nature microbiology

Article

https://doi.org/10.1038/s41564-022-01252-3

Methanotrophy by a *Mycobacterium* species that dominates a cave microbial ecosystem

In the format provided by the authors and unedited

Supplementary information

Methanotrophy by a *Mycobacterium* species that dominates a cave microbial ecosystem

S1 Methane monooxygenase

The genes encoding sMMO in the M. MAG 1 draft genome are clustered (Supplementary Table 5, cluster 1, ORFs 4010–4016). Compared to the operons from *M. capsulatus* Bath and *M. trichosporium* OB3b [1, 2], which have the gene order *mmoXYBZDC*, the order in *Candidatus* M. methanotrophicum is slightly shuffled (*mmoXYBDCZ*, Extended Data Fig. 2a). The transcriptional activator of the *mmo* gene cluster in *M. capsulatus* Bath and *M. trichosporium* OB3b is MmoR [3]. Its activity is dependent on MmoG, which is a chaperonin. A homologue of the *mmoR* gene is present close to the *mmo* gene cluster in *Candidatus* M. methanotrophicum, but *mmoG* is absent from this region. In close vicinity of the *mmo* genes is a gene encoding phosphoenolpyruvate carboxykinase (ORF 4007), which may serve in gluconeogenesis. It catalyses the GTP-dependent conversion of oxaloacetate (OAA) to phosphoenolpyruvate (PEP) before the production of glucose from precursors derived from the citric acid cycle [4].

When plotted in a phylogenetic tree of distantly related monooxygenases, alkene monooxygenases, toluene monooxygenases, and phenol hydroxylases [5], the branch with MmoX of the *Mycobacteriaceae* is distinct from the one where MmoX proteins of the Alpha- and Gammaproteobacteria are positioned (Supplementary Figure 4). Also, members of the *Mycobacteriaceae* are much more dispersed as opposed to those of the Proteobacteria. Note that the mycobacterial sequences do not have a conserved Cys151 (relative to the sequence of *M. trichosporium*), but a Thr at that position (Supplementary Figure 5). Surprisingly, the tree includes two members of the Betaproteobacteria, HGW *Betaproteobacterium* and *Glaciimonas* sp., whereas it was thought previously that only species from the Alpha- and Gammaproteobacteria and Verrucomicrobia were able to grow methanotrophically [6]. This view may need to be revisited considering that the similarity of MmoX in Beta-, Alpha-, and Gammaproteobacteria is substantial, including the identity of the amino acids for ligation and oxygen activation (see main text, section Methane metabolism in *Candidatus* M. methanotrophicum).

Another interesting observation is the positioning of the Alphaproteobacterial MmoX from an *Oleomonas* species within the mycobacterial branch (Supplementary Figure 4). It suggests a more ancient exchange of genetic material between members of these phyla. An additional branch in the trees include MmoX proteins from *Solimonas aquatica*, *Brachymonas petroleovorans*, *Azoarcus* sp. and *Thauera butanivorans*, which seem to have Cys151 (relative to the sequence of *M. trichosporium*) replaced

by an alternative amino acid (Ser or Asn). Other members of this monooxygenase family, the more distantly related alkene monooxygenases, toluene monooxygenases, and phenol hydroxylases [5], are further away in the tree in separate branches. The conserved Leu110, Cys151 and Cys211 residues in MmoX from *M. trichosporium* and other species of the proteobacterial branch are different in the latter types of protein.

A bioinformatic study on bacterial multicomponent monooxygenases (BMMs) showed a strictly conserved glutamate residue in BMMs active on methane, with E240 in *M. capsulatus* Bath as a reference, as opposed to a fully conserved glutamine in all other members of this family [7]. Although MmoX from *Candidatus* M. methanotrophicum has a glutamine at this position, it shares the conserved isoleucine residue at position 217 in BMMs active on methane as opposed to a phenylalanine in all other BMMs. Both E240 and I217 are in a stretch of 37 consecutive amino acids in the central region of BMMs active on methane that is characterized by the presence of 33 positionally conserved amino acids [7]. The counterpart in *Candidatus* M. methanotrophicum shares 27 of these amino acids, indicating that MmoX of *Candidatus* M. methanotrophicum resembles the MmoX of BMMs active on methane in that respect.

Using AlphaFold, we generated template-free structural predictions of sMMO subunits from *Candidatus* M. methanotrophicum. Their analysis showed that, despite showing only 23% sequence similarity and no clear evolutionary links, ORF 4012 of *Candidatus* M. methanotrophicum was predicted to encode a protein with a virtually identical fold (pLLDT 92.7/100 and RMSD 2.3 Å – below 1 Å for >90% of residues) to MmoD proteins associated with sMMOs (e.g., 6d7k Protein Data Bank) [8]. In the co-crystal structure of sMMO-MmoD from *Methylosinus trichosporium*, only the core part of MmoD (residues 12–75 out of 111) is structured, whereas the remaining residues, which are less conserved, are disordered. The MmoD from *Candidatus* M. methanotrophicum is considerably shorter (62 residues) and only contains the core domain and not the unstructured N- and C-termini. Moreover, AlphaFold Multimer predicts that this MmoD binds to a similar site on MmoX as previously determined for co-crystals of other MmoX-MmoD structures [8].

S2 Alcohol dehydrogenase

We hypothesize that alcohol dehydrogenase D (AdhD) in *Candidatus* M. methanotrophicum is functional as a methanol dehydrogenase since it is the second most abundant protein in the biofilm while a more canonical MxaF-type methanol dehydrogenase was not identified (Table 1). The phylogenetic tree of AdhD from *Candidatus* M. methanotrophicum along with AdhD from close and more distinct members includes Mno from *M. smegmatis* (Supplementary Figure 8). Mno was shown to be responsible for methanol oxidation by *M. smegmatis* [9], however it only shares 15% identical residues with AdhD from *Candidatus* M. methanotrophicum.

ORF 2892 encoding AdhD in *Candidatus* M. methanotrophicum is in a gene cluster with so-called *mft* genes encoding proteins proposed to be responsible for the co-factor termed mycofactocin. In the gene map of the *mft* gene cluster, the cluster from *Candidatus* M. methanotrophicum has *mftABCD* genes, while the clusters from other members of the *Mycobacteriaceae* contain additional *mftEF* genes (Extended Data Fig. 3). The counterparts of these genes are found elsewhere in the genome of *Candidatus* M. methanotrophicum, namely in a *mftRAD'D'EF* gene cluster (ORFs 3482–3477), where *mftBC* genes are missing, *mftD* is split in 5' and 3' parts, and the regulator *mftR* is located upstream of *mftA*. MftA from *Candidatus* M. methanotrophicum shares the positionally conserved amino acid residues found in MftAs from other species with this *mft* gene cluster (Supplementary Figure 9). More downstream of *mftF* is ORF 3471, which is highly similar to the *adh* genes of M. MAG 2 and *M. tuberculosis* H37Rv (Extended Data Fig. 3). We hypothesize that the *mftEF* genes in this cluster complement the *mftABCD* genes located in the cluster with *adhD*.

S3 Formaldehyde and formate dehydrogenases

Many methylotrophic organisms, such as *M. extorquens* AM1, make use of tetrahydromethanopterin and/or tetrahydrofolate for one-carbon transfer reactions and formaldehyde oxidation [10, 11]. The key enzymes of the formaldehyde oxidation pathway that involves tetrahydromethanopterin include formylmethanofuran dehydrogenase, formylmethanofurane tetrahydromethanopterin formyltransferase, methenyltetrahydromethanopterin cyclohydrolase, methylenetetrahydrofolate dehydrogenase, and 5,6,7,8-tetrahydromethanopterin hydro-lyase. According to the KEGG database, genes encoding these proteins

are also found in *M. capsulatus* Bath and *M. trichosporium* OB3b, but they are all absent in *Candidatus* M. methanotrophicum. The gene encoding formate dehydrogenase (FDH) is clustered (Supplementary Table 5, cluster 3, ORFs 1085–1090) with those encoding a biotin synthase, an ATP-dependent dethiobiotin synthetase, 8-amino-7-oxononanoate synthase, and adenosylmethionine-8-amino-7-oxononanoate aminotransferase, all of which participate in biotin formation [12]. Biotin is an essential co-factor of pyruvate carboxylases [13]. We hypothesize that expression of FDH and biotin formation are coordinated to activate pyruvate carboxylase for central metabolism in *Candidatus* M. methanotrophicum, thereby balancing catabolic and anabolic fluxes.

S4 Ribulose monophosphate pathway for formaldehyde incorporation

Key enzymes of the ribulose monophosphate (RuMP) pathway for the incorporation of formaldehyde into central metabolism in *Candidatus* M. methanotrophicum include (see Supplementary Table 5): hexulose-6-phosphate isomerase (H6PI, ORF 837), glucose-6-phosphate isomerase (GPI, two copies, ORFs 378 and 4307), and glucose-6-phosphate dehydrogenase (G6PDH, four copies, ORFs 214, 636, 2183 and 4311). ORFs 837 and 838 encoding H6PI and H6PS, respectively, are in gene cluster 4. Genes encoding additional copies of GPI, 6PGDH and G6PDH are clustered as well (cluster 5, ORFs 4307, 4308 and 4311, respectively). ORFs 2182 and 2183 encoding 6PGDH and G6PDH, respectively, are in gene cluster 7. A further description of clusters 4, 5 and 7 is in the parts on the pentose phosphate pathway and glyoxylate shunt. The genes for H6PI and H6PS were also found in a facultative methylotrophic *Mycobacterium gastri* MB19 [14]. A BLAST search revealed that these two genes are relatively rare within the *Mycobacteriaceae*, although they are present in *M. brisbanense*, *M. dioxanotrophicus*, *M. aquaticum*, *M. aubagnense*.

S5 Pentose-P-pathway and gluconeogenesis

6-phosphogluconolactonase (6-PGL, ORF 0184) is essential in the oxidative phase of the pentose-P pathway as it catalyses the hydrolysis of 6-phosphogluconolactone to 6-phosphogluconic acid. Interestingly, in the vicinity of ORF 0184 are the genes encoding proteins involved in central metabolism, i.e., PEP carboxylase (ORF 0182), phosphoglycerate kinase (ORF 0173), glyceraldehyde-3-phosphate dehydrogenase (ORF 0172), and triosephosphate isomerase (ORF 4304). The first three are also among

the most abundant proteins in the biofilm (Table 1). Genes encoding important enzymes of the RuMP pathway (GPI, 6PGDH and G6PDH, ORFs 4307, 4308 and 4311) are in cluster 5 with a suite of genes encoding enzymes of the pentose-P pathway and one gene for gluconeogenesis. Genes encoding the enzymes in the pentose-P pathway are also in this region, with ORFs 4305 and 4306 encoding second copies of transketolase and transaldolase, respectively, ORF 4309 encoding a second 6-PGL, and ORF 4310 encoding a ribose-5-phosphate isomerase (R5PI). The latter catalyses the reversible isomerization of ribose-5-P (R5P) into ribulose-5-P (Ru5P), which serves as a formaldehyde acceptor in the RuMP pathway. Cluster 4 has three genes of the RuMP pathway encoding G6PDH, H6PI and H6PS, and two genes encoding a third set of transketolase and transaldolase (ORFs 843 and 844, respectively). However, cluster 4 seems to be less transcribed than cluster 6 (ORFs 0170 to 0188) as none of the proteins encoded by the genes in this cluster is abundant in the biofilm.

S6 Hydrogenases

Genes encoding a putative O_2-tolerant uptake [Ni-Fe]-hydrogenases are in the vicinity of the sMMO genes of *Candidatus* M. methanotrophicum in cluster 1. Hydrogenases-3 are membrane-associated respiratory hydrogenases with nickel and iron-sulphur clusters. The small subunit of these enzymes (HycG) contains three iron-sulphur clusters that are used to transfer electrons to the large subunit (HycE). The latter contains a nickel-iron centre, which is the active site for the oxidation of H_2 into two protons and two electrons [15, 16]. The enzyme has two additional hydrophilic subunits (HycB and HycF), both of which are iron-sulphur proteins. They are believed to guide the electrons to low-potential multiheme cytochromes [15]. It has also two inner membrane subunits, HycC and HycD [17, 18, 19]. *Candidatus* M. methanotrophicum has ORFs 4005 and 4000 encoding the large and small subunits of such a Ni-Fe type-3 hydrogenase, respectively, ORFs 4001 and 4004 encoding the accompanying HycB and HycF subunits, respectively, and ORFs 4002 and 4003 encoding the two inner membrane subunits, HycC and HycD, respectively [17, 18, 19].

Candidatus M. methanotrophicum has quite a few different types of hydrogenase that may serve in additional energy transducing systems, perhaps similar as in members of Verrucomicrobia growing mixotrophically on both CH_4 and H_2 with the concomitant increase of their growth rate as compared to

growth solely on CH_4 [20]. Verrucomicrobia-dominated soil communities from an acidic geothermal field in Rotokawa, New Zealand, rapidly oxidized CH_4 and H_2 simultaneously [20]. Apparently, H_2 oxidation is particularly important for the adaptation to CH_4 limitation. H_2 oxidation may well be a general metabolic strategy during mixotrophic growth as genes encoding the subunits of several types of hydrogenases were also identified in publicly available genomes of other methanotrophs.

S7 Respiration and ATP synthesis

The subunits of ATP synthase are also among the most abundant proteins in the biofilm (Table 1). This finding emphasizes their importance in energy metabolism of *Candidatus* M. methanotrophicum. The membrane potential is built up by the respiratory enzymes that couple electron transfer to proton translocation from the inside to the outside of the membrane. These respiratory enzymes include (the corresponding ORFs encoding them in *Candidatus* M. methanotrophicum are given in parentheses): a proton translocating NADH dehydrogenase (ORFs 2955–2968), the cytochrome bc_1-complex (ORFs 2856–2858), cytochrome *c* oxidase subunit I (the gene has four copies: ORFs 870, 1303, 1446, and 3787), the remaining subunits of the oxidase (in a cluster with ORFs 2852–2859), and a succinate dehydrogenase (ORFs 503–506). Notably, the genes encoding the cytochrome bc_1-complex are sandwiched between those encoding subunits of cytochrome *c* oxidase. This arrangement has been noted in other *Mycobacteriaceae* as well, apparently to coregulate the expression of the subunits that make up a super complex of the cytochrome bc_1-complex and the aa_3-type oxidase [21]. Apart from the cytochrome *c* oxidase, *Candidatus* M. methanotrophicum has the potential to make a *bd*-type quinol oxidase (from ORFs 885–887), which has a high affinity for O_2. The *bd*-type quinol oxidase is therefore believed to be operational under conditions of low O_2. In pathogenic species, it seems to be a survival oxidase [22]. Genes encoding the subunits of ATPase are clustered in cluster 8 (ORFs 1779–1786), while the genes encoding the subunits of the cytochrome bc_1-complex and those encoding subunits of cytochrome *c* oxidase are in gene cluster 9 (Supplementary Table 5). Although one may expect that the rate of CH_3OH oxidation is lower with AdhD as compared to MxaF-type enzymes, the efficiency of AdhD concerning energy transduction is likely higher as it makes use of NAD^+ as an electron acceptor rather than PQQ in MxaF-type enzymes. This difference in efficiency is because oxidation of reduced PQQ contributes with 2 charge separations

per electron transferred to O_2, while it is 5 when NADH is oxidized by NADH dehydrogenase via the cytochrome bc_1-complex and the aa_3-type oxidase. Hence, this observation suggests a trade-off between the rate and efficiency of the CH_3OH oxidation step. The stoichiometry is different when the *bd*-type quinol oxidase is active in respiration rather than the aa_3-type cytochrome *c* oxidase, as in the former case the total charge separation per electron transferred from NADH to O_2 would be 3 rather than 5. This difference is because the latter pathway bypasses the electrogenic cytochrome bc_1-complex on the one hand, while cytochrome *bd* is half as efficient in energy transduction as compared to the aa_3-type cytochrome *c* oxidase on the other hand [22].

S8 Gene clustering of key metabolic genes

Most of the genes encoding key enzymes for CH_4 oxidation for energy transduction and pentose-P pathway, RuMP pathway, gluconeogenesis, and the glyoxylate shunt for carbon metabolism are grouped in mosaic-like gene clusters with different combinations of genes encoding enzymes for the different pathways. These clusters are scattered throughout the genome of *Candidatus* M. methanotrophicum. The clustering suggests coordinated expression of the enzymes in all crucial pathways of central metabolism of *Candidatus* M. methanotrophicum, perhaps to tune their concentration and activity towards one another thereby preventing the accumulation of toxic intermediates (e.g., formaldehyde) on the one hand, while yielding relatively high fluxes on the other hand.

S9 Methylotrophy and MxaF in *Mycobacteriaceae*

A number of different mycobacteria are capable of growing on methanol, including *Mycobacterium* strain JC1, *M. flavescens*, *M. gastri*, *M. neoaurum*, *M. parafortuitum*, *M. peregrinum*, *M. phlei*, *M. smegmatis*, *M. vaccae*, *M. petroleophilum* (previously called *M. cuneatum*), and *Mycobacterium* ID-Y [23, 24, 25]. Seeing this variety, we expect many more mycobacteria to be capable of methylotrophy if they are studied for this trait. Methylotrophy is best described for *Mycobacterium* strain JC1, which employs a ribulose-1,5-bisphosphate carboxylase for CO_2 fixation [23, 24]. *M. gastri* uses the RuMP pathway for formaldehyde fixation and shares the possession of homologues of hexulose-6-phosphate synthetase and hexulose-6-phosphate isomerase with *Candidatus* M. methanotrophicum [26]. *M. smegmatis* uses a dedicated type of p-nitroso-N,N-dimethylaniline (NDMA)-dependent alcohol dehydrogenase, called

Mno, to grow on methanol [9]. This enzyme is, however, only 15% identical to AdhD from *Candidatus* M. methanotrophicum.

Most methanotrophic species or CH_3OH oxidizers use a PQQ-dependent MxaF-like methanol dehydrogenase [27, 28], but *Candidatus* M. methanotrophicum does not have *mxaF*-like genes, nor genes for PQQ biosynthesis. BLAST searches using the MxaF, PqqB, and PqqC protein sequences of *Methylorubrum extorquens* AM1 yielded 7 hits in the family of *Mycobacteriacaea* that have homologues of these proteins, including *Mycolicibacterium goodie*, *Mycolicibacterium smegmatis*, and *Mycolicibacterium dioxanotrophicus*, which are all mycolicibacteria. To our knowledge, there are no studies reported on the involvement of MxaF in methylotrophic growth in mycobacteria. *M. smegmatis* has a gene encoding a copper-containing protein upstream of the *mxaF* gene. Its *pqqABCDE* genes are downstream of the *mxaF* gene and in that order. In many methylotrophs, *mxaI* and *mxaG* genes encode the small subunit of MxaF and the electron-accepting cytochrome *c*, respectively [27, 28]. But a BLAST search with these two genes from *Methylorubrum extorquens* AM1 [27] did not reveal sequence homology in any of the *Mycobacteriaceae*.

S10 Sulphur metabolism

The list of abundant proteins in the biofilm (Table 1) contains a ferredoxin-dependent sulphite reductase. Its allocated gene, called *sirA*, is adjacent to *cysH*, encoding a phosphoadenosine phosphosulfate reductase, and *sirB* encoding a sirohydrochlorin ferrochelatase. Such a gene cluster is quite common within the *Mycobacteriaceae* and the proteins encoded by this gene cluster are involved in the reductive incorporation of sulphur from sulphate into O-acetyl-l-serine to yield cysteine [29]. The genome of *Candidatus* M. methanotrophicum does not contain genes encoding the key enzymes for the microbial oxidation of sulphur either via the sulfide:quinone oxidoreductase pathway [30, 31] or the Sox pathway [32, 33].

Supplementary Table 1. Basic statistics of the crude mycobacterial metagenome-assembled genomes (MAGs) from the Sulfur Cave biofilm and of the complete genome sequence obtained from the cultured *Candidatus* M. methanotrophicum. Relative abundance is calculated based on MAGs DNA read coverage using the metagenome of the biofilm sample. NA, not applicable.

	Refined M. MAG 1	Refined M. MAG 2	Refined M. MAG 3	*Candidatus* M. methanotrophicum
Assembly Length (bp)	4,409,390	4,592,045	5,112,202	5,002,371
Contigs	169	197	467	2
N50 (bp)	65,873	155,066	37,850	4,719,641
Completeness (%)	99.17	98.42	93.56	99.62
Contamination (%)	1.15	1.32	3.02	1.61
tRNA	54	52	62	55
rDNA	3	2	0	6
CDS	4,576	4,239	5,275	5,194
Relative abundance (%)	18.68	2.22	0.74	NA

Supplementary Table 2. Sequence of the 5' 16S rRNA gene fragment in *Candidatus* M. methanotrophicum from position 157 to 224, including the unique indel at positions 180 to 199 (underlined): CATGCATGTGCTGTGGTGGAAAGTGCGGTGTGTAATAGCACTGTGCGGTGT-GGGATGGGCCCGCGGC. A megablast BLASTN search yielded a total of 52 similar sequences from diverse locations. N is the number of similar sequences at the published location. D1: clone hbdenovo2332; D2: isolate MOTU28, clone HIGH12_BAC3F8; D3: clone L3::G9RA0RH03FR3GD; D4: clone EV21.

Description	Max Score	Total Score	Query Cover	E value	Identity	Accession	Location	N
D1	108	108	100%	3.00E-20	95.59%	MH104493.1	Hawaii basalts, Hawaii, USA [34]	2
D2	108	108	100%	3.00E-20	95.52%	HF952435.1	Soils with elevated CO_2, Slovenia [35]	1
D3	104	104	97%	4.00E-19	95.38%	JX739216.1	Lava Beds National Monument, California, USA [36]	47
D4	97.1	97.1	95%	7.00E-17	93.85%	AF379035.1	Corroded concrete sewer system, Oostende, Belgium [37]	2

Supplementary Table 3. Thirty most abundant proteins in a laboratory-grown enrichment culture of *Candidatus* M. methanotrophicum, as determined by tandem mass spectrometry. Only the proteins from *Candidatus* M. methanotrophicum are shown.

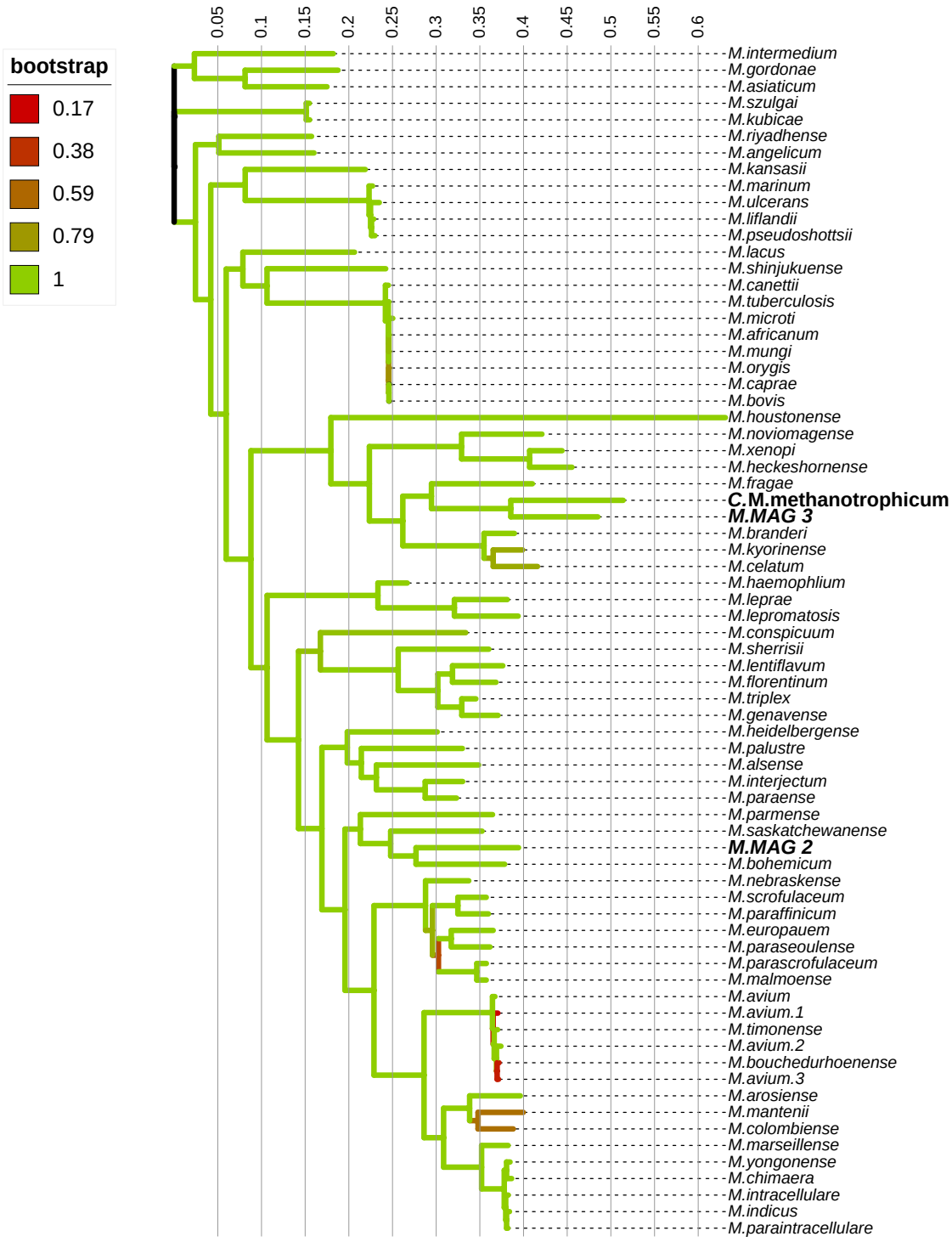
Protein IDs	Functional annotation	Unique peptides	Sequence coverage [%]	Mol. weight [kDa]	LFQ intensity [log10]	MS/MS count
WP_219067859.1	chaperonin GroEL	66	76.2	56.959	10.76	2739
WP_219067899.1	NDMA-dependent alcohol dehydrogenase AdhD	29	76.2	39.973	10.10	815
WP_219066027.1	type VII secretion protein EccA	52	43.4	143.88	9.77	641
WP_219067482.1	chaperonin GroEL	38	60.3	56.101	9.88	631
WP_219069711.1	MmoZ	22	87.6	21.354	9.72	480
WP_219069719.1	MmoX	22	72.1	60.428	9.39	419
WP_219068480.1	catalase/peroxidase HPI	30	49	81.34	9.82	368
WP_219066869.1	transketolase	17	60.6	75.576	9.39	357
WP_219066029.1	type VII secretionprotein EccB	17	33	54.419	9.28	311
WP_219067853.1	mycofactocin biosynthesis FMN-dependent deaminase MftD	16	56.8	41.106	9.45	302
WP_219066870.1	transaldolase	19	61.6	40.007	9.27	295
WP_219069311.1	DoxX family membrane protein	26	56.8	31.213	9.56	280
WP_219066387.1	3-hexulose-6-phosphate synthase	11	73.9	20.888	9.66	261
WP_219069713.1	MmoC	12	45.5	38.4	9.29	242
WP_219069217.1	30S ribosomal protein S7	8	60.9	17.581	9.17	231
WP_219067481.1	co-chaperone GroES	8	92	10.788	9.31	225
WP_219067194.1	HU family DNA-binding protein	10	45.2	20.649	9.14	220
WP_219069717.1	MmoY	16	59.2	44.79	9.23	216
WP_219069221.1	elongation factor Tu	14	60.9	43.678	9.14	212
WP_219066031.1	type VII secretion AAA-ATPase EccA	15	26.9	66.724	9.38	210
WP_219066135.1	NAD-dependent formate dehydrogenase	16	52.3	42.089	9.41	203
WP_219065959.1	molecular chaperone DnaK	17	46.1	66.249	9.12	174
WP_219070925.1	DUF5666 domain-containing protein	11	48.7	22.223	9.19	169
WP_219065672.1	heparin-binding hemagglutinin	12	44.8	20.8	9.37	156
WP_219066077.1	fumarate reductase/succinate dehydrogenase flavoprotein subunit	14	42.4	69.984	8.79	131
WP_219067040.1	AAA family ATPase	15	33.9	93.73	8.98	131
WP_219066881.1	typeI glyceraldehyde-3-phosphate dehydrogenase	9	52.9	35.863	8.71	106
WP_219069715.1	MmoB	5	44	17.499	8.62	102
WP_219066421.1	pyruvate kinase	10	49.6	50.515	8.37	70
WP_219069685.1	nitrite/sulfite reductase	10	55.8	62.396	8.58	50

Supplementary Table 4. Characteristics of samples collected from sites on the Puturosu Mountain close to and in Sulfur Cave. Samples A were collected on the way from the hotel Bálványos (46.11753° N, 25.94374° E) to Sulfur Cave (46.12027° N, 25.94872° E), samples B were collected at a site close to the hotel Apor Baths Bálványos (46.11524° N, 25.94958° E). **Abbreviations:** Gas, volcanic gas flowing through the sampling site; pH, pH measured with a pH paper; DNA, DNA concentration after extraction; *C.* M. meth %, percentage of total OTUs present in the sample belonging to *Candidatus* M. methanotrophicum, as judged by the identity of these OTUs to the V3–V4 region of the 16S rRNA gene of *Candidatus* M. methanotrophicum; ND, not determined.

Sample	Location	Gas	Type	pH	DNA (ng/μl)	*C.* M. meth. %
A1	Hill top	yes	soil	< 2	0.8	79.2
A2	Hill left A	yes	soil	< 2	5.2	96.0
A3	Hill left B	yes	soil	< 2	5.3	94.0
A4	Hill right B	yes	soil	< 2	0.7	81.0
A5	Hill control	no	soil	6	0.2	3.7
A6	Hill control	no	soil	5	0.5	7.6
A7	Small Cave A	yes	wall interface	< 1	1.5	98.5
A8	Small Cave B	yes	wall interface	< 2	0.3	78.4
A9	Alum Cave A	yes	wall black spot	ND	10.6	20.8
A10	Alum Cave B	yes	wall black spot	ND	1.0	47.2
A11	Sulfur Cave A	yes	wall grey interface	< 1	1.2	98.6
A12	Sulfur Cave B	yes	wall white interface	< 1	0.4	98.1
A13	Sulfur Cave control	no	wall entrance bench	ND	0.2	6.6
B1	Red pond downhill	yes	red biofilm	5	0.6	0.1
B2	Red pond downhill	no	red biofilm	5	1.1	0.3
B3	White pond downhill	yes	white biofilm	5	1.0	0.7
B4	White pond downhill	yes	white biofilm	5	0.2	0.8
B6	Pond uphill right	yes	waterish, muddy	< 2	0.1	95.2
B8	Pond uphill left	yes	waterish, muddy	< 2	0.13	95.1

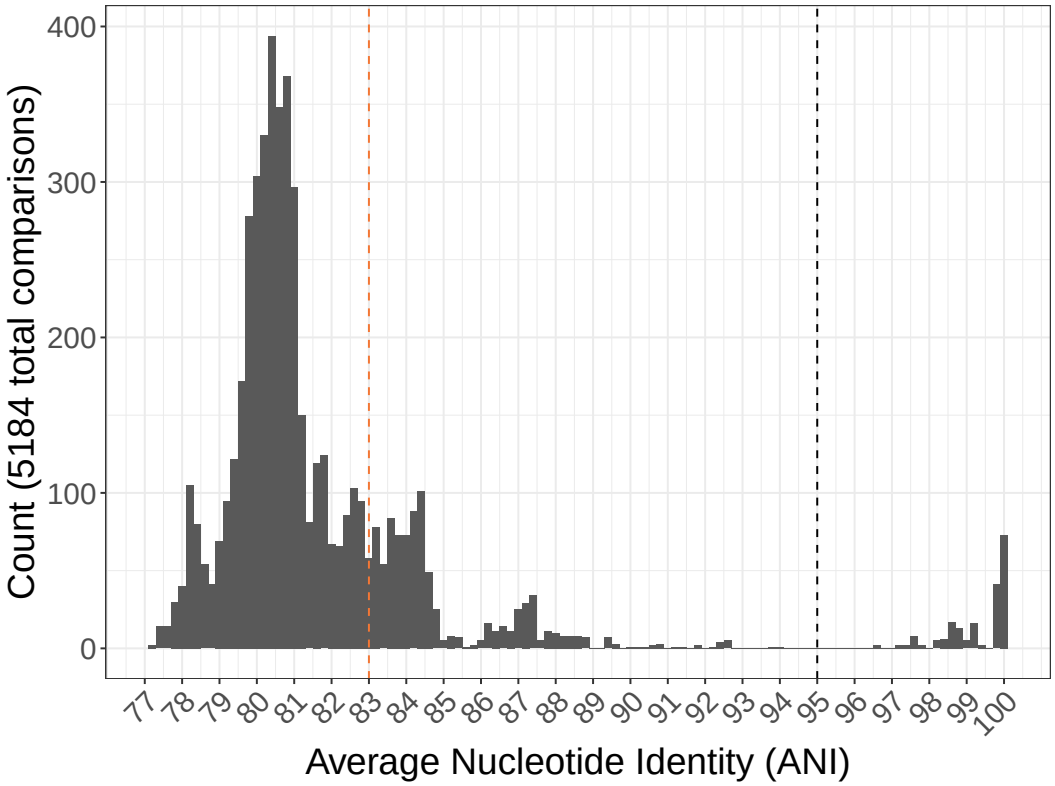
Supplementary Table 5. Gene clusters discussed in the Supplementary Material. Shown are the ORF numbers, the pathway or the enzymes, and additional ORFs encoding enzymes involved in the RuMP cycle. **Abbreviations:** sMMO, soluble methane monooxygenase; ADH, NDMA-dependent alcohol dehydrogenase; FyDH, formaldehyde dehydrogenase; FDH, formate dehydrogenase; H6PI, hexulose-6-phosphate isomerase; GPI, glucose phosphate isomerase; G6PDH, glucose-6-phosphate dehydrogenase; 6PGDH, 6-phosphogluconate dehydrogenase; H6PS, hexulose-6-phosphate synthetase; Hyd/Hyc, hydrogenase; RuMP, ribulose-5-P pathway; Pent-P, pentose-P-pathway.

Cluster	ORFs in the cluster	Pathway or enzyme	RuMP enzymes
1	4010–4016	sMMO+Hyd+PEP kinase	
2	2892–2905	ADH	
3	1085–1090	FDH	
4	831–853	RuMP	H6PS, H6PI
5	4305–4311	RuMP	6PGDH, GPI, G6PDH
6	170–188	Pent-P	
7	2182–2221	Glyoxylate, RuMP, Hyc	6PGDH (2), G6PDH
8	1782–1785	ATPase	
9	2852–2859	Cytochromes bc_1 and aa_3	
Additional ORFs	ORFs position	Pathway	RuMP enzymes
1	214	RuMP	G6PDH
2	378	RuMP	GPI
3	636	RuMP	G6PDH

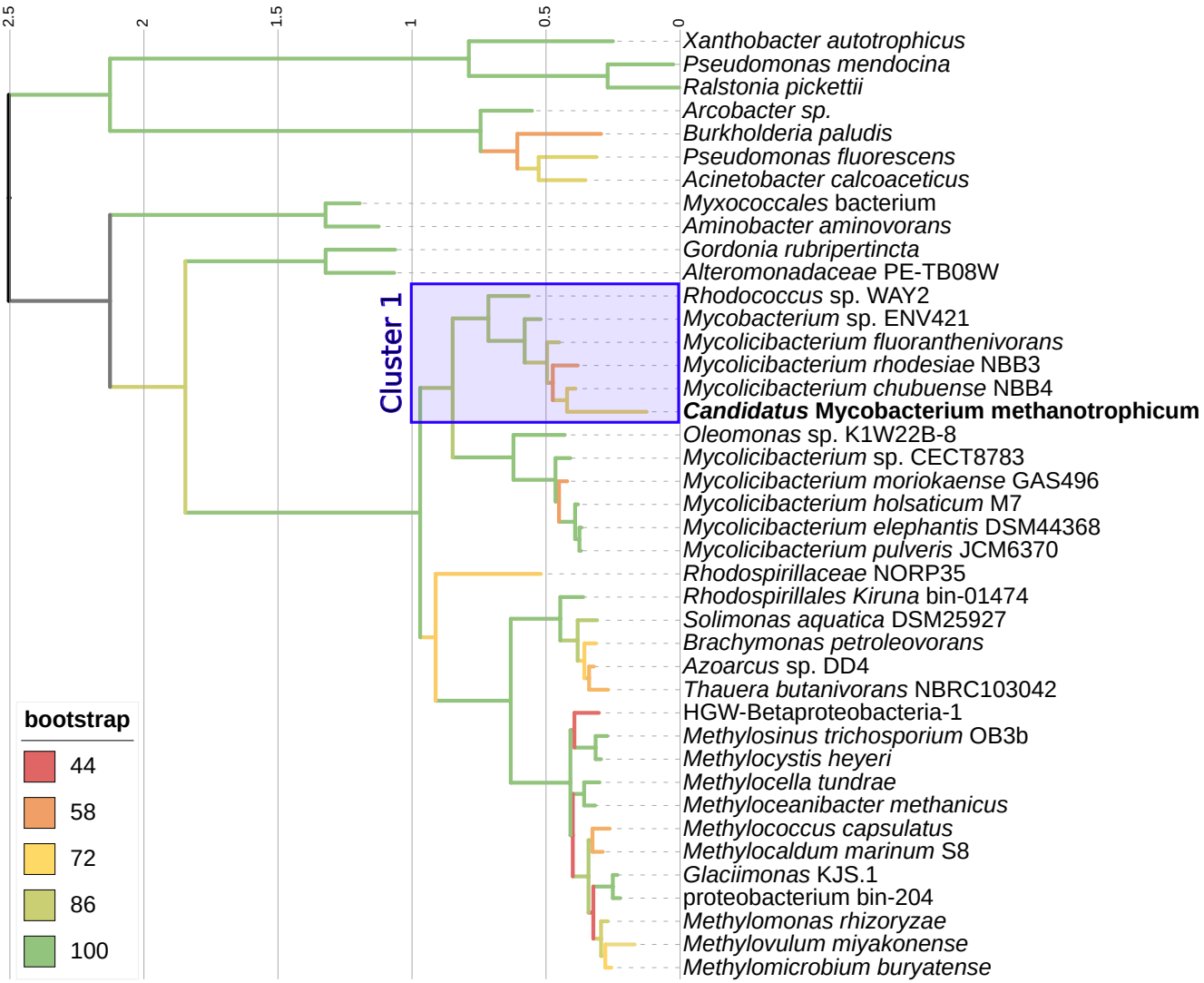


Supplementary Figure 1. Phylogenetic tree of the genus *Mycobacterium*. The tree excludes the fast-growing mycobacteria of the other four genera of the *Mycobacteriaceae*. The tree was constructed using 316 monocore gene markers as described in Methods. Bootstrap values are displayed with the color scale from red to green.

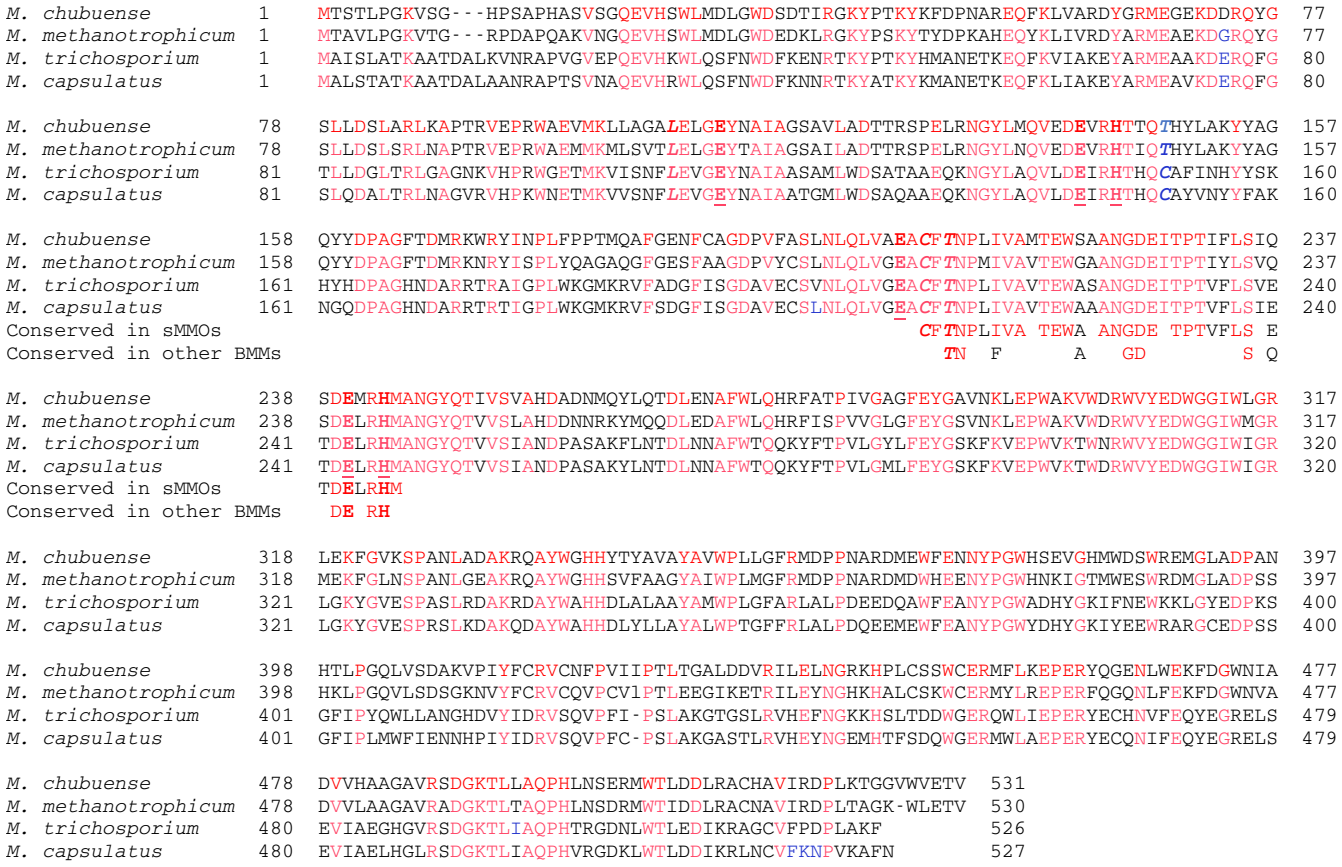
Supplementary Figure 2. Comparison of phylogenetic trees of the genus *Mycobacterium*. The trees were constructed using 27 (left) and 316 (right) monocore gene markers. The tree labels and tips of *Candidatus* M. methanotrophicum (M. MAG 1), M. MAG 2, and M. MAG 3 are coloured in green, orange, and blue, respectively. The entanglement coefficient between the trees is 0.23 (1 corresponds to a full entanglement, whereas a low value corresponds to a good tree alignment). The entanglement lines connect each of the mycobacterial species in the two trees.



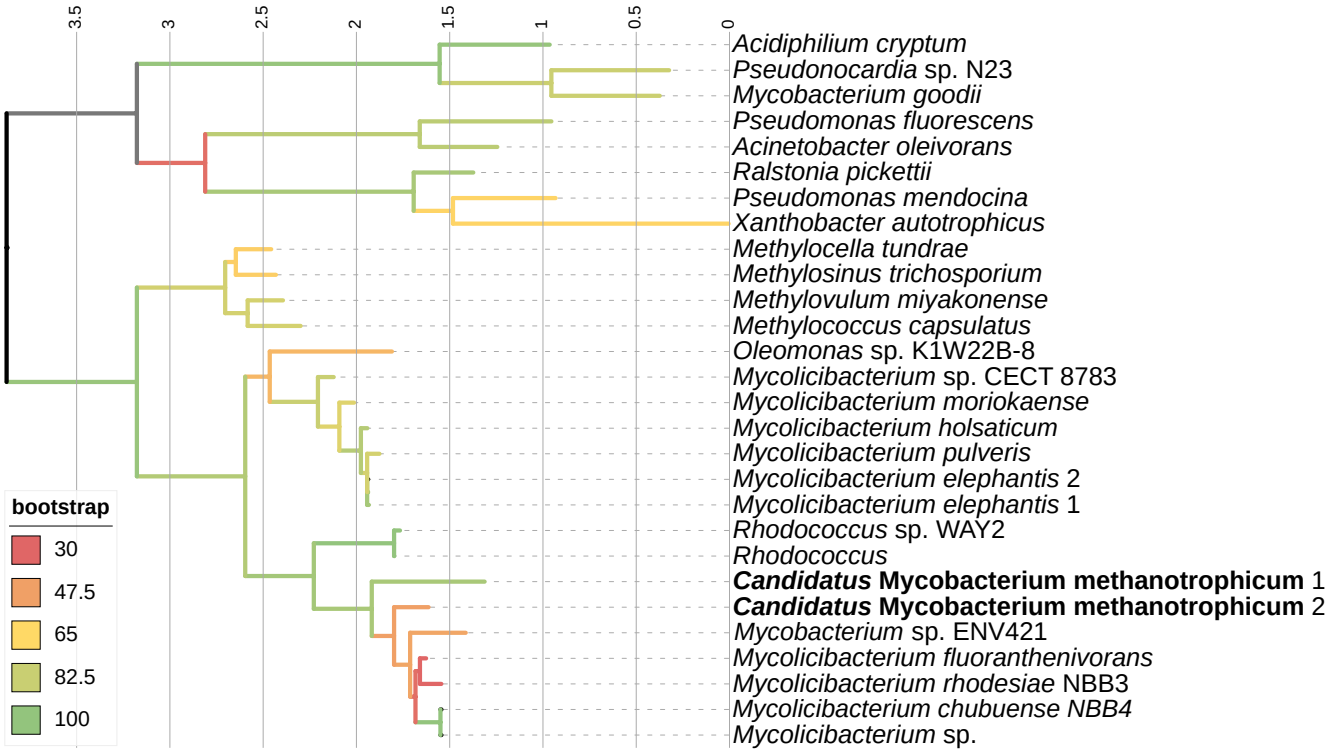
Supplementary Figure 3. Histogram of 5184 pairwise ANI comparisons of 72 publicly available *Mycobacterium* genomes, including the three novel genomes from this study. Dashed lines indicate the intra-species (black, $> 95\%$) and inter-species (orange, $< 83\%$) boundaries according to the ANI values [38].



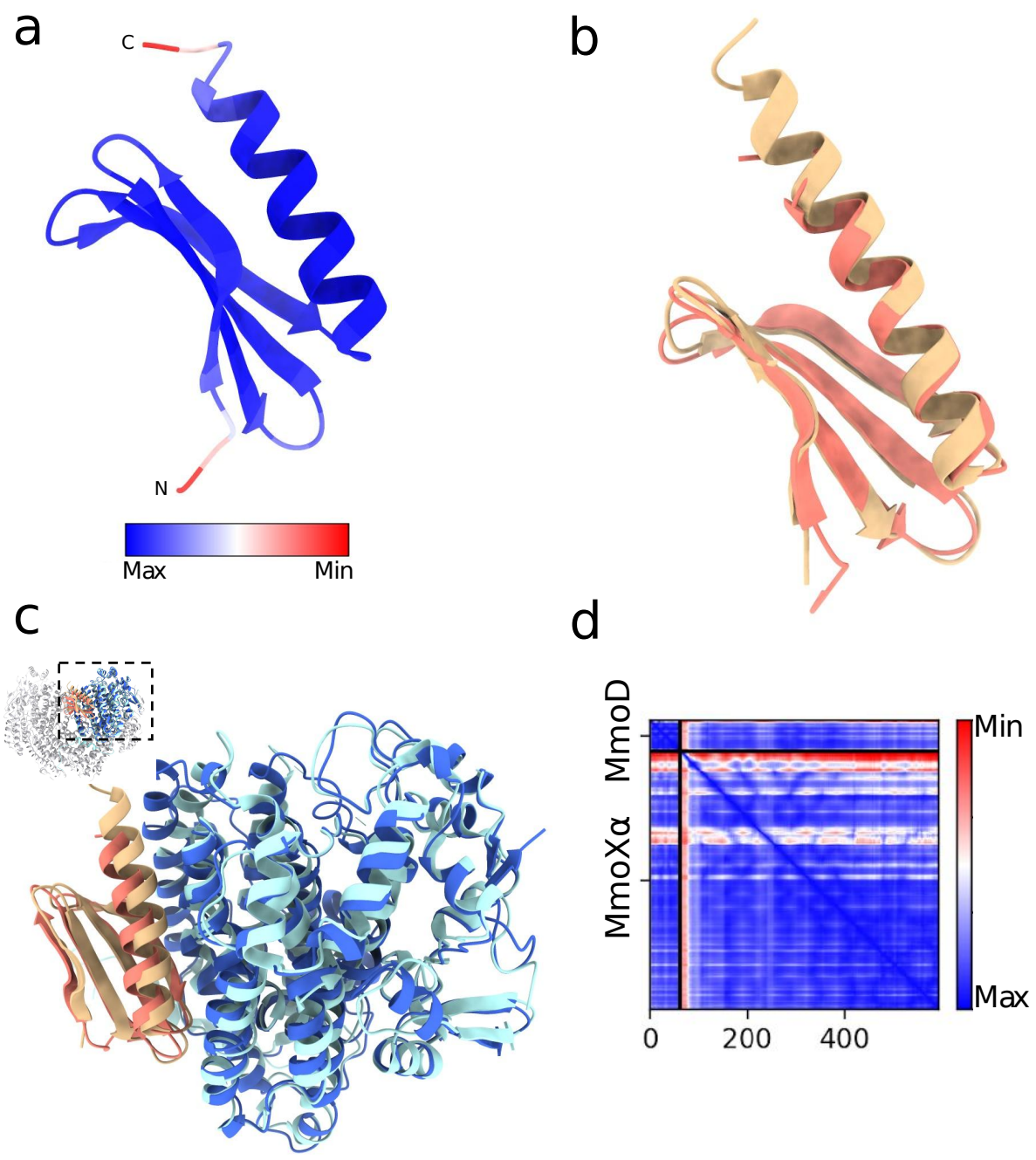
Supplementary Figure 4. Phylogenetic tree of the α subunit of methane monooxygenases (MmoX), including those from *Candidatus* M. methanotrophicum and close homologues. Grey colored branches are not supported by bootstrap values. Cluster 1 (in blue) marks the species with the closest homologues of MmoX from *Candidatus* M. methanotrophicum.



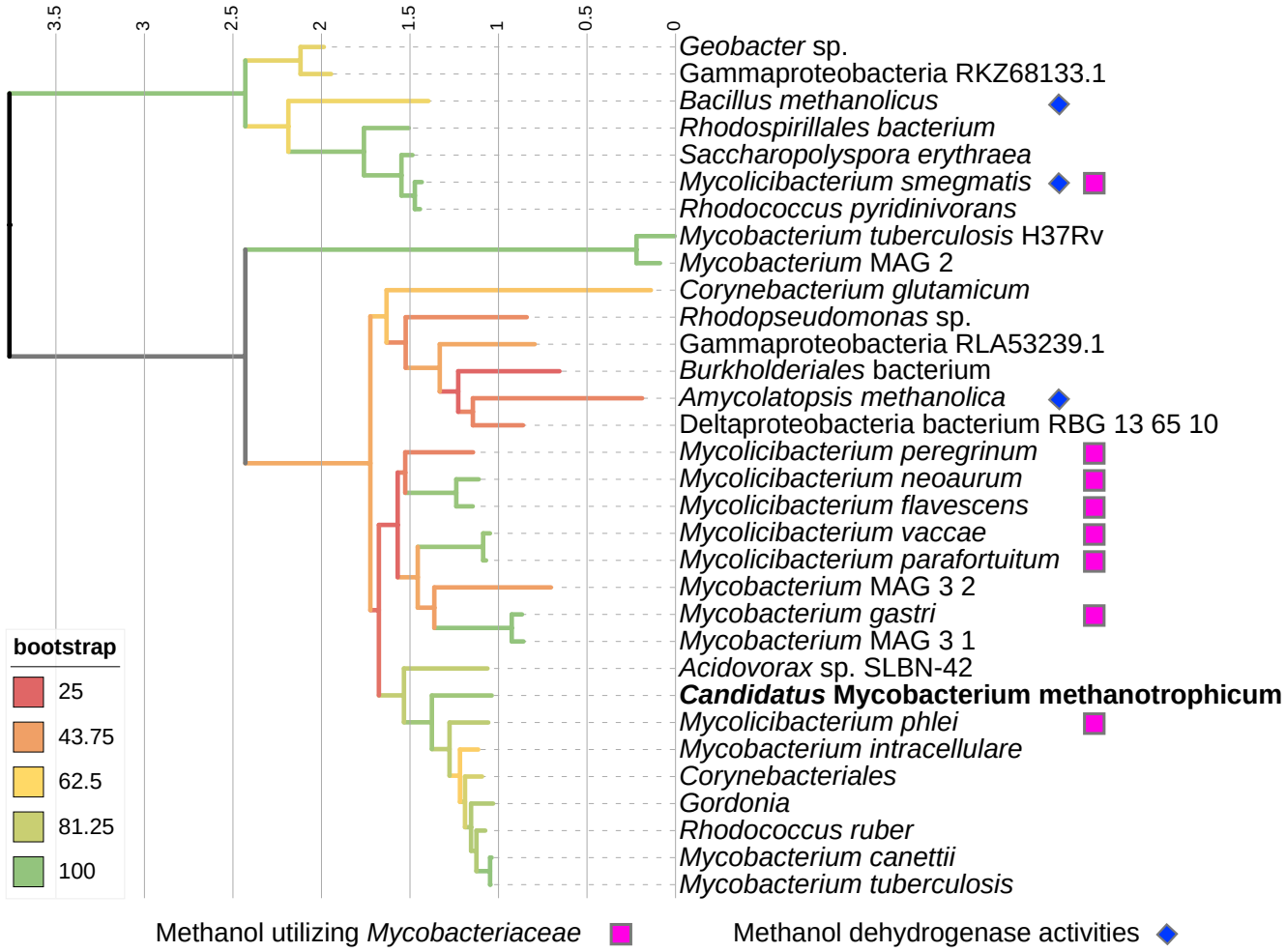
Supplementary Figure 5. Alignment of MmoX protein sequences from *Mycolicibacterium chubuense* NBB4, *Candidatus* M. methanotrophicum, *Methylosinus trichosporium* OB3b, and *Methylococcus capsulatus* Bath. Amino acids in red are positionally identical. The numbering of the amino acids is according to that of the *M. capsulatus* sequence. His (His147 to Fe1 and His246 to Fe2) and Glu (Glu 114, Glu144, Glu209, and Glu 243) residues that bind the Fe atoms in sMMOs are in bold and underlined, Leu110 (the gating residue), Cys151, Cys211, and Thr213-involved in oxygen activation are in bold and italics. Note the Cys151 to Thr substitution in the sequence of *Candidatus* M. methanotrophicum (in blue). Halfway of the sequence alignments are the conserved sequences of BMM hydroxylase α-subunits differentiated between those with known methane reactivity (conserved in sMMOs) and those that are inactive towards methane (conserved in other BMMs) [7].



Supplementary Figure 6. Phylogenetic tree of methane monooxygenase regulatory protein B (MmoB) including two copies encoded by *mmoB* genes from *Candidatus* M. methanotrophicum and close homologues. Grey colored branches are not supported by bootstrap values.



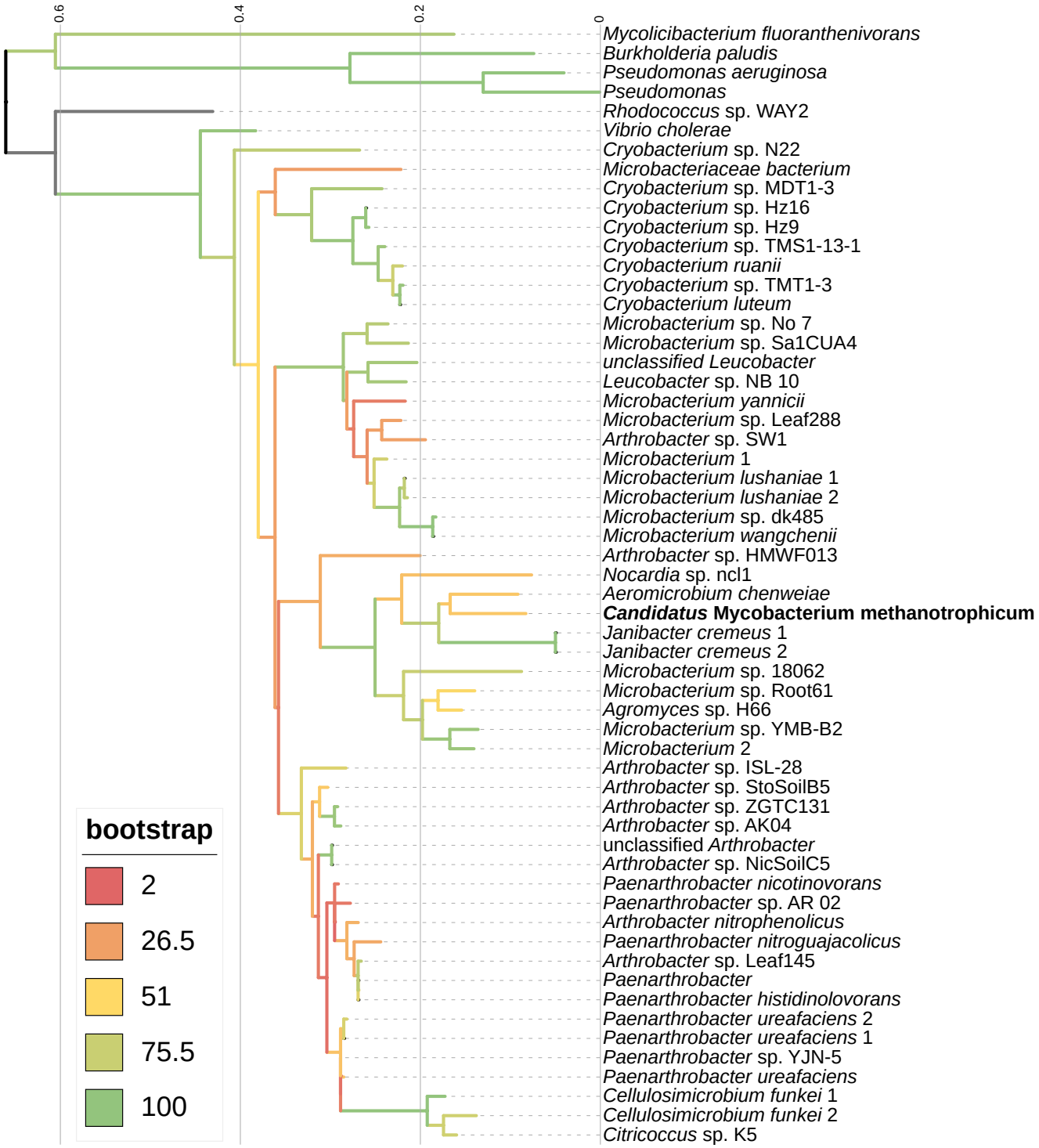
Supplementary Figure 7. AlphaFold prediction of ORF 4012 from *Candidatus* M. methanotrophicum. **(a)** AlphaFold template free prediction of ORF 4012, colored by pLDDT. pLDDT confidence range from 51.5 (red) to 98.1 (blue). **(b)** Superposition of ORF 4012 prediction, in pink, to MmoD from *Methylosinus trichosporium* (6d7k), chain D, in light brown. **(c)** AlphaFold Multimer template free prediction of the α-subunit together with MmoD, in dark blue and pink, respectively, superimposed onto the *Methylosinus trichosporium* sMMO-MmoD complex (6d7k), chain A-α-subunit, in light brown, and chain D-MmoD, in light blue. **(d)** AlphaFold Multimer PAE heat map of *Candidatus* M. methanotrophicum MmoD with MmoX. Note that the confidence score for the position of the N-terminus of MmoX is low. The overall pTM score is 0.8.



Supplementary Figure 8. Phylogenetic tree of alcohol dehydrogenases in assays dependent on p-nitroso-N,N-dimethylaniline (NDMA) as an electron acceptor, including AdhD from *Candidatus* M. methanotrophicum and close homologues, Mno from *M. smegmatis* [9] and close homologues, and two alcohol dehydrogenases from M. MAG 2 and *M. tuberculosis* H37Rv, the genes of which are in the *mft* gene clusters shown in Extended Data Fig. 3. Members of the *Mycobacteriaceae* able to grow on methanol are marked with pink squares. Gram-positive bacteria with AdhD proteins that were purified and shown to exhibit methanol dehydrogenase activities are marked with blue diamonds.

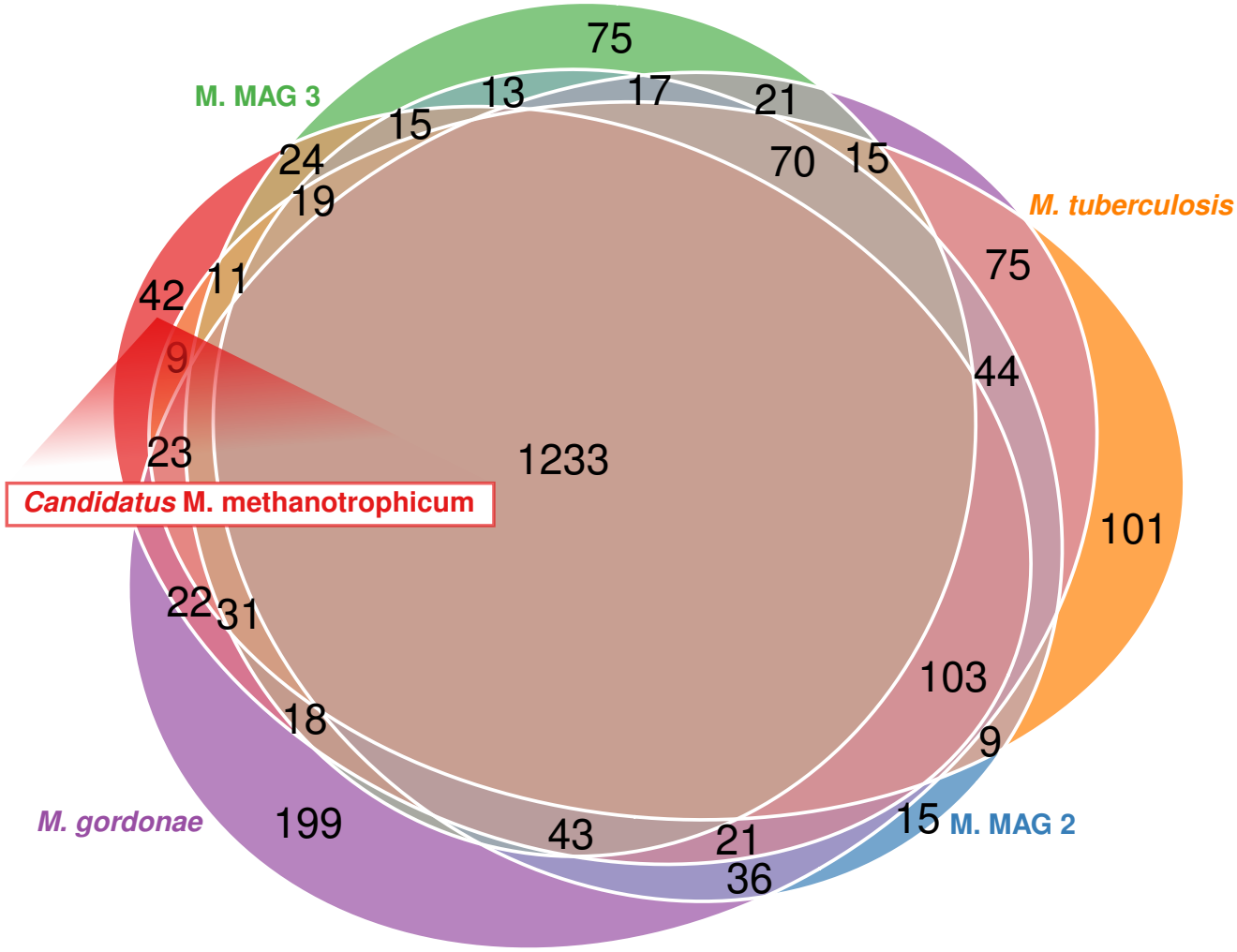
Rhodococcus	9	VESDLIEESLVEEVSIDGMCGVY	31
Frankia discariae	31	QEPSVEEELLIEEVSIDGMCGVY	53
Nocardioides	12	IASDVATEDLIEEVSIDGMCGVY	34
Haloterrigena	33	ETPEIEEDLVSEELRIDGICGVY	55
Actinoplanes	20	VDSLVETDELIEEVSIDGMCGVY	42
Thermomonospora	17	LTEEAAVESLIEEVSIDGMCGVY	39
Nocardia	132	DNSAVVDETLIEEVSIDGMCGVY	154
M. meth 2896	12	TAEQLISETLVEELSIDGMCGVY	34
Consensus		EE IDG CGVY	

Supplementary Figure 9. Alignment of the MftA sequence from *Candidatus* M. methanotrophicum with that from other species of the Actinobacteria and of the Archaeal species *Haloterrigena turkmenica*.



Supplementary Figure 10. Phylogenetic tree of formaldehyde dehydrogenases including FyDH from *Candidatus* M. methanotrophicum and close homologues.

Supplementary Figure 11. Examples of the different kinds of pools on the Puturosu Mountain. The pools were sampled to explore the diversity of sites with abundant *Mycobacteria*. Intense ebullition of volcanic gases occurs in most of the pools (**a**–**d**). The water in most of the pools is turbid (greyish) to a variable degree. Leafs in the pools are covered with a greyish biofilm (**b**).



Supplementary Figure 12. Venn diagram of the KO (KEGG ortholog) genome annotations of the three mycobacterial MAGs from the Sulfur Cave biofilms along with those of *M. tuberculosis* and *M. gordonae*. Out of the total of 2339 KOs, 1233 are shared by all five mycobacteria.

Supplementary References

1. Stainthorpe A, Lees V, Salmond GP, Dalton H, and Murrell J. The methane monooxygenase gene cluster of *Methylococcus capsulatus* (Bath). Gene 1990 Jan; 91:27–34. DOI: 10.1016/0378-1119(90)90158-n. Available from: https://doi.org/10.1016/0378-1119(90)90158-n

2. Cardy DLN, Laidler V, Salmond GPC, and Murrell JC. The methane monooxygenase gene cluster of *Methylosinus trichosporium*: cloning and sequencing of the *mmoC* gene. Archives of Microbiology 1991 Nov; 156:477–83. DOI: 10.1007/bf00245395. Available from: https://doi.org/10.1007/bf00245395

3. Stafford GP, Scanlan J, McDonald IR, and Murrell JC. *rpoN*, *mmoR* and *mmoG*, genes involved in regulating the expression of soluble methane monooxygenase in *Methylosinus trichosporium* OB3b. Microbiology 2003 Jul; 149:1771–84. DOI: 10.1099/mic.0.26060-0. Available from: https://doi.org/10.1099/mic.0.26060-0

4. Mukhopadhyay B, Concar EM, and Wolfe RS. A GTP-dependent vertebrate-type phosphoenolpyruvate carboxykinase from *Mycobacterium smegmatis*. Journal of Biological Chemistry 2001 Feb; 276:16137–45. DOI: 10.1074/jbc.m008960200. Available from: https://doi.org/10.1074/jbc.m008960200

5. Leahy JG, Batchelor PJ, and Morcomb SM. Evolution of the soluble diiron monooxygenases. FEMS Microbiology Reviews 2003 Oct; 27:449–79. DOI: 10.1016/s0168-6445(03)00023-8. Available from: https://doi.org/10.1016/s0168-6445(03)00023-8

6. Guerrero-Cruz S, Vaksmaa A, Horn MA, Niemann H, Pijuan M, and Ho A. Methanotrophs: Discoveries, environmental relevance, and a perspective on current and future applications. Frontiers in Microbiology 2021 May; 12. DOI: 10.3389/fmicb.2021.678057. Available from: https://doi.org/10.3389/fmicb.2021.678057

7. Osborne CD and Haritos VS. Beneath the surface: Evolution of methane activity in the bacterial multicomponent monooxygenases. Molecular Phylogenetics and Evolution 2019 Oct; 139:106527.

DOI: 10.1016/j.ympev.2019.106527. Available from: https://doi.org/10.1016/j.ympev.2019.106527

8. Kim H, An S, Park YR, Jang H, Yoo H, Park SH, Lee SJ, and Cho US. MMOD-induced structural changes of hydroxylase in soluble methane monooxygenase. Science Advances 2019 Oct; 5. DOI: 10.1126/sciadv.aax0059. Available from: https://doi.org/10.1126/sciadv.aax0059

9. Dubey AA, Wani SR, and Jain V. Methylotrophy in mycobacteria: Dissection of the methanol metabolism pathway in *Mycobacterium smegmatis*. Journal of Bacteriology 2018 Jun; 200. Ed. by Metcalf WW. DOI: 10.1128/jb.00288-18. Available from: https://doi.org/10.1128/jb.00288-18

10. Chistoserdova L. C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. Science 1998 Jul; 281:99–102. DOI: 10.1126/science.281.5373.99. Available from: https://doi.org/10.1126/science.281.5373.99

11. Vorholt JA, Chistoserdova L, Lidstrom ME, and Thauer RK. The NADP-dependent methylene tetrahydromethanopterin dehydrogenase in *Methylobacterium extorquens* AM1. Journal of Bacteriology 1998 Oct; 180:5351–6. DOI: 10.1128/jb.180.20.5351-5356.1998. Available from: https://doi.org/10.1128/jb.180.20.5351-5356.1998

12. Marquet A, Bui BTS, and Florentin D. Biosynthesis of biotin and lipoic acid. *Vitamins & Hormones*. Elsevier, 2001 :51–101. DOI: 10.1016/s0083-6729(01)61002-1. Available from: https://doi.org/10.1016/s0083-6729(01)61002-1

13. Jitrapakdee S, St Maurice M, Rayment I, Cleland WW, Wallace JC, and Attwood PV. Structure, mechanism and regulation of pyruvate carboxylase. Biochemical Journal 2008 Jul; 413:369–87. DOI: 10.1042/bj20080709. Available from: https://doi.org/10.1042/bj20080709

14. Mitsui R, Sakai Y, Yasueda H, and Kato N. A novel operon encoding formaldehyde fixation: the ribulose monophosphate pathway in the Gram-positive facultative methylotrophic bacterium *Mycobacterium gastri* MB19. Journal of Bacteriology 2000 Feb; 182:944–8. DOI: 10.1128/jb.

182.4.944-948.2000. Available from: https://doi.org/10.1128/jb.182.4.944-948.2000

15. Jugder BE, Welch J, Aguey-Zinsou KF, and Marquis CP. Fundamentals and electrochemical applications of [Ni–Fe]-uptake hydrogenases. RSC Advances 2013; 3:8142. DOI: 10.1039/c3ra22668a. Available from: https://doi.org/10.1039/c3ra22668a

16. Liebgott PP, Leroux F, Burlat B, Dementin S, Baffert C, Lautier T, Fourmond V, Ceccaldi P, Cavazza C, Meynial-Salles I, Soucaille P, Fontecilla-Camps JC, Guigliarelli B, Bertrand P, Rousset M, and Léger C. Relating diffusion along the substrate tunnel and oxygen sensitivity in hydrogenase. Nature Chemical Biology 2009 Dec; 6:63–70. DOI: 10.1038/nchembio.276. Available from: https://doi.org/10.1038/nchembio.276

17. Vignais PM. Classification and phylogeny of hydrogenases. FEMS Microbiology Reviews 2001 Aug; 25:455–501. DOI: 10.1016/s0168-6445(01)00063-8. Available from: https://doi.org/10.1016/s0168-6445(01)00063-8

18. Vignais PM. Hydrogenases and H^+-reduction in primary energy conservation. *Bioenergetics*. Springer Berlin Heidelberg, 2007 :223–52. DOI: 10.1007/400_2006_027. Available from: https://doi.org/10.1007/400_2006_027

19. Hedderich R and Forzi L. Energy-converting [NiFe] hydrogenases: More than just H_2 activation. Journal of Molecular Microbiology and Biotechnology 2005; 10:92–104. DOI: 10.1159/000091557. Available from: https://doi.org/10.1159/000091557

20. Carere CR, Hards K, Houghton KM, Power JF, McDonald B, Collet C, Gapes DJ, Sparling R, Boyd ES, Cook GM, Greening C, and Stott MB. Mixotrophy drives niche expansion of verrucomicrobial methanotrophs. The ISME Journal 2017 Aug; 11:2599–610. DOI: 10.1038/ismej.2017.112. Available from: https://doi.org/10.1038/ismej.2017.112

21. Gong H, Li J, Xu A, Tang Y, Ji W, Gao R, Wang S, Yu L, Tian C, Li J, Yen HY, Lam SM, Shui G, Yang X, Sun Y, Li X, Jia M, Yang C, Jiang B, Lou Z, Robinson CV, Wong LL, Guddat LW, Sun F, Wang Q, and Rao Z. An electron transfer path connects subunits of a mycobacterial respiratory

supercomplex. Science 2018 Oct; 362:eaat8923. DOI: 10.1126/science.aat8923. Available from: https://doi.org/10.1126/science.aat8923

22. Lu P, Heineke M, Koul A, Andries K, Cook GM, Lill H, Spanning RJ van, and Bald D. Cytochrome *bd* as survival factor in mycobacteria. Biochimica et Biophysica Acta (BBA) - Bioenergetics 2016 Aug; 1857:e120. DOI: 10.1016/j.bbabio.2016.04.257. Available from: https://doi.org/10.1016/j.bbabio.2016.04.257

23. Park SW, Hwang EH, Jang HS, Lee JH, Kang BS, Oh JI, and Kim YM. Presence of duplicate genes encoding a phylogenetically new subgroup of form I ribulose 1, 5-bisphosphate carboxylase/oxygenase in *Mycobacterium* sp. strain JC1 DSM 3803. Research in Microbiology 2009 Mar; 160:159–65. DOI: 10.1016/j.resmic.2008.12.002. Available from: https://doi.org/10.1016/j.resmic.2008.12.002

24. Park SW, Hwang EH, Park H, Kim JA, Heo J, Lee KH, Song T, Kim E, Ro YT, Kim SW, and Kim YM. Growth of mycobacteria on carbon monoxide and methanol. Journal of Bacteriology 2003 Jan; 185:142–7. DOI: 10.1128/jb.185.1.142-147.2003. Available from: https://doi.org/10.1128/jb.185.1.142-147.2003

25. Reed WM and Dugan PR. Isolation and characterization of the facultative methylotroph *Mycobacterium* ID-Y. Microbiology 1987 May; 133:1389–95. DOI: 10.1099/00221287-133-5-1389. Available from: https://doi.org/10.1099/00221287-133-5-1389

26. Song Z, Orita I, Yin F, Yurimoto H, Kato N, Sakai Y, Izui K, Li K, and Chen L. Overexpression of an HPS/PHI fusion enzyme from *Mycobacterium gastri* in chloroplasts of geranium enhances its ability to assimilate and phytoremediate formaldehyde. Biotechnology Letters 2010 Jun; 32:1541–8. DOI: 10.1007/s10529-010-0324-7. Available from: https://doi.org/10.1007/s10529-010-0324-7

27. Chistoserdova L, Chen SW, Lapidus A, and Lidstrom ME. Methylotrophy in *Methylobacterium extorquens* AM1 from a genomic point of view. Journal of Bacteriology 2003 May; 185:2980–7. DOI: 10.1128/jb.185.10.2980-2987.2003. Available from: https://doi.org/10.1128/jb.185.10.2980-2987.2003

28. Harms N and Spanning RJM van. C1 metabolism in *Paracoccus denitrificans*: Genetics of *Paracoccus denitrificans*. Journal of Bioenergetics and Biomembranes 1991 Apr; 23:187–210. DOI: 10.1007/bf00762217. Available from: https://doi.org/10.1007/bf00762217

29. Pinto R, Harrison JS, Hsu T, Jacobs WR, and Leyh TS. Sulfite Reduction in Mycobacteria. Journal of Bacteriology 2007 Sep; 189:6714–22. DOI: 10.1128/jb.00487-07. Available from: https://doi.org/10.1128/jb.00487-07

30. Yamanaka T, Fukumori Y, and Okunuki K. Preparation of subunits of flavocytochromes c derived from Chlorobium limicola f. thiosulfatophilum and Chromatium vinosum. Analytical Biochemistry 1979 May; 95:209–13. DOI: 10.1016/0003-2697(79)90207-0. Available from: https://doi.org/10.1016/0003-2697(79)90207-0

31. Beller HR, Chain PSG, Letain TE, Chakicherla A, Larimer FW, Richardson PM, Coleman MA, Wood AP, and Kelly DP. The Genome Sequence of the Obligately Chemolithoautotrophic, Facultatively Anaerobic Bacterium *Thiobacillus denitrificans*. Journal of Bacteriology 2006 Feb; 188:1473–88. DOI: 10.1128/jb.188.4.1473-1488.2006. Available from: https://doi.org/10.1128/jb.188.4.1473-1488.2006

32. Kelly DP, Shergill JK, Lu WP, and Wood AP. Antonie van Leeuwenhoek 1997; 71:95–107. DOI: 10.1023/a:1000135707181. Available from: https://doi.org/10.1023/a:1000135707181

33. Friedrich CG, Rother D, Bardischewsky F, Quentmeier A, and Fischer J. Oxidation of Reduced Inorganic Sulfur Compounds by Bacteria: Emergence of a Common Mechanism? Applied and Environmental Microbiology 2001 Jul; 67:2873–82. DOI: 10.1128/aem.67.7.2873-2882.2001. Available from: https://doi.org/10.1128/aem.67.7.2873-2882.2001

34. Cockell CS, Harrison JP, Stevens AH, Payler SJ, Hughes SS, Nawotniak SEK, Brady AL, Elphic R, Haberle CW, Sehlke A, Beaton KH, Abercromby AF, Schwendner P, Wadsworth J, Landenmark H, Cane R, Dickinson AW, Nicholson N, Perera L, and Lim DS. A low-diversity microbiota inhabits extreme terrestrial basaltic terrains and their fumaroles: implications for the exploration of Mars.

Astrobiology 2019 Mar; 19:284–99. DOI: 10.1089/ast.2018.1870. Available from: https://doi.org/10.1089/ast.2018.1870

35. Šibanc N, Dumbrell AJ, Mandić-Mulec I, and Maček I. Impacts of naturally elevated soil CO_2 concentrations on communities of soil archaea and bacteria. Soil Biology and Biochemistry 2014 Jan; 68:348–56. DOI: 10.1016/j.soilbio.2013.10.018. Available from: https://doi.org/10.1016/j.soilbio.2013.10.018

36. Lavoie KH, Winter AS, Read KJH, Hughes EM, Spilde MN, and Northup DE. Comparison of bacterial communities from lava cave microbial mats to overlying surface soils from Lava Beds National Monument, USA. PLOS ONE 2017 Feb; 12. Ed. by Brusetti L:e0169339. DOI: 10.1371/journal.pone.0169339. Available from: https://doi.org/10.1371/journal.pone.0169339

37. Vincke E, Boon N, and Verstraete W. Analysis of the microbial communities on corroded concrete sewer pipes; a case study. Applied Microbiology and Biotechnology 2001 Dec; 57:776–85. DOI: 10.1007/s002530100826. Available from: https://doi.org/10.1007/s002530100826

38. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, and Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nature Communications 2018 Nov; 9. DOI: 10.1038/s41467-018-07641-9. Available from: https://doi.org/10.1038/s41467-018-07641-9