represents the largest clade, consisting of 5 sub-clades (A1- A5, Figure 5). The species within each sub-clade had relatively high pairwise ANI and dDDH values (Table 6-7). The GBDP tree showed some inconsistency with the 16S rRNA gene tree especially for clade B phylogeny. Unlike the 16S rRNA gene phylogeny, in which a clear separation between *Methylobacterium* and *Methylobacterium* species was seen, the GBDP showed that clade B nested with some *Methylobacterium* species. These species were referred to as sub-clade B2. The sub-clade B2 included five species, *Mb. organophilum*, the type species of the genus *Methylobacterium*. On the other hand, the members in clade C clustered separately from the other two clades. The clade C had a higher GC%, larger genome size, and more CDS counts than the other two clades. The GBDP is constructed from the interpretation of the pairwise whole-genome similarity, which counts paralogs and other regions in the genomes in addition to orthologous genes (Henz et al., 2005; Meier-Kolthoff & Göker, 2019). After identifying the high score segment pairs (HSPs) with genome based blastn, a computational algorithm determines distance between species, and the whole genome-based phylogeny is created (Henz et al., 2005). This method has been used widely to evaluate the taxonomic states of many species (Meier-Kolthoff et al., 2013), making it more reliable to show the true taxonomic states (Meier-Kolthoff & Göker, 2019).

3.6. Pan-genome analysis

3.6.1. The PPFM pan and core genes count

The pan-genome of 62 PPFM strains contains 55,128 proteins and is still open, suggesting that the contents of the PPFM genomes are very diverse. The PPFM core genome includes 711 proteins, the accessory genome (genes common to two or more strains but not to all strains) includes 2886-5500 proteins, and the unique genome (genes found exclusively in one strain) includes 147-1764 genes (Table 8). Most of the strains showed low numbers (0-82) of exclusively absent proteins.

3.6.2. The PPFM pan and core phylogeny

The pan-genome phylogenetic tree, based on all gene presence and absence matrix, showed consistency with the TYGS tree. Clade A was separated into several sub-clades. Clade B clustered with sub-clade B2. Clade C members clustered distinctly (Figure 6A).

The core phylogenetic tree is based on the alignment of 711 core gene amino acid sequences. Six clear clades were seen, A5, A3-A4, A2, B2-B, A1, and C. Similar to the pan-genome phylogeny, clade A members maintained their clustering as those shown in the previous trees (Figure 5 and Figure 6A). The sub-clade B2 clustered with clade B. Clade C was phylogenetically farther, and the sub-clade A1 was the closest one to it (Figure 6B).

3.6.3. KEGG assignments and functional annotation of the PPFM pan-genome

KEGG functional annotation and distribution analysis for the core, accessory, and unique genes is shown in Figure 7A. The PPFM genomes, as a whole, are relatively highly enriched in the metabolism process. The core genome was enriched in genetic information processing. The accessory and unique genes were enriched in environmental information processing. More detailed information is shown in (Figure 7B). The core genes are noticeably enriched in the biosynthesis of secondary metabolites, and metabolism of cofactors and vitamins. The accessory and unique genes are enriched in motility, and interestingly also in the xenobiotics biodegradation and infectious diseases. Because this analysis gives only general information, I investigated more as below.

3.7. Pathway and module analysis

More differentiated information on functional pathways and modules was obtained using the Genomaple pipeline (Figure 8). To see the PPFM-specific functions, the other members in *Methylobacteriaceae* were also included in the analysis. The clustering result based on the presence/absence of the modules clearly separated *Microvirga* and PPFMs. The apparent major differences in the other members were the absence of nitrate reduction and assimilation module, and anoxygenic photosystem II, and the presence of

glyoxylate cycle, degradation systems for D-glucuronate, and transport systems for D-xylose, raffinose/stachyose/melibiose, thiamine, and spermidine/putrescine.

Within PPFMs, the members in clade B and C formed distinct clusters except for *Mb. soli* KCTC 22810^T (sub-clade A3) that clustered with clade C. Some of the members in clade A sub-clades clustered with those in different sub-clades, which may mean that they have acquired or lost some modules keeping their core genes less modified, during evolution and adaptation to their habitats. Exclusively, clade C lacks modules for sulfate/thiosulfate transport and assimilation system, Entner-Doudoroff pathway, biotin synthesis, and tungstate transport system, but carries those for transporters for ribose, L-amino acid, glycerol, tricarboxylic acid, dipeptide, and biotin, in addition to reductive pentose phosphate cycle, anoxygenic photosynthesis (purple bacteria), and degradation pathways for catechol and benzoate. Clade B did not have exclusively lacking or existing modules compared to the other members in clade A, except that only riboflavin biosynthesis pathway was present. Together with the members in sub-clade A2 and A3, clade B members lack transport systems for polar amino acid, methionine, and sugars, pathways for GABA biosynthesis, and pyrimidine degradation. Of note is that the nitrogen fixation module was found only in *Mb. nodulans* ORS 2060^T and some members of *Microvirga* species.

3.8. Methyлотrophy genes cluster

Although Genomape analysis indicated that the “C1-unit interconversion” module is conserved only in clade B, this cannot be true given that most of the PPFMs are methyлотrophic. Therefore, I investigated the genes for the C1 assimilation pathway manually. The genes involved in methyлотrophy (Table 2) were extracted from model strains of *Mr. extorquens* AM1, *Mr. extorquens* PA1, *dichloromethanicum* subsp. *dichloromethanicum* DM4, and *Mb. aquaticum* MA-22A. Cluster completeness was also calculated based on the total number of genes present in each cluster and recorded as a percentage (Table 9). The detailed data are presented in Table 10. A simplified depiction of the presence of all clusters is provided in Figure 9.

Methylamine utilization genes were examined with methylamine dehydrogenase (MaDH) cluster of *Mr. extorquens* AM1 and the N-methylglutamate (NMG) pathway of *Mr. extorquens* PA1 as queries. Only one strain from clade A (*Mb. oxalidis* DSM 24028^T) and three clade B strains (*Mr. rhodesianum* DSM 5687^T, *Mr. podarium* DSM 15083^T, and *Mr. aminovorans* NBRC 15686^T) have the whole MaDH gene cluster. The complete NMG pathway genes are conserved in seven species of clade A, nine strains of clade B species, three strains of sub-clade B2, and only two strains of clade C (*Mb. oryzihabitans* TER-1^T and *Mb. ajmalii* IF7SW-B2^T). Only one gene of NMG pathway was missing in other four species (*Mb. marchantiae* DSM 21328^T and *Mr. salsuginis* CGMCC 1.6474^T, *Mb. crusticola* KCTC 52305^T, *Mb. aquaticum* DSM 16371^T). This result may indicate that the NMG pathway is more abundant than the MaDH pathway in PPFMs. Previous studies have pointed out that the NMG pathway is highly abundant not only in methylotrophic bacteria (Kröber et al., 2021) but can also be found in non-methylotrophic bacteria, which rely on it as a nitrogen source (Chen et al., 2010). In a metagenomic study on soybean and rice symbionts, the NMG pathway was more abundantly detected than the *mau* cluster in *Methylobacterium* (Minami et al., 2016).

The Ca²⁺-dependent MDH gene cluster (*mx*a cluster) consisting of 14 genes was found to be conserved in 27 out of 32 species in clade A (A and A1-A5), all of which have all of the essential genes in the cluster (Figure 9). The remaining five type strains (*Mb. tardum* NBRC 103632^T, *Mb. longum* DSM 23933^T, *Mb. persicinum* NBRC 103628^T, *Mb. komagatae* DSM 19563^T, and *Mb. iners* DSM 19015^T) lack *mx*a genes; this absence has no correlation with sub-clade assignment. All strains in clades B and B2 have the conserved *mx*a cluster except for *Mr. thiocyanatum* JCM 10893^T, in which *mx*aK involved in Ca²⁺ insertion into MxaF is missing, yet this strain can grow on methanol in the absence of La³⁺ (shown below). In clade C, *mx*aFI encoding Mxa-MDH subunits were found only in five strains of clade C1 (*Mb. oryzihabitans* TER-1^T, *Mb. indicum* SE2.11^T, *Mb. ajmalii* IF7SW-B2^T, *Mb. tarhaniae* DSM 25844^T, and *Mb. terricola* 17Sr1-39^T) and in one strain of clade C2 (*Mb. nodulans* ORS 2060^T). *Mb. aquaticum* DSM 16371^T has *mx*aF with an identity of 90% but is missing many of the other *mx*a genes. The strain AM1 genome encodes two genes/regions for Ln³⁺-dependent XoxF-MDH: the *xoxFGJ* cluster and *xoxF2*

(Skovran et al., 2011; Vu et al., 2016). These have 87% amino acid sequence identity with each other and 50% identity to the Mx_AF of the same strain. To differentiate among *xoxF*-homologues in PPFM genomes, I manually checked whether the *xoxF1* was within a cluster with *xoxG* and *xoxJ*. This is the case in 59 of the 62 PPFM strains; the exceptions are *Mb. jeotgali* LMG 23639^T, *Mb. oxalidis* DSM 24028^T, and *Mr. suomiense* DSM 14458^T. In clade A, *Mb. jeotgali* LMG 23639^T *xoxF1* (AOPFMNJM_03121), whose identity to AM1 *xoxF1* and *xoxF2* is 93% and 89%, respectively, is not clustered with *xoxG* and *xoxJ*. *Mb. oxalidis* DSM 24028^T has only one *xoxF*-like gene (LDDCCGHA_05356) with low identity of 46.9% and 47.3% to AM1 *xoxF1* and *xoxF2*, respectively; in addition, its *xoxG* and *xoxJ* are distantly encoded. In clade B, *Mr. suomiense* DSM 14458^T has an *xoxF* homologue encoded with *xoxG* and *xoxJ*, but its *xoxF* is a pseudogene encoding a short polypeptide of only 63 amino acids. Interestingly, the *xoxFs* in clade C1 (with the sole exception of *Mb. oryzihabitans* TER-1^T) have higher identity to AM1 *xoxF2* than to AM1 *xoxF1*. We also found that some species possess two highly homologous types of XoxF, which is discussed in detail below.

The Ln³⁺ uptake and utilization (*lut*) cluster was recently identified in strains PA1 and AM1. It plays an important role in Ln³⁺ uptake and the regulation of XoxF1 in the presence of Ln³⁺ ((Ochsner et al., 2019; Roszczenko-Jasińska et al., 2020). The cluster consists of genes encoding ABC transporter, TonB-dependent receptor, and the Ln³⁺-binding protein lanmodulin (LanM). The analysis showed that almost all *lut* genes except for *lutH* (TonB-dependent receptor) are conserved in all PPFM genomes. LutH homologues with an identity greater than 50% with AM1 LutH are exclusively present in clade B members and two strains from clade A (*Mb. bullatum* DSM 21893^T and *Mb. marchantiae* DSM 21328^T). *lutH* is reported to be necessary for *xoxF*-dependent methylotrophic growth in strain AM1 (Roszczenko-Jasińska et al., 2020). The strains that do not harbor *lutH* may utilize another TonB-dependent receptor.

Dichloromethane utilization genes (*dmcR* and *dmcA*) are exclusively found only in the *Methylobacterium dichloromethanicum* subsp. *dichloromethanicum* DM4 genome. These genes are flanked by IS256 family transposase genes in its chromosome, suggesting that the cluster is mobile among bacterial genomes.

The formaldehyde assimilation pathway was also investigated. The H₄MPT pathway genes are conserved in all PPFMs. The glutathione-dependent formaldehyde dehydrogenase pathway genes (*gfa*, *hgd*, and *gfh*) were found mostly in members of clade C and in some members of clade A, but were not found in any members of clade B.

Strain AM1 harbors four formate dehydrogenases (FDH1 to FDH4). Most of the clade B strains and many of the sub-clade A2 and A5 strains conserve these four modules. Most of the clade C strains lack the FDH1 and FDH 4 clusters, while about half of the clade A strains lack the FDH1 and FDH3 clusters.

The H₄F, folate synthesis, serine cycle, and ethylmalonyl-CoA pathways are indispensable for methanol dissimilation in PPFMs; accordingly, the genes for these pathways are all well conserved. The recently discovered factors TtmR and EfgA, which are involved in methylotrophy regulation and formaldehyde binding, respectively, in strain PA1 are also well conserved (Bazurto et al., 2020, 2021), but the latter is missing in most clade C strains (Table 10). Regarding the other species in the *Methylobacteriaceae* family, *Microvirga massiliensis* JC119^T, an isolate from human feces (Caputo et al., 2016), has some of the methylotrophy gene clusters. The complete *xoxFGJ* cluster and two of the GSH pathway genes are also present in the genomes of some *Microvirga* species. Their methylotrophy should be experimentally verified.

3.9. MDH and alcohol dehydrogenase (ADH) homologues

I examined which types of ADHs are present in *Methylobacteriaceae*, using blastP and 49 collected representative amino acid sequences as queries (Table 3). MDHs (MxaF and XoxF) belong to a larger PQQ-dependent periplasmic alcohol dehydrogenase (ADH) family. PQQ-ADHs family can be separated into three types. Type I contains MDHs (XoxF and MxaF), type II ADHs are quinoxemoproteins, and type III ADHs are only found in acetic acid bacteria. Type I and II ADHs are separated into MDHs and many other types (PQQ-ADH types 1 to 9) (Keltjens et al., 2014). XoxF-type MDHs are encoded widely in many bacterial genomes, and phylogenetically separated into five families XoxF1~XoxF5 (Chistoserdova, 2011). XoxFs from alpha-, beta-, and gamma-proteobacteria mainly belong to XoxF1,

XoxF3, and XoxF5 families (Chistoserdova, 2011; Huang et al., 2018). XoxF5 family is the most abundant among all. Multiple genes for ADHs coexist in the PPFM genomes. For example, *Mr. extorquens* AM1 contain *xoxF1* and *xoxF2*, both of which belong to the XoxF5 family. Another example is ExaF, a Ln^{3+} -dependent ethanol dehydrogenase (Good et al., 2016).

Each PPFM genome encodes 2~7 PQQ-ADH and MDH orthologs. The phylogenetic analysis revealed that the orthologs could be divided into 8 major and several minor groups (Figure 10). The first group, represented by the well-characterized XoxF1 from strain *Mr. extorquens* AM1, is shared in clade A and B members. Of note, many genomes encode duplicated genes of this group. The second cluster represented by XoxF2 from strain AM1 is shared mostly in clades A and C members. The third group contains XoxF1 from *Mb. aquaticum* strain 22A that is an Ln^{3+} -dependent MDH also participating in the regulation of *mxoF* in the strain (Masuda et al., 2018; Yanpirat et al., 2020). This group is found distinctively in clade C members. Most XoxF-type MDHs found within PPFMs are from the XoxF5 family. The sequences within these groups conserve Asp³⁰¹; the position corresponds to that of mature XoxF polypeptide from *Methylophilum fumarioricum* SolV Mfum_190005, (Pol et al., 2014), which determines metal specificity of XoxF proteins to Ln^{3+} . Group 4 contains MxoF from the strain AM1 as a representative and orthologs from clade A, B, and some members in clade C. This group conserves Ala³⁰¹, which determines metal specificity of MxoF proteins to Ca^{2+} (Pol et al., 2014). Group 5 (XoxF3 family) is occupied mainly by orthologs from *Microvirga* species, clade C2, and the only XoxF found in *Mb. oxalidis* DSM 24028^T. This group had one different amino acid, Trp³⁰⁰ instead of Tyr³⁰⁰ or Phe³⁰⁰, compared to XoxF5 family. Group 6 contains ExaF from *Mb. aquaticum* strain 22A that contributes a little for methylotrophic growth (Yanpirat et al., 2020), and is found only in clade C1 members. Group 7 includes PQQ-ADH type 6a, and most of the members are from clade C1. Group 8 PQQ-ADH type 2a, a calcium dependent ADH (Keltjens et al., 2014), consisted of a small group with members from clade A, B, and *Meganema perideroedes*. This group had a conserved Ser³⁰¹. Group 9 includes ExaF from *Mr. extorquens* strain AM1, and all orthologs are from clades A and B. The characteristic double cysteine (CC, position 131-132) encoded in PQQ-ADHs are recently shown to be important for the catalytic activity of

MDH (Chan et al. 2021). This CC is conserved in almost all orthologs, except for some minor members of PQQ-ADH type 7, 8, and 9.

Thus, multiple copies of PQQ-ADH orthologs encoded in different clades of PPFMs are phylogenetically different. The presence of multiple copies may suggest its importance for their survival to adapt to different environments. This analysis also showed that there are some minor groups that have never been characterized in terms of their substrate specificity, induction, and metal-requirements. Most XoxFs found in the PPFMs were from the XoxF5 family. The XoxF3 family was underrepresented. *Mr. rhodinum* DSM 2163^T was the only strain that has one of the XoxF1 family. XoxFs from other species within *Methylobacteriaceae* family (*Microvirga* species) are mainly XoxF1 and XoxF5 families. the XoxF4 family was not present in the PPFMs, this correlates with the fact that XoxF4 type is abundant only in betaproteobacteria (Huang et al., 2018). Although the amino acid Asp³⁰¹ is supposed to determine the specificity to Ln³⁺ (Pol et al., 2014), it is worth mentioning in a special case Ca²⁺ was found in the purified *xoxF*-type MDH from *Candidatus Methyloirabilis oxyfera* (Wu et al., 2015).

3.10. Phenotyping of PPFMs

3.10.1. PPFM growth assay on single carbons sources

3.10.1.1. PPFM growth assay on methanol

To examine the correlation between the PPFM genotypic data and their phenotypes, a growth test in various conditions was done (Table 13). Growth on methanol (in the absence of Ln^{3+}) has been a hallmark of PPFMs, although some strains have been known to be unable to grow. The type strains in clade A, that are lacking *mxoF* (*Mb. tardum* NBRC 103632^T, *Mb. longum* DSM 23933^T, *Mb. persicinum* NBRC 103628^T, *Mb. komagatae* NBRC 103627^T, and *Mb. iners* DSM 19015^T) did not grow on methanol. All strains in clade B were able to grow well on methanol. Some strains in clade C did not grow or showed very weak growth on methanol. These strains lack the whole *mxo* cluster as described above, although some of such strains (*Mb. terrae* KCTC 52904^T, *Mb. crusticola* MIMD6^T, and *Mb. currus* PR1016AT^T) were reported to grow on methanol. *Mb. cerastii* DSM 23679^T also did not grow on methanol, though it had the whole *mxo* cluster, this strain showed very weak growth on various test throughout the whole analysis.

3.10.1.2. Growth assay on methanol supplemented with LaCl_3 (Ln^{3+} -dependent methylotrophy)

To check how PPFM species react to the presence of La^{3+} , I did a growth assay where I compared their growth in liquid culture. I used two conditions; Methanol MM medium and Methanol MM medium supplemented with 30 μM LaCl_3 . I tested 13 strains (mostly from clade C1) ability to utilize methanol with and without the La^{3+} supplementation. Without La^{3+} , only five strains were able to grow (Figure 11A). This result correlated with the presence and absence of *mxoF* in the genomes of these strain (Table 9). *Mb. aquaticum* DSM 16371^T, which has an *MxoF* homologue with 90% amino acid identity to that of strain AM1 but does not have the full cluster, could grow without La^{3+} . On the other hand, all strains were able to grow with La^{3+} (Figure 11B). This was a result of the conservation of *oxoF* in almost all strains'

genome. Interestingly, a few strains that were able to grow without the La^{3+} supplementation grew much stronger with the supplementation. These strains include *Mb. tarhaniae* DSM 25844^T, *Mb. terricola* NBRC 112874^T. The strain *Mb. aquaticum* DSM 16371^T growth was not affected by the presence of La^{3+} .

As noted already, most strains had XoxF type MDH that is activated by Ln^{3+} (Hibi et al., 2011; Pol et al., 2014). When tested on methanol in the presence of La^{3+} , almost all strains, including those that showed weak or no growth in the absence of La^{3+} , were able to grow. *Mb. nonmethylophilum* KCTC 42760^T was reported as a non-methylophilic bacterium lacking the *mx* cluster (Feng et al., 2020), but it is indeed an Ln^{3+} -dependent methylophilic strain (Figure 11B).

All strains were also tested on MMM+ La^{3+} agar medium (Table 13). *Mb. tardum* NBRC 103632^T, *Mb. longum* DSM 23933^T, *Mb. persicinum* NBRC 103628^T, *Mb. komagatae* NBRC 103627^T, and *Mb. iners* DSM 19015^T required La^{3+} to grow on methanol but their growth was not as strong as Clade C1 members. Of note, the clade C1 strains grew better in the presence of La^{3+} . It is possible that one of the reasons is that their XoxF type is phylogenetically different compared to the other clades (Figure 10). This note will be left for future research to support it.

These results indicate that Ln^{3+} -dependent methylophilicity is of primary importance compared to Ca^{2+} - and MxaF-dependent methylophilicity in PPFMs. But the conservation of the latter pathway in many of PPFMs also suggests that they are adapted to the environments where Ln^{3+} is not readily available, such as the phyllosphere.

3.10.1.3. PPFM growth assay on methylamine

I found that 11 strains in clade A could utilize methylamine, and all of them had NMG pathway, while *Mb. oxalidis* DSM 24028^T had both. All clade B strains grew on methylamine, owing to NMG genes but not to *mau* cluster. Most of C1 and C2 strains were not able to utilize methylamine. This result is almost consistent with the previous description on the three clades (Green and Ardley, 2018), except that a

number of strains from clade A were able to use methylamine. The occurrence of methylamine growth in clade A members did not correlate with their phylogeny (Table 13).

3.10.1.4. PPFM growth assay on dichloromethane

After ten days of incubating of all the PPFM type strains in the presence of dichloromethane, I compared the strains growth with the one growing on no carbon condition. Only one strain was able to grow strongly in the presence of dichloromethane. This strain was *Mr. extorquens* subsp. *dichloromethanicum* strain DM4 and it was the only strain found capable of utilizing dichloromethane.

3.10.2. PPFM growth assay on different carbons and alcohols

I tested the PPFM ability to grow on different alcohols such as ethanol and glycerol. Almost all strains grew well on ethanol irrespective of the presence of La^{3+} , however, many of C1 strains showed weak growth in the absence of La^{3+} .

I also tested their ability to grow on different carbon sources, such as fructose, sucrose, and glucose. The occurrence of sugar utilization in clade A was almost random, *Mr. extorquens* strains exceptionally show no growth within clade B, and almost all clade C1 strains grow well on the tested sugars. Glycerol can support growth of most strains, and it was used to examine their vitamin requirement. Almost all the strains grew well on PDA medium, except for a few strains. Most species from clade B and C grew on NB medium, while strains from clade A and sub-clade B2 showed a weaker growth (Table 13). Strains from sub-clade A4 did not grow on sucrose medium. The species close to strain AM1 grew weakly on different carbon sources, suggesting that these strains could be stricter methylotrophs.

3.10.3. Growth under different stress conditions

All strains were tested for three different stresses, growth in the presence of 1% NaCl or ampicillin (25 and 50 mg/L). In clade A the growth under these three conditions did not correlate with the sub-clades except for the sub-clade A5 members that grew well on R2A medium regardless of the ampicillin and

NaCl. All the strains from clade B tolerated the NaCl, while a few grew in high temperature conditions (37°C). Interestingly, the strains from clade C1 and C2 were not able to grow under 1% NaCl condition, but most showed resistance to the high temperature and ampicillin resistance (Table 13).

3.10. 6. API20NE test result

Phenotypic analysis was performed by means of the API 20NE test. After 24 hours of incubation, nitrate reduction, indole production, and glucose fermentation were recorded. In general, as shown in table 14, all PPFM strains were negative for indole production and glucose fermentation. Nine species out of the 60 tested were positive for reduction of nitrates to nitrites; these were mostly members of clades A, B, and C2. All species were negative for β -glucosidase, protease, β -galactosidase, p-nitrophenyl- β -D-galactopyranosidase, and assimilation of capric acid. Most of the strains were urease-positive, but noticeably weak urease production was detected for some members of clade C1. On the other hand, there was a clear variation regarding the assimilation of sugars and organic acids. The majority of species in sub-clades A1, A2, and A3 were positive for sugar assimilation but weak or negative for the assimilation of organic acids. Most strains from sub-clades A4 and A5 were positive for the assimilation of all sugars and organic acids except for adipic acid. Some clade B strains showed weak assimilation of sugars, but most strains were positive for the assimilation of N-acetyl-glucosamine and organic acids. In the sub-clade B2, some members showed weak growth on both tested sugars and organic acids. Most strains in clade C1 were positive for the assimilation of sugars. Clade C2 members were negative or weak for most of the tests. All data is summarized in table 14.

Some species (*Mb. iners* DSM 19015^T, *Mb. thuringiense* DSM 23674, *Mb. trifolii* DSM 23632, *Mb. cerastii* DSM 23679^T, and *Mb. crusticola* KCTC 52305^T) were negative for most of the tests, consistently showing weak growth although they belong to different sub-clades.

3.11. Secondary metabolites clusters

Secondary metabolites clusters were identified using the antiSMASH pipeline (Table 11). The most abundant cluster in the PPFM genomes was terpene cluster. All genomes had two or more terpene clusters, one of which was a carotenoid cluster, while the others were unknown or malleobactin synthesis clusters with low similarity value. Homoserine lactone cluster was found in most genomes, and especially members of clade C had two or more clusters. Type I polyketide synthase, non-alpha poly-amino acids like e-polylysine, and redox-cofactor (PQQ) clusters were found in almost all genomes too. Some strains had non-ribosomal peptide synthetase clusters (NRPS) and NRPS-like clusters. The clade C members, compared to the others, contain relatively larger numbers of clusters for various secondary metabolites.

3.12. Plant growth-promoting genes

PPFMs can induce plant growth in a direct way by producing phytohormones such as auxin and cytokinin. The complete L-tryptophan pathway for IAA biosynthesis in *Mb. nodulans* ORS 2060^T (L-tryptophan decarboxylase, monoamine oxidase, and aldehyde dehydrogenase), was also found mostly in members of clade C and some members of clade A (Table 12). Monoamine oxidase and aldehyde dehydrogenase were also found in sub-clades A4 and A5 and clade C. tRNA delta (2)-isopentenylpyrophosphate transferase gene (*miaA*), which is responsible of cytokinin synthesis through trans-zeatin pathway (Koenig et al., 2002), was conserved in almost all the genomes. Bacterial 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase (*AcdS*) is involved in lowering ethylene levels in plants. The importance of bacterial D-cysteine desulfhydrase (*DcyD*), which is structurally homologous to ACC deaminase, is unknown but it may be involved in production of hydrogen sulfide that also stimulates plant growth (Ekimova et al., 2018). The analysis showed that *acdS* was found in clade A and C members, but not in clade B members. *dcyD* was found mostly in members of clade B and also in a few species of clade A but not in clade C. Chitinase gene was found in a few strains in clade A, and pectinase gene was found in a few species of clade B. A number of antioxidant enzymes (glutathione peroxidase and glutathione oxidase) were found to be conserved in the genome of all PPFMs. Ergothioneine is an unusual amino acid involved in UV

resistance in *Methylobacterium* (Alamgir et al., 2015). The genes for ergothioneine synthesis are conserved across all PPFMs (Table 12).

3.13. Indole acetic acid (IAA) production through the L-tryptophan pathway

Most of the tested strains produced IAA, of which concentration ranged within 8-29 ug/ml (IAA). The pathogenic bacteria *Pseudomonas syringae* DC 3000 produced high amount of IAA (57 ug/ml, data not shown). The strains from clade A (especially *Mb. gnaphalii* DSM 24027^T, *Mb. oxalidis* DSM 24028^T, and *Mb. dankookense* DSM 22415^T and clade B were potent producers compared to those from clade C despite that clade C strains had the complete L-tryptophan pathway. This result correlates with the fact that PPFMs can enhance plant growth directly by the production of phytohormone as one of the important direct growth promoting mechanisms (Subhaswaraj et al., 2017; Vadivukkarasi & Bhai, 2020), and the addition of tryptophan induces IAA production (Omer et al., 2004). Thus, despite that the tested genomes did not have the whole L-tryptophan pathway of *Mb. nodulans* ORS 2060^T, these strains were able to produce IAA. These strains may have the genes with identity lower than 50% or they employ a different pathway to produce IAA. The blast hits for L-tryptophan decarboxylase indeed had very low identity values, and mostly missing. This was also seen in *Mb. oryzae* CBMB20^T genome, where L-tryptophan decarboxylase is missing (Kwak et al., 2014).

3.14. Antifungal activity

I tested 50 PPFMs strains for antifungal activity, which can be an indirect means to promote plant growth. The sub-clade A4 that includes *Mb. radiotolerans* IAM 12098^T, *Mb. oryzae* DSM 18207^T, and its closely related strains, showed a strong activity (Table 13), and some of these strains carry chitinase genes (Table 12). Also, the sub-clade A5 showed a strong activity. The activity in clade B strains was moderate except for *Mr. rhodesianum* DSM 5687^T that showed the strongest activity. Most of the strains in clade C1 had a clear inhibition effect on *Fusarium* growth. The genome mining data showed that chitinase, cellulase, pectinase, and a number of secondary metabolites clusters were conserved across PPFM strains. The siderophore production and NRPS may be involved in the antifungal activity, however, more investigation is still needed to identify which secondary metabolite is active toward particular pathogens. Figure 12 shows the PPFM strains strongest against the fungi: the fungi colony spread less widely on the half plate where the PPFM strains were growing.

3.16. Elicitation of rice cells

I employed a method of rice cell elicitation assay to monitor plant cell reaction against PPFMs. As shown in figure 13, all PPFM strains and the negative control did not cause any detectable coloration of the rice cells nor increased turbidity of the medium. The pathogenic bacteria *P. syringae* DC3000 and the *Serratia marcescens* subsp. *marcescens* strains 1-7 and 4-24 (Wari et al., 2019) caused clear browning coloration in rice cells, which is an indicative for the activation of secondary metabolites production, in addition to the rapid (in less than 20 h) microbial growth that increased the turbidity of the media. These results suggested that all the studied PPFMs were recognized as non-pathogenic bacteria to rice. This is supported by the fact that PPFMs are abundant endophytic and epiphytic bacteria (Delmotte et al., 2009; Knief et al., 2010; Roodi et al., 2020). A negative effect of rice leaf bleaching by inoculation of a strain of *Methylobacterium indicum* had been reported, however (Lai et al., 2020).

➤ 4. Discussion

4.1. General discussion

Methylobacterium species are environmentally important bacteria as a phyllospheric symbiont, as well as a platform for fine chemical synthesis using methanol as a cheap feedstock. Their diverse metabolic capacity on not only methanol but also other C1 compounds invokes much interest in the genetics and biochemistry of methylotrophy. In a taxonomical point of view, currently *Methylobacterium* species belong to the family *Methylobacteriaceae*, within which the other genera *Microvirga*, *Enterovirga* exist. The *Methylobacterium* species had been classified into three major clades, and clade B species had been classified into *Methylorubrum* (Green & Ardley, 2018). The main purpose of creating the new genera *Methylorubrum* was to create a more homogeneous bacterial group from the divergent *Methylobacterium* species. Recent large survey (Hördt et al., 2020) however, proposed unionizing *Methylobacterium* and *Methylorubrum*. Besides, as many as 62 species have been described, some other species have been reclassified into the existing species, such as *Mb. lusitanum* (*Mb. rhodesianum*), and *Mr. chloromethanicum* and *Mr. dichloromethanicum* (subspecies under *Mr. extorquens*). *Mr. organophilum* and *Mr. populi* had been reported to be able to grow on methane but no following research reproduced the growth (Dedysh et al., 2004). The strains in clades A and B are relatively phenotypically similar to each other, while the strains in clade C show phylogenetically distinct positions apart from the other clades. They were proposed as species *incertae sedis*, and further evidence to clarify their taxonomic position had been required (Green & Ardley, 2018). Although *Methylobacterium* species are often called PPFMs, some species had been reported incapable of methanol utilization and non-pigmented. These backgrounds motivated me to clarify the genotype and phenotype of all species in PPFMs. Also, by comparing with the other genera in the same family I aimed to emboss the common characteristics shared only in PPFMs.

The 16S rRNA gene sequences have been used for the delineation of species-species boundaries. The PCR-amplified sequences have shown to be problematic if repeated in different laboratory conditions (Kato et al., 2005). Although usually intragenomic heterogeneity of 16S rRNA genes should be within the

species threshold, contamination of phenotypically and morphologically similar strains may cause ambiguous determination of the sequences. Besides the PCR problem, I also found problems in the 16S rRNA gene sequences from the PPFM genomes due to copy number, contamination, missing gene, or assembly quality, as reported already (Hördt et al., 2020). Therefore, it is more desirable to employ a phylogenetic analysis based on the whole genomes. In this study I sequenced the genomes of 29 strains to cover all known species in *Methylobacterium*/*Methylobacterium*. The genome data provided precise and detailed insights into the genomics and phylogeny.

The GBP tree showed that clade A and B nested together, while clade C was distinct. My research with newly sequenced genomes showed that clade A can be divided into more homogeneous sub-clades A1-A5. Each sub-clade shares a high pairwise dDDH and FastANI values, and in some cases their phenotype profile was highly similar too. The pairwise dDDH values evidenced that some species were classified poorly based on 16S rRNA sequences as discussed below. Unlike the 16S rRNA tree that relies only on one gene sequence identity, I present here more accurate phylogeny of PPFMs based on their genome sequences. These approaches (GBDP, core and pan phylogeny) used in this thesis correlated well showing that the two genera (*Methylobacterium* and *Methylobacterium*) are phylogenetically close. One of the closest species to clade B was the type species of the genus *Methylobacterium*, *Mb. organophilum* NBRC 15689^T.

The utilization of methanol is an important strategy and ability for the PPFMs in the environment. Methanol emitted from plants is reported to affect the distribution of PPFM species in soil (Macey et al., 2020). Although several PPFM species can utilize other C1 compounds such as methylamine and chloromethane, all species have at least one MDH cluster that guarantees their thrive on methanol as also evidenced by this study. Other methylotrophy gene sets (PQQ synthesis, H₄MPT pathway, serine cycle, *lut* genes, H₄F pathway, and EMC pathway) were conserved in most PPFM species but not in the other species in the *Methylobacteriaceae*. Since glyoxylate necessary for the serine cycle is supplied by the EMC pathway (Schneider et al., 2012), the glyoxylate cycle was missing in the PPFM species, but not in the other species of the *Methylobacteriaceae* family. Among the methylamine utilizers in PPFMs, the

NMG pathway seems to be preferred. The NMG pathway is believed to be the major methylamine utilization pathway in soybean symbionts (Minami et al., 2016).

PPFM genomes encode several homologues of ADH and MDH proteins. Nine phylogenetically distinct groups were observed. XoxF types mainly belonging to XoxF5 family (Chistoserdova, 2011) were the dominant group. Phylogenetically ExaF types were divided into two groups (PQQ ADH type 2b and PQQ ADH type 6a), both ADH types are a putative Ln^{3+} -containing protein (Keltjens et al., 2014). Also, a small group of putative calcium dependent ADH PQQ type 2a were found. Generally, PPFM genomes encode a variety of ADH and MDH proteins which are, in most cases, Ln^{3+} -dependent types.

Methylobacterium species are highly abundant in and adapted to the phyllosphere that is a harsh environment, with limited nutrients, intense UV light, and severe fluctuation of weather conditions (Knief et al., 2010; Minami et al., 2016; Yoshida et al., 2017). My data have shown that they are well prepared with a UV damage repair gene, antioxidative stress genes (ergothioneine and peroxidase genes), and the carotenoid synthesis genes. All *Methylobacterium* species had the full carotenoid cluster except for non-pigmented strains of *Mb. nodulans* ORS 2060^T and *Mb. jeotgali* LMG 23639^T (data not shown). These two species have a unique niche; *Mb. nodulans* ORS 2060^T nodulates in *Crotalaria podocarpa* (Jourand et al., 2004), and *Mb. jeotgali* LMG 23639^T was found in a Korean salted seafood (Aslam et al., 2007). With no beneficial function for these species to survive in their niche, the genes for carotenoid synthesis had been lost during evolution and adaptation. Thus, though there are some such non-pigmented PPFMs as exceptions, most of them are pink pigmented to different intensity.

KEGG functional annotation has shown that core genes of PPFM species were enriched with cofactors and vitamins biosynthesis genes, and they are important for their survival and symbiotic interaction with their hosts. Some microalgae species are reported to be dependent on *Methylobacterium* for the supplementation of vitamins (Krug et al., 2020). This analysis also showed that the core genome was rich in secondary metabolites clusters, which was confirmed by antiSMASH analysis. These clusters are valuable for plant growth promotion, solubilization of nutrients and plant protection (Vadivukkarasi &

Bhai, 2020). Nitrate reduction and assimilation were also unique features for *Methylobacterium* species. These characteristics provide yet another beneficial advantage for their host (Calatrava et al., 2018; M. Madhaiyan et al., 2010). The anoxygenic photosystem II, was also a unique feature of the PPFMs. This photosystem involves a bacterial chlorophyll gene that clusters close to the carotenoid gene cluster. It enables the PPFMs to gain ATP but not to fix carbon or to produce oxygen (Zervas et al., 2018). Not only promoting plant growth and utilizing methanol, as a few research pointed previously (Eevers et al., 2015; Li et al., 2020; Ventrino et al., 2014) they are also involved in xenobiotics biodegradation for which the genes are enriched in their accessory and unique genes.

This study found that all the studied strains had some PGP traits in general, i.e. conservation of PGP genes (*miaA*, *acdS*, and *dcyD*), IAA production, and their antifungal activity. The sub-clades A4, A5, and clade C members, showed a strong antifungal activity. Most clade A4 and A5 species were isolated from the plants. Development of plant-growth promoting, antifungal biocontrol agents containing *Methylobacterium* species will be possible. However, it is also known that only specific species can be isolated from specific plants (Minami et al., 2016; Tani et al., 2015), suggesting that there is a species-species specificity. By taking a comparative genome-based approach on the isolates from plants, I may also be able to find genes or the factors that determine the specificity.

4.2. Taxonomic consequences: synonymous species, new combinations and emendation of the genus

Through the use of dDDH and ANI to delineate species boundaries, several synonyms have been found within the genus *Methylobacterium* (Hördt et al., 2020) and the family *Rhizobiales* (Volpiano et al., 2021). My results support the consideration of *Mb. oryzae* and *Mb. phyllosphaerae* as the same species (Green & Ardley, 2018; Hördt et al., 2020; Kröber et al., 2021). Here I also propose reclassifying *Mb. phyllosphaerae* (Madhaiyan et al., 2009) and *Mb. oryzae* (Madhaiyan et al., 2007) as later heterotypic synonyms of *Mb. fujisawaense* (Green et al., 1988). In addition, as *Mr. thiocyanatum* (Wood et al., 1998) and *Mr. populi* (Van Aken et al., 2004) have >95% ANI and >80% dDDH, both of which are higher than the species-species separation thresholds, *Mr. populi* is a later heterotypic synonym of *Mr. thiocyanatum*.

Mb. dichloromethanicum (Doronina et al., 2000) and *Mb. chloromethanicum* (McDonald et al., 2001) are later heterotypic synonyms of *Mb. extorquens* as reported previously (Kato et al., 2005). Hördt et al. (2020) have recently suggested uniting the genera *Methylobacterium* and *Methylobacterium*. Due to these groups' high dDDH values (70~79%), those authors also proposed classifying *Mb. chloromethanicum* as *Mb. dichloromethanicum* subsp. *chloromethanicum*, subsp. nov. and *Mr. extorquens* as *Mb. dichloromethanicum* subsp. *extorquens*, subsp. nov. In a recent announcement of newly validated names (Validation list no. 194, DOI 10.1099/ijsem.0.004244), the name *Mb. dichloromethanicum* subsp. *chloromethanicum* was validated, but *Mb. dichloromethanicum* subsp. *extorquens*, subsp. nov. was not. Instead, the name *Mb. dichloromethanicum* subsp. *dichloromethanicum* was also created automatically according to a taxonomical rule in the same validation list. The fact that the name *Mb. extorquens* (Bousfield & Green, 1985; Urakami & Komagata, 1984) had priority was ignored. Therefore, I propose classifying the names as *Mb. extorquens* subsp. *dichloromethanicum* subsp. nov., *Mb. extorquens* subsp. *chloromethanicum* subsp. nov., and *Mb. extorquens* subsp. *extorquens* subsp. nov.

Emended description of *Methylobacterium* Patt, Cole, and Hanson 1976; emend. Green and Bousfield 1983 (Approved Lists 1980)

The description is as given before (Green & Bousfield, 1983; Patt et al., 1976) with the following additions. Some strains exhibit Ln^{3+} -dependent growth on methanol. The genomic G+C content is 66-72.8%. The type species is *Methylobacterium organophilum* and its approximate genome size is 5.07 Mbp.

Description of *Methylobacterium extorquens* subsp. *dichloromethanicum* subsp. nov. comb. nov

Basonym: *Methylobacterium dichloromethanicum* subsp. *dichloromethanicum* (Doronina et al. 2000) Hördt et al. 2020, *Methylobacterium dichloromethanicum* (Doronina et al., 2000).

M. ex.tor'quens. subsp. dichlo.ro.me.tha'ni.cum (di.chlor'o'meth.an'ic.um. M.L. n. chlor; N.L. neut. n. methanum, methane; L. pref. di-, two; N.L. neut. adj. *dichloromethanicum*, dichloromethane utilizing). The description is as given for *Methylobacterium dichloromethanicum* (Doronina et al., 2000). The type

strain is DM4 = VKM B-2191 = DSM 6343 = CIP 106787.

Description of *Methylobacterium extorquens* subsp. *chloromethanicum* subsp. nov. comb. nov

Basonym: *Methylobacterium dichloromethanicum* subsp. *chloromethanicum* (McDonald et al. 2001)

Hördt et al. 2020, *Methylobacterium chloromethanicum* (McDonald et al., 2001),

M. ex.tor'quens. subsp. *chlo.ro.me.tha'ni.cum* (N.L. neut. n. *chloromethanicum*, chloromethane-utilizing).

The description is as given for *Methylobacterium chloromethanicum* (McDonald et al., 2001). The type

strain is CM4 = VKM B-2223 = NCIMB 13688.

Description of *Methylobacterium extorquens* subsp. *extorquens* subsp. nov., comb. nov.

Basonym *Methylobacterium extorquens* (Green and Ardley, 2018), *Methylobacterium extorquens* (Urakami and Komagata 1984) Bousfield and Green 1985, *Protomonas extorquens* (ex Bassalik 1913) Urakami and Komagata 1984. Automatically created based on Rule 40d. and lowering to subspecies rank from *Methylobacterium extorquens* (Green and Ardley, 2018).

M. ex.tor'quens. subsp. *ex.tor'quens* subsp.nov (ex.tor'quens. L. part. adj. *extorquens*, twisting out). The description is as given for *Protomonas extorquens* (ex Bassalik 1913) Urakami and Komagata 1984, Bousfield and Green 1985. The type strain is TK 0001 = ATCC 43645 = CCUG 2084 = DSM 1337 = JCM 2802 = NBRC 15687 = VKM B-2064.

➤ 5. Conclusion

The Pink pigmented facultative methylotrophs PPFMs are a highly important bacterial group in nature. They participate in methanol cycle and transfer this simple and abundant C1 compound into energy forms and other carbons. The abundance of methanol on the plant phyllosphere permitted for a strong association between these bacteria and the plant. The PPFMs can increase the plant tolerance to biotic and abiotic stresses and at the same time produce hormones which can enhance the plant growth. Phylogenetically the PPFMs are three groups of heterogeneous species (A, B, and C) and two genera *Methylobacterium* and *Methylobacterium*. Many researches have pointed out recently the importance of investigating the clear boundaries between these two genera. All the previous facts and the curiosity to look beyond the traditional genotyping and phenotyping studies, encouraged me to do this comparative genomics based on whole genome data.

We completed whole-genome information for the type strains of PPFMs by newly sequencing the type strains of 29 species in this study. The genome sequence of the type species of the genus *Methylobacterium*, *Mb. organophilum* NBRC 15689^T, is announced in this study. Using whole-genome data, we were able to classify the 62 PPFM species into eight homogeneous groups. Species from *Methylobacterium* and *Methylobacterium* were nested together, indicating the importance of reconsidering their separation into two genera. Only one of the eight groups (Clade C) was uniquely separated from all other species; the members of this clade are also phenotypically different from the others. Methylotrophy gene clusters were one of the unique features common to the PPFMs as opposed to the other genera in the *Methylobacteriaceae*. All PPFM genomes had at least one cluster of methanol dehydrogenase (MDH), and Ln^{3+} -dependent MDH was conserved in almost all species. This discovery was verified by phenotyping, which revealed that almost all of the PPFMs were able to grow on methanol with La^{3+} supplementation. These findings highlight the importance of Ln^{3+} as one of the main environmental factors assuring the survival of PPFMs. Screening for other factors related to plant symbiosis revealed that all PPFMs produce IAA and trigger no immune response in rice, and that some can inhibit pathogenic fungal growth *in vivo*.

➤ 6. Acknowledgements

This work is supported by Japan Society for the Promotion of Science KAKENHI (18H02129 and 21H02105 to AT, and 16H06279 (PAGS)), and Ohara foundation. I thank Drs. Tomonori Shinya and David Wari for the rice cell experiment, and Dr. Nobuhiro Suzuki and Mr. Masashi Matsusaki for *Fusarium* strain, Dr. Yoshitoshi Ogura for assessing the assembly quality and helping in the submission process, Dr. Hideto Takami for his help and assistance in Genomape analysis.

I would also like to express my full gratitude and acknowledge for help of Dr. Nurettin Sahin throughout the whole analysis. I highly appreciate the support and the guidance that I have received from Dr. Akio Tani and our lab technician Mrs. Yoshiko Fujitani.

➤ 7. Figure captions

Figure 1. Methanol assimilation pathway.

Figure 2. TYGS- 16S rRNA tree. This tree was created from the 16S rRNA gene sequences extracted from the assemblies. The tree shows 59 PPFM species and an outgroup represented by *Microvirga brassicacearum*^T.

Figure 3. The correlation of pairwise PCR-derived 16S rRNA gene identity, dDDH (digital DNA-DNA hybridization), and ANI (average nucleotide identity) values among all combinations of PPFMs. (A) The correlation between dDDH and ANI, (B) the correlation between 16S rRNA and ANI, and (C) the correlation between 16S rRNA and dDDH. The lower panels show plots of species combinations with high correlation values.

Figure 4. Heat map created with the genome-genome pairwise similarity values calculated with FastANI. Arrows point out the species with similarity values higher than the species-species separation values 95-96%. Rectangles point out the homogenous clusters indicated as A1 to A5, B, B2, and C1.

Figure 5. TYGS-genome blast distance phylogeny (GBDP) of 62 PPFM type strains and an outgroup represented by *Microvirga brassicacearum*^T. Coloured column bars (left to right) indicate each strain's clade, genome GC% content, genome size in kb, and number of coding sequences. Branch lengths are scaled based on GBDP formula d5 and numbers above the branch are bootstrap support values from 100 replications.

Figure 6. Phylogenetic trees based on (A) pan matrix, binary gene presence and absence matrix (1/0) and (B) concatenated sequences of core 711 proteins from the 62 PPFMs. Both trees are generated with BPGA pipeline, and presented with iTOL.

Figure 7. The presence and absence of functional modules and pathways found in PPFMs and the other *Methylobacteriaceae* strains analysed using the Genomagle pipeline. The presence and absence matrix

(1/0) is presented in black and white, respectively. KEGG modules ID and pathways names are shown on the right. Colours: pink, specific to PPFM species; blue, specific for *Microvirga* spp.; red, specific to A3 and A4; purple, specific to B, B2 and some species from A sub-clades; green, specific to clade C among PPFMs; and grey, nitrogen fixation and nodulation genes found in *Mb. nodulans* ORS 2060^T.

Figure 8. Essential pathways related to C1 metabolism in PPFMs. These pathways include methanol, methylamine, and dichloromethane metabolism. The conservation of gene clusters in each *Methylobacterium* species is indicated by colored dots. *Methylobacterium* type strains are arranged in sub-clades as shown in Table S9. A gene cluster is considered conserved when more than 75% of its genes are present. Black dots indicate a missing cluster or the presence of less than 75% of the gene cluster.

Figure 9. Maximum likelihood phylogeny of PQQ-ADH homologues found in *Methylobacteriaceae* genomes. The tree was built based on the alignment of 252 homologues and 49 reference sequences. The tree was created with FastTree v2.1, and the local support values (>0.7) on the branches shown as a percentage. The reference sequences are shown in red. The highlighted clades labelled G1-G9 represent the following groups: G1, XoxF5-type MDH (XoxF1 in AM1); G2, XoxF5-type MDH (XoxF2 in AM1); G3, XoxF5-type MDH (XoxF1 in 22A); G4, MxaF-type MDH; G5, XoxF3-type MDH; G6, ExaF-type ADH (ExaF in 22A); G7, PQQ-ADH type 6a; G8, PQQ-ADH type 2a; and G9, PQQ-ADH type 2b (ExaF in AM1). Circles around the tree indicate PPFM clades and conserved amino acid residues that are important for the coordination of metals (Ca^{2+} or Ln^{3+}) in the active sites of PQQ-ADHs. The positions of amino acid residues corresponding to the mature XoxF from *Methylacidiphilum fumariolicum* SolV (C104, C105, E172, N256, D269, D299, Y300, and D301) are presented in the outer circles from the inside to the outside in that order.

Figure 10. Functional annotation and distribution of core, accessory, and unique genes of the PPFMs within KEGG main groups. (A) KEGG general category, and (B) more specific details of KEGG category. The gene count in each category is shown as a percentage of the total count in each group (core, accessory, and unique). * The overview category includes carbon metabolism (ko01200), oxocarboxylic acid

metabolism (ko01210), fatty acid metabolism (ko01212), degradation of aromatic compounds (ko01220), and biosynthesis of amino acids (ko01230).

Figure 11. Growth assay of some *Methylobacterium* type strains on (A) Methanol medium, (B) Methanol supplemented by La^{3+} . The experiment was run in technical replicates. The data are shown as the average $\text{OD}^{600} \pm \text{SD}$ (n=3).

Figure 12. Antifungal activity of PPFMs against *Fusarium oxysporum*. Among the tested strains, only those with strong inhibition activity are shown. Each PPFM was inoculated onto half of a PDA plate, and the fungus was allowed to grow from the centre. The plates were incubated at 25°C for 7-10 days.

Figure 13. Rice cell elicitation assay. The PPFMs were inoculated into rice cell suspension, then incubated for 24 h. The numbers represent the tested species listed in Table S13. DC3000, *Pseudomonas syringae* DC3000; 1-7 and 4-24, *Serratia marcescens* subsp. *marcescens* isolates 1-7 and 4-24; GlcNAc, chitin oligomer (10 nM); and NC, negative control (saline solution).

➤ 8. Tables captions

Table 1. Genome information used in this study

Table 2. Reference sequences used in the analysis

Table 3. ADH and MDH reference sequences used as query.

Table S4. Minimum medium components.

Table 5. Details of the newly sequenced assemblies. *1, CLC-GW, CLC Genomics Workbench.

Table 6. Pairwise dDHH values. Values were calculated with TYGS server. The Threshold for species-species separation is less than 70%. Values highlighted in dark red indicated combination with pairwise dDDH values higher than the threshold.

Table 7. Pairwise ANI values. Values were calculated using FastANI. The threshold for species-species separation is less than 95 %. Values highlighted in dark green indicated combination with pairwise ANI values higher than the threshold.

Table 8. Core, accessory, unique, and exclusively absent gene counts in each type strain.

Table 9. Conservation of methylotrophy genes. Values represent pathway completeness percentage.

Table 10. Detailed information of methylotrophy genes conservation in the PPFMs and the other *Methylobacteriaceae* species. Values represent percentage of identity against the query proteins.

Table 11. Detected gene clusters for secondary metabolites. Values represent clusters counted in the genome of each type strain.

Table 12. Plant growth-promoting genes. Values represent percentage of identity against the query proteins.

Table 13. Phenotype analysis of PPFMs. The symbols represent the growth strength of the strain; +, Positive; w, weak growth; vw, very weak growth; -, negative growth. *, -, no activity; nd, not determined, values represent inhibition percentage compared to the control.

Table 14. API 20NE test result. Explanation of each test: NO₃, reduction of nitrates to nitrites; TRP, indole production; GLU, fermentation of glucose; ADH, arginine dihydrolase; URE, urease; ESC, β -glucosidase (esculin); GEL, protein hydrolysis (gelatin); PNPG, β -galactosidase (p-nitrophenyl- β -D-

galactopyranosidase); GLU, glucose; ARA, arabinose; MNE, mannose; MAN, mannitol; NAG, N-acetylglucosamine; MAL, maltose; GNT, gluconate; CAP, capric acid; ADI, adipic acid; MLT, malate; CIT, citrate; and PAC, phenylacetic acid. The values represent the growth strength of the strain; 3, strong; 2, weak; 1, very weak; and -, not tested.

➤ 9. References

- Abanda-Nkpwatt, D., Müsch, M., Tschiersch, J., Boettner, M., and Schwab, W. (2006). Molecular interaction between *Methylobacterium extorquens* and seedlings: Growth promotion, methanol consumption, and localization of the methanol emission site. *J. Exp. Bot.* 57, 4025–4032. doi:10.1093/jxb/erl173.
- Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. doi:10.1093/nar/gky379.
- Alamgir, K. M., Masuda, S., Fujitani, Y., Fukuda, F., and Tani, A. (2015). Production of ergothioneine by *Methylobacterium* species. *Front. Microbiol.* 6:1185. doi:10.3389/fmicb.2015.01185.
- Alessa, O., Ogura, Y., Fujitani, Y., Takami, H., Hayashi, T., Sahin, N., et al. (2021). Comprehensive comparative genomics and phenotyping of *Methylobacterium* Species. *Front. Microbiol.* 12:740610 doi: 10.3389/fmicb.2021.740610.
- Anderson, D. J., Morris, C. J., Nunn, D. N., Anthony, C., and Lidstrom, M. E. (1990). Nucleotide sequence of the *Methylobacterium extorquens* AM1 *moxF* and *moxJ* genes involved in methanol oxidation. *Gene* 90, 173–176. doi:10.1016/0378-1119(90)90457-3.
- Anthony, C. (1983). The biochemistry of methylotrophs. *Trends Biochem. Sci.* 160, 2–3. doi:10.1016/0968-0004(83)90116-0.
- Ardanov, P., Sessitsch, A., Häggman, H., Kozyrovska, N., and Pirttilä, A. M. (2012). *Methylobacterium*-induced endophyte community changes correspond with protection of plants against pathogen attack. *PLoS One* 77:e46802.. doi:10.1371/journal.pone.0046802.
- Aslam, Z., Lee, C. S., Kim, K. H., Im, W. T., Ten, L. N., and Lee, S. T. (2007). *Methylobacterium jeotgali* sp. nov., a non-pigmented, facultatively methylotrophic bacterium isolated from jeotgal, a traditional Korean fermented seafood. *Int. J. Syst. Evol. Microbiol.* 57, 566–571. doi:10.1099/ijs.0.64625-0.

- Auch, A. F., von Jan, M., Klenk, H. P., and Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genomic Sci.* 2, 117–134. doi:10.4056/sigs.531120.
- Bazurto, J. V., Bruger, E. L., Lee, J. A., Lambert, L. B., and Marx, C. J. (2021). Formaldehyde-responsive proteins TtmR and EfgA reveal a trade-off between formaldehyde resistance and efficient transition to methylotrophy in *Methylobacterium extorquens*. *J. Bacteriol.* 203:e00589-20. doi: 10.1128/JB.00589-20.
- Bazurto, J. V., Nayak, D. D., Ticak, T., Davlieva, M., Lee, J. A., Lambert, L. B., et al. (2020). EfgA is a conserved formaldehyde sensor that halts bacterial translation in response to elevated formaldehyde. *BioRxiv [Preprint]*. doi: 10.1101/2020.10.16.343392.
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, 81–87. doi:10.1093/nar/gkz310.
- Bousfield, I. J., and Green, P. N. (1985). Reclassification of bacteria of the genus *protomonas* Urakami and Komagata 1984 in the genus *Methylobacterium* (Patt, Cole, and Hanson) emend. Green and Bousfield 1983. *Int. J. Syst. Bacteriol.* 35, 209–209. doi:10.1099/00207713-35-2-209.
- Bruhn, D., Møller, I. M., Mikkelsen, T. N., and Ambus, P. (2012). Terrestrial plant methane production and emission. *Physiol. Plant.* 144, 201–209. doi:10.1111/j.1399-3054.2011.01551.x.
- Calatrava, V., Hom, E. F. Y., Llamas, Á., Fernández, E., and Galván, A. (2018). OK, thanks! A new mutualism between *Chlamydomonas* and *methylobacteria* facilitates growth on amino acids and peptides. *FEMS Microbiol. Lett.* 365, 1–9. doi:10.1093/femsle/fny021.
- Caputo, A., Lagier, J.-C., Azza, S., Robert, C., Mouelhi, D., Fournier, P.-E., et al. (2016). *Microvirga massiliensis* sp. nov., the human commensal with the largest genome. *Microbiologyopen* 5, 307–322. doi:https://doi.org/10.1002/mbo3.329.
- Chan, S. I., Chuankhayan, P., Reddy Nareddy, P. K., Tsai, I. K., Tsai, Y. F., Chen, K. H. C., et al. (2021). Mechanism of pyrroloquinoline quinone-dependent hydride transfer Chemistry from Spectroscopic and High-Resolution X-ray Structural Studies of the Methanol Dehydrogenase from

- Methylococcus capsulatus* (Bath). *J. Am. Chem. Soc.* 143, 3359–3372. doi:10.1021/jacs.0c11414.
- Chaudhari, N. M., Gupta, V. K., and Dutta, C. (2016). BPGA-an ultra-fast pan-genome analysis pipeline. *Sci. Rep.* 6, 1–10. doi:10.1038/srep24373.
- Chen, Y., McAleer, K. L., and Colin Murrell, J. (2010). Monomethylamine as a nitrogen source for a nonmethylophilic bacterium, *agrobacterium tumefaciens*. *Appl. Environ. Microbiol.* 76, 4102–4104. doi:10.1128/AEM.00469-10.
- Chistoserdov, A. Y., Chistoserdova, L. V., McIntire, W. S., and Lidstrom, M. E. (1994). Genetic organization of the mau gene cluster in *Methylobacterium extorquens* AM1: Complete nucleotide sequence and generation and characteristics of mau mutants. *J. Bacteriol.* 176, 4052–4065. doi:10.1128/jb.176.13.4052-4065.1994.
- Chistoserdova, L. (2011). Modularity of methylophilicity, revisited. *Environ. Microbiol.* 13, 2603–2622. doi:10.1111/j.1462-2920.2011.02464.x.
- Chistoserdova, L., and Lidstrom, M. E. (1997). Molecular and mutational analysis of a DNA region separating two methylophilicity gene clusters in *Methylobacterium extorquens* AM1. *Microbiology* 143, 1729–1736. doi:10.1099/00221287-143-5-1729.
- Cuccuru, G., Orsini, M., Pinna, A., Sbardellati, A., Soranzo, N., Travaglione, A., et al. (2014). Orione, a web-based framework for NGS analysis in microbiology. *Bioinformatics* 30, 1928–1929. doi:10.1093/bioinformatics/btu135.
- Dedysh, S. N., Dunfield, P. F., and Trotsenko, Y. A. (2004). Methane utilization by *Methylobacterium* species: New evidence but still no proof for an old controversy [3]. *Int. J. Syst. Evol. Microbiol.* 54, 1919–1920. doi:10.1099/ijms.0.63493-0.
- Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., et al. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 106, 16428–16433. doi:10.1073/pnas.0905240106.
- Doronina, N. V., Torgonskaya, M. L., Fedorov, D. N., and Trotsenko, Y. A. (2015). Aerobic *methylobacteria* as promising objects of modern biotechnology (Review). *Appl. Biochem. Microbiol.* 51, 125–134. doi:10.1134/S0003683815020052.

- Doronina, N. V, Sokolov, A. P., and Trotsenko, Y. A. (1996). Isolation and initial characterization of aerobic chloromethane-utilizing bacteria. *FEMS Microbiol. Lett.* 142, 179–183.
doi:10.1111/j.1574-6968.1996.tb08427.x.
- Doronina, N. V, Trotsenko, Y. A., Tourova, T. P., Kuznetsov, B. B., and Leisinger, T. (2000). *Methylophila helvetica* sp. nov. and *Methylobacterium dichloromethanicum* sp. nov. novel aerobic facultatively methylotrophic bacteria utilizing dichloromethane. *Syst. Appl. Microbiol.* 23, 210–218. doi:10.1016/S0723-2020(00)80007-7.
- Eevers, N., Van Hamme, J. D., Bottos, E. M., Weyens, N., and Vangronsveld, J. (2015). Draft genome sequence of *Methylobacterium radiotolerans*, a DDE-degrading and plant growth-promoting strain isolated from Cucurbita pepo. *Genome Announc.* 3, 4–5. doi:10.1128/genomeA.00488-15.
- Ekimova, G. A., Fedorov, D. N., Tani, A., Doronina, N. V., and Trotsenko, Y. A. (2018). Distribution of 1-aminocyclopropane-1-carboxylate deaminase and d-cysteine desulphydrase genes among type species of the genus *Methylobacterium*. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 111, 1723–1734. doi:10.1007/s10482-018-1061-5.
- Feng, G. Da, Chen, W., Zhang, X. J., Zhang, J., Wang, S. N., and Zhu, H. (2020). *Methylobacterium nonmethylotrophicum* sp. Nov., isolated from tungsten mine tailing. *Int. J. Syst. Evol. Microbiol.* 70, 2867–2872. doi:10.1099/ijsem.0.004112.
- Gallego, V., García, M. T., and Ventosa, A. (2005). *Methylobacterium hispanicum* sp. nov. and *Methylobacterium aquaticum* sp. nov., isolated from drinking water. *Int. J. Syst. Evol. Microbiol.* 55, 281–287. doi:10.1099/ijs.0.63319-0.
- Good, N. M., Vu, H. N., Suriano, C. J., Subuyuj, G. A., Skovran, E., and Martinez-Gomez, N. C. (2016). Pyrroloquinoline quinone ethanol dehydrogenase in *Methylobacterium extorquens* AM1 extends lanthanide-dependent metabolism to multicarbon substrates. *J. Bacteriol.* 198, 3109–3118. doi:10.1128/JB.00478-16.
- Green, P. N., and Ardley, J. K. (2018). Review of the genus *Methylobacterium* and closely related organisms: A proposal that some *Methylobacterium* species be reclassified into a new genus, *Methylorubrum* gen. nov. *Int. J. Syst. Evol. Microbiol.* 68, 2727–2748.

doi:10.1099/ijsem.0.002856.

- Green, P. N., and Bousfield, I. J. (1982). A taxonomic study of some Gram-negative facultatively methylotrophic bacteria. *J. Gen. Microbiol.* 128, 623–638. doi:10.1099/00221287-128-3-623.
- Green, P. N., Bousfield, I. J., and Hood, D. (1988). Three New *Methylobacterium* Species: *M. rhodesianum* sp. nov., *M. zatmanii* sp. nov., *M. fujisawaense* sp. npv. *Int. J. Sytematic Bacteriol.* 38, 124–127.
- Grossi, C. E. M., Fantino, E., Serral, F., Zawoznik, M. S., Fernandez Do Porto, D. A., and Ulloa, R. M. (2020). *Methylobacterium* sp. 2A is a plant growth-promoting rhizobacteria that has the potential to improve potato crop yield under adverse conditions. *Front. Plant Sci.* 11:71. doi:10.3389/fpls.2020.00071.
- Gruffaz, C., Muller, E. E. L., Louhichi-Jelail, Y., Nelli, Y. R., Guichard, G., and Bringel, F. (2014). Genes of the N-Methylglutamate pathway are essential for growth of *Methylobacterium extorquens* DM4 with monomethylamine. *Appl. Environ. Microbiol.* 80, 3541–3550. doi:10.1128/AEM.04160-13.
- Hedlund, B. P., Jiao, J.-Y., Zhou, E.-M., Wang, C. C. C., Venkateswaran, K., Bijlani, S., et al. (2021). *Methylobacterium ajmalii* sp. nov., isolated from the International Space Station. *Front. Microbiol.* 12:639396. doi: 10.3389/fmicb. 2021.639396
- Henz, S. R., Huson, D. H., Auch, A. F., Nieselt-Struwe, K., and Schuster, S. C. (2005). Whole-genome prokaryotic phylogeny. *Bioinformatics* 21, 2329–2335. doi:10.1093/bioinformatics/bth324.
- Hibi, Y., Asai, K., Arafuka, H., Hamajima, M., Iwama, T., and Kawai, K. (2011). Molecular structure of La³⁺-induced methanol dehydrogenase-like protein in *Methylobacterium radiotolerans*. *J. Biosci. Bioeng.* 111, 547–549. doi:10.1016/j.jbiosc.2010.12.017.
- Hood, D. W., Dow, C. S., and Green, P. N. (1987). DNA:DNA hybridization studies on the pink-pigmented facultative methylotrophs. *J. Gen. Microbiol.* 133, 709–720. doi:10.1099/00221287-133-3-709.
- Hördt, A., López, M. G., Meier-Kolthoff, J. P., Schleuning, M., Weinhold, L. M., Tindall, B. J., et al. (2020). Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of

- alphaproteobacteria. *Front. Microbiol.* 11:468. doi: 10.3389/fmicb.2020.00468
- Huang, J., Yu, Z., and Chistoserdova, L. (2018). Lanthanide-dependent methanol dehydrogenases of XoxF4 and XoxF5 clades are differentially distributed among methylotrophic bacteria and they reveal different biochemical properties. *Front. Microbiol.* 9:1366. doi:10.3389/fmicb.2018.01366.
- Ivanova, E. G., Doronina, N. V., and Trotsenko, Y. A. (2001). Aerobic *Methylobacteria* Are Capable of Synthesizing Auxins. 70, 392–397.
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9, 1–8. doi:10.1038/s41467-018-07641-9.
- Johnson, J. L. (1973). Use of nucleic acid homologies in the taxonomy of anaerobic bacteria. *Int. J. Syst. Bacteriol.* 23, 308–315. doi:10.1099/00207713-23-4-308.
- Jorge, G. L., Kisiala, A., Morrison, E., Aoki, M., Nogueira, A. P. O., and Emery, R. J. N. (2019). Endosymbiotic *Methylobacterium oryzae* mitigates the impact of limited water availability in lentil (*Lens culinaris* Medik.) by increasing plant cytokinin levels. *Environ. Exp. Bot.* 162, 525–540. doi:10.1016/j.envexpbot.2019.03.028.
- Jourand, P., Giraud, E., Béna, G., Sy, A., Willems, A., Gillis, M., et al. (2004). *Methylobacterium nodulans* sp. nov., for a group of aerobic, facultatively methylotrophic, legume root-nodule-forming and nitrogen-fixing bacteria. *Int. J. Syst. Evol. Microbiol.* 54, 2269–2273. doi:10.1099/ijs.0.02902-0.
- Kato, Y., Asahara, M., Arai, D., Goto, K., and Yokota, A. (2005). Reclassification of *Methylobacterium chloromethanicum* and *Methylobacterium dichloromethanicum* as later subjective synonyms of *Methylobacterium extorquens* and of *Methylobacterium lusitanum* as a later subjective synonym of *Methylobacterium rhodesianum*. *J. Gen. Appl. Microbiol.* 51, 287–299. doi:10.2323/jgam.51.287.
- Kayser, M. F., Ucurum, Z., and Vuilleumier, S. (2002). Dichloromethane metabolism and C1 utilization genes in *Methylobacterium* strains. *Microbiology* 148, 1915–1922. doi:10.1099/00221287-148-6-1915.
- Keltjens, J. T., Pol, A., Reimann, J., and Op Den Camp, H. J. M. (2014). PQQ-dependent methanol

- dehydrogenases: Rare-earth elements make a difference. *Appl. Microbiol. Biotechnol.* 98, 6163–6183. doi:10.1007/s00253-014-5766-8.
- Kim, J., Chhetri, G., Kim, I., Kim, H., Kim, M. K., and Seo, T. (2019). *Methylobacterium terrae* sp. nov., a radiation-resistant bacterium isolated from gamma ray-irradiated soil. *J. Microbiol.* 57, 959–966. doi:10.1007/s12275-019-9007-9.
- Knief, C., Frances, L., and Vorholt, J. A. (2010). Competitiveness of diverse *Methylobacterium* strains in the phyllosphere of *Arabidopsis thaliana* and identification of representative models, including *M. extorquens* PA1. *Microb. Ecol.* 60, 440–452. doi:10.1007/s00248-010-9725-3.
- Koenig, R. L., Morris, R. O., and Polacco, J. C. (2002). tRNA is the source of low-level trans-zeatin production in *Methylobacterium* spp. *J. Bacteriol.* 184, 1832–1842. doi:10.1128/JB.184.7.1832-1842.2002.
- Kröber, E., Cunningham, M. R., Peixoto, J., Spurgin, L., Wischer, D., Kruger, R., et al. (2021a). Comparative genomics analyses indicate differential methylated amine utilization trait within members of the genus *Gemmobacter*. *Environ. Microbiol. Rep.* 13, 195–208. doi:https://doi.org/10.1111/1758-2229.12927.
- Kröber, E., Wende, S., Kanukollu, S., Buchen-Tschiskale, C., Besaury, L., Keppler, F., et al. (2021b). ¹³C-chloromethane incubations provide evidence for novel bacterial chloromethane degraders in a living tree fern. *Environ. Microbiol.* n/a. doi:https://doi.org/10.1111/1462-2920.15638.
- Krug, L., Morauf, C., Donat, C., Müller, H., Cernava, T., and Berg, G. (2020). Plant growth-promoting methylobacteria selectively increase the biomass of biotechnologically relevant microalgae. *Front. Microbiol.* 11:427. doi:10.3389/fmicb.2020.00427.
- Kwak, M. J., Jeong, H., Madhaiyan, M., Lee, Y., Sa, T. M., Oh, T. K., et al. (2014). Genome information of *Methylobacterium oryzae*, a plant-probiotic methylotroph in the phyllosphere. *PLoS One* 9:e106704. doi:10.1371/journal.pone.0106704.
- Lai, K., Nguyen, N. T., Miwa, H., Yasuda, M., Nguyen, H. H., and Okazaki, S. (2020). Diversity of *Methylobacterium* spp. in the rice of the vietnamese mekong delta. *Microbes Environ.* 35, 3–7. doi:10.1264/jsme2.ME19111.

- Latypova, E., Yang, S., Wang, Y. S., Wang, T., Chavkin, T. A., Hackett, M., et al. (2010). Genetics of the glutamate-mediated methylamine utilization pathway in the facultative methylotrophic beta-proteobacterium *Methyloversatilis universalis* FAM5. *Mol. Microbiol.* 75, 426–439. doi:10.1111/j.1365-2958.2009.06989.x.
- Lee, I., Chalita, M., Ha, S. M., Na, S. I., Yoon, S. H., and Chun, J. (2017). ContEst16S: An algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *Int. J. Syst. Evol. Microbiol.* 67, 2053–2057. doi:10.1099/ijsem.0.001872.
- Li, X., Wang, J., Jia, Y., Reheman, A., and Yan, Y. (2020). The genome analysis of *Methylobacterium populi* yc-xj1 with diverse xenobiotics biodegrading capacity and degradation characteristics of related hydrolase. *Int. J. Mol. Sci.* 21, 1–18. doi:10.3390/ijms21124436.
- Lidstrom, M. E., and Chistoserdova, L. (2002). Plants in the Pink : Cytokinin Production by *Methylobacterium*. 184, 9193. doi:10.1128/JB.184.7.1818.2002.
- Lidstrom, M. E., and Stirling, D. I. (1990). Methylotrophs: genetics and commercial applications. *Annu. Rev. Microbiol.* 44, 27–58.
- MacDonald, R. C., and Fall, R. (1993). Detection of substantial emissions of methanol from plants to the atmosphere. *Atmos. Environ. Part A. Gen. Top.* 27, 1709–1713. doi:https://doi.org/10.1016/0960-1686(93)90233-O.
- Macey, M. C., Pratscher, J., Crombie, A. T., and Murrell, J. C. (2020). Impact of plants on the diversity and activity of methylotrophs in soil. *Microbiome* 8, 1–17. doi:10.1186/s40168-020-00801-4.
- Madhaiyan, M., Kim, B. Y., Poonguzhali, S., Kwon, S. W., Song, M. H., Ryu, J. H., et al. (2007). *Methylobacterium oryzae* sp. nov., an aerobic, pink-pigmented, facultatively methylotrophic, 1-aminocyclopropane-1-carboxylate deaminase-producing bacterium isolated from rice. *Int. J. Syst. Evol. Microbiol.* 57, 326–331. doi:10.1099/ijs.0.64603-0.
- Madhaiyan, M., Poonguzhali, S., Kang, B. G., Lee, Y. J., Chung, J. B., and Sa, T. M. (2010). Effect of co-inoculation of methylotrophic *Methylobacterium oryzae* with *Azospirillum brasilense* and *Burkholderia pyrrocinia* on the growth and nutrient uptake of tomato, red pepper and rice. *Plant Soil* 328, 71–82. doi:10.1007/s11104-009-0083-1.

- Madhaiyan, M., Poonguzhali, S., Kwon, S. W., and Sa, T. M. (2009). *Methylobacterium phyllosphaerae* sp. nov., a pink-pigmented, facultative methylotroph from the phyllosphere of rice. *Int. J. Syst. Evol. Microbiol.* 59, 22–27. doi:10.1099/ij.s.0.001693-0.
- Madhaiyan, M., Poonguzhali, S., Senthilkumar, M., Lee, J. S., and Lee, K. C. (2011). *Methylobacterium gossipiicola* sp. nov., a pink-pigmented, facultatively methylotrophic bacterium isolated from the cotton phyllosphere. *Int. J. Syst. Evol. Microbiol.* 62, 162–167. doi:10.1099/ij.s.0.030148-0.
- Maeng, S., Kim, D. U., Lim, S., Lee, B. H., Lee, K. eun, Kim, M., et al. (2021). *Methylobacterium radiodurans* sp. nov., a novel radiation-resistant *Methylobacterium*. *Arch. Microbiol.* 203, 3435–3442. doi:10.1007/s00203-021-02293-8.
- Masuda, S., Suzuki, Y., Fujitani, Y., Mitsui, R., Nakagawa, T., Shintani, M., et al. (2018). Lanthanide-dependent regulation of methylotrophy in *Methylobacterium aquaticum* strain 22A. *mSphere* 3:e00462-17. doi: 10.1128/ mSphere.00462-17.
- Mccarthy, B. J., and Bolton, E. T. (1963). An approach to the measurement of genetic relatedness among organisms. *Proc. Natl. Acad. Sci. U. S. A.* 50, 156–164. doi:10.1073/pnas.50.1.156.
- McDonald, I. R., Doronina, N. V, Trotsenko, Y. A., McAnulla, C., and Murrell, J. C. (2001). *Hyphomicrobium chloromethanicum* sp. nov. and *Methylobacterium chloromethanicum* sp. nov., chloromethane-utilizing bacteria isolated from a polluted environment. *Int. J. Syst. Evol. Microbiol.* 51, 119–122. doi:10.1099/00207713-51-1-119.
- Medema, M. H., Blin, K., Cimermancic, P., De Jager, V., Zakrzewski, P., Fischbach, M. A., et al. (2011). AntiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39(Suppl. 2), W339–W346. doi: 10.1093/nar/ gkr466.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. doi:10.1186/1471-2105-14-60.
- Meier-Kolthoff, J. P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* 10:2182. doi:10.1038/s41467-019-10210-3.

- Minami, T., Anda, M., Mitsui, H., Sugawara, M., Kaneko, T., Sato, S., et al. (2016). Metagenomic analysis revealed methylamine and ureide utilization of soybean-associated *Methylobacterium*. *Microbes Environ.* 31, 268–278. doi:10.1264/jsme2.ME16035.
- Mizuno, M., Yurimoto, H., Iguchi, H., Tani, A., and Sakai, Y. (2013). Dominant colonization and inheritance of *Methylobacterium* sp. strain OR01 on perilla plants. *Biosci. Biotechnol. Biochem.* 77, 1533–1538. doi:10.1271/bbb.130207.
- Nayak, D. D., and Marx, C. J. (2014). Methylamine utilization via the N-methylglutamate pathway in *Methylobacterium extorquens* PA1 involves a novel flow of carbon through C1 assimilation and dissimilation pathways. *J. Bacteriol.* 196, 4130–4139. doi:10.1128/JB.02026-14.
- Nemecek-Marshall, M., MacDonald, R. C., Franzen, J. J., Wojciechowski, C. L., and Fall, R. (1995). Methanol emission from leaves (Enzymatic detection of gas-phase methanol and relation of methanol fluxes to stomatal conductance and leaf development). *Plant Physiol.* 108, 1359–1368. doi:10.1104/pp.108.4.1359.
- Nunn, D. N., Day, D., and Anthony, C. (1989). The second subunit of methanol dehydrogenase of *Methylobacterium extorquens* AM1. *Biochem. J.* 260, 857–862. doi:10.1042/bj2600857.
- Nunn, D. N., and Lidstrom, M. E. (1986). Isolation and complementation analysis of 10 methanol oxidation mutant classes and identification of the methanol dehydrogenase structural gene of *Methylobacterium* sp. strain AM1. *J. Bacteriol.* 166, 581–590. doi:10.1128/jb.166.2.581-590.1986.
- Ochsner, A. M., Hemmerle, L., Vonderach, T., Nüssli, R., Bortfeld-Miller, M., Hattendorf, B., et al. (2019). Use of rare-earth elements in the phyllosphere colonizer *Methylobacterium extorquens* PA1. *Mol. Microbiol.* 111, 1152–1166. doi:10.1111/mmi.14208.
- Omer, Z. S., Tombolini, R., Broberg, A., and Gerhardson, B. (2004). Indole-3-acetic acid production by pink-pigmented facultative methylotrophic bacteria. *Plant Growth Regul.* 43, 93–96. doi:10.1023/B:GROW.0000038360.09079.ad.
- Oren, A., and Garrity, G. (2020). List of new names and new combinations previously effectively, but not validly, published. *Int. J. Syst. Evol. Microbiol.* 70, 4043–4049. doi:10.1099/ijsem.0.004244.
- Park, C., Lee, Y. S., Park, S. Y., and Park, W. (2018). *Methylobacterium currus* sp. nov., isolated from a

- car air conditioning system. *Int. J. Syst. Evol. Microbiol.* 68, 3621–3626.
doi:10.1099/ijsem.0.003045.
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi:10.1101/gr.186072.114.
- Patt, T. E., Cole, G. C., and Hanson, R. S. (1976). *Methylobacterium*, a New Genus of Facultatively Methylophilic Bacteria. *Int. J. Syst. Bacteriol.* 26, 226–229. doi:10.1099/00207713-26-2-226.
- Pirttilä, A. M., Laukkanen, H., Pospiech, H., Myllylä, R., and Hohtola, A. (2000). Detection of intracellular bacteria in the buds of scotch pine (*Pinus sylvestris* L.) by In situ hybridization. *Appl. Environ. Microbiol.* 66, 3073–3077. doi:10.1128/AEM.66.7.3073-3077.2000.
- Pol, A., Barends, T. R. M., Dietl, A., Khadem, A. F., Eygensteyn, J., Jetten, M. S. M., et al. (2014). Rare earth metals are essential for methanotrophic life in volcanic mudpots. *Environ. Microbiol.* 16, 255–264. doi:10.1111/1462-2920.12249.
- Roodi, D., Millner, J. P., McGill, C., Johnson, R. D., Jauregui, R., and Card, S. D. (2020). *Methylobacterium*, a major component of the culturable bacterial endophyte community of wild *Brassica* seed. *PeerJ* 8:e9514. doi: 10.7717/peerj. 9514.
- Roszczenko-Jasińska, P., Vu, H. N., Subuyuj, G. A., Crisostomo, R. V., Cai, J., Lien, N. F., et al. (2020). Gene products and processes contributing to lanthanide homeostasis and methanol metabolism in *Methylobacterium extorquens* AM1. *Sci. Rep.* 10, 1–15. doi:10.1038/s41598-020-69401-4.
- Schildkraut, C. L., Marmur, J., and Doty, P. (1961). The formation of hybrid DNA molecules and their use in studies of DNA homologies. *J. Mol. Biol.* 3, 595–617. doi:10.1016/S0022-2836(61)80024-7.
- Schneider, K., Peyraud, R., Kiefer, P., Christen, P., Delmotte, N., Massou, S., et al. (2012). The ethylmalonyl-CoA pathway is used in place of the glyoxylate cycle by *Methylobacterium extorquens* AM1 during growth on acetate. *J. Biol. Chem.* 287, 757–766.
doi:10.1074/jbc.M111.305219.
- Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069.

doi:10.1093/bioinformatics/btu153.

- Shinya, T., Hojo, Y., Desaki, Y., Christeller, J. T., Okada, K., Shibuya, N., et al. (2016). Modulation of plant defense responses to herbivores by simultaneous recognition of different herbivore-associated elicitors in rice. *Sci. Rep.* 6, 1–13. doi:10.1038/srep32537.
- Skovran, E., Palmer, A. D., Rountree, A. M., Good, N. M., and Lidstrom, M. E. (2011). XoxF is required for expression of methanol dehydrogenase in *Methylobacterium extorquens* AM1. *J. Bacteriol.* 193, 6032–6038. doi:10.1128/JB.05367-11.
- Subhaswaraj, P., Jobina, R., Parasuraman, P., and Siddhardha, B. (2017). Plant growth promoting activity of pink pigmented facultative methylotroph - *Methylobacterium extorquens* MM2 on *Lycopersicon esculentum* L. *J. Appl. Biol. Biotechnol.* 5, 042–046. doi:10.7324/jabb.2017.50107.
- Sy, A., Timmers, A. C. J., Knief, C., and Vorholt, J. A. (2005). Methylotrophic metabolism is advantageous for *Methylobacterium extorquens* during colonization of *Medicago truncatula* under competitive conditions. *Appl. Environ. Microbiol.* 71, 7245–7252. doi:10.1128/AEM.71.11.7245-7252.2005.
- Tani, A., Ogura, Y., Hayashi, T., and Kimbara, K. (2016). Complete genome sequence of *Methylobacterium aquaticum* strain 22A, isolated from a *Racomitrium japonicum* moss. *Genome Announc.* 3, 2–3. doi:10.1128/genomeA.00266-15.
- Tani, A., Sahin, N., Fujitani, Y., Kato, A., Sato, K., and Kimbara, K. (2015). *Methylobacterium* species promoting rice and barley growth and interaction specificity revealed with whole-cell matrix-assisted laser desorption / ionization-time-of-flight mass spectrometry (MALDI-TOF/ MS) analysis. *PLoS One* 10:e0129509. doi:10.1371/journal.pone.0129509.
- Tani, A., Takai, Y., Suzukawa, I., Akita, M., Murase, H., and Kimbara, K. (2012). Practical application of methanol-mediated mutualistic symbiosis between *Methylobacterium* species and a roof greening moss, *racomitrium japonicum*. *PLoS One* 7:e33800. doi:10.1371/journal.pone.0033800.
- Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., et al. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome.” *Proc. Natl. Acad. Sci. U. S. A.* 102, 13950–13955.

doi:10.1073/pnas.0506758102.

- Tsuji, K., Tsien, H. C., Hanson, R. S., DePalma, S. R., Scholtz, R., and LaRoche, S. (1990). 16S ribosomal RNA sequence analysis for determination of phylogenetic relationship among methylotrophs. *J. Gen. Microbiol.* 136, 1–10. doi:10.1099/00221287-136-1-1.
- Urakami, T., and Komagata, K. (1984). *Protomonas*, a new genus of facultatively methylotrophic bacteria. *Int. J. Syst. Evol. Microbiol.* 34, 188–201.
- Vadivukkarasi, P., and Bhai, R. S. (2020). Phyllosphere-associated *Methylobacterium*: a potential biostimulant for ginger (*Zingiber officinale* Rosc.) cultivation. *Arch. Microbiol.* 202, 369–375. doi:10.1007/s00203-019-01753-6.
- Van Aken, B., Peres, C. M., Doty, S. L., Yoon, J. M., and Schnoor, J. L. (2004). *Methylobacterium populi* sp. nov., a novel aerobic, pink-pigmented, facultatively methylotrophic, methane-utilizing bacterium isolated from poplar trees (*Populus deltoides* x *nigra* DN34). *Int. J. Syst. Evol. Microbiol.* 54, 1191–1196. doi:10.1099/ijms.0.02796-0.
- Ventorino, V., Sannino, F., Piccolo, A., Cafaro, V., Carotenuto, R., and Pepe, O. (2014). *Methylobacterium populi* VP2: Plant growth-promoting bacterium isolated from a highly polluted environment for polycyclic aromatic hydrocarbon (PAH) biodegradation. *Sci. World J.* 2014:931793. doi:10.1155/2014/931793.
- Verginer, M., Siegmund, B., Cardinale, M., Müller, H., Choi, Y., Míguez, C. B., et al. (2010). Monitoring the plant epiphyte *Methylobacterium extorquens* DSM 21961 by real-time PCR and its influence on the strawberry flavor. *FEMS Microbiol. Ecol.* 74, 136–145. doi:10.1111/j.1574-6941.2010.00942.x.
- Volpiano, C. G., Sant’Anna, F. H., Ambrosini, A., de São José, J. F. B., Beneduzi, A., Whitman, W. B., et al. (2021). Genomic metrics applied to *Rhizobiales* (*Hyphomicrobiales*): species reclassification, identification of unauthentic genomes and false type strains. *Front. Microbiol.* 12:661. doi:10.3389/fmicb.2021.614957.
- Vu, H. N., Subuyuj, G. A., Vijayakumar, S., Good, N. M., Martinez-Gomez, N. C., and Skovran, E. (2016). Lanthanide-dependent regulation of methanol oxidation systems in *Methylobacterium*

- extorquens* AM1 and their contribution to methanol growth. *J. Bacteriol.* 198, 1250–1259.
doi:10.1128/JB.00937-15.
- Vuilleumier, S., Chistoserdova, L., Lee, M. C., Bringel, F., Lajus, A., Yang, Z., et al. (2009). *Methylobacterium* genome sequences: A reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. *PLoS One* 4:e5584.
doi:10.1371/journal.pone.0005584.
- Wari, D., Kabir, M. A., Mujiono, K., Hojo, Y., Shinya, T., Tani, A., et al. (2019). Honeydew-associated microbes elicit defense responses against brown planthopper in rice. *J. Exp. Bot.* 70, 1683–1696.
doi:10.1093/jxb/erz041.
- Weisburg, W. G., Barns, S. M., Pelletier, D. A., and Lane, D. J. (1991). Weisburg 1991. 173, 697–703.
- Wellner, S., Lodders, N., Glaeser, S. P., and Kämpfer, P. (2013). *Methylobacterium trifolii* sp. nov. and *Methylobacterium thuringiense* sp. nov., methanolutilizing, pink-pigmented bacteria isolated from leaf surfaces. *Int. J. Syst. Evol. Microbiol.* 63, 2690–2699. doi:10.1099/ij.s.0.047787-0.
- William, S., Feil, H., and Copeland, A. (2012). Bacterial genomic DNA isolation using CTAB. *Sigma* 50.
- Wood, A. P., Kelly, D. P., McDonald, I. R., Jordan, S. L., Morgan, T. D., Khan, S., et al. (1998). A novel pink-pigmented facultative methylotroph, *Methylobacterium thiocyanatum* sp. nov., capable of growth on thiocyanate or cyanate as sole nitrogen sources. *Arch. Microbiol.* 169, 148–158.
doi:10.1007/s002030050554.
- Wu, M. L., Wessels, H. J. C. T., Pol, A., Op den Camp, H. J. M., Jetten, M. S. M., van Niftrik, L., et al. (2015). XoxF-type methanol dehydrogenase from the anaerobic methanotroph “*Candidatus Methyloirabilis oxyfera*.” *Appl. Environ. Microbiol.* 81, 1442–1451. doi:10.1128/AEM.03292-14.
- Yanpirat, P., Nakatsuji, Y., Hiraga, S., Fujitani, Y., Izumi, T., Masuda, S., et al. (2020). Lanthanide-dependent methanol and formaldehyde oxidation in *methylobacterium aquaticum* strain 22A. *Microorganisms* 8, 1–17. doi:10.3390/microorganisms8060822.
- Yoshida, S., Hiradate, S., Koitabashi, M., Kamo, T., and Tsushima, S. (2017). Phyllosphere *Methylobacterium* bacteria contain UVA-absorbing compounds. *J. Photochem. Photobiol. B Biol.*

167, 168–175. doi:10.1016/j.jphotobiol.2016.12.019.

Zervas, A., Zeng, Y., Madsen, A. M., and Hansen, L. H. (2018).
Culturomics of aerobic anoxygenic phototrophic bacteria in wheat phyllosphere revealed a divergent
evolutionary pattern of photosynthesis genes in *Methylobacterium* spp. bioRxiv[Preprint].
doi:10.1101/488148.

Table 1. Genome information used in this study

Species name	Strain	Source	Accession ID	Contigs No.	Bases count	CDS	tRNA	GC%	Completeness (%)	Contamination (%)	Strain heterogeneity (%)	16S rRNA gene copies	Identity to PCR-determined sequence
<i>Enterovirus rhinocerotis</i>	DSM 25903T	NCBI	GCA_004363955.1	20	5,166,770	4854	55	68.54	99.37	1.25	0	1	99.72
<i>Mb. adhaesivum</i>	DSM 17169T	This study	SAMD00388900	61	4,677,674	4,344	50	68.5	99.69	0.63	66.67	1	99.93
<i>Mb. aerolatum</i>	DSM 19013T	This study	SAMD00388901	67	4,877,645	4,448	47	69.8	99.69	0	0	1	99.92
<i>Mb. ajmalii</i>	IF75W-B2T	NCBI	GCA_016613415.1	192	6,802,552	6,284	89	71.07	99.69	0.78	0	missing	-
<i>Mb. aquaticum</i>	DSM 16371T	NCBI	GCA_001043915.1	512	7,425,331	6,736	67	69.89	97.66	0.74	0	1	100
<i>Mb. brachiatum</i>	TX0642T	NCBI	GCA_003697185.1	4	6,339,577	6,103	61	69.52	99.69	0	0	5	99.79-100
<i>Mb. brachythecii</i>	DSM 24105T	NCBI	GCA_014196345.1	29	5,409,070	5,107	51	66.6	99.37	0	0	1	100
<i>Mb. bullatum</i>	DSM 21893T	This study	SAMD00388906	54	4,956,045	4,603	52	67.07	99.84	0.31	0	1	100
<i>Mb. cerastii</i>	DSM 23679T	This study	SAMD00388907	150	5,388,584	5,128	56	69.56	99.37	0.31	0	1	100
<i>Mb. crusticola</i>	KCTC 52305T	This study	SAMD00388908	790	7,680,383	6,839	77	72.79	100	1.57	0	1	99.46
<i>Mb. currus</i>	PR1016AT	NCBI	GCA_003058325.1	5	7,745,135	7,109	111	70.19	100	0.78	0	12	94.69-100
<i>Mb. dankookense</i>	DSM 22415T	This study	SAMD00388910	273	6,046,408	5,737	60	70.39	98.12	0.94	0	1	100
<i>Mb. durans</i>	17SD2-17T	NCBI	GCA_003173715.1	1	6,788,375	6,772	68	68.74	98.76	0	0	4	100
<i>Mb. frigidaeris</i>	JCM 32048T	This study	SAMD00388912	513	7,909,330	7,217	73	69.88	100	1.33	0	1	100
<i>Mb. fujisawaense</i>	DSM 5686T	NCBI	GCA_014138435.1	25	5,971,667	5,545	63	70.03	100	0	0	1	100
<i>Mb. gnaphalii</i>	DSM 24027T	This study	SAMD00388914	195	4,782,631	4,706	48	65.96	99.27	1.04	80	1	99.93
<i>Mb. goesingense</i>	DSM 21331T	This study	SAMD00388915	263	5,334,262	5,006	49	68.8	96.97	0.16	100	1	100
<i>Mb. gossipicola</i>	Gh-105T	NCBI	GCA_900113485.1	77	4,524,655	4,171	54	68.7	99.37	0	0	1	100
<i>Mb. gregans</i>	NBRC 103626T	This study	SAMD00388917	209	5,744,513	5,353	54	70.27	99.37	0.34	100	1	100
<i>Mb. haplocladii</i>	DSM24195T	This study	SAMD00388918	132	4,615,811	4,355	55	67.48	96.94	0.63	100	1	99.86
<i>Mb. hispanicum</i>	DSM 16372T	This study	SAMD00388919	122	6,668,779	6,311	67	71.04	100	2.93	9.09	1	100
<i>Mb. indicum</i>	SE2.11T	NCBI	GCA_001043895.1	506	6,921,653	6,303	69	70.43	96.26	0.31	0	1	100
<i>Mb. iners</i>	DSM 19015T	This study	SAMD00388921	175	5,495,114	5,277	51	67.79	99.69	0	0	1	100
<i>Mb. isiliense</i>	DSM 17168T	This study	SAMD00388922	167	6,971,527	6,614	59	71.11	100	0.16	0	1	100
<i>Mb. jeotgali</i>	LMG 23639T	This study	SAMD00388923	112	4,459,488	4,271	55	71.04	98.71	0.31	0	1	100
<i>Mb. komagatae</i>	DSM 19563T	JGI	MetkomDSM19563	46	5,502,493	5,284	51	67.47	99.37	1.25	0	1	100
<i>Mb. longum</i>	DSM 23933T	This study	SAMD00388925	213	6,606,771	6,324	57	69.33	100	0	0	1	100
<i>Mb. marchantiae</i>	DSM 21328T	This study	SAMD00388926	82	4,614,373	4,279	52	66.59	100	0	0	1	100
<i>Mb. mesophilicum</i>	NBRC 15688T	This study	SAMD00388927	220	5,902,708	5,583	61	69.12	99.69	0.63	0	1	100
<i>Mb. nodulans</i>	ORS 2060T	NCBI	GCA_000022085.1	8	8,839,022	8,622	81	68.44	99.84	0.42	0	7	100
<i>Mb. nonmethylotrophicum</i>	6HR-1T	NCBI	GCA_004745635.1	102	7,831,372	7,033	82	70.97	100	0.42	66.67	1	99.85
<i>Mb. organophilum</i>	NBRC 15689T	This study	SAMD00388930	114	5,070,401	4,671	55	69.99	100	0.73	0	1	100
<i>Mb. oryzae</i>	CBMB20T	NCBI	GCA_000757795.1	5	6,524,597	6,029	71	69.53	100	0	0	4	100
<i>Mb. oryzihabitans</i>	TER-1T	NCBI	GCA_004004555.2	204	6,328,127	5,609	65	71.86	99.69	0.48	0	2	100
<i>Mb. oxalidis</i>	DSM 24028T	This study	SAMD00388933	297	6,301,862	6,117	55	69.91	98.12	0.31	100	1	100
<i>Mb. persicinum</i>	NBRC 103628T	This study	SAMD00388934	95	5,271,491	4,888	49	68.57	99.06	0.34	50	1	100
<i>Mb. phyllosphaerae</i>	CBMB27T	NCBI	GCA_900113465.1	4	6,316,624	5,774	73	69.57	100	0	0	4	99.04-99.78
<i>Mb. phyllostachyos</i>	BL47T	NCBI	GCA_900103445.1	47	6,015,046	5,765	54	68.71	100	0.63	50	1	100
<i>Mb. planium</i>	YIM 132548T	NCBI	GCA_008806345.1	80	5,875,670	5,537	53	70.19	100	0	0	1	99.75
<i>Mb. platani</i>	PMB02T	NCBI	GCA_001653715.1	150	7,023,890	6,403	89	71.24	99.79	0.31	0	1	100
<i>Mb. pseudosasicola</i>	BL36T	NCBI	GCA_900114535.1	129	6,845,277	6,695	51	68.37	99.69	0.94	60	1	100
<i>Mb. radiodurans</i>	17Sr1-43T	NCBI	GCA_003173735.1	1	5,539,695	5,165	71	70.45	100	0.42	0	6	100
<i>Mb. radiotolerans</i>	JCM 2831T	NCBI	GCA_000019725.1	9	6,899,110	6,495	72	71.04	100	0.16	0	6	99.86-100
<i>Mb. segetis</i>	17J42-1T	NCBI	GCA_004348265.1	256	5,591,199	5,322	58	70.19	100	2.31	100	missing	-
<i>Mb. soli</i>	KCTC 22810T	This study	SAMD00388943	827	6,167,034	6,138	49	68.38	93.32	0.65	50	1	99.72
<i>Mb. symbioticum</i>	SB0023/3T	NCBI	GCA_902141845.1	247	6,086,597	5,665	57	70.08	99	0.16	0	1	100
<i>Mb. tardum</i>	NBRC 103632T	This study	SAMD00388945	147	6,890,653	6,555	60	69.17	100	0.31	0	1	100
<i>Mb. tarhaniae</i>	DSM 25844T	NCBI	GCA_001043955.1	483	6,743,422	6,060	75	70.43	95.19	0.08	100	1	100
<i>Mb. terrae</i>	17Sr1-28T	NCBI	GCA_003173755.1	1	6,162,702	5,612	100	71.56	98.8	0.16	0	9	100-98.36
<i>Mb. terricola</i>	17Sr1-39T	NCBI	GCA_006151805.1	103	7,873,283	7,190	86	70.42	100	0	0	1	99.5
<i>Mb. thuringiense</i>	DSM 23674	This study	SAMD00388949	47	4,707,776	4,388	60	67.51	99.37	0	0	1	99.93
<i>Mb. trifolii</i>	DSM 23632	This study	SAMD00388950	424	5,163,180	4,887	56	69.74	96.27	0.63	50	1	99.85
<i>Mb. variabile</i>	DSM 16961T	NCBI	GCA_001043975.1	562	7,433,592	6,858	70	70.75	94.69	0.55	66.67	1	100
<i>Mr. aminovorans</i>	NBRC 15686T	This study	SAMD00388952	120	6,087,464	5,673	52	68.12	99.69	0	0	1	100
<i>Mr. extorquens</i>	TK 0001T	NCBI	GCA_900234795.1	1	5,715,512	5,422	67	68.29	98.4	0	0	5	100
<i>Mr. podarium</i>	DSM 15083T	This study	SAMD00388954	732	5,300,219	4,952	47	69.98	91.87	0.83	20	1	99.63
<i>Mr. populi</i>	BJ001T	NCBI	GCA_000019945.1	3	5,848,997	5,436	64	69.36	99.69	0.47	0	5	99.86-100
<i>Mr. rhodesianum</i>	DSM 5687T	NCBI	GCA_014199985.1	52	5,873,985	5,446	51	69.13	99.58	0.47	0	1	100
<i>Mr. rhodinum</i>	DSM 2163T	NCBI	GCA_014199935.1	58	5,406,054	4,978	53	69.93	99.69	0	0	1	100
<i>Mr. salsuginis</i>	CGMCC 1.6474T	NCBI	GCA_900114375.1	64	5,323,773	4,986	57	69.63	99.37	0	0	missing	-
<i>Mr. suomiense</i>	DSM 14458T	This study	SAMD00388959	57	5,547,687	5,241	55	68.07	99.37	0.94	0	1	100
<i>Mr. thiocyanatum</i>	JCM 10893T	This study	SAMD00388960	61	5,637,136	5,250	54	69.26	99.69	0	0	1	100
<i>Mr. zatmanii</i>	LMG 6087T	NCBI	GCA_014845115.1	392	5,229,025	4835	58	68.95	96.34	0.31	0	1	100
<i>Microvirga aerophila</i>	NBRC 106136T	NCBI	GCA_007991675.1	629	7,363,484	7514	61	61.64	98.9	3.21	12.5	1	100
<i>Microvirga brassicacearum</i>	CDVBN77T	NCBI	GCA_008757455.1	82	5,220,401	4969	48	62.27	99.22	0.63	0	1	100
<i>Microvirga calopogonii</i>	CCBAU 65841T	NCBI	GCA_003347665.1	249	7,413,381	7295	73	62.48	99.69	2.62	42.86	5	100-78
<i>Microvirga flocculans</i>	ATCC BAA-817T	NCBI	GCA_000518665.1	30	4,029,781	3821	55	64.27	98.9	0.63	0	1	100
<i>Microvirga guangxiensis</i>	CGMCC 1.7666T	NCBI	GCA_900102135.1	34	4,721,732	4516	58	61.42	98.12	1	0	1	100
<i>Microvirga lotononidis</i>	WSM3557T	NCBI	GCA_000262405.1	18	7,082,538	6981	62	63.08	99.69	0.85	0	2	100
<i>Microvirga lupini</i>	AT3.9	NCBI	GCA_014191055.1	17	4,945,970	4733	52	62.93	99.06	0	0	1	98.3
<i>Microvirga makkahensis</i>	KCTC 23863T	NCBI	GCA_009830105.1	94	5,580,646	5361	51	63.1	99.37	1.57	0	1	100
<i>Microvirga massiliensis</i>	JC119T	NCBI	GCA_001006805.1	50	9,172,134	9054	55	63.28	99.22	2.27	9.09	1	100
<i>Microvirga ossetica</i>	V5/3MT	NCBI	GCA_002741015.1	6	9,631,051	9583	70	62.67	99.37	2.36	46.15	4	100
<i>Microvirga pakistanensis</i>	NCCP-1258T	NCBI	GCA_004458765.1	723	5,764,046	5302	67	63.32	98.75	3.47	78.26	1	99.76
<i>Microvirga subterranea</i>	DSM 14364T	NCBI	GCA_003350535.1	36	5,147,802	4874	70	65.14	99.06	0.38	0	1	100
<i>Microvirga thermotolerans</i>	HR1T	NCBI	GCA_009363855.1	1	3,823,049	3556	57	67.71	99.06	0.63	0	2	100
<i>Microvirga tunisiensis</i>	Lmie10	NCBI	GCA_009296175.1	1022	8,834,183	9215	63	61.76	99.53	2.36	35.71	1	nd
<i>Microvirga vignae</i>	BR3299T	NCBI	GCA_001011715.1	165	6,472,445	6526	57	61.07	99.22	2.12	11.11	1	100

Table 2. Reference sequences used in the analysis				
Pathway name	Strain name	Gene name	Query ID	Function
Methylamine, NMG pathway	PA1	<i>mgdD</i>	Mext_1651	Methylglutamate dehydrogenase subunit D
Methylamine, NMG pathway	PA1	<i>mgdC</i>	Mext_1652	Methylglutamate dehydrogenase subunit C
Methylamine, NMG pathway	PA1	<i>mgdB</i>	Mext_1653	Methylglutamate dehydrogenase subunit B
Methylamine, NMG pathway	PA1	<i>mgdA</i>	Mext_1654	Methylglutamate dehydrogenase subunit A
Methylamine, NMG pathway	PA1	<i>mgsA</i>	Mext_1656	N-methyltransferase subunit A
Methylamine, NMG pathway	PA1	<i>mgsB</i>	Mext_1657	N-methyltransferase subunit B
Methylamine, NMG pathway	PA1	<i>mgsC</i>	Mext_1658	N-methyltransferase subunit C
Methylamine, NMG pathway	PA1	<i>gmaS</i>	Mext_1659	Glutamine synthetase
Methylamine dehydrogenase	AM1	<i>mauF</i>	MexAM1_META1p2769	Unknown
Methylamine dehydrogenase	AM1	<i>mauB</i>	MexAM1_META1p2770	Methylamine dehydrogenase large subunit
Methylamine dehydrogenase	AM1	<i>mauE</i>	MexAM1_META1p2771	Essential for small subunit maturation
Methylamine dehydrogenase	AM1	<i>mauD</i>	MexAM1_META1p2772	Essential for small subunit maturation
Methylamine dehydrogenase	AM1	<i>mauA</i>	MexAM1_META1p2773	Methylamine dehydrogenase small subunit
Methylamine dehydrogenase	AM1	<i>mauC</i>	MexAM1_META1p2774	Amicyanin
Methylamine dehydrogenase	AM1	<i>mauJ</i>	MexAM1_META1p2775	Unknown
Methylamine dehydrogenase	AM1	<i>mauG</i>	MexAM1_META1p2776	Unknown
Methylamine dehydrogenase	AM1	<i>mauI</i>	MexAM1_META1p2777	Unknown
Methylamine dehydrogenase	AM1	<i>mauM</i>	MexAM1_META1p2778	Ferredoxin
Methylamine dehydrogenase	AM1	<i>mauN</i>	MexAM1_META1p2779	Ferredoxin
Ca-dependent MDH, Mxa cluster	AM1	<i>mxkB</i>	MexAM1_META1p4525	Transcriptional regulator
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhH</i>	MexAM1_META1p4526	Unknown
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhE</i>	MexAM1_META1p4527	Unknown
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhD</i>	MexAM1_META1p4528	Unknown
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhL</i>	MexAM1_META1p4529	Essential for Ca ²⁺ insertion into MDH
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhK</i>	MexAM1_META1p4530	Essential for Ca ²⁺ insertion into MDH
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhC</i>	MexAM1_META1p4531	Essential for Ca ²⁺ insertion into MDH
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhA</i>	MexAM1_META1p4532	Essential for Ca ²⁺ insertion into MDH
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhS</i>	MexAM1_META1p4533	Unknown
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhR</i>	MexAM1_META1p4534	Unknown
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhI</i>	MexAM1_META1p4535	Methanol dehydrogenase, small subunit
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhG</i>	MexAM1_META1p4536	Cytochrome c550
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhJ</i>	MexAM1_META1p4537	MxaJ, possible chaperone
Ca-dependent MDH, Mxa cluster	AM1	<i>mxhF</i>	MexAM1_META1p4538	Methanol dehydrogenase, large subunit
MDH regulation	AM1	<i>mxhM</i>	MexAM1_META1p1752	Transcriptional regulator
MDH regulation	AM1	<i>mxhD</i>	MexAM1_META1p1753	Sensor kinase
MDH regulation	AM1	<i>mxhQ</i>	MexAM1_META1p4896	Sensor kinase
MDH regulation	AM1	<i>mxhE</i>	MexAM1_META1p4897	Transcriptional regulator
PQQ synthesis	AM1	<i>pqqE</i>	MexAM1_META1p1748	PQQ synthesis protein, PqqE
PQQ synthesis	AM1	<i>pqqC/D</i>	MexAM1_META1p1749	PQQ synthesis protein, PqqC/D
PQQ synthesis	AM1	<i>pqqB</i>	MexAM1_META1p1750	PQQ synthesis protein, PqqB
PQQ synthesis	AM1	<i>pqqA</i>	MexAM1_META1p1751	PQQ synthesis protein, PqqA
PQQ synthesis	AM1	<i>pqqF</i>	MexAM1_META1p2330	PQQ synthesis protein, PqqF
PQQ synthesis	AM1	<i>pqqG</i>	MexAM1_META1p2331	PQQ synthesis protein, PqqG
Ln-dependent MDH, XoxF1 cluster	AM1	<i>xoxF1</i>	MexAM1_META1p1740	Ln-dependent MDH, XoxF1
Ln-dependent MDH, XoxF1 cluster	AM1	<i>xoxG</i>	MexAM1_META1p1741	Cytochrome c
Ln-dependent MDH, XoxF1 cluster	AM1	<i>xoxJ</i>	MexAM1_META1p1742	Periplasmic binding proteins
Ln-dependent MDH, XoxF2	AM1	<i>xoxF2</i>	MexAM1_META1p2757	Ln-dependent MDH, XoxF2
Dichloromethane oxidation	DM4	<i>dcmR</i>	METDI2655	Transcriptional regulator of DCM dehalogenase
Dichloromethane oxidation	DM4	<i>dcmA</i>	METDI2656	Dichloromethane dehalogenase (DCM dehalogenase)
Ln-dependent EDH, ExaF	AM1	<i>exaF</i>	MexAM1_META1p1139	Ln-dependent ethanol dehydrogenase
Ln transport and utilization (lut) genes	AM1	<i>lutA</i>	MexAM1_META1p1778	ABC transporter ATPase
Ln transport and utilization (lut) genes	AM1	<i>lutB</i>	MexAM1_META1p1779	Putative periplasmic protein
Ln transport and utilization (lut) genes	AM1	<i>lutC</i>	MexAM1_META1p1780	Hypothetical protein
Ln transport and utilization (lut) genes	AM1	<i>lutD</i>	MexAM1_META1p1781	Hypothetical protein
Ln transport and utilization (lut) genes	AM1	<i>lutE</i>	MexAM1_META1p1782	ABC transporter ATPase
Ln transport and utilization (lut) genes	AM1	<i>lutF</i>	MexAM1_META1p1783	multidrug ABC transporter permease
Ln transport and utilization (lut) genes	AM1	<i>lutG</i>	MexAM1_META1p1784	Putative periplasmic protein
Ln transport and utilization (lut) genes	AM1	<i>lutH</i>	MexAM1_META1p1785	TonB-dependent receptor
Ln transport and utilization (lut) genes	AM1	<i>lanM</i>	MexAM1_META1p1786	Lanmodulin
Ln transport and utilization (lut) genes	AM1	<i>lutI</i>	MexAM1_META1p1787	Hypothetical protein

Table 2. Reference sequences used in the analysis				
Pathway name	Strain name	Gene name	Query ID	Function
H4MPT pathway	AM1	<i>fhcC</i>	MexAM1_META1p1755	Formyltransferase/hydrolase complex,gamma subunit
H4MPT pathway	AM1	<i>fhcD</i>	MexAM1_META1p1756	Formyltransferase/hydrolase complex, delta subunit
H4MPT pathway	AM1	<i>fhcA</i>	MexAM1_META1p1757	Formyltransferase/hydrolase complex, alpha subunit
H4MPT pathway	AM1	<i>fhcB</i>	MexAM1_META1p1758	Formyltransferase/hydrolase complex,beta subunit
H4MPT pathway	AM1	<i>mptG</i>	MexAM1_META1p1760	Ribofuranosylaminobenzene, 5'-phosphate synthase
H4MPT pathway	AM1	<i>mtdB</i>	MexAM1_META1p1761	Methylene-H4MPT dehydrogenase
H4MPT pathway	AM1	<i>orfY</i>	MexAM1_META1p1762	Unknown
H4MPT pathway	AM1	<i>mch</i>	MexAM1_META1p1763	Methenyl-H4MPT cyclohydrolase
H4MPT pathway	AM1	<i>orf5</i>	MexAM1_META1p1764	H4MPT biosynthesis
H4MPT pathway	AM1	<i>orf7</i>	MexAM1_META1p1765	Unknown
H4MPT pathway	AM1	<i>fae</i>	MexAM1_META1p1766	Formaldehyde activating enzyme
H4MPT pathway	AM1	<i>orf17</i>	MexAM1_META1p1767	Unknown
H4MPT pathway	AM1	<i>orf9</i>	MexAM1_META1p1768	H4MPT biosynthesis
H4MPT pathway	AM1		MexAM1_META1p1770	Homolog of mxaE
H4MPT pathway	AM1		MexAM1_META1p1771	Homolog of mxaD
H4MPT pathway	AM1		MexAM1_META1p1772	Homolog of mxaD
H4MPT pathway	AM1	<i>orf19</i>	MexAM1_META1p1773	H4MPT biosynthesis
H4MPT pathway	AM1	<i>orf20</i>	MexAM1_META1p1774	H4MPT biosynthesis
H4MPT pathway	AM1	<i>orf21</i>	MexAM1_META1p1775	H4MPT biosynthesis
H4MPT pathway	AM1	<i>orf22</i>	MexAM1_META1p1776	H4MPT biosynthesis
GSH-dependent formaldehyde dehydrogenase	22A	<i>gfa</i>	Maq22A_1p37180	Glutathione-dependent formaldehyde-activating enzyme
GSH-dependent formaldehyde dehydrogenase	22A	<i>hgd</i>	Maq22A_c21490	S-hydroxymethyl glutathione dehydrogenase
GSH-dependent formaldehyde dehydrogenase	22A	<i>fgh</i>	Maq22A_c21495	S-formylglutathione hydrolase
Formate dehydrogenase 1	AM1	<i>fdh1B</i>	MexAM1_META1p5031	Tungsten-dependent formate dehydrogenase,beta subunit
Formate dehydrogenase 1	AM1	<i>fdh1A</i>	MexAM1_META1p5032	Tungsten-dependent formate dehydrogenase, alpha subunit
Formate dehydrogenase 2	AM1	<i>fdh2C</i>	MexAM1_META1p4846	Molybdenum-dependent formate dehydrogenase, gamma subunit
Formate dehydrogenase 2	AM1	<i>fdh2B</i>	MexAM1_META1p4847	Molybdenum-dependent formate dehydrogenase, beta subunit
Formate dehydrogenase 2	AM1	<i>fdh2A</i>	MexAM1_META1p4848	Molybdenum-dependent formate dehydrogenase, alpha subunit
Formate dehydrogenase 2	AM1	<i>fdh2D</i>	MexAM1_META1p4849	Molybdenum-dependent formate dehydrogenase, delta subunit
Formate dehydrogenase 3	AM1	<i>fdh3A</i>	MexAM1_META1p0303	Cytochrome-linked formate dehydrogenase alpha subunit
Formate dehydrogenase 3	AM1	<i>fdh3B</i>	MexAM1_META1p0304	Cytochrome-linked formate dehydrogenase beta subunit
Formate dehydrogenase 3	AM1	<i>fdh3C</i>	MexAM1_META1p0305	Cytochrome-linked formate dehydrogenase gamma subunit
Formate dehydrogenase 4	AM1	<i>fdh4B</i>	MexAM1_META1p2093	Protein associated with expression of formate dehydrogenase 4
Formate dehydrogenase 4	AM1	<i>fdh4A</i>	MexAM1_META1p2094	Formate dehydrogenase 4
Folate synthesis	AM1	<i>folK</i>	MexAM1_META1p1743	2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase
Folate synthesis	AM1	<i>folB</i>	MexAM1_META1p1744	Dihydroneopterin aldolase
Folate synthesis	AM1	<i>folP</i>	MexAM1_META1p1745	Dihydropteroate synthase
Folate synthesis	AM1	<i>pabB</i>	MexAM1_META1p4284	para-Aminobenzoic acid synthase component
Folate synthesis	AM1	<i>folC</i>	MexAM1_META1p4888	Dihydrofolate synthase
H4F pathway	AM1	<i>ftfL</i>	MexAM1_META1p0329	Formyl-tetrahydrofolate ligase
H4F pathway	AM1	<i>mtdA</i>	MexAM1_META1p1728	Bifunctional methylene-H4MPT/ methylene-H4F dehydrogenase
H4F pathway	AM1	<i>fch</i>	MexAM1_META1p1729	Methenyl-H4F cyclohydrolase
H4F pathway	AM1	<i>folA</i>	MexAM1_META1p2852	Dihydrofolate reductase
Methylotrophy regulator	PA1	<i>ttmR</i>	Mext_0925	formaldehyde-responsive transcription factor
Formaldehyde-binding protein	PA1	<i>efgA</i>	Mext_4158	formaldehyde-sensing protein
Serine cycle	AM1	<i>qscR</i>	MexAM1_META1p0756	Serine cycle transcriptional regulator
Serine cycle	AM1	<i>mdh</i>	MexAM1_META1p1537	Malate dehydrogenase
Serine cycle	AM1	<i>sga</i>	MexAM1_META1p1726	Serine glyoxylate aminotransferase
Serine cycle	AM1	<i>hpr</i>	MexAM1_META1p1727	Hydroxypyruvate reductase
Serine cycle	AM1	<i>mtkA</i>	MexAM1_META1p1730	Malate thiokinase, alpha subunit
Serine cycle	AM1	<i>mtkB</i>	MexAM1_META1p1731	Malate thiokinase, beta subunit
Serine cycle	AM1	<i>ppc</i>	MexAM1_META1p1732	PEP carboxylase
Serine cycle	AM1	<i>mcl</i>	MexAM1_META1p1733	Malyl-CoA lyase/beta-malylyl-CoA lyase
Serine cycle	AM1	<i>gck</i>	MexAM1_META1p2944	Glycerate kinase
Serine cycle	AM1	<i>eno</i>	MexAM1_META1p2984	Enolase
Serine cycle	AM1	<i>glyA</i>	MexAM1_META1p3384	Serine hydroxymethyltransferase
Serine / EMC pathway	AM1	<i>phaA</i>	MexAM1_META1p3700	Beta-ketothiolase
TCA cycle	AM1	<i>sdhC</i>	MexAM1_META1p3859	Succinate dehydrogenase, cytochrome b556 subunit
TCA cycle	AM1	<i>sdhD</i>	MexAM1_META1p3860	Succinate dehydrogenase, membrane anchor subunit
TCA cycle	AM1	<i>sdhA</i>	MexAM1_META1p3861	Succinate dehydrogenase, flavoprotein subunit
TCA cycle	AM1	<i>sdhB</i>	MexAM1_META1p3863	Succinate dehydrogenase,iron-sulfur subunit
EMC pathway	AM1	<i>pccB</i>	MexAM1_META1p0172	Propionyl-CoA carboxylase beta subunit
EMC pathway	AM1	<i>ccr</i>	MexAM1_META1p0178	Crotonyl-CoA carboxylase/reductase
EMC pathway	AM1	<i>ecm</i>	MexAM1_META1p0180	Ethylmalonyl-CoA mutase
EMC pathway	AM1	<i>meaB</i>	MexAM1_META1p0188	Methylmalonyl-CoA mutase accessory protein
EMC pathway	AM1	<i>epi</i>	MexAM1_META1p0839	Ethylmalonyl-CoA/methylmalonyl-CoA epimerase
EMC pathway	AM1	<i>mcmB</i>	MexAM1_META1p2390	Methylmalonyl-CoA mutase, beta subunit
EMC pathway	AM1	<i>pccA</i>	MexAM1_META1p3203	Propionyl-CoA carboxylase, alpha subunit
EMC pathway	AM1	<i>croR</i>	MexAM1_META1p3675	Crotonase
EMC pathway	AM1	<i>phaB</i>	MexAM1_META1p3701	Acetoacetyl-CoA reductase
EMC pathway	AM1	<i>mcd</i>	MexAM1_META1p4153	Mesaconyl-CoA hydratase
EMC pathway	AM1	<i>mcmA</i>	MexAM1_META1p5251	Methylmalonyl-CoA mutase, alpha subunit
EMC pathway	AM1	<i>msd</i>	MexAM1_META1p2223	Methylsuccinyl-CoA dehydrogenase

Table 2. Reference sequences used in the analysis				
Pathway name	Strain name	Gene name	Query ID	Function
ACC deaminase	CBMB20T	<i>acdS</i>	MOC_1898	ACC deaminase
D-cysteine desulphydrase	AM1	<i>dcyD</i>	MexAM1_META1p2360	D-cysteine desulphydrase
Cellulase	AM1	<i>celC</i>	MexAM1_META1p1169	Cellulase
Cellulase	22A	<i>bglC</i>	Maq22A_1p38150	Endoglucanase, cellulase
Cellulase	22A		Maq22A_2p41770	Cellulase
Cellulase	22A	<i>celA</i>	Maq22A_2p42730	EndoglucanaseY
Antifungal	CBMB20T		MOC_0504	Chitinase
Cytokinin synthesis	CBMB20T	<i>miaA</i>	MOC_5190	Isopentenyl tRNA transferase
IAA synthesis	ORS 2060		Mnod_1238	L-tryptophan decarboxylase
IAA synthesis	ORS 2060		Mnod_0239	Aldehyde dehydrogenase
IAA synthesis	ORS 2060		Mnod_4579	Monoamine oxidase
IAA synthesis	AM1		MexAM1_META1p4817	Monoamine oxidase
IAA synthesis	CBMB20T		MOC_0996	Monoamine oxidase
IAA synthesis	CBMB20T		MOC_5534	Aldehyde dehydrogenase
Pectin metabolism	AM1		MexAM1_META1p4514	Pectate lyase
ROS protection	CBMB20T		MOC_4484	Glutathione peroxidase
ROS protection	CBMB20T		MOC_5157	Glutathione disulfide reductase
ROS protection	CBMB20T		MOC_2537	UV damage repair
ROS protection	CBMB20T		MOC_1064	Glutathione peroxidase
ROS protection	CBMB20T		MOC_3063	Glutathione disulfide reductase
ROS protection	22A	<i>egtB</i>	Maq22A_1p31895	Hercynine oxygenase, ergothioneine synthesis
ROS protection	22A	<i>egtD</i>	Maq22A_1p31890	Histidine N-alpha-methyltransferase, ergothioneine synthesis

Table 3. ADH and MDH reference sequences used as query.		
Reference sequence source	Accession No.	MDH and ADH family type
<i>Methylophaga lonarensis</i> MPL	EMR11880	MxaF
<i>Methylobacterium extorquens</i> AM1	META1p4538	MxaF
<i>Paracoccus</i>	WP 011749265.1	MxaF
<i>Methylobacterium methanica</i>	WP 013820880.1	MxaF
<i>Bradyrhizobium oligotrophicum</i>	WP 015666256.1	PQQ ADH type 1
<i>Pseudomonas aeruginosa</i>	CAA08896	PQQ ADH type 2a
<i>Pseudomonas butanovora</i>	AF326086	PQQ ADH type 2a Butanol dehydrogenase
<i>Methylobacterium petroleiphilum</i> PM1	YP 001019670	PQQ ADH type 2b
<i>Methylobacterium extorquens</i> AM1	WP 003602767.1	PQQ ADH type 2b ExaF
<i>Pseudomonas putida</i> NBRC 14164	BAN55012.1	PQQ ADH type 2b PedH
<i>Pseudomonas putida</i> KT2440	NP 744823.1	PQQ ADH type 2b PedH
<i>Bradyrhizobium diazoefficiens</i> USDA 11	NP 772847	PQQ ADH type 3
<i>Thauera butanivorans</i>	AAK27220	PQQ ADH type 4 quinohemoproteins
<i>Pseudomonas putida</i>	BAC15559	PQQ ADH type 4 quinohemoproteins
<i>Methylobacterium petroleiphilum</i>	WP 011829190	PQQ ADH type 5
<i>Bradyrhizobium diazoefficiens</i> USDA 11	NP 772860	PQQ ADH type 6a
<i>Bradyrhizobium</i> sp. BTAi1	WP 012042238.1	PQQ ADH type 6b
<i>Paraburkholderia graminis</i> C4D1M	EDT10345	PQQ ADH type 7
<i>Candidatus Solibacter usitatus</i>	WP 011684813.1	PQQ ADH type 8
<i>Methylosinus trichosporium</i> OB3b	EFH04530.1	PQQ ADH type 9
<i>Bradyrhizobium diazoefficiens</i> USDA 11	NP 774279	PQQ ADH type 9
<i>Paraburkholderia xenovorans</i>	WP 011490939.1	PQQ ADH type 9
<i>Methylophaga thiooxidans</i> DMS010	ZP 05104584.1	PQQ ADH type 9
<i>Methylocella silvestris</i>	WP 012592127.1	XoxF1
<i>Candidatus Methylobacterium oxyfera</i>	WP 012814374.1	XoxF1
<i>Verrucomicrobium</i> sp. 3C	WP 020494391.1	XoxF1
uncultured bacterium BAC10-4	AAAY96669.1	XoxF2
<i>Methylobacterium fumariolicum</i> SolV	ACM16493	XoxF2
<i>Candidatus Methylobacterium oxyfera</i>	WP 012814383.1	XoxF2
<i>Methylobacteriaceae</i>	WP 012251923.1	XoxF3
<i>Methylocella silvestris</i>	WP 012592359.1	XoxF3
<i>Methylobacillus flagellatus</i>	WP 011480535.1	XoxF4
<i>Methylobacterium versatilis</i>	WP 013147907.1	XoxF4
<i>Methylobacterium mobilis</i>	WP 015832985.1	XoxF4
<i>Methylobacterium extorquens</i> AM1	META1p2757	XoxF5
<i>Methylobacterium extorquens</i> AM1	META1p1740	XoxF5
<i>Paraburkholderia graminis</i> C4D1M	EDT12687	XoxF5
<i>Methylosinus trichosporium</i> OB3b	EFH02991.1	XoxF5
<i>Bradyrhizobium diazoefficiens</i> USDA 11	NP 772853	XoxF5
<i>Methylobacterium methanica</i>	WP 013818199.1	XoxF5
<i>Methylobacterium nodulans</i>	WP 015927174.1	XoxF5
<i>Hyphomicrobium denitrificans</i>	YP 007913380	XoxF5
<i>Methylophaga thiooxidans</i> DMS010	ZP 05104902	XoxF5
<i>Methylobacterium aquaticum</i> 22A	Maq22A_1p33165	MxaF (gene name)
<i>Methylobacterium aquaticum</i> 22A	Maq22A_c05215	XoxF1 (gene name)
<i>Methylobacterium aquaticum</i> 22A	Maq22A_c27990	XoxF2 (gene name)
<i>Methylobacterium aquaticum</i> 22A	Maq22A_c07235	ExaF (gene name)
<i>Methylobacterium aquaticum</i> 22A	Maq22A_1p32165	ADH4 (gene name)
<i>Methylobacterium aquaticum</i> 22A	Maq22A_1p30675	ADH6 (gene name)

Table 4. Minimum medium components		
Components	Chemicals	in L
Mineral salts	NH ₄ Cl	8.1 g
(Autoclave)	MgSO ₄ ·7H ₂ O	1.0 g
	Water	1000 ml
Buffer pH 7.1	K ₂ HPO ₄	8.0 g
(Autoclave)	NaH ₂ PO ₄ ·2H ₂ O	3.6 g
	Water	
Iron solution	FeSO ₄ ·7H ₂ O	13.0 g
(Filter sterilization)	1N HCl	
	Water	1000 ml
Trace elements	ZnSO ₄ ·7H ₂ O	4.5 g
(Filter sterilization)	CoCl ₂ ·6H ₂ O	3 g
	MnCl ₂ ·4H ₂ O	0.64 g
	H ₃ BO ₃	1 g
	Na ₂ MoO ₄ ·2H ₂ O	0.4 g
	CuSO ₄ ·5H ₂ O	0.3 g
	CaCl ₂ ·2H ₂ O	3 g
	Water	1000 ml
Vitamin solution	Calcium pantothenate	0.4 g
(Filter sterilization)	Inositol	0.2 g
	Niacin	0.4 g
	p-Aminobenzoic acid	0.2 g
	Pyridoxine hydrochloride	0.4 g
	Thiamin hydrochloride	0.4 g
	Biotin	0.2 g
	Vitamin B12	0.2 g
	Water	1000 ml
Medium preparation	Mineral salts	200 ml
	Buffer	300 ml
	Iron solution	1 ml
	Trace elements	1 ml
	Vitamin solution	10 ml
	Carbon source	X ml or g
	Water	fill up to 1 L

Table 5. Details of the newly sequenced assemblies.													
Source species	Strain name	Sequencer type	Assembler *1	Coverage	Total sequences	# contigs (>= 0 bp)	# contigs (>= 1000 bp)	Largest contig	N50	N75	L50	L75	# N's per 100 kbp
<i>Methylobacterium adhaesivum</i>	DSM 17169T	MiSeq	CLC-GW 20.0.3	251.32	1,175,569,639	61	61	554,752	176,763	75,734	7	17	0
<i>Methylobacterium aerolatum</i>	DSM 19013T	MiSeq	CLC-GW 20.0.4	217.62	1,061,494,012	67	67	488,770	154,739	98,285	10	20	0
<i>Methylobacterium bullatum</i>	DSM 21893T	MiSeq	CLC-GW 20.0.4	204.23	1,012,193,062	54	54	647,375	206,168	129,631	7	14	0
<i>Methylobacterium cerastii</i>	DSM 23679T	MiSeq	CLC-GW 20.0.4	178.22	960,363,617	150	150	269,374	98,709	44,125	19	42	0
<i>Methylobacterium crusticola</i>	KCTC 52305T	DNBSeq	SPAdes (v 3.12.0)	108.64	834,414,600	790	67	571,168	270,721	146,929	10	19	7.7
<i>Methylobacterium dankookense</i>	DSM 22415T	MiSeq	CLC-GW 11.0.1	36.01	217,703,948	273	273	159,818	36,678	21,302	53	108	0
<i>Methylobacterium frigidaeris</i>	JCM 32048T	DNBSeq	SPAdes (v 3.12.0)	157.67	1247051400	513	123	374,537	157,499	62,096	16	37	11.22
<i>Methylobacterium gnaphalii</i>	DSM 24027T	MiSeq	CLC-GW 11.0.1	23.85	114,064,900	195	195	189,613	50,299	25,169	27	60	0
<i>Methylobacterium goesingense</i>	DSM 21331T	MiSeq	CLC-GW 11.0.1	19.55	104,265,987	263	263	108,376	38,529	20,181	43	89	0
<i>Methylobacterium gregans</i>	NBRC 103626T	MiSeq	CLC-GW 11.0.1	27.13	155,857,359	209	209	175,977	41,432	24,374	41	86	0
<i>Methylobacterium haplocladii</i>	DSM 24195T	MiSeq	CLC-GW 11.0.1	32.08	148,068,690	132	132	189,906	69,314	40,487	21	43	0
<i>Methylobacterium hispanicum</i>	DSM 16372T	DNBSeq	Shovill (v 1.1.0)	257.40	1,716,528,300	122	82	455,122	163,091	100,045	13	26	0
<i>Methylobacterium iners</i>	DSM 19015T	MiSeq	CLC-GW 20.0.3	202.47	1,112,605,802	175	175	173,123	71,746	38,147	24	49	0
<i>Methylobacterium isbiliense</i>	DSM 17168T	MiSeq	CLC-GW 20.0.3	114.44	797,815,728	167	167	370,219	85,536	42,906	22	51	0
<i>Methylobacterium jeotgali</i>	LMG 23639T	MiSeq	CLC-GW 11.0.1	35.92	160,194,073	112	112	171,843	64,725	41,150	23	44	0
<i>Methylobacterium longum</i>	DSM 23933T	DNBSeq	Shovill (v 1.1.0)	196.11	1,295,742,300	213	144	398,795	124,885	52,545	15	36	0
<i>Methylobacterium marchantiae</i>	DSM 21328T	MiSeq	CLC-GW 20.0.3	165.08	761,744,906	82	82	247,709	112,542	62,699	16	30	0
<i>Methylobacterium mesophilicum</i>	NBRC 15688T	DNBSeq	SPAdes (v 3.12.0)	42.39	250,239,000	220	84	512,232	157,383	82,260	12	25	6.88
<i>Methylobacterium organophilum</i>	NBRC 15689T	DNBSeq	SPAdes (v 3.12.0)	88.69	449,685,900	114	24	581,052	350,043	199,175	6	11	0
<i>Methylobacterium oxalidis</i>	DSM 24028T	MiSeq	CLC-GW 20.0.4	45.38	285,978,100	297	297	153,404	40,713	23,402	48	100	0
<i>Methylobacterium persicinum</i>	NBRC 103628T	MiSeq	CLC-GW 20.0.5	164.01	864,581,162	95	95	336,710	105,200	73,201	14	29	0
<i>Methylobacterium soli</i>	KCTC 22810T	MiSeq	CLC-GW 20.0.6	12.30	75,824,364	827	827	62,114	11,257	6,346	175	353	0
<i>Methylobacterium tardum</i>	NBRC 103632T	DNBSeq	Shovill (v 1.1.0)	521.91	3,596,441,400	147	104	591,219	165,545	100,683	12	24	0
<i>Methylobacterium thuringiense</i>	DSM 23674	MiSeq	CLC-GW 20.0.3	221.72	1,043,797,930	47	47	457,393	262,820	140,824	8	13	0
<i>Methylobacterium trifolii</i>	DSM 23632	MiSeq	CLC-GW 11.0.1	16.31	84,231,470	424	424	80,601	20,059	10,243	78	164	0
<i>Methylorubrum aminovorans</i>	NBRC 15686T	DNBSeq	Shovill (v 1.1.0)	230.29	1,402,044,000	120	90	1,039,568	224,166	75,056	9	20	0
<i>Methylorubrum podarium</i>	DSM 15083T	MiSeq	CLC-GW 11.0.1	12.27	65,014,733	732	732	43,487	10,390	6,249	158	325	0
<i>Methylorubrum suomiense</i>	DSM 14458T	DNBSeq	Shovill (v 1.1.0)	400.86	2,223,976,800	57	46	689,385	333,922	166,553	7	14	0
<i>Methylorubrum thiocyanatum</i>	JCM 10893T	MiSeq	CLC-GW 20.0.3	210.37	1,185,864,130	61	61	475,724	176,685	115,295	10	19	0
*1, CLC-GW, CLC Genomics Workbench													

Table 6. Pairwise dDHH values.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23	A24	A25	A26	A27	A28	A29	A30	A31	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21	B22	B23	B24	B25	B26	B27	B28	B29	B30	B31	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30	C31	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D23	D24	D25	D26	D27	D28	D29	D30	D31	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	E21	E22	E23	E24	E25	E26	E27	E28	E29	E30	E31	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30	F31	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	G25	G26	G27	G28	G29	G30	G31	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24	H25	H26	H27	H28	H29	H30	H31	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	I21	I22	I23	I24	I25	I26	I27	I28	I29	I30	I31	J1	J2	J3	J4	J5	J6	J7	J8	J9	J10	J11	J12	J13	J14	J15	J16	J17	J18	J19	J20	J21	J22	J23	J24	J25	J26	J27	J28	J29	J30	J31	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20	K21	K22	K23	K24	K25	K26	K27	K28	K29	K30	K31	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	L16	L17	L18	L19	L20	L21	L22	L23	L24	L25	L26	L27	L28	L29	L30	L31	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	M25	M26	M27	M28	M29	M30	M31	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24	N25	N26	N27	N28	N29	N30	N31	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11	O12	O13	O14	O15	O16	O17	O18	O19	O20	O21	O22	O23	O24	O25	O26	O27	O28	O29	O30	O31	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28	P29	P30	P31	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27	Q28	Q29	Q30	Q31	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24	R25	R26	R27	R28	R29	R30	R31	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30	S31	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20	T21	T22	T23	T24	T25	T26	T27	T28	T29	T30	T31
Mb mves 190101	A1	21.5	21.5	21.5	22.1	A1	A4	22.3	23.7	A24	A3	21.3	21.1	A29	22.7	A24	A27	22.8	22.5	21.6	A4	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5	A5																																																																																																																																																																									

[illegible][illegible]

Table 8. Core, accessory, unique, and exclusively absent gene counts in each type strain.					
Organism name	Sub-clade	No. of core genes	No. of accessory genes	No. of unique genes	No. of exclusively absent genes
Mb. cerastii DSM 23679T	A	711	3566	535	2
Mb. jeotgali LMG 23639T	A	711	2886	567	10
Mb. trifolii DSM 23632	A	711	3412	553	15
Mb. adhaesivum DSM 17169T	A1	711	3110	357	1
Mb. bullatum DSM 21893T	A1	711	3306	419	1
Mb. goesingense DSM 21331T	A1	711	3465	604	30
Mb. gossypicola Gh-105T	A1	711	3013	277	2
Mb. iners DSM 19015T	A1	711	3231	988	5
Mb. marchantiae DSM 21328T	A1	711	3114	342	1
Mb. durans 17SD2-17T	A2	711	4360	615	2
Mb. oxalidis DSM 24028T	A2	711	4358	498	18
Mb. planium YIM 132548T	A2	711	4049	532	2
Mb. segetis 17J42-1T	A2	711	3860	390	1
Mb. soli KCTC 22810T	A2	711	3984	965	64
Mb. aerolatum DSM 19013T	A3	711	3263	344	3
Mb. komagatae DSM 19563T	A3	711	3870	397	0
Mb. persicinum NBRC 103628T	A3	711	3552	378	3
Mb. brachiatum TX0642T	A4	711	4625	424	0
Mb. fujisawaense DSM 5686T	A4	711	4437	168	0
Mb. longum DSM 23933T	A4	711	4530	598	2
Mb. mesophilicum NBRC 15688T	A4	711	4309	291	4
Mb. oryzae CBMB20T	A4	711	4783	174	0
Mb. phyllosphaerae CBMB27T	A4	711	4609	147	0
Mb. phyllostachyos BL47T	A4	711	4253	486	0
Mb. pseudosasicola BL36T	A4	711	4825	567	1
Mb. radiotolerans JCM 2831T	A4	711	5006	353	1
Mb. tardum NBRC 103632T	A4	711	4747	609	2
Mb. dankookense DSM 22415T	A5	711	4316	379	26
Mb. gregans NBRC 103626T	A5	711	4046	341	20
Mb. hispanicum DSM 16372T	A5	711	4262	943	3
Mb. radiodurans 17Sr1-43T	A5	711	3866	317	2
Mb. symbioticum SB0023/3T	A5	711	4366	311	2
Mb. brachythecii DSM 24105T	B2	711	3645	535	3
Mb. gnaphalii DSM 24027T	B2	711	3262	545	2
Mb. haplocladii DSM24195T	B2	711	3064	408	17
Mb. organophilum NBRC 15689T	B2	711	3300	513	2
Mb. thuringiense DSM 23674	B2	711	3175	329	6
Mr. aminovorans NBRC 15686T	B	711	4168	472	2
Mr. extorquens TK 0001T	B	711	3944	355	2
Mr. podarium DSM 15083T	B	711	3591	485	82
Mr. populi BJ001T	B	711	4242	192	2
Mr. rhodesianum DSM 5687T	B	711	4216	227	1
Mr. rhodinum DSM 2163T	B	711	3696	405	1
Mr. salsuginis CGMCC 1.6474T	B	711	3706	319	1
Mr. suomiense DSM 14458T	B	711	3839	431	0
Mr. thiocyanatum JCM 10893T	B	711	4121	152	0
Mr. zatmanii LMG 6087T	B	711	3582	330	20
Mb. oryzihabitans TER-1T	C	711	4031	676	5
Mb. ajmalii IF7SW-B2T	C1	711	4864	284	2
Mb. aquaticum DSM 16371T	C1	711	4873	682	22
Mb. crusticola KCTC 52305T	C1	711	4825	895	6
Mb. curus PR1016AT	C1	711	5347	465	2
Mb. frigidarieris JCM 32048T	C1	711	5500	461	3
Mb. indicum SE2.11T	C1	711	4771	475	26
Mb. nonmethylophilum 6HR-1T	C1	711	5394	495	2
Mb. platani PMB02T	C1	711	5010	384	7
Mb. tarhaniae DSM 25844T	C1	711	4658	426	27
Mb. terrae 17Sr1-28T	C1	711	4464	231	8
Mb. terricola 17Sr1-39T	C1	711	5397	568	0
Mb. variabile DSM 16961T	C1	711	5108	501	58
Mb. isbiliense DSM 17168T	C2	711	4621	859	5
Mb. nodulans ORS 2060T	C2	711	4679	1764	12

Table 9. Conservation of methylotrophy genes.

Pathway		Methylamine, NMG pathway	Methylamine dehydrogenase	Co-dependent MDH, Mxa cluster	MDH regulation	PQQ synthesis	La-dependent MDH, XoxF1 cluster	La-dependent EDH, ExaF	Lut genes	HAMP pathway	GSH-dep. formaldehyde dehydrogenase	Formate dehydrogenase 1	Formate dehydrogenase 2	Formate dehydrogenase 3	Formate dehydrogenase 4	Folate synthesis	H4F pathway	Methylotrophy regulator	Formaldehyde de-binding protein	Serine cycle	EMC pathway
Number of genes		8	11	14	4	6	3	1	10	20	34	2	4	3	2	5	4	1	1	11	12
Gene symbol		<i>mgd, mgs</i>	<i>mxa</i>	<i>mxa</i>	<i>mxb, mxc</i>	<i>pqq</i>	<i>xox</i>	<i>exaF</i>	<i>lut, lum</i>	<i>fhc, fnc, etc.</i>	<i>gld, gld, fgh</i>	<i>fhl1</i>	<i>fhl2</i>	<i>fhl3</i>	<i>fhl4</i>	<i>fol, pab</i>	<i>fhl, fhlA, fhlB, fhlC</i>	<i>trnR</i>	<i>elgA</i>	<i>sga, mtk, etc.</i>	<i>pcdB, ccr, etc.</i>
Sequence source	Sub-clone	PA1	PA1	AM1	AM1	AM1	AM1	AM1	AM1	AM1	22A	AM1	AM1	AM1	AM1	AM1	AM1	PA1	PA1	AM1	AM1
Mb. jeotgali LMG 23639T	A	100	0	100	100	83.3	100	0	90	95	0	0	100	0	0	100	100	100	100	100	100
Mb. trifolii DSM 23632	A	0	0	100	100	100	100	100	90	95	0	0	100	0	100	100	100	100	100	100	91.7
Mb. cerastii DSM 23679T	A	0	0	92.9	100	100	100	0	90	90	66.7	0	100	0	50	100	100	100	100	100	100
Mb. iners DSM 19015T	A1	0	0	14.3	100	100	100	100	80	100	0	100	100	100	50	100	100	100	100	100	100
Mb. marchantiae DSM 21328T	A1	87.5	0	100	100	100	100	0	90	100	66.7	0	100	100	100	100	100	100	100	100	100
Mb. bullatum DSM 21893T	A1	100	0	100	100	100	100	0	90	95	66.7	0	100	100	100	100	100	100	100	100	100
Mb. goessingense DSM 21331T	A1	0	0	100	100	100	100	100	90	95	66.7	100	100	100	100	100	100	100	100	81.8	83.3
Mb. gossipicola Gh-105T	A1	12.5	0	100	100	100	100	0	90	95	66.7	0	100	100	100	100	100	100	100	100	100
Mb. adhaesivum DSM 17169T	A1	0	0	100	100	100	100	0	90	95	66.7	0	100	100	100	100	100	100	100	100	100
Mb. soli KCTC 22810T	A2	100	0	85.7	75	100	100	100	70	85	0	100	75	100	50	80	100	100	100	90.9	75
Mb. planium YIM 132548T	A2	100	0	92.9	100	100	100	100	90	95	0	100	100	100	100	100	100	100	100	100	100
Mb. oxalidis DSM 24028T	A2	100	90.9	92.9	100	100	66.7	100	90	90	0	100	100	100	100	100	75	100	100	100	100
Mb. segetis 17J42-1T	A2	0	18.2	92.9	100	100	100	100	90	100	0	100	50	100	100	100	100	100	100	100	100
Mb. durans 17SD2-17T	A2	100	0	100	100	100	100	100	90	100	0	100	100	100	100	100	100	100	100	100	100
Mb. aerolatum DSM 19013T	A3	0	0	100	100	100	100	100	90	100	0	0	100	0	50	100	100	100	100	100	100
Mb. persicinum NBRC 103628T	A3	0	0	14.3	100	100	100	100	90	100	0	0	100	0	50	100	100	100	100	100	100
Mb. komagatae DSM 19563T	A3	0	0	14.3	100	100	100	100	90	100	0	0	100	0	100	100	100	100	100	100	100
Mb. mesophilum NBRC 15688T	A4	0	9.09	100	100	100	100	0	90	100	66.7	0	100	0	100	100	100	100	100	100	100
Mb. pseudosacchara BL36T	A4	62.5	0	100	100	100	100	100	80	100	100	0	100	0	100	100	75	100	100	100	100
Mb. brachiatum TX0642T	A4	0	9.09	100	100	100	100	0	90	100	66.7	0	100	0	100	100	100	100	100	100	100
Mb. phyllostachyos BL47T	A4	50	0	100	100	100	100	100	90	100	0	0	100	0	100	100	100	100	100	100	100
Mb. longum DSM 23933T	A4	0	9.09	7.14	100	100	100	100	80	100	66.7	100	100	0	100	100	100	100	100	100	100
Mb. tardum NBRC 103632T	A4	50	0	7.14	100	100	100	100	90	100	0	100	100	100	100	100	100	100	100	100	100
Mb. radiotolerans JCM 2831T	A4	62.5	0	100	100	100	100	100	90	100	0	100	100	100	100	100	100	100	100	100	100
Mb. oryzae CBMB20T	A4	0	0	100	100	100	100	100	90	100	66.7	0	100	0	100	100	100	100	100	100	100
Mb. phyllosphaerae CBMB27T	A4	0	0	100	100	100	100	100	90	100	66.7	0	100	0	100	100	100	100	100	100	100
Mb. fujisawaense DSM 5686T	A4	0	0	100	100	100	100	100	90	100	66.7	0	100	0	100	100	100	100	0	100	100
Mb. symbioticum SB0023/3T	A5	0	0	92.9	100	100	100	100	90	95	0	100	100	0	100	100	100	100	100	90.9	100
Mb. dankookense DSM 22415T	A5	100	0	100	100	100	100	100	90	100	0	100	100	100	50	100	75	100	100	100	100
Mb. radiodurans 17Sr1-43T	A5	12.5	0	100	100	100	100	100	90	100	0	100	100	100	100	100	100	100	100	100	100
Mb. gregans NBRC 103626T	A5	0	0	92.9	50	83.3	100	100	90	100	0	100	100	100	100	100	100	100	100	100	100
Mb. hispanicum DSM 16372T	A5	0	9.09	100	100	100	100	100	90	100	0	100	100	100	100	100	100	100	100	100	100
Mb. gnaphalii DSM 24027T	B2	100	0	92.9	100	100	100	0	90	95	0	100	100	100	100	100	100	100	100	90.9	100
Mb. brachythecii DSM 24105T	B2	100	0	100	100	100	100	100	90	100	0	0	100	0	100	100	75	100	100	100	100
Mb. thuringiense DSM 23674	B2	0	0	100	100	100	100	0	90	100	66.7	0	100	0	100	100	100	100	100	100	100
Mb. haplocladii DSM24195T	B2	0	0	100	100	100	100	100	90	95	0	0	100	0	100	100	100	100	100	90.9	100
Mb. organophilum NBRC 15689T	B2	100	0	100	100	100	100	100	90	95	0	0	100	0	100	100	100	100	100	100	100
Mr. suomiense DSM 14458T	B	100	0	100	100	100	66.7	100	100	95	0	100	100	100	100	100	100	100	100	100	100
Mr. rhodinum DSM 2163T	B	100	0	100	100	100	100	100	100	90	0	100	100	100	100	100	100	100	100	100	100
Mr. salsuginis CGMCC 1.6474T	B	87.5	9.09	100	100	100	100	100	100	95	0	100	100	100	100	100	100	100	100	100	100
Mr. zatmanii LMG 6087T	B	100	0	100	100	100	100	100	100	95	0	0	100	100	100	100	75	100	100	90.9	91.7
Mr. extorquens TK 0001T	B	100	0	100	100	100	100	100	100	95	0	100	100	100	100	100	100	100	100	100	100
Mr. populi BJ001T	B	100	0	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100
Mr. thiocyanatum JCM 10893T	B	87.5	0	92.9	100	100	100	100	100	95	0	100	100	100	100	100	100	100	100	100	100
Mr. aminovorans NBRC 15686T	B	100	100	100	100	100	100	0	100	90	0	100	100	100	100	100	100	100	100	100	100
Mr. podarium DSM 15083T	B	100	100	100	100	83.3	100	0	80	85	0	100	50	100	50	100	75	100	100	100	91.7
Mr. rhodesianum DSM 5687T	B	100	100	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100
Mb. oryzae TER-1T	C	100	0	92.9	100	83.3	100	0	90	95	100	0	100	0	0	100	75	0	0	100	100
Mb. crusticola KCTC 52305T	C1	87.5	0	14.3	100	83.3	100	0	90	100	100	0	100	100	50	80	100	100	0	100	100
Mb. variable DSM 16961T	C1	0	9.09	14.3	100	83.3	100	0	90	100	100	0	100	100	50	80	100	100	0	90.9	83.3
Mb. nonmethylotrophicum 6HR-1T	C1	50	0	14.3	100	83.3	100	0	80	100	100	0	100	100	50	100	100	100	0	100	91.7
Mb. indicum SE2.11T	C1	62.5	0	78.6	75	83.3	100	0	90	100	100	0	100	66.7	0	100	75	100	100	90.9	91.7
Mb. ajmalii IF7SW-B2T	C1	100	0	78.6	100	83.3	100	0	90	100	100	0	100	100	0	100	100	100	0	100	100
Mb. terricola 17Sr1-39T	C1	12.5	0	85.7	100	83.3	100	0	90	95	100	0	100	100	50	100	100	100	0	100	100
Mb. terrae 17Sr1-28T	C1	12.5	0	14.3	100	83.3	100	0	90	95	33.3	0	100	100	0	100	100	100	0	100	100
Mb. platani PMB02T	C1	12.5	0	7.14	100	83.3	100	0	90	100	100	0	100	100	0	100	100	100	0	100	100
Mb. aquaticum DSM 16371T	C1	87.5	0	42.9	100	83.3	100	0	80	90	100	0	100	100	0	100	100	100	0	100	91.7
Mb. tarhaniae DSM 25844T	C1	12.5	0	85.7	100	83.3	100	0	90	90	100	0	100	100	0	100	100	100	0	100	100
Mb. curus PR1016AT	C1	0	9.09	14.3	100	83.3	100	0	90	90	100	0	100	100	0	100	100	100	100	100	100
Mb. frigidaeris JCM 32048T	C1	12.5	9.09	21.4	100	83.3	100	0	90	95	100	0	100	100	50	100	100	100	0	100	100
Mb. isbileni DSM 17168T	C2	0	0	14.3	100	83.3	100	0	80	95	33.3	100	50	100	50	100	100	100	0	100	91.7
Mb. nodulans ORS 2060T	C2	12.5	0	92.9	100	83.3	100	0	70	100	33.3	100	100	100	0	100	100	100	100	100	100
Enterovirus rhinocerotis DSM 25903T		0	0	0	0	66.7	0	0	0	0	33.3	0	100	0	0	80	50	0	0	54.5	100
Microvira lupini																					

Table 10. Detailed information of methylotrophy genes conservation in the PPFMs and the other Methylobacteriaceae species.

Pathway	Function	Accession ID	Gene name	Gene source	Enterovirga rhinocerotis DSM 25903T	Microvirga lupini AT3.9	Microvirga flocculans ATCC BAA-817T	Microvirga guangxiensis CGMCC 1.7666T	Microvirga subterranea DSM 14364T	Microvirga thermotolerans HRIT	Microvirga brassicaeum CDOBN77T	Microvirga aerophila NBRC 106136T	Microvirga massiliensis ICI19T	Microvirga loononidis WSM3557T	Microvirga tunisiensis Lmie10	Microvirga calopogonii CCBAU 65941T	Microvirga ossetica V5/3MT	Microvirga vignae BR3299T	Microvirga makkahensis KCTC 23863T	Microvirga pakistanensis NCCP-1258T
Methylamine, NMG pathway	Methylglutamate dehydrogenase subunit D	Mext_1651	mgdD	PA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Methylamine, NMG pathway	Methylglutamate dehydrogenase subunit C	Mext_1652	mgdC	PA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Methylamine, NMG pathway	Methylglutamate dehydrogenase subunit B	Mext_1653	mgdB	PA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Methylamine, NMG pathway	Methylglutamate dehydrogenase subunit A	Mext_1654	mgdA	PA1	0	0	0	0	0	0	0	0	54	0	0	0	0	52	0	0
Methylamine, NMG pathway	N-methyltransferase subunit A	Mext_1656	mgsA	PA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Methylamine, NMG pathway	N-methyltransferase subunit B	Mext_1657	mgsB	PA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Methylamine, NMG pathway	N-methyltransferase subunit C	Mext_1658	mgsC	PA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Methylamine, NMG pathway	Glutamine synthetase	Mext_1659	gmsA	PA1	0	0	0	0	0	0	0	0	77	0	0	0	0	0	0	0
Methylamine dehydrogenase	Unknown	MexAM1_META1p2769	mauF	AM1	0	0	0	0	0	0	0	0	61	0	0	0	0	0	0	0
Methylamine dehydrogenase	Methylamine dehydrogenase large subunit	MexAM1_META1p2770	mauB	AM1	0	0	0	0	0	0	0	0	71	0	0	0	0	0	0	0
Methylamine dehydrogenase	Essential for small subunit maturation	MexAM1_META1p2771	mauE	AM1	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0
Methylamine dehydrogenase	Essential for small subunit maturation	MexAM1_META1p2772	mauD	AM1	0	0	0	0	0	0	0	0	74	0	0	0	0	0	0	0
Methylamine dehydrogenase	Methylamine dehydrogenase small subunit	MexAM1_META1p2773	mauA	AM1	0	0	0	0	0	0	0	0	88	0	0	0	0	0	0	0
Methylamine dehydrogenase	Amicyanin	MexAM1_META1p2774	mauC	AM1	0	0	0	0	0	0	0	0	55	0	0	0	0	0	0	0
Methylamine dehydrogenase	Unknown	MexAM1_META1p2775	mauJ	AM1	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0
Methylamine dehydrogenase	Unknown	MexAM1_META1p2776	mauG	AM1	0	0	0	0	0	0	0	0	73	0	0	50	0	0	50	0
Methylamine dehydrogenase	Unknown	MexAM1_META1p2777	mauI	AM1	0	0	0	0	0	0	0	0	59	0	0	0	0	0	0	0
Methylamine dehydrogenase	Ferredoxin	MexAM1_META1p2778	mauM	AM1	0	0	0	0	0	0	0	0	84	0	0	0	0	0	0	0
Methylamine dehydrogenase	Ferredoxin	MexAM1_META1p2779	mauN	AM1	0	0	0	0	0	0	0	0	74	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Transcriptional regulator	MexAM1_META1p4525	mxsA	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Unknown	MexAM1_META1p4526	mxsH	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Unknown	MexAM1_META1p4527	mxsE	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Unknown	MexAM1_META1p4528	mxsD	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Essential for Ca2+ insertion into MDH	MexAM1_META1p4529	mxsL	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Essential for Ca2+ insertion into MDH	MexAM1_META1p4530	mxsK	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Essential for Ca2+ insertion into MDH	MexAM1_META1p4531	mxsC	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Essential for Ca2+ insertion into MDH	MexAM1_META1p4532	mxsA	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Unknown	MexAM1_META1p4533	mxsS	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Unknown	MexAM1_META1p4534	mxsR	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Methanol dehydrogenase, small subunit	MexAM1_META1p4535	mxsA	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Cytochrome c550	MexAM1_META1p4536	mxsG	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	MxaJ, possible chaperone	MexAM1_META1p4537	mxsJ	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ca-dependent MDH, Mxa cluster	Methanol dehydrogenase, large subunit	MexAM1_META1p4538	mxsF	AM1	0	0	0	0	0	0	0	0	51	0	0	0	0	51	51	0
MDH regulation	Transcriptional regulator	MexAM1_META1p1752	mxsB	AM1	0	0	0	0	0	0	0	0	67	0	0	0	0	0	0	0
MDH regulation	Sensor kinase	MexAM1_META1p1753	mxsD	AM1	0	0	0	0	0	0	0	0	55	0	0	0	0	0	0	0
MDH regulation	Sensor kinase	MexAM1_META1p4896	mxsQ	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MDH regulation	Transcriptional regulator	MexAM1_META1p4897	mxsE	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PQQ synthesis	PQQ synthesis protein, PqqE	MexAM1_META1p1748	pqqE	AM1	58	0	0	0	0	0	0	0	78	0	0	0	0	0	72	71
PQQ synthesis	PQQ synthesis protein, PqqC/D	MexAM1_META1p1749	pqqC/D	AM1	61	0	0	0	0	0	0	0	79	0	0	0	0	0	71	71
PQQ synthesis	PQQ synthesis protein, PqqB	MexAM1_META1p1750	pqqB	AM1	0	0	0	0	0	0	0	0	68	0	0	0	0	0	59	62
PQQ synthesis	PQQ synthesis protein, PqqA	MexAM1_META1p1751	pqqA	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PQQ synthesis	PQQ synthesis protein, PqqF	MexAM1_META1p2330	pqqF	AM1	73	79	77	78	77	79	77	78	77	78	78	78	78	78	77	78
PQQ synthesis	PQQ synthesis protein, PqqG	MexAM1_META1p2331	pqqG	AM1	67	72	71	72	70	71	71	71	67	71	71	71	72	71	71	71
ROS protection	Glutathione peroxidase	MOC_1064		CBMB20T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ROS protection	Glutathione disulfide reductase	MOC_3063		CBMB20T	73	74	75	73	73	72	73	72	73	75	75	75	74	71	74	74
Ln-dependent MDH, XoxF1 cluster	Ln-dependent MDH, XoxF1	MexAM1_META1p1740	xoxF1	AM1	0	0	0	0	0	0	0	0	81	72	71	0	72	0	79	79
Ln-dependent MDH, XoxF1 cluster	Cytochrome c	MexAM1_META1p1741	xoxG	AM1	0	0	0	0	0	0	0	0	61	57	59	0	61	0	56	57
Ln-dependent MDH, XoxF1 cluster	Periplasmic binding proteins	MexAM1_META1p1742	xoxJ	AM1	0	0	0	0	0	0	0	0	55	0	0	0	0	0	52	57
Methylotrophy regulator	formaldehyde-responsive transcription factor	Mext_0925	itmR	PA1	0	61	64	0	58	0	0	58	0	65	58	61	0	0	0	59
Formaldehyde-binding protein	formaldehyde-sensing protein	Mext_4158	efgA	PA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ln-dependent EDH, ExaF	Ln-dependent ethanol dehydrogenase	MexAM1_META1p1139	exaF	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ln transport and utilization (lut) genes	ABC transporter ATPase	MexAM1_META1p1778	lutA	AM1	0	0	0	0	0	0	0	0	65	63	63	0	61	0	62	60
Ln transport and utilization (lut) genes	Putative periplasmic protein	MexAM1_META1p1779	lutB	AM1	0	56	0	0	56	53	0	0	56	76	71	72	0	74	0	71
Ln transport and utilization (lut) genes	Hypothetical protein	MexAM1_META1p1780	lutC	AM1	0	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0
Ln transport and utilization (lut) genes	Hypothetical protein	MexAM1_META1p1781	lutD	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ln transport and utilization (lut) genes	ABC transporter ATPase	MexAM1_META1p1782	lutE	AM1	0	52	0	0	51	54	0	0	55	70	55	51	0	57	53	62
Ln transport and utilization (lut) genes	multidrug ABC transporter permease	MexAM1_META1p1783	lutF	AM1	0	64	0	0	64	60	0	0	62	80	65	66	0	65	63	77
Ln transport and utilization (lut) genes	Putative periplasmic protein	MexAM1_META1p1784	lutG	AM1	0	0	0	0	0	0	0	0	62	0	0	0	0	0	54	53
Ln transport and utilization (lut) genes	TonB-dependent receptor	MexAM1_META1p1785	lutH	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ln transport and utilization (lut) genes	Lanmodulin	MexAM1_META1p1786	lanM	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ln transport and utilization (lut) genes	Hypothetical protein	MexAM1_META1p1787	lutI	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Formate dehydrogenase 1	Tungsten-dependent formate dehydrogenase, beta	MexAM1_META1p5031	fdh1B	AM1	0	0	0	0	0	0	0	0	68	0	0	0	0	0	0	0
Formate dehydrogenase 1	Tungsten-dependent formate dehydrogenase, alpha	MexAM1_META1p5032	fdh1A	AM1	0	0	0	0	0	0	0	0	71	0	0	0	0	0	0	0
Formate dehydrogenase 2	Molybdenum-dependent formate dehydrogenase, MexAM1_META1p4846	fdh2C	AM1	57	0	0	0	61	61	0	0	0	55	55	61	54	58	64	55	55
Formate dehydrogenase 2	Molybdenum-dependent formate dehydrogenase, MexAM1_META1p4847	fdh2B	AM1	68	0	0	0	68	66	0	0	0	67	66	69	68	68	67	66	66
Formate dehydrogenase 2	Molybdenum-dependent formate dehydrogenase, MexAM1_META1p4848	fdh2A	AM1	80	0	0	0	80	80	0	0	0	81	81	82	81	81	81	81	82
Formate dehydrogenase 2	Molybdenum-dependent formate dehydrogenase, MexAM1_META1p4849	fdh2D	AM1	53	0	0	0	0	56	0	0	0	52	52	52	52	52	0	51	51
Formate dehydrogenase 3	Cytochrome-linked formate dehydrogenase alpha	MexAM1_META1p0303	fdh3A	AM1	0	0	0	0	0	0	0	0	88	0	0	0	0	0	0	0
Formate dehydrogenase 3	Cytochrome-linked formate dehydrogenase beta	MexAM1_META1p0304	fdh3B	AM1	0	0	0	0	0	0	0	0	92	0	0	0	0	0	0	0
Formate dehydrogenase 3	Cytochrome-linked formate dehydrogenase gamma	MexAM1_META1p0305	fdh3C	AM1	0	0	0	0	0	0	0	0	63	0	0	0	0	0	0	0
Formate dehydrogenase 4	Protein associated with expression of formate dehydrogenase	MexAM1_META1p2093	fdh4B	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Formate dehydrogenase 4	Formate dehydrogenase 4	MexAM1_META1p2094	fdh4A	AM1	0	0	0	0	0	0	0	0	52	0	51	0	0	0	0	0
GSH-dependent formaldehyde dehydrogenase	Glutathione-dependent formaldehyde-activating cysteine	Maq22A_1p37180	gfa	22A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GSH-dependent formaldehyde dehydrogenase	S-hydroxymethyl glutathione dehydrogenase	Maq22A_c21490	hgd	22A	91	89	90	88	90	88	87	89	86	90	90	90	89	90	89	89
GSH-dependent formaldehyde dehydrogenase	S-formylglutathione hydrolase	Maq22A_c21495	fgh	22A	0	59	61	60	60	0	0	0	59	58	62	59	61	57	59	0

Table 10. Detailed information of methylotrophy genes conservation in the PPFMs and the other Methylobacteriaceae species.

Pathway	Function	Accession ID	Gene name	Gene source	Enterovirga rhinocerotis DSM 25903T	Microvirga lupini AT3.9	Microvirga flocculans ATCC BAA-817T	Microvirga guangxiensis CGMCC 1.7666T	Microvirga subterranea DSM 14364T	Microvirga thermotolerans HRIT	Microvirga brassicaearum CDVBN77T	Microvirga aerophila NBRC 106136T	Microvirga massiliensis ICI19T	Microvirga lotonoidis WSM3557T	Microvirga tunisiensis Lmie10	Microvirga calopogonii CCBBAU 65841T	Microvirga ossetica V5/3MT	Microvirga vignae BR3299T	Microvirga makkahensis KCTC 23663T	Microvirga pakistanensis NCCP-1258T
Folate synthesis	2-amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine synthase	MexAM1_META1p1743	folK	AM1	60	0	0	0	0	0	0	65	54	55	50	0	55	55	0	
Folate synthesis	Dihydroneopterin aldolase	MexAM1_META1p1744	folB	AM1	61	0	0	0	0	0	0	57	57	55	57	0	57	55	0	
Folate synthesis	Dihydropterotate synthase	MexAM1_META1p1745	folP	AM1	69	0	0	0	0	0	0	71	63	65	63	0	63	64	0	
Folate synthesis	para-Aminobenzoic acid synthase component	MexAM1_META1p4284	pabB	AM1	0	0	0	0	0	0	0	55	50	54	52	0	54	0	0	
Folate synthesis	Dihydrofolate synthase	MexAM1_META1p4888	folC	AM1	70	76	74	75	75	75	75	76	73	76	76	75	76	77	76	
H4F pathway	Formyl-tetrahydrofolate ligase	MexAM1_META1p0329	frfL	AM1	71	0	0	0	0	0	0	67	0	0	0	0	0	0	0	
H4F pathway	Bifunctional methylene-H4MPT/ methylene-H4F	MexAM1_META1p1728	mttA	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H4F pathway	Methenyl-H4F cyclohydrolase	MexAM1_META1p1729	fch	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H4F pathway	Dihydrofolate reductase	MexAM1_META1p2852	folA	AM1	55	57	60	56	59	58	55	57	51	59	57	57	58	59	60	
H4MPT pathway	Formyltransferase/hydrolase complex, gamma subunit	MexAM1_META1p1755	fhcC	AM1	0	0	0	0	0	0	0	56	0	0	0	0	0	0	0	
H4MPT pathway	Formyltransferase/hydrolase complex, delta subunit	MexAM1_META1p1756	fhcD	AM1	0	0	0	0	0	0	0	73	0	0	0	0	0	0	0	
H4MPT pathway	Formyltransferase/hydrolase complex, alpha subunit	MexAM1_META1p1757	fhcA	AM1	0	0	0	0	0	0	0	62	0	0	0	0	0	0	0	
H4MPT pathway	Formyltransferase/hydrolase complex, beta subunit	MexAM1_META1p1758	fhcB	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H4MPT pathway	Ribofuransylaminobenzene, 5'-phosphate synthase	MexAM1_META1p1760	mptG	AM1	0	0	0	0	0	0	0	57	0	0	0	0	0	0	0	
H4MPT pathway	Methylene-H4MPT dehydrogenase	MexAM1_META1p1761	mttB	AM1	0	0	0	0	0	0	0	67	0	0	0	0	0	0	0	
H4MPT pathway	Unknown	MexAM1_META1p1762	orfY	AM1	0	0	0	0	0	0	0	52	0	0	0	0	0	0	0	
H4MPT pathway	Methenyl-H4MPT cyclohydrolase	MexAM1_META1p1763	mch	AM1	0	0	0	0	0	0	0	68	0	0	0	0	0	0	0	
H4MPT pathway	H4MPT biosynthesis	MexAM1_META1p1764	orf5	AM1	0	0	0	0	0	0	0	70	0	0	0	0	0	0	0	
H4MPT pathway	Unknown	MexAM1_META1p1765	orf7	AM1	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	
H4MPT pathway	Formaldehyde activating enzyme	MexAM1_META1p1766	fae	AM1	0	0	0	0	0	0	0	83	0	0	0	0	0	0	0	
H4MPT pathway	Unknown	MexAM1_META1p1767	orf17	AM1	0	0	0	0	0	0	0	53	0	0	0	0	0	0	0	
H4MPT pathway	H4MPT biosynthesis	MexAM1_META1p1768	orf9	AM1	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0	
H4MPT pathway	Homolog of mxaE	MexAM1_META1p1770		AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H4MPT pathway	Homolog of mxaD	MexAM1_META1p1771		AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H4MPT pathway	Homolog of mxaD	MexAM1_META1p1772		AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
H4MPT pathway	H4MPT biosynthesis	MexAM1_META1p1773	orf19	AM1	0	0	0	0	0	0	0	63	0	0	0	0	0	0	0	
H4MPT pathway	H4MPT biosynthesis	MexAM1_META1p1774	orf20	AM1	0	0	0	0	0	0	0	74	0	0	0	0	0	0	0	
H4MPT pathway	H4MPT biosynthesis	MexAM1_META1p1775	orf21	AM1	0	0	0	0	0	0	0	55	0	0	0	0	0	0	0	
H4MPT pathway	H4MPT biosynthesis	MexAM1_META1p1776	orf22	AM1	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	
Serine / EMC pathway	Beta-ketothiolase	MexAM1_META1p3700	phaA	AM1	80	79	81	80	79	79	80	79	0	0	79	79	80	79	79	
Serine cycle	Serine cycle transcriptional regulator	MexAM1_META1p0756	qscR	AM1	0	0	0	0	0	0	0	69	0	0	0	0	0	0	0	
Serine cycle	Malate dehydrogenase	MexAM1_META1p1537	mdh	AM1	82	86	85	86	87	84	86	85	87	85	86	86	86	86	86	
Serine cycle	Serine glyoxylate aminotransferase	MexAM1_META1p1726	sga	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Serine cycle	Hydroxypyruvate reductase	MexAM1_META1p1727	hpr	AM1	0	0	0	52	0	0	0	53	53	0	56	56	53	53	0	
Serine cycle	Malate thiokinase, alpha subunit	MexAM1_META1p1730	mtkA	AM1	0	0	50	0	0	50	50	0	0	0	50	0	50	50	0	
Serine cycle	Malate thiokinase, beta subunit	MexAM1_META1p1731	mtkB	AM1	52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Serine cycle	PEP carboxylase	MexAM1_META1p1732	ppc	AM1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Serine cycle	Malyl-CoA lyase/beta-malyl-CoA lyase	MexAM1_META1p1733	mcl	AM1	53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Serine cycle	Glycerate kinase	MexAM1_META1p2944	gck	AM1	69	0	0	0	0	51	0	50	0	0	0	0	0	0	0	
Serine cycle	Enolase	MexAM1_META1p2984	eno	AM1	80	82	83	81	82	81	81	81	84	83	83	83	82	83	83	84
Serine cycle	Serine hydroxymethyltransferase	MexAM1_META1p3384	glyA	AM1	78	77	78	77	79	79	76	77	78	77	77	77	78	78	78	
TCA / EMC cycle	Fumarase C	MexAM1_META1p2857	fumC	AM1	83	78	78	78	81	80	79	80	77	77	77	77	77	77	78	78
TCA cycle	Succinate dehydrogenase, cytochrome b556 subunit	MexAM1_META1p3859	sdhC	AM1	62	64	65	68	66	68	67	61	63	64	64	64	65	67	65	
TCA cycle	Succinate dehydrogenase, membrane anchor subunit	MexAM1_META1p3860	sdhD	AM1	59	68	67	63	67	70	66	69	64	67	68	67	66	65	69	
TCA cycle	Succinate dehydrogenase, flavoprotein subunit	MexAM1_META1p3861	sdhA	AM1	88	87	88	88	88	88	88	88	88	88	88	88	88	88	89	
TCA cycle	Succinate dehydrogenase, iron-sulfur subunit	MexAM1_META1p3863	sdhB	AM1	78	80	82	80	79	81	78	80	80	81	81	80	81	79	79	
EMC pathway	Propionyl-CoA carboxylase beta subunit	MexAM1_META1p0172	pccB	AM1	85	88	88	88	88	87	87	90	84	88	89	88	87	88	0	88
EMC pathway	Crotonyl-CoA carboxylase/reductase	MexAM1_META1p0178	ccr	AM1	85	0	0	0	0	0	0	86	0	0	0	0	0	0	0	0
EMC pathway	Ethylmalonyl-CoA mutase	MexAM1_META1p0180	ecm	AM1	77	0	0	0	0	0	0	77	0	0	0	0	0	0	0	0
EMC pathway	Methylmalonyl-CoA mutase accessory protein	MexAM1_META1p0188	meaB	AM1	70	76	76	76	77	77	74	76	74	72	72	73	76	73	76	74
EMC pathway	Ethylmalonyl-CoA/methylmalonyl-CoA epimerase	MexAM1_META1p0839	epi	AM1	75	70	69	68	75	74	72	71	77	67	68	69	68	69	70	72
EMC pathway	Methylmalonyl-CoA mutase, beta subunit	MexAM1_META1p2390	mcmB	AM1	59	55	58	59	56	59	53	57	54	54	61	54	53	58	57	59
EMC pathway	Propionyl-CoA carboxylase, alpha subunit	MexAM1_META1p3203	pccA	AM1	81	81	80	81	82	82	80	79	77	81	81	81	81	0	0	81
EMC pathway	Crotonase	MexAM1_META1p3675	croR	AM1	76	0	0	0	0	0	0	84	0	0	0	65	0	0	0	0
EMC pathway	Acetoacetyl-CoA reductase	MexAM1_META1p3701	phaB	AM1	65	82	78	79	79	76	74	0	78	81	0	80	82	82	80	79
EMC pathway	Mesaconyl-CoA hydratase	MexAM1_META1p4153	mcd	AM1	68	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0
EMC pathway	Methylmalonyl-CoA mutase, alpha subunit	MexAM1_META1p5251	mcmA	AM1	82	83	83	83	84	86	84	84	82	84	83	84	84	83	0	83
EMC pathway	Methylsuccinyl-CoA dehydrogenase	MexAM1_META1p2223	msd	AM1	72	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0
Values represent percentage of identity against the query proteins.																				

Values represent percentage of identity against the query proteins.

Table 10. Detailed information of methylotrophy genes conservation in the PPFMs and the other Methylobacteriaceae species.

[illegible]

Values represent percentage of identity against the query proteins.

Table 11. Detected gene clusters for secondary metabolites.

[illegible]

Values represent clusters counted in the genome of each strain.

Table 12. Plant growth-promoting genes.									
Function	Query ID	Gene function	Gene ID	Source strain	A	A	A	A	A
ACC deaminase	MOC_1898	Acc deaminase	acdS	CBMB20T	0	0	90	0	90
D-cysteine desulphydrase	MexAM1_META1p2360	D-cysteine desulphydrase	dcdD	AM1	0	75	21	0	0
Cellulase	MexAM1_META1p1169	Cellulase	celC	AM1	89	0	65	0	69
Cellulase	Maq22A_1p38150	Endoglucanase, cellulase	bglC	22A	0	0	0	0	0
Cellulase	Maq22A_2p41770	Cellulase	22A	22A	57	0	56	0	51
Cellulase	Maq22A_2p42730	EndoglucanaseY	celA	22A	0	0	0	0	0
Antifungal	MOC_0504	putative chitinase	CBMB20T	0	0	0	0	0	0
Cytokinin synthesis	MOC_5190	Isopentenyl tRNA transferase	miaA	71	78	76	70	73	71
IAA synthesis	Mnod_1238	L-tryptophan decarboxylase	ORS 2060	0	0	0	0	0	0
IAA synthesis	MexAM1_META1p4817	Monoamine oxidase	AM1	0	82	0	0	0	0
IAA synthesis	Mnod_0239	Aldehyde dehydrogenase	ORS 2060	0	0	0	0	0	0
IAA synthesis	Mnod_4579	Monoamine oxidase	ORS 2060	0	59	56	0	0	0
IAA synthesis	MOC_0996	Monoamine oxidase	CBMB20T	0	69	65	0	0	0
IAA synthesis	MOC_5534	Aldehyde dehydrogenase	CBMB20T	0	0	0	0	0	0
Pectin metabolism	MexAM1_META1p4514	Pectate lyase	AM1	0	0	0	0	0	0
ROS protection	MOC_4484	Glutathione peroxidase	CBMB20T	73	86	89	67	70	66
ROS protection	MOC_5157	Superoxide dismutase	CBMB20T	92	97	93	93	87	89
ROS protection	Maq22A_1p31895	Hercynine oxygenase, ergothioneine synthesis	egtB	74	71	73	73	69	69
ROS protection	Maq22A_1p31890	Histidine N-alpha-methyltransferase, ergothioneine sy	egtD	68	69	70	70	68	67
UV damage repair	MOC_2537	UV DNA damage endonuclease	CBMB20T	71	81	75	70	72	73
PHB synthesis	MexAM1_META1p3699	PHB synthesis regulator/acetyl-CoA flux	phoR	79	78	74	80	81	78
Values represent percentage of identity against the query proteins.									

Table 13. Phenotype analysis of PPFMs.																							
Strains code NO.	Strains	strain name	Clade	Methanol	Methanol +La	Methylamine	Dichloromethane	Ethanol	Ethanol+La	D(-)Fructose	Sucrose	D(+Glucose	Glycerol	Glycerol+La	Glycerol+1%NaCl	Amp (25 mg/l) on R2A	Amp (50 mg/l) on R2A	PDA	37°C on R2A	NB solid	IAA production (µg/ml)	Fusarium inhibition % *	
				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
17	<i>Mb. jeotgali</i>	LMG 23639T	A	+	+	+	-	+	+	-	+	-	+	+	+	+	+	+	+	+	+	13 ± 0.86	-
33	<i>Mb. trifolii</i>	DSM 23632	A	+	+	-	-	+	+	w	-	+	+	-	+	-	vw	vw	+	-	-	7.8 ± 0.1	-
55	<i>Mb. cerastii</i>	DSM 23679T	A	-	vw	-	-	w	nd	-	+	-	-	-	nd	-	-	+	nd	-	-	10.6 ± 0.71	22±26
27	<i>Mb. gossypicola</i>	NRRL B 51692T	A1	+	+	-	-	+	+	+	+	+	w	+	+	+	vw	-	-	+	+	14.9 ± 0.6	nd
34	<i>Mb. goesingense</i>	DSM 21331T	A1	+	+	-	-	+	+	+	+	+	+	+	-	+	+	+	+	-	-	11.7 ± 0.1	12±11
47	<i>Mb. adhaesivum</i>	DSM 17169T	A1	+	+	-	-	+	+	+	+	+	+	vw	-	vw	+	-	w	-	-	18 ± 0.78	-
53	<i>Mb. iners</i>	DSM 19015T	A1	-	vw	-	-	w	+	-	w	-	w	+	nd	vw	vw	-	-	-	-	8.8 ± 0.16	nd
54	<i>Mb. marchantiae</i>	DSM 21328T	A1	+	+	+	-	+	+	+	+	+	+	+	+	+	vw	w	vw	w	-	16.9 ± 1.4	nd
56	<i>Mb. bullatum</i>	DSM 21893T	A1	+	+	+	-	+	+	+	+	+	+	+	vw	+	+	+	vw	w	-	20.2 ± 2.04	nd
26	<i>Mb. soli</i>	KCTC 22810T	A2	+	+	+	-	+	+	-	-	+	+	+	-	-	+	+	-	-	-	12.7 ± 0.39	11±26
41	<i>Mb. oxalidis</i>	DSM 24028T	A2	+	+	+	-	+	+	+	-	+	+	+	w	+	+	+	-	w	-	27.4 ± 2.66	-
62	<i>Mb. durans</i>	KCTC 52908T	A2	+	+	+	-	+	+	+	+	+	+	+	-	-	-	+	-	+	+	14.7 ± 0.46	nd
63	<i>Mb. segetis</i>	KCTC 62267T	A2	+	+	w	-	+	+	+	+	+	+	+	-	+	-	+	vw	+	+	17 ± 0.61	-
67	<i>Mb. planium</i>	NBRC 114056T	A2	w	+	+	-	+	+	+	+	+	+	+	-	-	-	-	-	+	+	9.1 ± 0.26	nd
13	<i>Mb. komagatae</i>	NBRC 103627T	A3	-	w	-	-	+	+	+	+	+	+	+	+	+	+	+	+	vw	-	9.7 ± 0.23	77±2
48	<i>Mb. persicinum</i>	NBRC 103628T	A3	-	+	-	-	+	+	-	-	-	+	+	w	+	+	+	-	-	-	13.3 ± 0.45	66±3
52	<i>Mb. aerolatum</i>	DSM 19013T	A3	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	-	+	11.7 ± 0.08	53±18
1	<i>Mb. radiotolerans</i>	IAM 12098T	A4	+	+	-	-	+	+	-	-	+	+	+	+	+	+	+	+	+	+	12.8 ± 0.18	70±5
10	<i>Mb. mesophilicum</i>	NBRC 15688T	A4	+	+	-	-	+	+	-	-	+	+	+	+	+	+	+	+	-	+	9.2 ± 0.36	12±3
14	<i>Mb. brachiatum</i>	NBRC 103629T	A4	+	+	-	-	w	+	-	+	+	+	+	+	+	+	+	+	+	+	9.3 ± 0.15	63±1
15	<i>Mb. tardum</i>	NBRC 103632T	A4	-	+	-	-	+	+	-	w	+	+	+	+	+	+	+	vw	w	-	9.1 ± 0.1	45±25
19	<i>Mb. oryzae</i>	DSM 18207T	A4	+	+	-	-	+	+	-	-	+	+	+	+	+	+	+	+	+	+	12 ± 0.57	65±7
20	<i>Mb. phyllosphaerae</i>	DSM 19779T	A4	+	+	-	-	+	+	-	-	+	+	+	+	+	+	+	+	+	+	12.3 ± 0.47	45±24
28	<i>Mb. longum</i>	DSM 23933T	A4	-	vw	-	-	+	+	-	-	-	+	+	w	+	-	+	-	-	-	8.9 ± 0.21	51±25
29	<i>Mb. fujisawaense</i>	DSM 5686T	A4	+	+	-	-	+	+	-	-	+	+	+	+	+	+	+	+	+	+	16.1 ± 0.51	44±23
44	<i>Mb. pseudosasicola</i>	NBRC 105203T	A4	+	+	-	-	+	+	+	-	+	+	+	w	+	+	+	-	vw	-	16.4 ± 1.57	64±4
46	<i>Mb. phyllostachyos</i>	NBRC 105206T	A4	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	-	w	-	12.9 ± 0.26	56±13
12	<i>Mb. gregans</i>	NBRC 103626T	A5	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	vw	+	+	13.7 ± 0.49	51±34
24	<i>Mb. dankookense</i>	DSM 22415T	A5	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	w	+	+	26.7 ± 1.36	71±3
50	<i>Mb. hispanicum</i>	DSM 16372T	A5	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	-	+	+	19.1 ± 0.34	70±5
2	<i>Mb. dichloromethanicum</i>	DSM 6343	B	+	+	+	+	+	+	-	-	-	+	+	+	+	+	+	-	+	+	14 ± 0.39	nd
3	<i>Mb. chloromethanicum</i>	NCIMB 13688	B	+	+	+	-	+	+	+	-	-	+	+	+	+	+	vw	+	w	+	10.6 ± 0.18	nd
4	<i>Mr. zatmanii</i>	DSM 5688T	B	+	+	+	-	+	+	+	-	-	+	+	+	+	+	+	-	+	+	13.2 ± 0.18	-
5	<i>Mr. lusitanum</i>	NCIMB 13779	B	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	11.3 ± 0.72	nd
8	<i>Mr. aminovorans</i>	NBRC 15686T	B	+	+	+	-	+	+	+	+	+	+	+	+	+	vw	+	w	+	+	13.8 ± 0.82	-
9	<i>Mr. extorquens</i>	NBRC 15687T	B	+	+	+	-	+	+	+	-	-	+	+	+	+	+	+	vw	+	+	11.1 ± 0.65	53±2
18	<i>Mr. suomiense</i>	DSM 14458T	B	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	vw	-	-	19.5 ± 1.98	12±5
21	<i>Mr. salsuginis</i>	NCCB 100140T	B	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	20.4 ± 1.23	-
23	<i>Mr. rhodinum</i>	DSM 2163T	B	+	+	+	-	+	+	+	+	+	+	+	+	+	vw	-	-	w	-	14.3 ± 0.59	nd
25	<i>Mr. extorquens</i>	AM1 ATCC14718	B	+	+	+	-	+	+	+	-	-	+	+	+	+	-	+	-	+	+	15.2 ± 0.13	nd
30	<i>Mr. rhodesianum</i>	DSM 5687T	B	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	23.7 ± 1.27	60±5
31	<i>Mr. lusitanum</i>	DSM 14457	B	+	+	+	-	+	+	+	+	+	+	+	+	+	-	+	+	w	-	16.6 ± 1.08	12±5
32	<i>Mr. podarium</i>	DSM 15083T	B	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	20.7 ± 0.6	37±20
49	<i>Mr. thiocyanatum</i>	JCM 10893T	B	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	20.2 ± 0.54	30±7
11	<i>Mb. organophilum</i>	NBRC 15689T	B2	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	12.3 ± 0.34	-
38	<i>Mb. gnaphalii</i>	DSM 24027T	B2	+	+	+	-	+	+	-	-	-	+	+	+	+	-	+	-	-	-	29.3 ± 2.48	-
42	<i>Mb. brachythecii</i>	DSM 24105T	B2	+	+	+	-	w	w	+	-	+	+	+	+	+	-	+	-	-	-	27.4 ± 0.36	-
43	<i>Mb. haplocladii</i>	DSM 24195T	B2	+	+	-	-	+	+	-	+	+	+	w	-	-	+	-	-	-	-	9 ± 0.68	-
57	<i>Mb. thuringiense</i>	DSM 23674	B2	+	+	-	-	+	+	+	-	+	w	w	nd	-	-	w	nd	-	-	14.3 ± 1.18	-
64	<i>Mb. oryzihabitans</i>	BCRC 81157T	C	+	+	w	-	+	+	+	+	+	+	+	-	-	+	+	vw	+	+	10.4 ± 0.34	nd
7	<i>Mb. variabile</i>	DSM 16961T	C1	-	+	-	-	w	+	+	+	+	+	+	+	+	+	+	+	+	+	9.1 ± 0.23	18±23
16	<i>Mb. platani</i>	JCM 14648T	C1	-	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	w	-	10.3 ± 0.29	47±15
22	<i>Mb. aquaticum</i>	DSM 16371T	C1	+	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	11.6 ± 0.29	39±5
35	<i>Mb. tarhaniae</i>	DSM 25844T	C1	+	+	w	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	12.5 ± 0.2	53±30
58	<i>Mb. frigidaeris</i>	JCM 32048T	C1	w	+	-	-	w	+	+	+	+	+	+	w	+	+	+	+	+	+	12.5 ± 0.15	64±8
59	<i>Mb. nonmethylotrophicum</i>	KCTC 42760T	C1	-	+	-	-	w	+	+	+	+	+	+	-	+	+	+	vw	-	-	11.6 ± 0.23	42±18
60	<i>Mb. terrae</i>	KCTC 52904T	C1	vw	w	-	-	w	+	+	+	+	+	+	-	+	+	+	+	+	+	12.3 ± 0.76	19±7
61	<i>Mb. crusticola</i>	KCTC 52305T	C1	vw	+	-	-	w	vw	-	-	-	w	+	-	+	+	-	-	-	-	10.1 ± 3.12	nd
65	<i>Mb. indicum</i>	JCM 30761T	C1	+	+	w	-	+	+	+	+	+	+	+	-	+	+	+	+	+	+	11 ± 0.28	32±21
66	<i>Mb. currus</i>	JCM 32670T	C1	vw	+	w	-	w	+	+	+	+	+	+	w	+	+	+	+	+	+	11.5 ± 0.2	nd
68	<i>Mb. terricola</i>	NBRC 112874T	C1	+	+	w	-	+	+	+	+	+	+	+	-	+	+	+	+	+	+	11.9 ± 0	54±27
6	<i>Mb. nodulans</i>	LMG 21967T	C2	+	+	-	-	+	+	-	-	+	+	+	+	+	+	+	+	+	+	8.8 ± 0.45	17±9
51	<i>Mb. isbiliense</i>	DSM 17168T	C2	-	w	-	-	w	+	-	-	-	+	w	w	+	+	-	+	-	-	10.4 ± 0.26	nd
Symbols; +, Positive; w, weak growth; vw, very weak growth; -, negative growth.																							
*; -, no activity; nd, not determined, values represent inhibition percentage compared to the control.																							

Species name	Strain name	clade	24 hours										7 days									
			Enzymes activity tests										Assimilation of sugar					Assimilation of acids				
			NO ₃	TRP	GLU	ADH	URE	ESC	GEL	PNPG	GLU	ARA	MNE	MAN	NAG	MAL	GNT	CAP	ADI	MLT	CIT	PAC
<i>Mb. jeotgali</i>	LMG 23639T	A	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mb. trifolii</i>	DSM 23632	A	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Mb. cerastii</i>	DSM 23679T	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Mb. gossypicola</i>	NRRL B 51692T	A1	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mb. goesingense</i>	DSM 21331T	A1	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	0	3	1
<i>Mb. adhaesivum</i>	DSM 17169T	A1	0	0	0	0	3	0	0	0	3	3	3	3	3	3	1	0	0	0	1	1
<i>Mb. iners</i>	DSM 19015T	A1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Mb. marchantiae</i>	DSM 21328T	A1	0	0	0	0	3	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0
<i>Mb. bullatum</i>	DSM 21893T	A1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Mb. soli</i>	KCTC 22810T	A2	0	0	0	0	2	0	0	0	3	3	3	3	3	3	3	0	0	0	0	0
<i>Mb. oxalidis</i>	DSM 24028T	A2	0	0	0	0	1	0	0	0	3	3	3	3	3	3	3	0	2	2	0	0
<i>Mb. durans</i>	KCTC 52908T	A2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Mb. segetis</i>	KCTC 62267T	A2	0	0	0	0	2	0	0	0	3	3	3	3	3	3	3	0	3	0	3	1
<i>Mb. planum</i>	NBRC 114056T	A2	0	0	0	0	2	0	0	0	3	3	3	3	3	3	2	0	0	0	0	0
<i>Mb. komagatae</i>	NBRC 103627T	A3	0	0	0	1	3	0	0	0	3	3	1	1	1	1	3	0	0	3	3	0
<i>Mb. persicinum</i>	NBRC 103628T	A3	0	0	0	3	3	0	0	0	2	3	0	0	0	0	3	0	2	0	1	0
<i>Mb. aerolatum</i>	DSM 19013T	A3	0	0	0	2	3	0	0	0	3	3	3	3	3	3	3	0	2	3	3	0
<i>Mb. radiotolerans</i>	IAM 12098T	A4	3	0	0	1	3	0	0	0	2	2	2	2	2	2	2	0	1	3	3	3
<i>Mb. mesophilicum</i>	NBRC 15688T	A4	0	0	0	0	3	0	0	0	1	2	2	2	2	2	2	0	2	3	3	2
<i>Mb. brachiatum</i>	NBRC 103629T	A4	0	0	0	2	3	0	0	0	3	3	3	3	3	3	3	0	0	3	3	3
<i>Mb. tardum</i>	NBRC 103632T	A4	0	0	0	2	3	0	0	0	3	3	0	0	0	0	3	0	2	3	0	0
<i>Mb. oryzae</i>	DSM 18207T	A4	3	0	0	2	3	0	0	0	3	3	3	3	3	3	3	0	3	3	2	2
<i>Mb. phyllosphaerae</i>	DSM 19779T	A4	0	0	0	2	3	0	0	0	2	3	2	2	3	3	3	0	3	3	2	2
<i>Mb. longum</i>	DSM 23933T	A4	0	0	0	0	3	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0
<i>Mb. fujisawaense</i>	DSM 5686T	A4	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	2	2
<i>Mb. pseudosasicola</i>	NBRC 105203T	A4	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	1	3	3	3
<i>Mb. phyllostachyos</i>	NBRC 105206T	A4	0	0	0	2	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mb. gregans</i>	NBRC 103626T	A5	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	0	3	3	2
<i>Mb. dankookense</i>	DSM 22415T	A5	3	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mb. hispanicum</i>	DSM 16372T	A5	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	0	3	3	3
<i>Mb. dichloromethanicum</i>	DSM 6343	B	0	0	0	0	3	0	0	0	1	1	1	1	1	1	1	0	1	1	3	0
<i>Mb. chloromethanicum</i>	NCIMB 13688	B	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	2	2
<i>Mr. zatmanii</i>	DSM 5688T	B	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mr. lusitanum</i>	NCIMB 13779T	B	2	0	0	0	3	0	0	0	3	3	2	2	2	2	3	0	3	3	1	2
<i>Mr. aminovorans</i>	NBRC 15686T	B	3	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	1
<i>Mr. extorquens</i>	NBRC 15687T	B	2	0	0	0	3	0	0	0	2	2	2	2	2	2	2	0	2	3	2	2
<i>Mr. suomiense</i>	DSM 14458T	B	0	0	0	0	1	0	0	0	3	3	3	3	3	3	3	0	3	3	3	2
<i>Mr. salsuginis</i>	NCCB 100140T	B	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mr. rhodinum</i>	DSM 2163T	B	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	1	3	3	3
<i>Mr. extorquens</i>	AM1 ATCC14718	B	0	0	0	0	3	0	0	0	1	1	1	1	0	1	1	0	2	3	3	2
<i>Mr. rhodesianum</i>	DSM 5687T	B	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mr. lusitanum</i>	DSM 14457	B	3	0	0	0	3	0	0	0	2	2	2	2	2	3	3	0	3	3	3	3
<i>Mr. podarium</i>	DSM 15083T	B	3	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mr. thiocyanatum</i>	JCM 10893T	B	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mb. organophilum</i>	NBRC 15689T	B2	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mb. gnaphalii</i>	DSM 24027T	B2	0	0	0	0	2	0	0	0	2	2	2	2	2	2	2	0	2	3	1	0
<i>Mb. brachytheticii</i>	DSM 24105T	B2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Mb. haplocladii</i>	DSM 24195T	B2	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
<i>Mb. thuringiense</i>	DSM 23674	B2	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Mb. oryzihabitans</i>	BCRC 81157T	C	0	0	0	0	2	0	0	0	3	3	3	3	3	3	3	0	3	2	3	1
<i>Mb. variabile</i>	DSM 16961T	C1	0	0	0	0	2	0	0	0	3	2	2	2	1	2	2	0	3	2	3	3
<i>Mb. platani</i>	JCM 14648T	C1	0	0	0	1	2	0	0	0	3	3	3	3	3	1	3	0	3	3	3	3
<i>Mb. aquaticum</i>	DSM 16371T	C1	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	0	1	3
<i>Mb. tarhaniae</i>	DSM 25844T	C1	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	3
<i>Mb. frigidaeris</i>	JCM 32048T	C1	0	0	0	0	3	0	0	0	3	3	3	3	1	2	3	0	3	3	3	3
<i>Mb. nonmethylophilum</i>	KCTC 42760T	C1	0	0	0	2	3	0	0	0	3	3	2	2	1	1	0	0	0	0	0	3
<i>Mb. terrae</i>	KCTC 52904T	C1	0	0	0	0	2	0	0	0	3	3	3	1	1	1	3	0	3	3	2	0
<i>Mb. crusticola</i>	KCTC 52305T	C1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Mb. indicum</i>	JCM 30761T	C1	0	0	0	2	3	0	0	0	3	3	3	3	3	3	3	0	3	3	3	0
<i>Mb. currus</i>	JCM 32670T	C1	0	0	0	1	3	0	0	0	3	3	3	2	2	2	3	0	3	3	3	3
<i>Mb. terricola</i>	NBRC 112874T	C1	0	0	0	0	3	0	0	0	3	3	3	3	3	3	3	0	3	2	3	3
<i>Mb. nodulans</i>	LMG 21967T	C2	3	0	0	0	3	0	0	0	1	2	1	1	1	1	1	0	2	0	0	1
<i>Mb. isbiliense</i>	DSM 17168T	C2	0	0	0	0	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Explanation of each test: NO₃, reduction of nitrates to nitrites; TRP, indole production; GLU, fermentation of glucose; ADH, arginine dihydrolase; URE, urease; ESC, β-glucosidase (esculin); GEL, protein hydrolysis (gelatin); PNPG, β-galactosidase (p-nitrophenyl-β-D-galactopyranoside); GLU, glucose; ARA, arabinose; MNE, mannose; MAN, mannitol; NAG, N-acetyl-glucosamine; MAL, maltose; GNT, gluconate; CAP, capric acid; ADI, adipic acid; MLT, malate; CIT, citrate; and PAC, phenylacetic acid.

The values represent the growth strength of the strain; 3, strong; 2, weak; 1, very weak; and -, not tested.

Figure 1

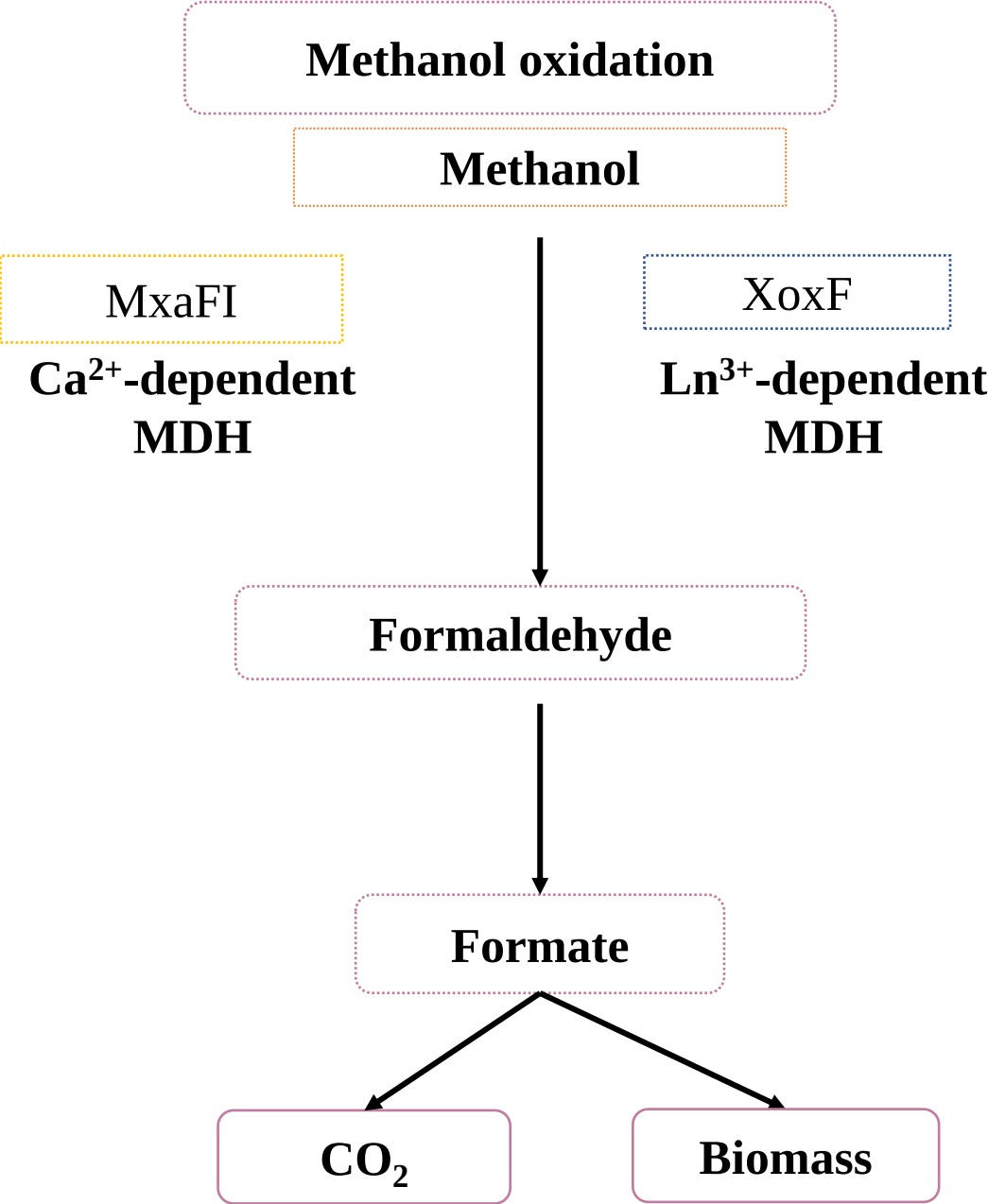


Figure 2

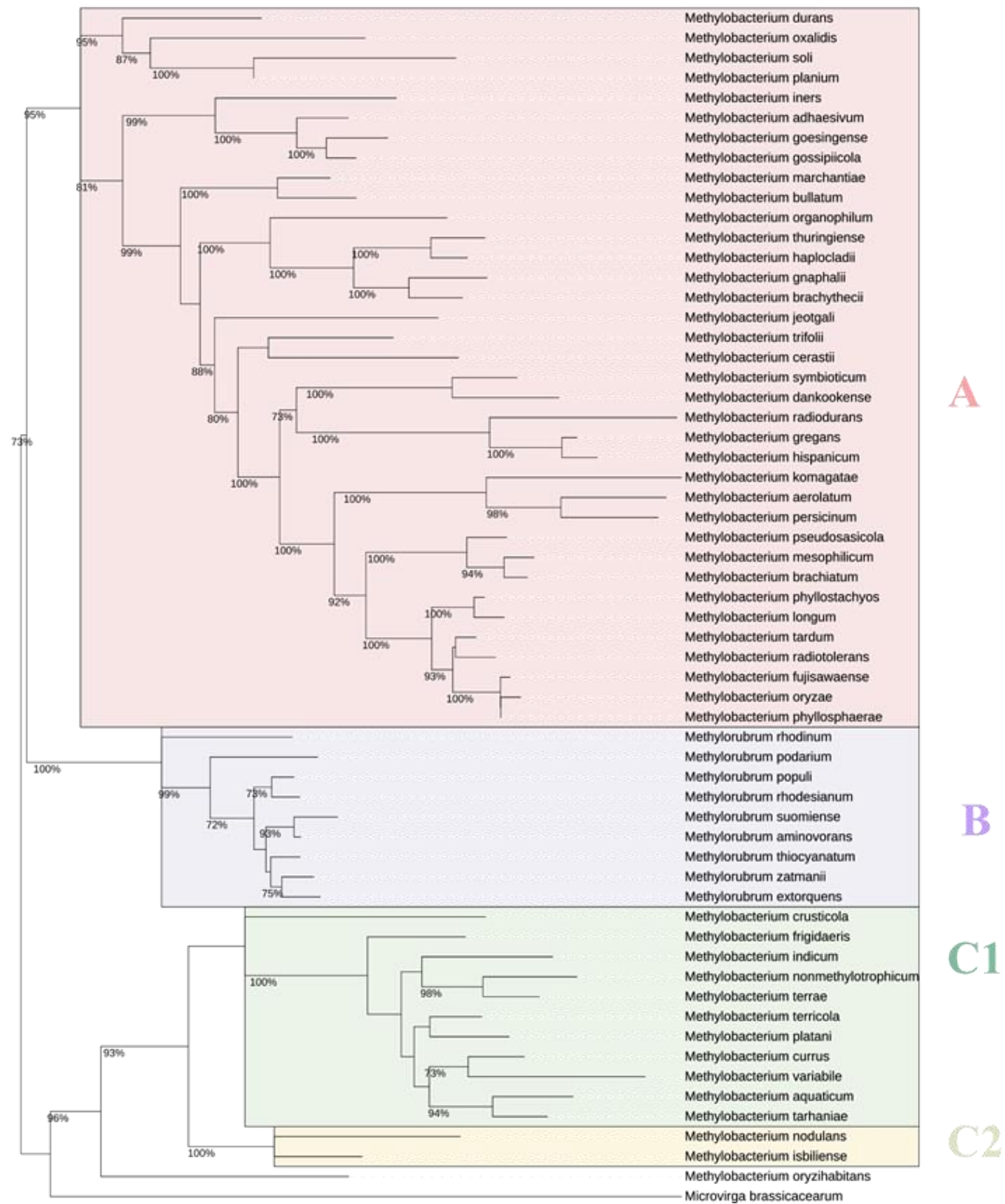


Figure 3

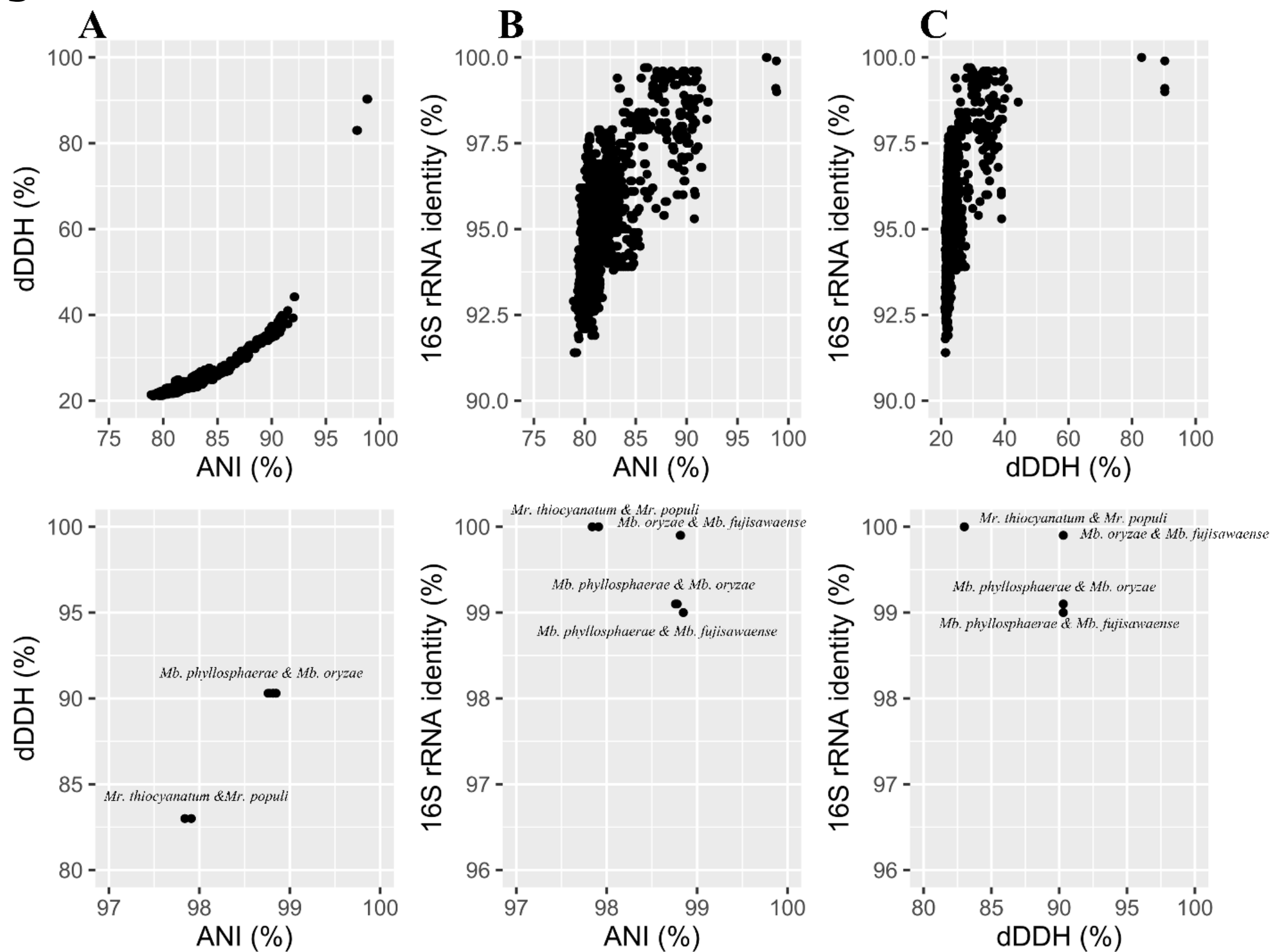


Figure 4

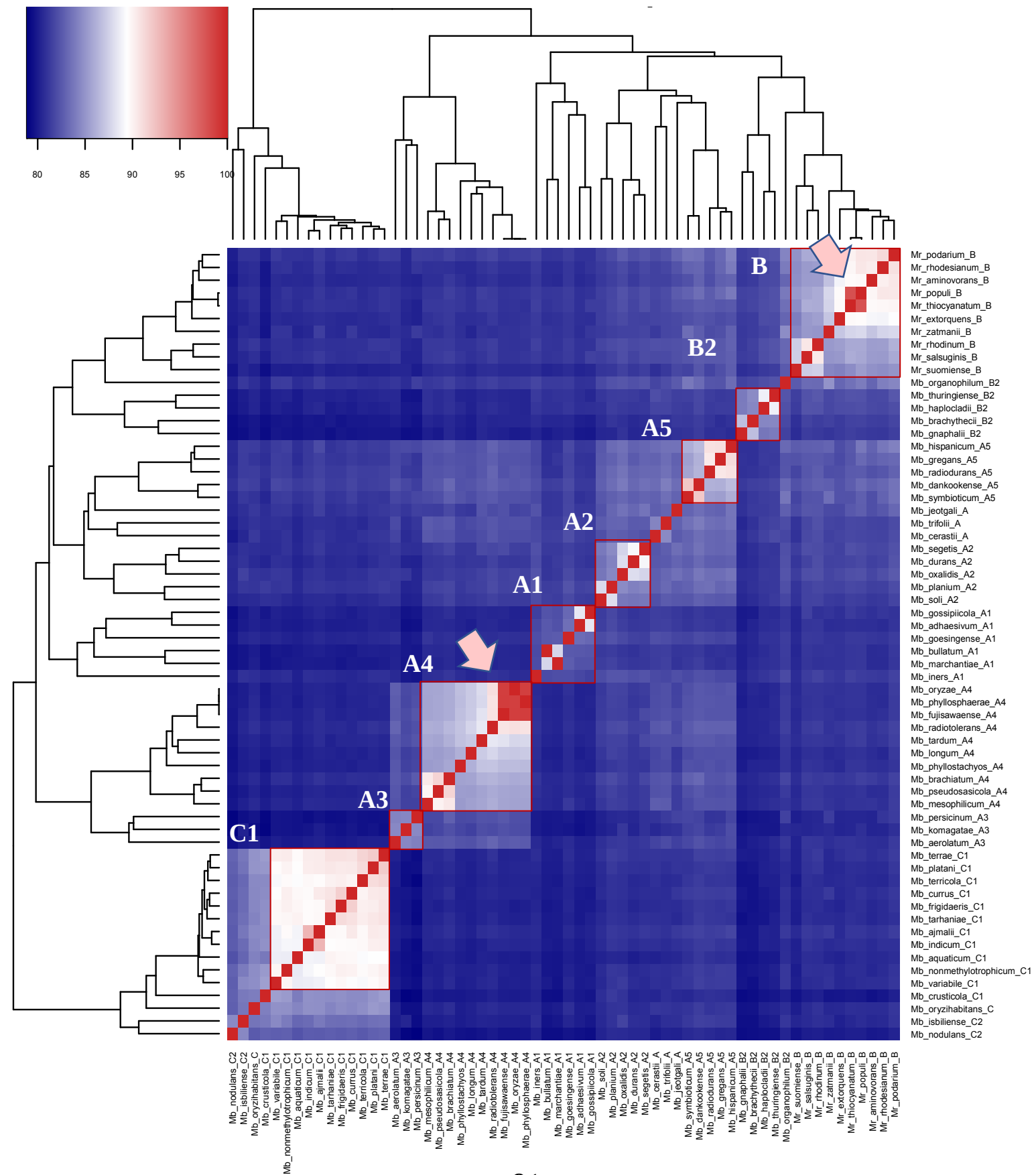


Figure 5

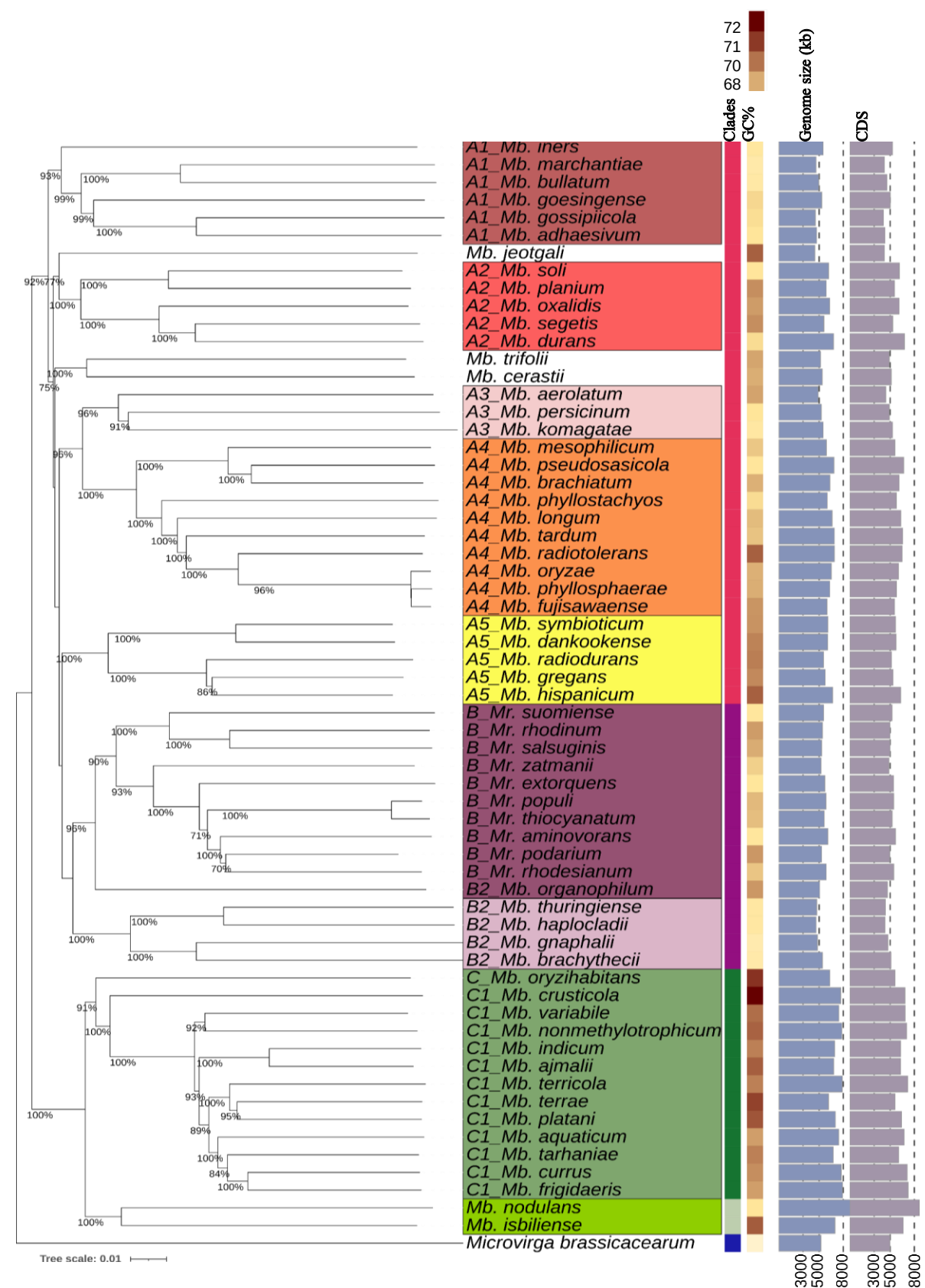


Figure 6

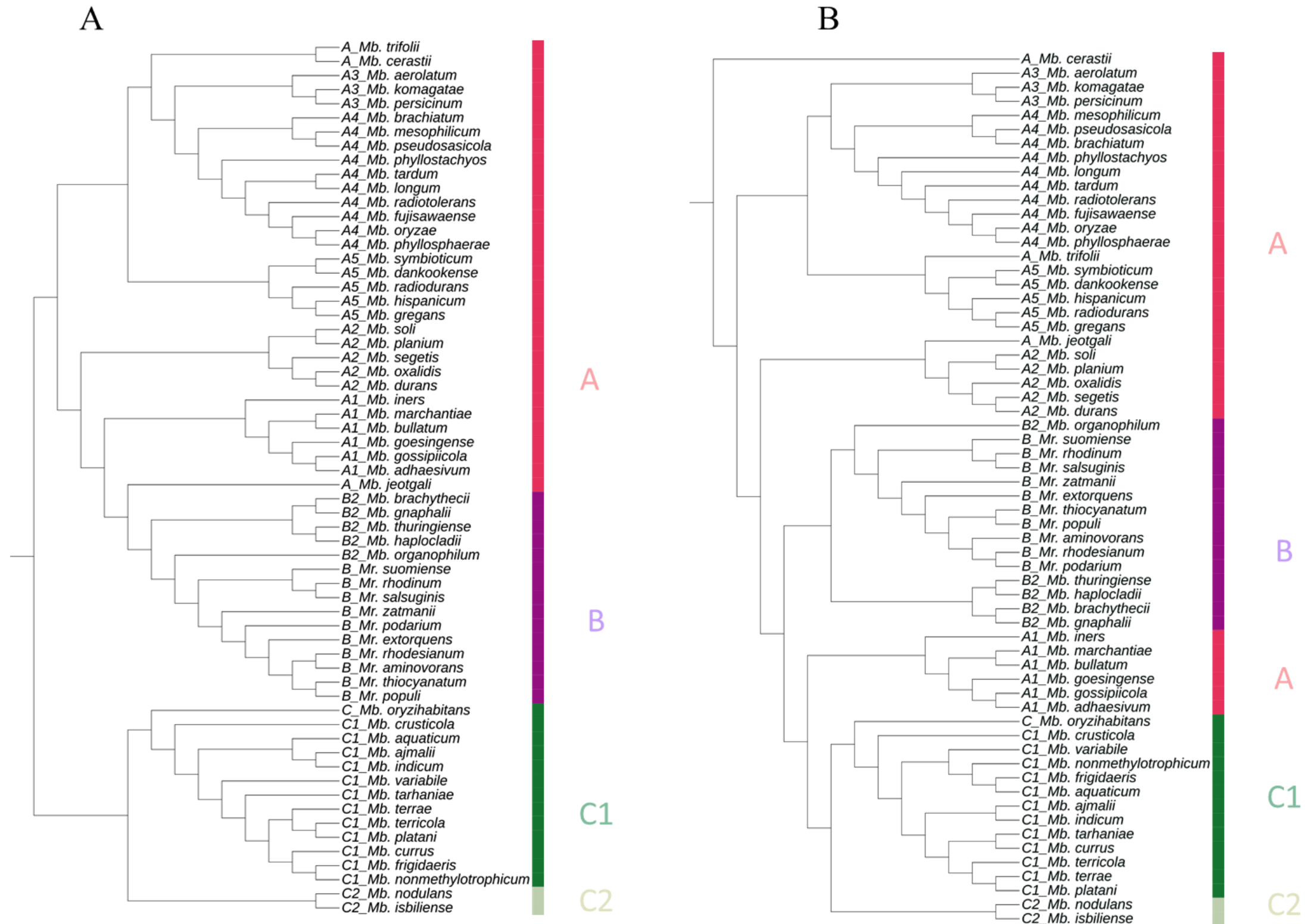


Figure 7

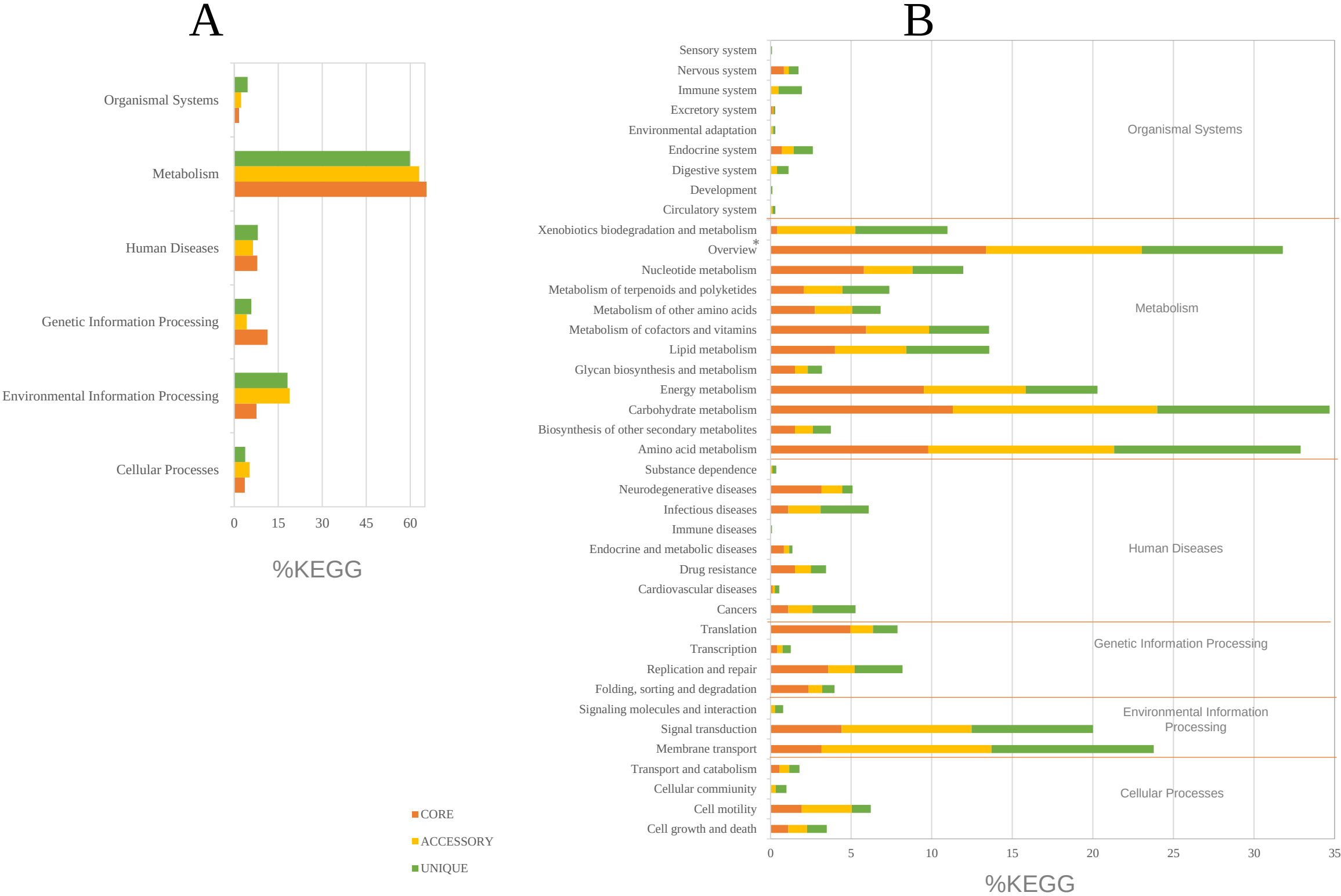


Figure 8

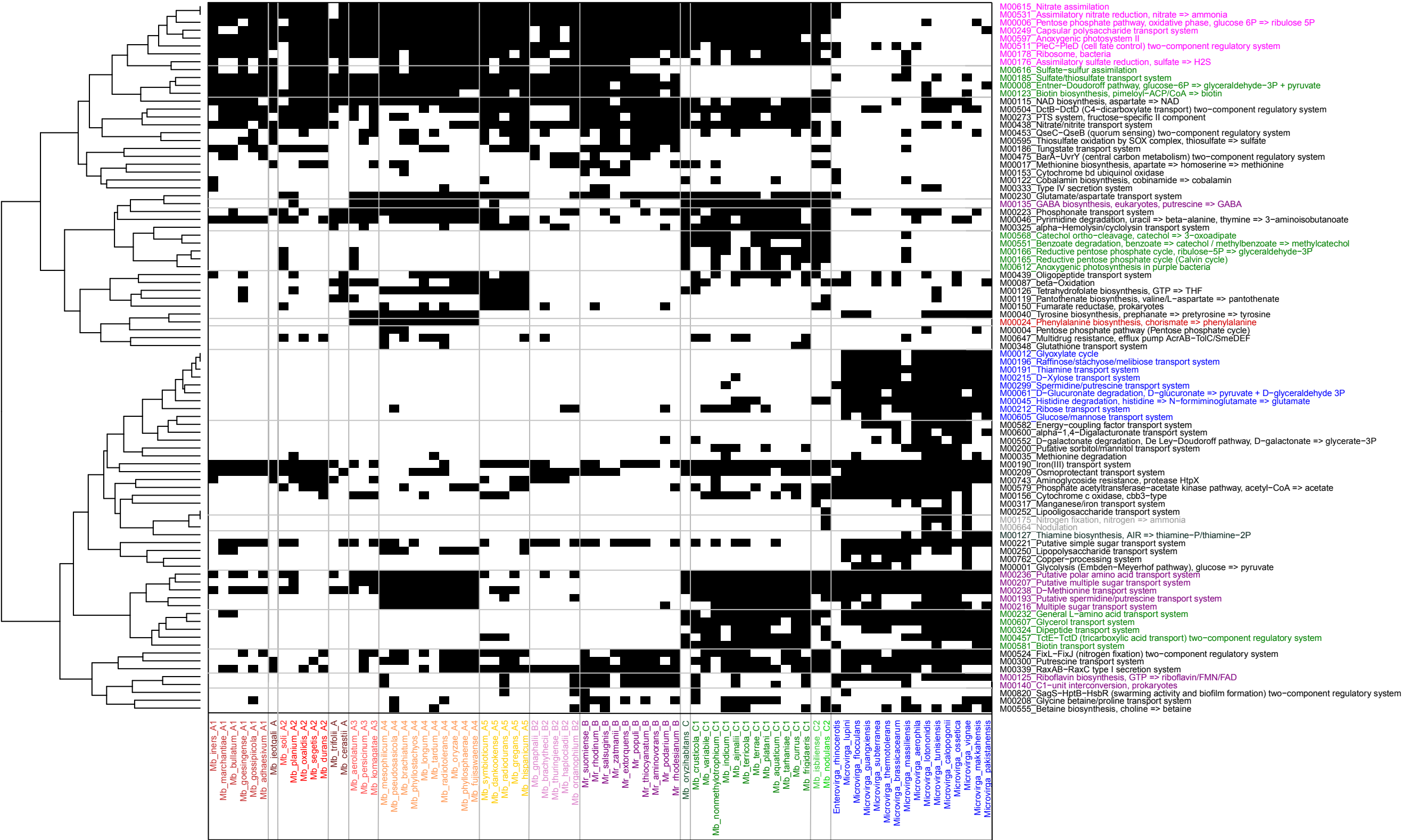


Figure 9

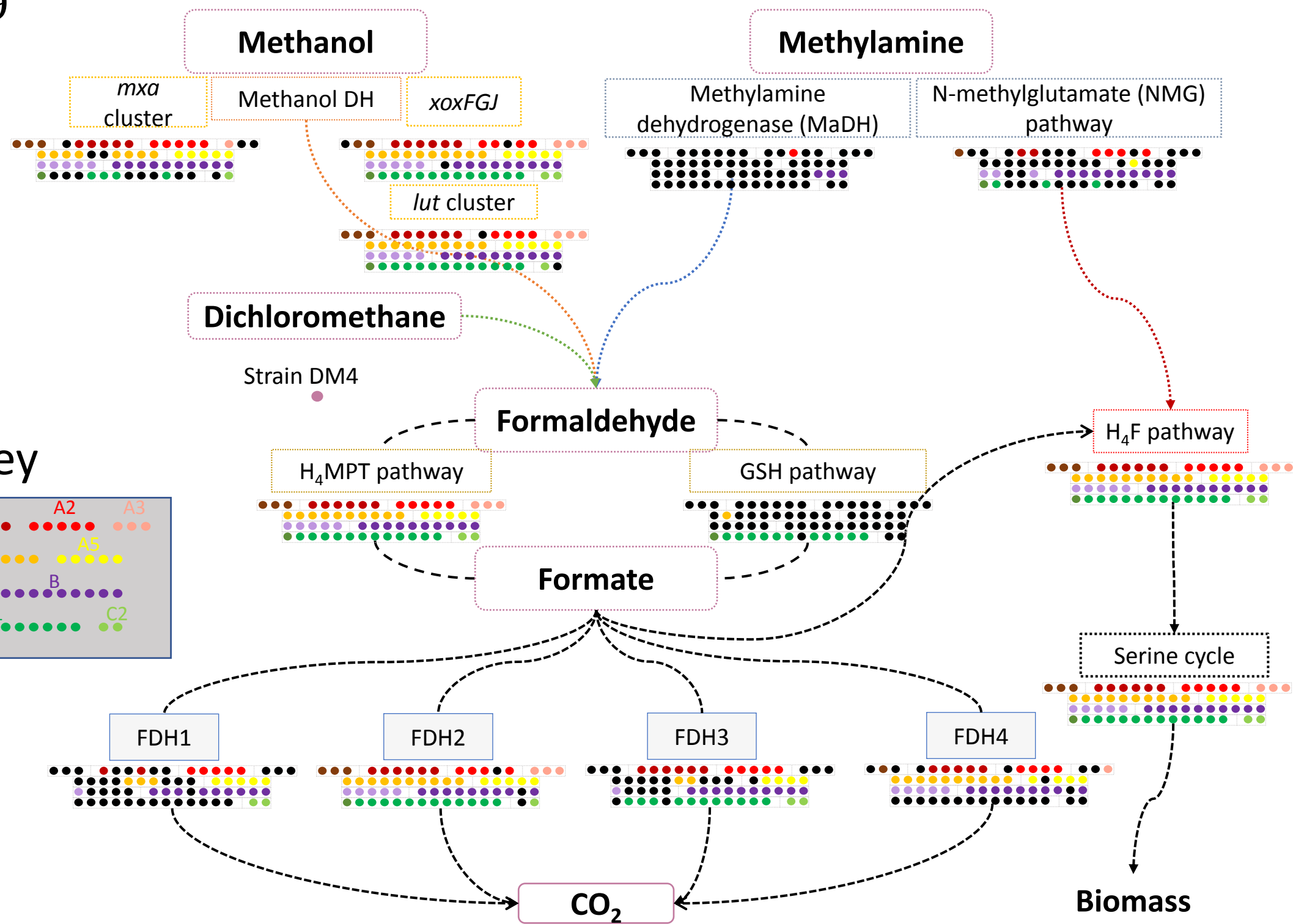


Figure 10

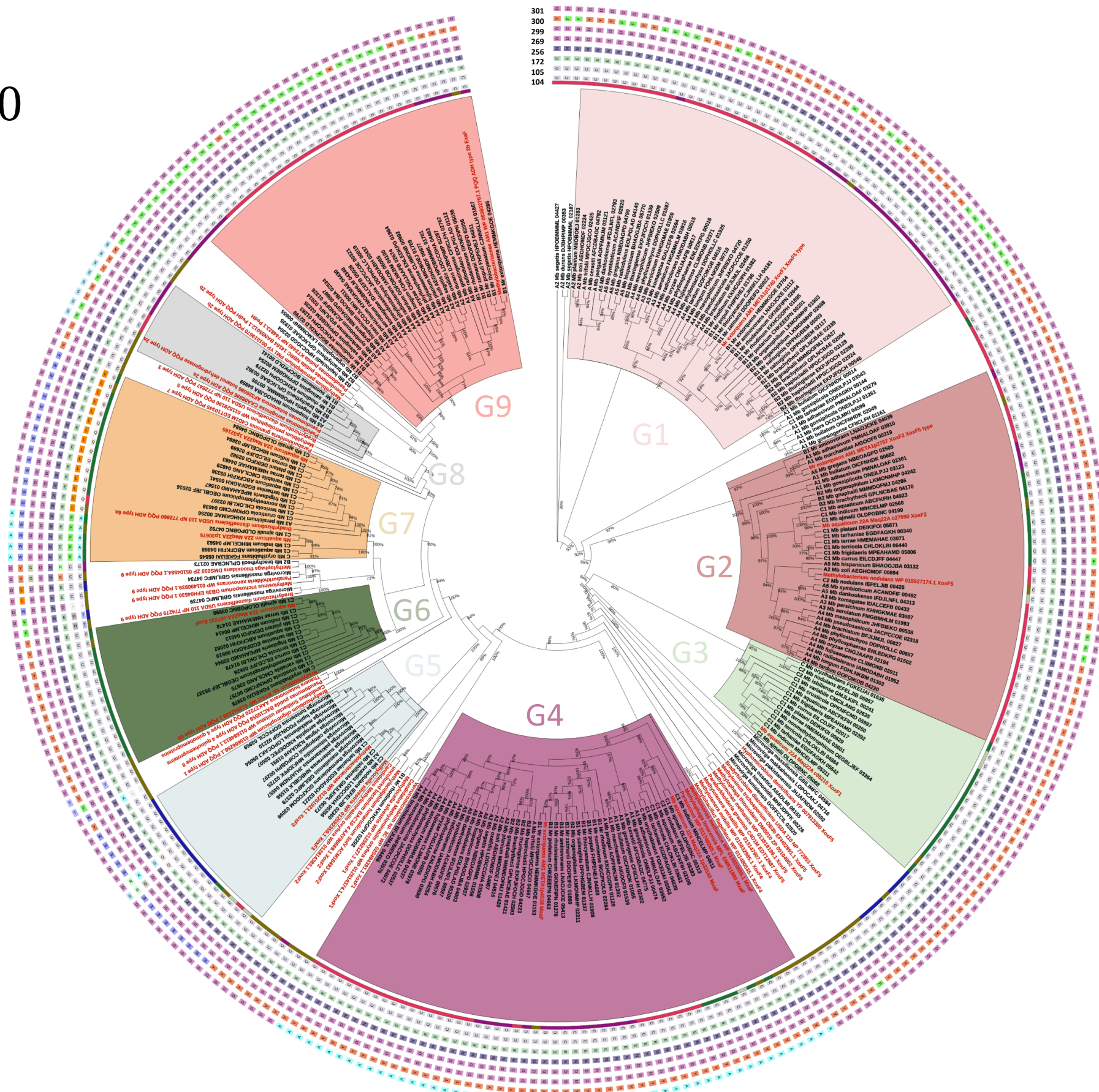
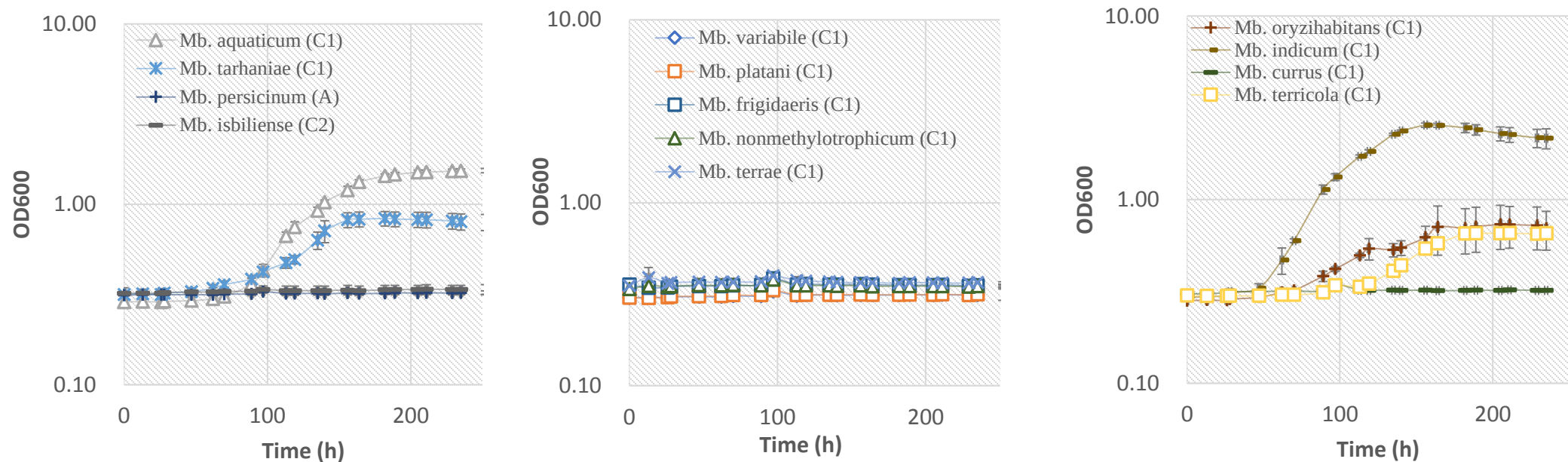


Figure 11

A



B

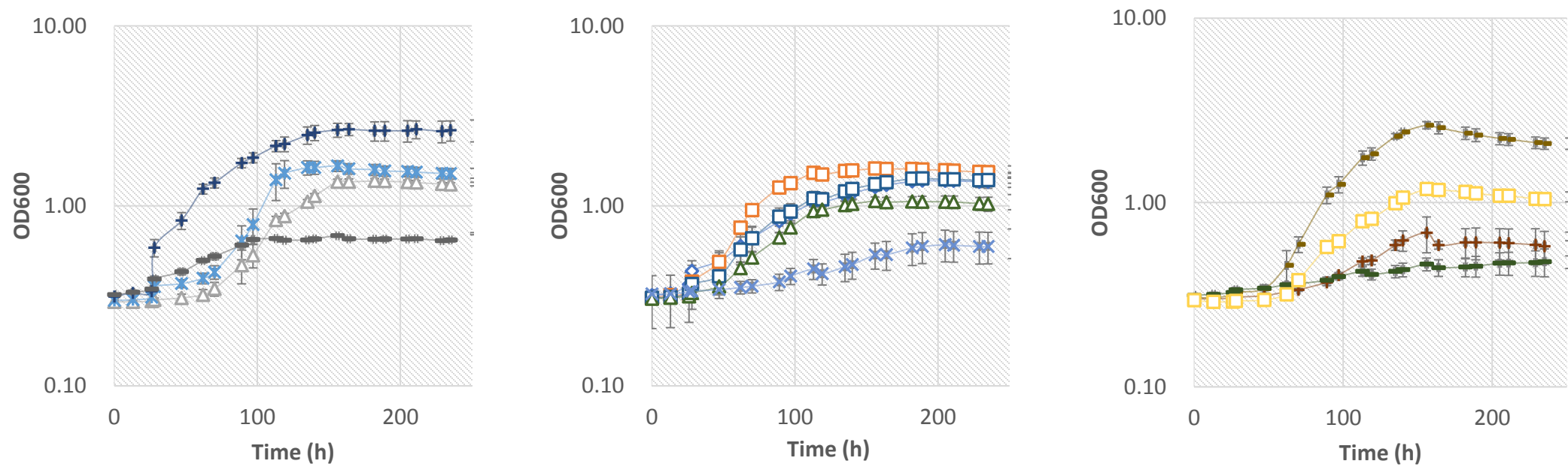
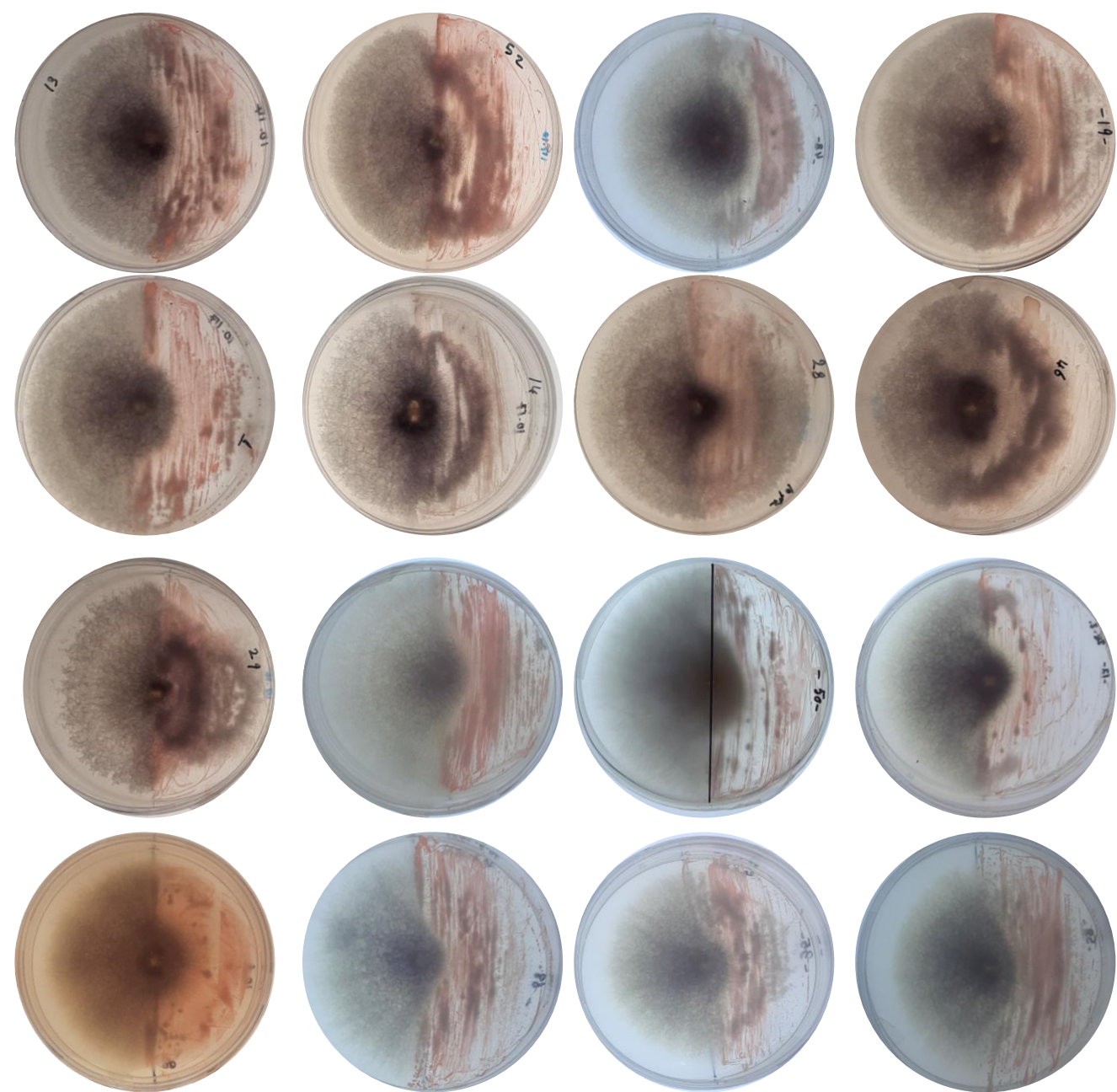


Figure 12



<i>Mb. komagatae</i> (T)	<i>Mb. aerolatum</i> (T)	<i>Mb. persicinum</i> (T)	<i>Mb. oryzae</i> (T)
<i>Mb. radiotolerans</i> (T)	<i>Mb. brachiatum</i> (T)	<i>Mb. longum</i> (T)	<i>Mb. phyllostachos</i> (T)
<i>Mb. fujisawaense</i> (T)	<i>Mb. phyllosphaerae</i> (T)	<i>Mb. hispanicum</i> (T)	<i>Mb. gregans</i> (T)
<i>Mr. rhodesianum</i> (T)	<i>Mb. terricola</i> (T)	<i>Mb. tarhaniae</i> (T)	<i>Mb. frigidaeris</i> (T)

Figure 13

