

Supplementary information for:

Hydrogenotrophic methanogenesis in archaeal phylum Verstraetearchaeota reveals the shared ancestry of all methanogens

Bojk A. Berghuis¹, Feiqiao Brian Yu^{1,2}, Frederik Schulz³, Paul Blainey⁴,
Tanja Woyke³, Stephen R. Quake^{1,2,5}

¹Department of Bioengineering, Stanford University, Stanford, United States

²Chan Zuckerberg Biohub, San Francisco, United States

³Department of Energy Joint Genome Institute, Walnut Creek, United States

⁴MIT Department of Biological Engineering and Broad Institute of Harvard and MIT

⁵Department of Applied Physics, Stanford University, Stanford, United States

List of figures

1. Obsidian Pool (mini-)metagenomics pipeline and dataset
2. Key genes in Obsidian pool genomes
3. Genes related to energy conservation in methanogenic and related orders (extension of Fig. 2).
4. Key methanogenesis/WLP genes in the clades of the Verstraetearchaeota, Ignisphaera, and Bathyarchaeota
5. KEGG-based comparison of Verstraetearchaeota orders
6. Verstraetearchaeota ANI analysis
7. Relative presence of WLP/methanogenesis genes in all methanogenic orders
8. Gene tree of MtrG
9. Gene tree of MtrB
10. Gene tree of MtrH
11. Gene tree of McrA
12. Gene tree of CdhD
13. Gene tree of CdhG
14. Gene tree of OP CDH gene clusters (concatenated alignment)
15. Gene tree of trimethylamine methyltransferase (MttB).

List of tables

1. Obsidian Pool sample information
2. Genes used for methanogenesis pathway reconstruction

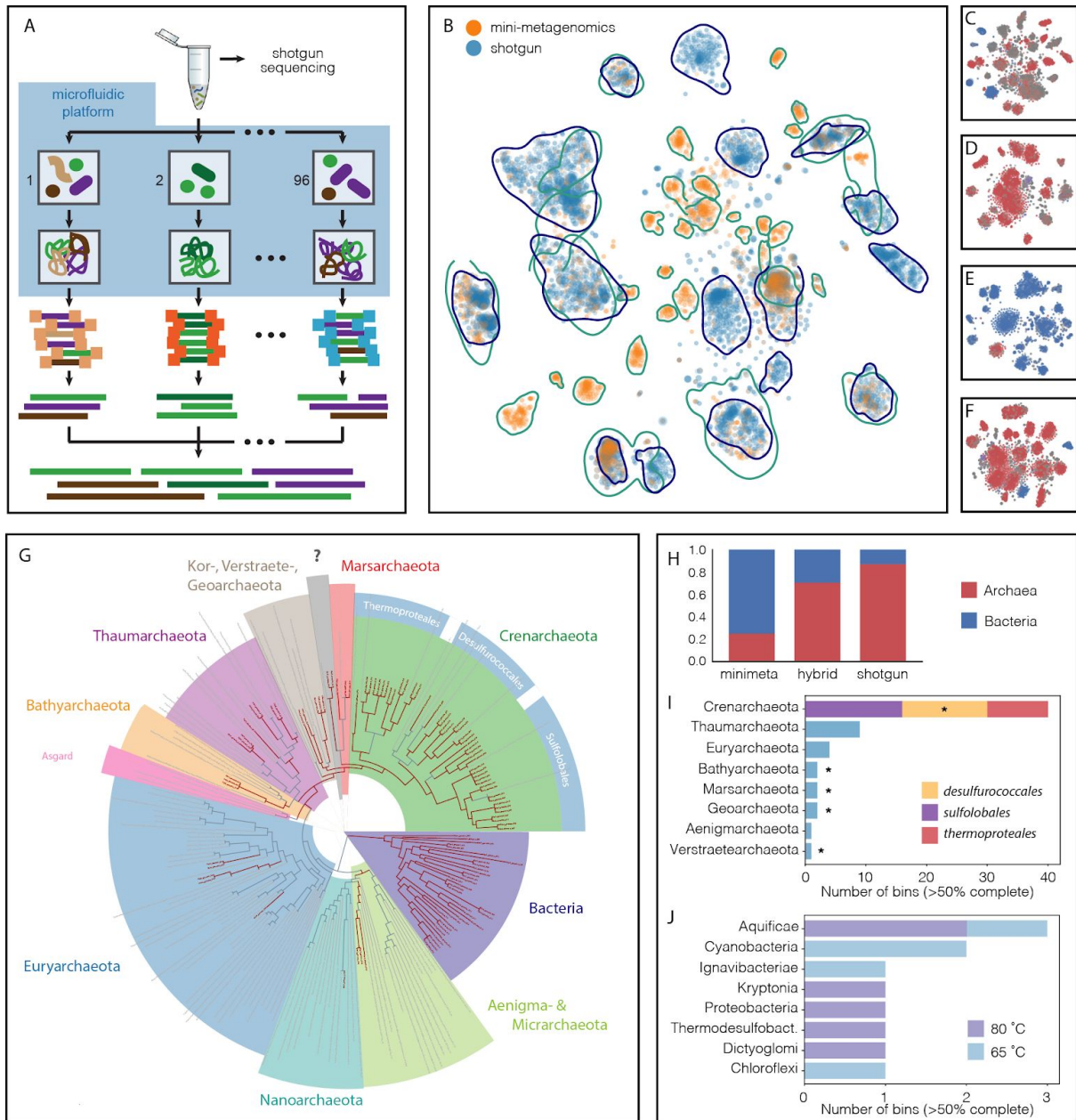


Figure S1. Obsidian Pool (mini-)metagenomics pipeline and dataset. (A) Mini-metagenomic sequencing. (B) 5-mer-based contig clustering with tSNE, obsidian 3; mini-metagenomic (orange) and metagenomic (blue) contigs shown. (C-F) Obsidian Pool samples 2,4,5,6 resp., showing contigs of archaeal (red), bacterial (blue) and unknown (gray) origin. (G) Phylogenomic tree of OP archaea (red) and publicly available reference genomes (gray) based on a concatenated alignment of 56 marker genes. (H) Fraction of archaeal (red) or bacterial (blue) bins based on sequencing method. (I,J) Phylogeny vs. number of high-quality bins, phyla containing novel phylogeny are indicated with (*).

Table S1. Yellowstone National Park, Obsidian Pool, sample information

	Sample #1	Sample #2	Sample #3	Sample #4	Sample #5
Sample area	Obsidian Pool, mud volcano area				
Sample type	sediment				
Location	44.61 N, 110.44 W				
Location type	hot spring				
Temperature (C)	79	79	79	65	78
pH	5	5	5	5	5
IMG genome ID	3300017482	3300017562	3300017469	3300017461	3300017696
Shotgun number of reads (millions)	44	20	49	6.56	104
Mini-metagenomic number of reads (millions)	323	148	146	191	151
reads per microfluidic chamber (millions)	1.68	1.61	1.59	2.08	1.64
Number of contigs mini-m (>5 kbp)	1318	1994	164	2427	16
Number of contigs total (>5 kbp)	6956	6010	4931	4401	7170

Table S2. List of genes used for methanogenesis pathway reconstruction.

gene	subunits	name	pathway
atp	ABCDEF	archaeal type H+ATPase	energy metabolism
mcr	ABGCD	methyl coenzyme M reductase	methanogenesis
mtr	ABCDEFGH	methyl-H4MPT:coenzyme M methyltransferase	methanogenesis
hdr	ABCDE	heterodisulfide reductase	methanogenesis
flp	D	methyl-viologen-reducing hydrogenase	methanogenesis
fae		formaldehyde activating enzyme	methanogenesis
cdh	ABGDE	CO dehydrogenase/acetyl-CoA synthase	WLP
frh	ABGD	F420-reducing hydrogenase	WLP
fwd	ABCDE	formylmethanofuran dehydrogenase	WLP
flr		formylmethanofuran:H4MPT formyltransferase	WLP
mch		Methenyltetrahydromethanopterin cyclohydrolase	WLP
mtd		F420-dependent methylenetetrahydromethanopterin dehydrogenase	WLP
mer		methylene tetrahydromethanopterin reductase	WLP
eha	ABCEFGHM	energy converting hydrogenase A	energy conservation (methanogenesis related)
ehb	ABCDHILMNPQ	energy converting hydrogenase B	energy conservation (methanogenesis related)
nfh	SLG	Ni,Fe hydrogenase III	energy conservation (methanogenesis related)
nuo	ABCDEFGHIJKLMN	NADH-ubiquinone oxidoreductase	ferredoxin regeneration (methanogenesis related)
mtmB		monomethylamine methyltransferase	methylotrophic methanogenesis related
mtbC		dimethylamine methyltransferase	methylotrophic methanogenesis related
mttB		Trimethylamine:corrinoid methyltransferase	methylotrophic methanogenesis related
mtaB		methanol-cobalamin methyltransferase	methylotrophic methanogenesis related
rgyr		reverse gyrase	thermophilicity marker
hco	I,II	heme-copper cytochrome oxidase	aerobic metabolism marker

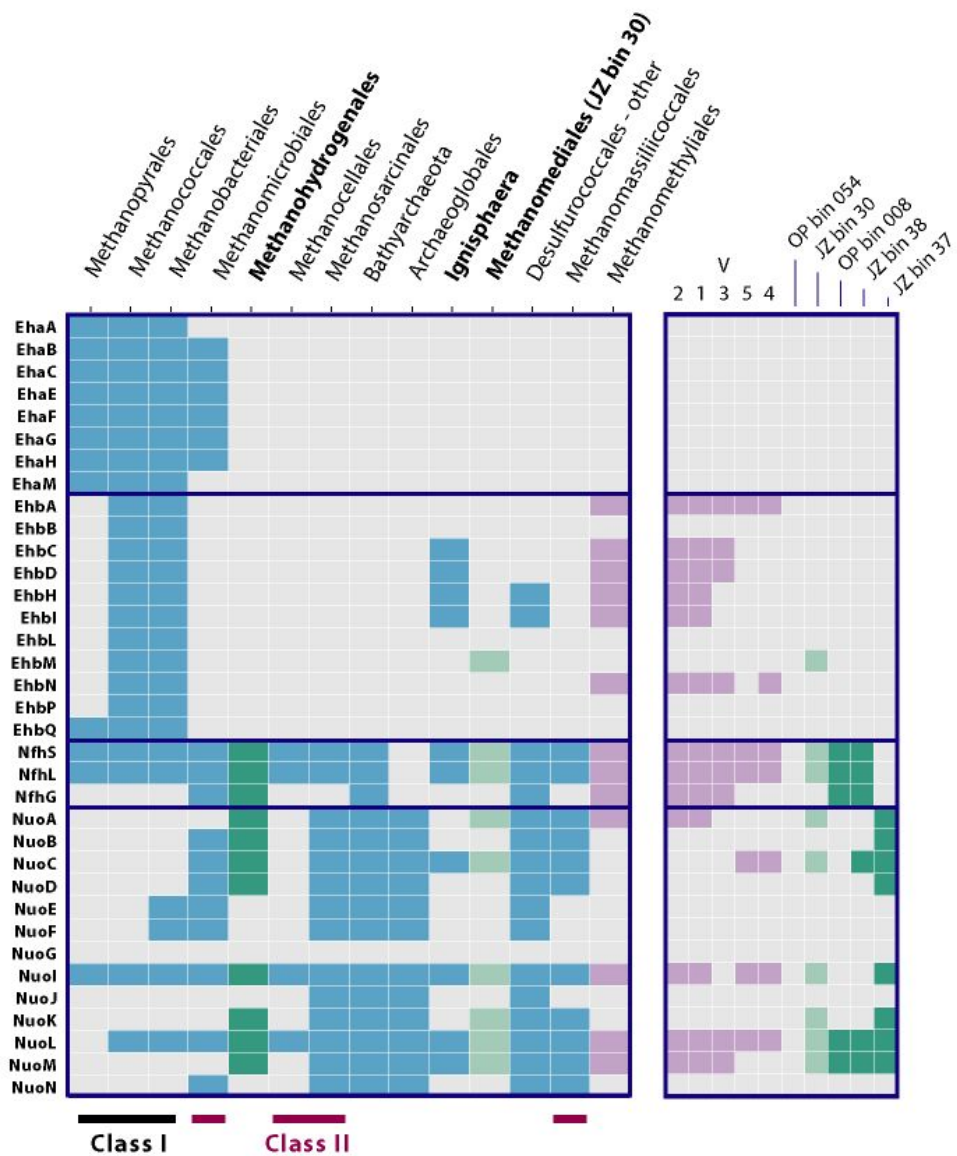


Figure S3. Genes related to energy conservation in methanogenic and related orders (extension of Fig. 2).

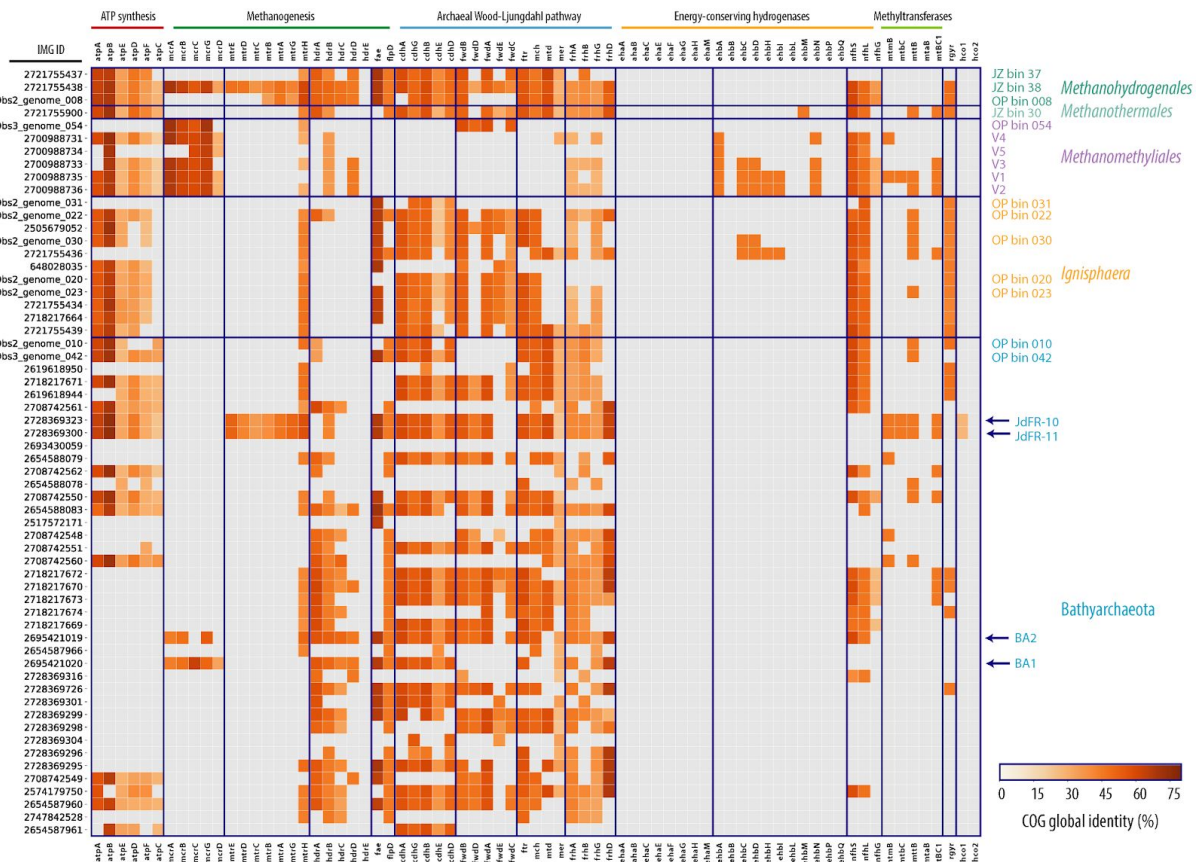


Fig. S4. Key WLP and methanogenesis genes in Obsidian Pool clades of the Bathyarchaeota, Ignisphaera and Verstraetearchaeota. The presence (orange) or absence (gray) of gene homologues and/or subunits of gene clusters involved in methanogenesis (*mcr*, *mtr*, *hdr*), the Wood-Ljungdahl pathway (*cdh*, *fwd*, *frt*, *mch*, *mtd*, *mer*, *frh*, *por*), related processes (*atp*, *ech*, methyltransferases), and others (see **Table 1** for details). Presence shaded by COG percent identity of the homologue. Non-Euryarchaeal methanogenic genomes reported indicated in the Bathyarchaeota (BA1,BA2; [Evans et al. 2015](#)) and Verstraetearchaeota (V1-5; [Vanwonterghem et al. 2016](#)) clades. Euryarchaeal genomes shown as reference.

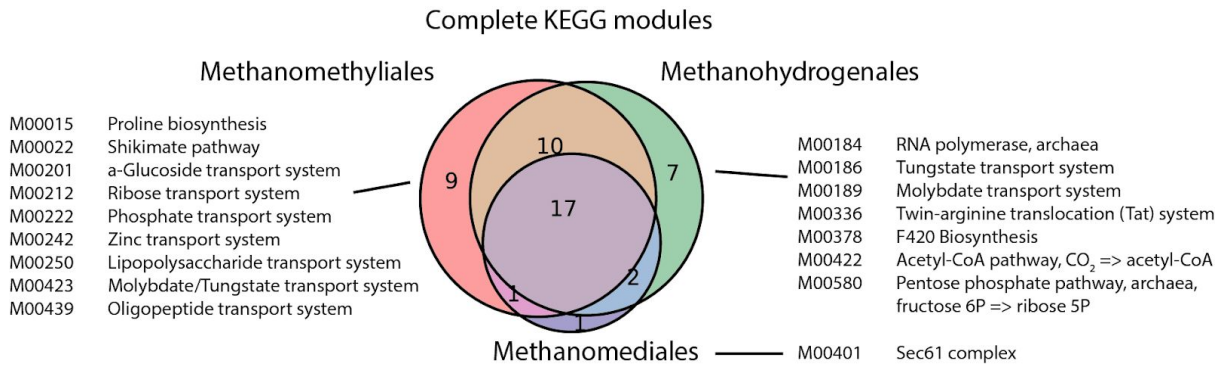


Figure S5. Complete KEGG modules unique to the 3 verstraetearchaeotal orders.

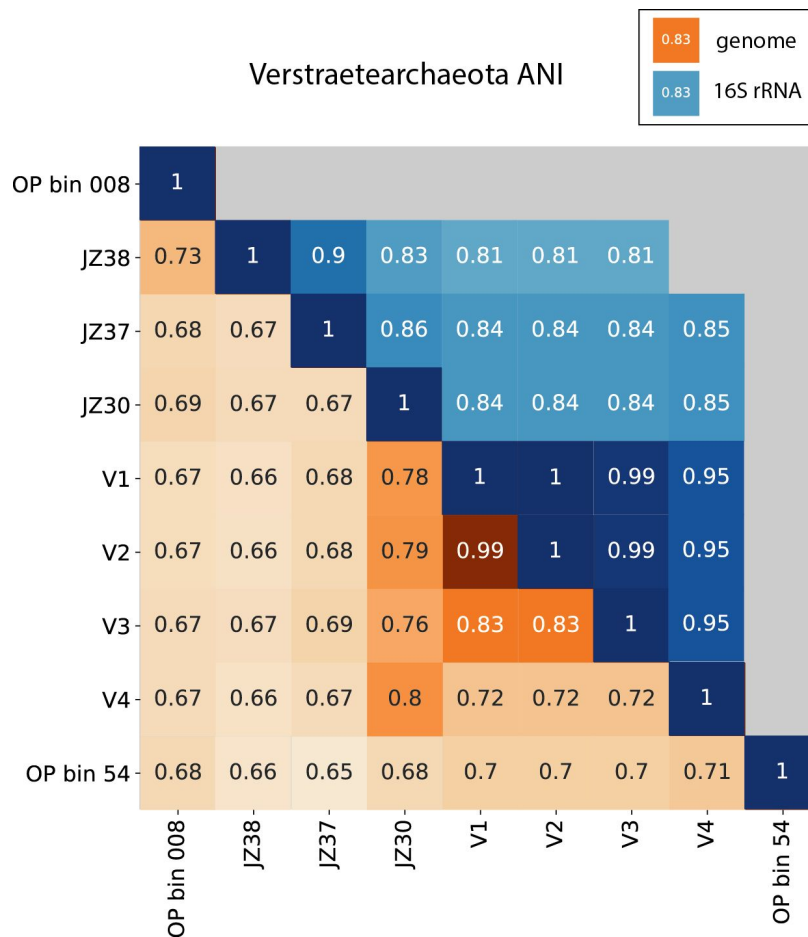


Fig S6. Pairwise average nucleotide identity (ANI) of the Verstraetearchaeota. 16S rRNA is not present in OP bin 008 and OP bin 54, partially present in JZ bin 38 (bp 1-512) and V4 (bp 890-1500). All other genomes have full length 16S rRNA.

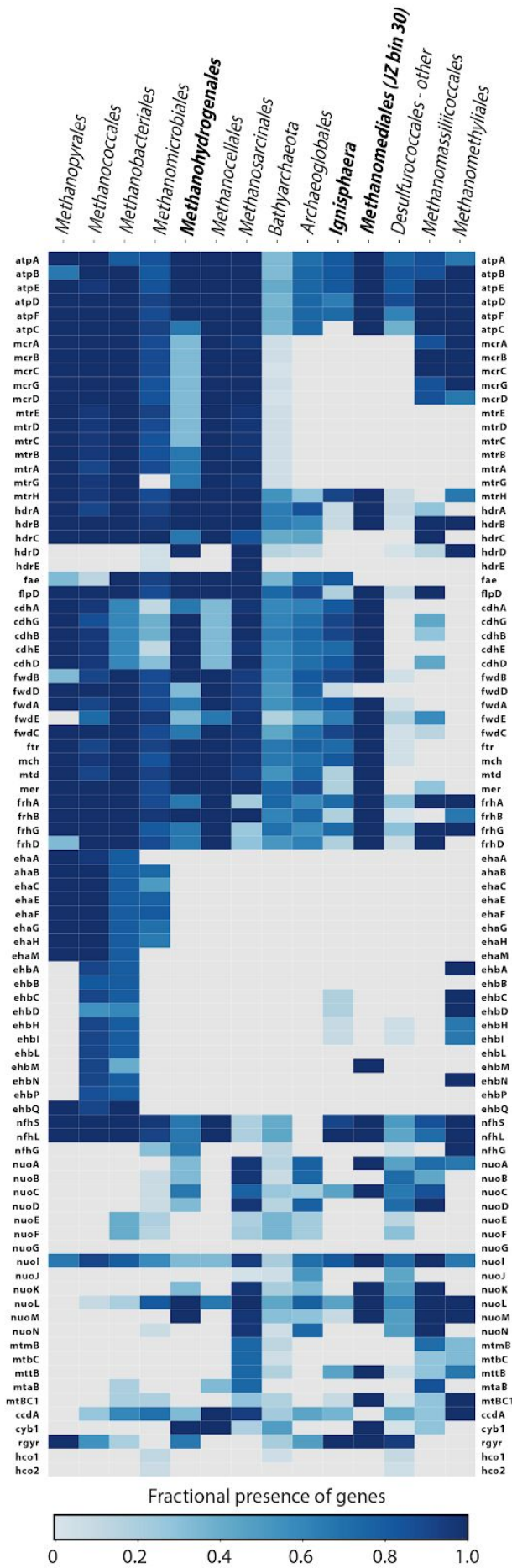


Figure S7. Presence of genes associated with methanogenesis and the WLP for all methanogenic orders, Archaeoglobales, Ignisphaera and other Desulfurococcales, shown as a fraction of the number of genomes in a clade containing the genes.

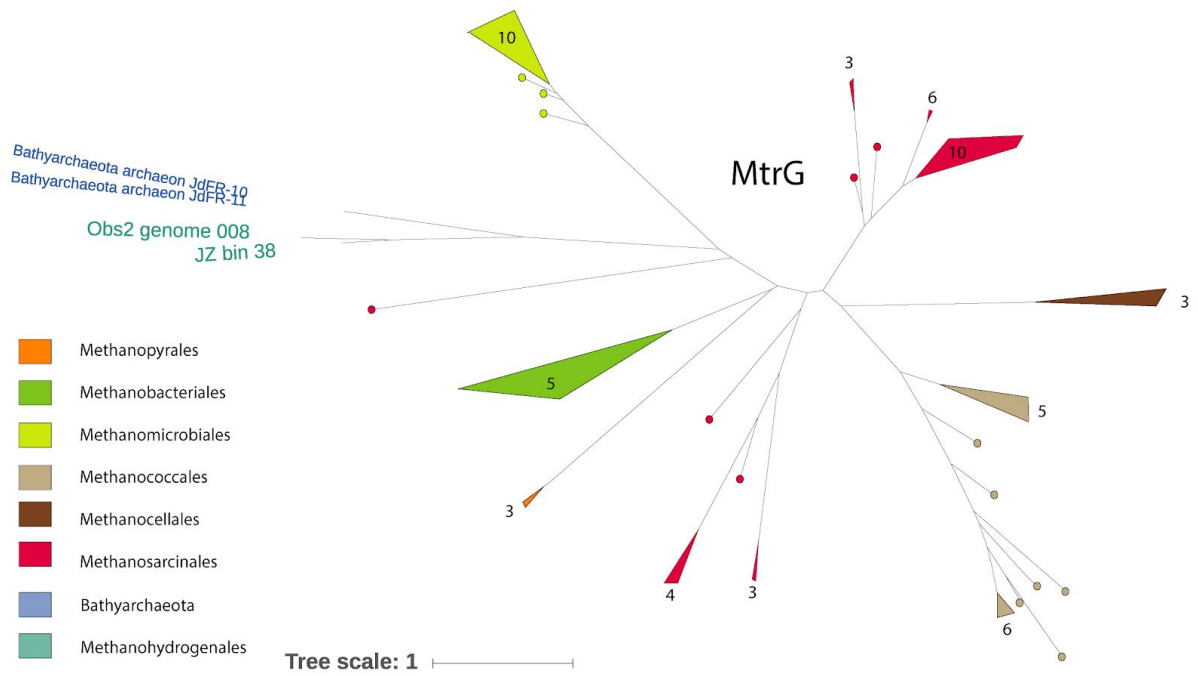


Fig S8. Gene tree of MtrG

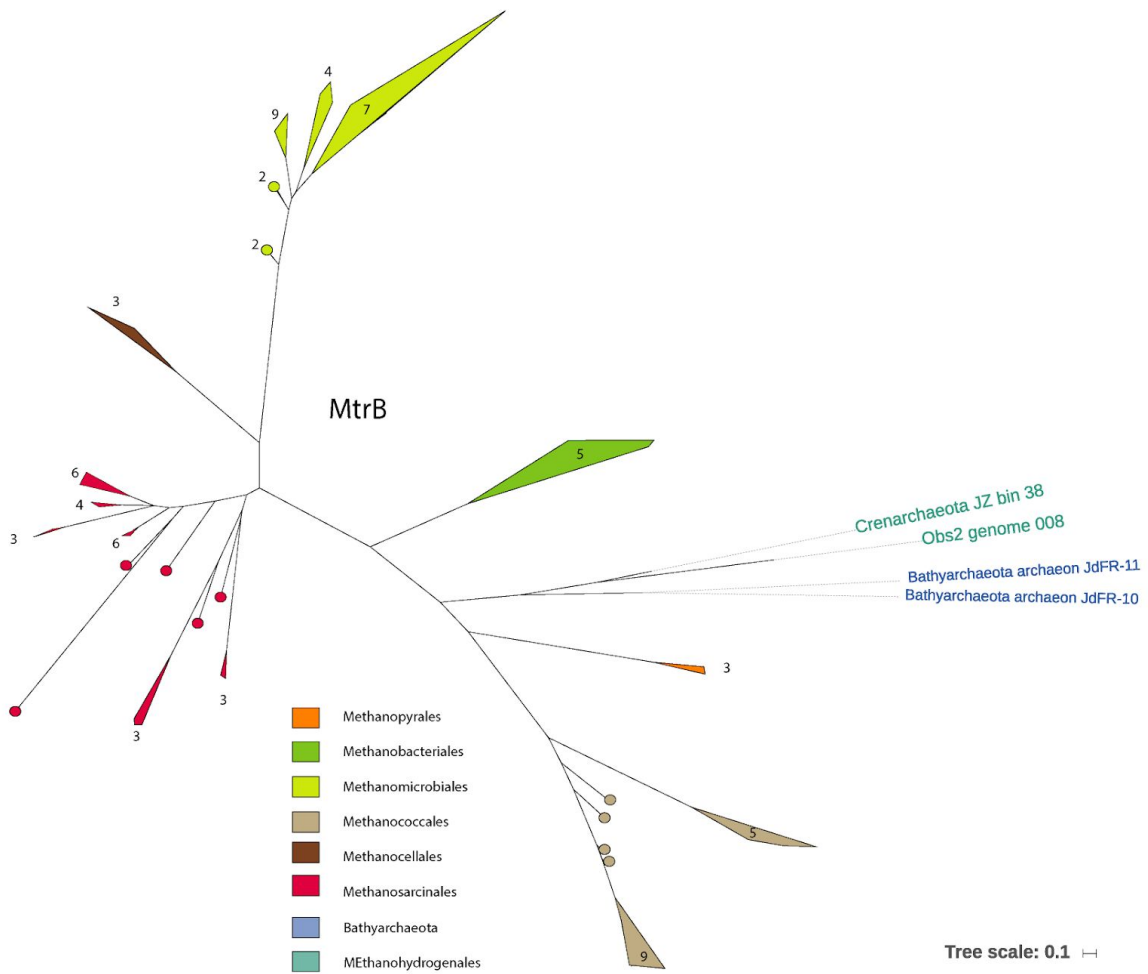


Figure S9. Phylogeny of MtrB.

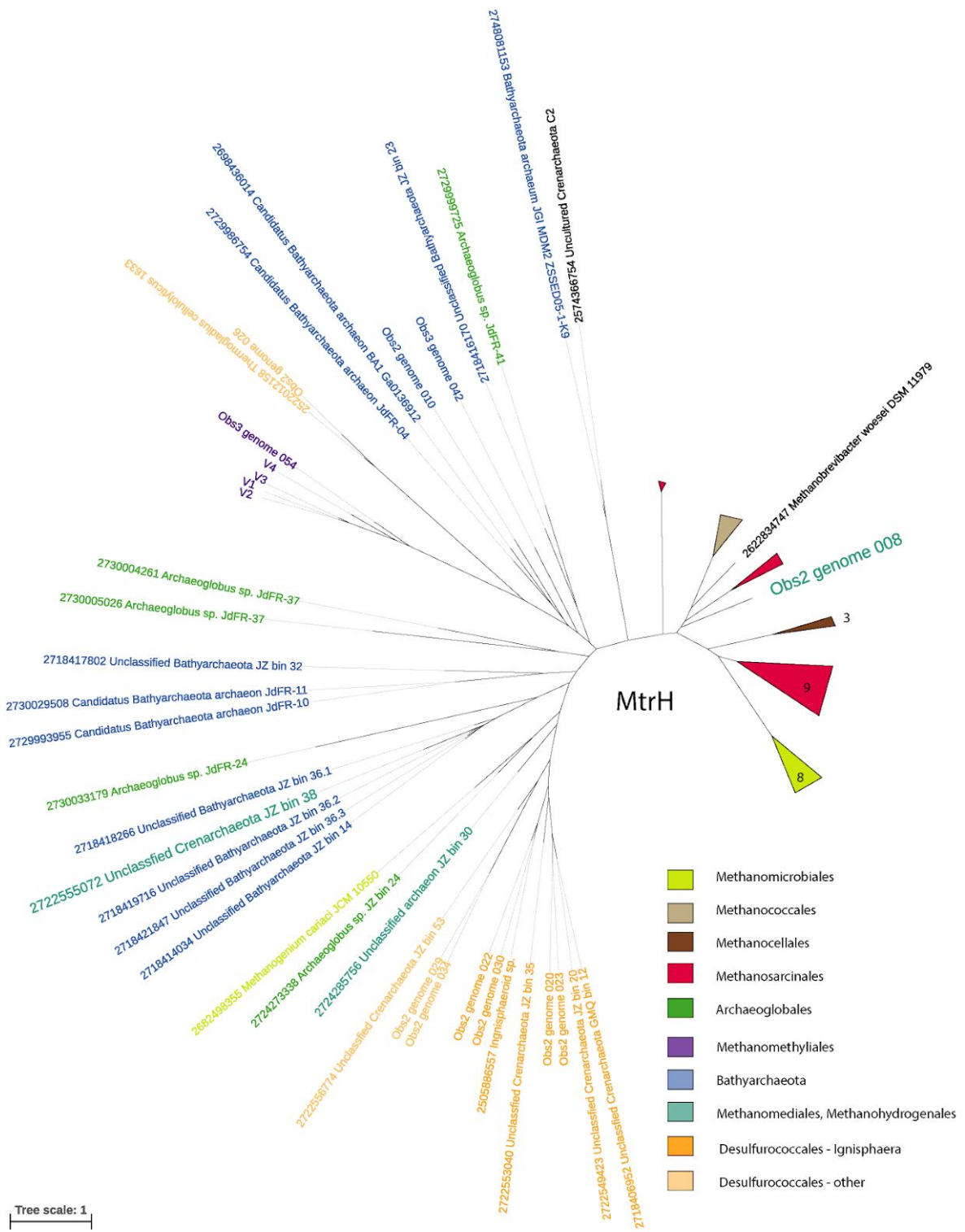


Figure S10. Phylogeny of mtrH

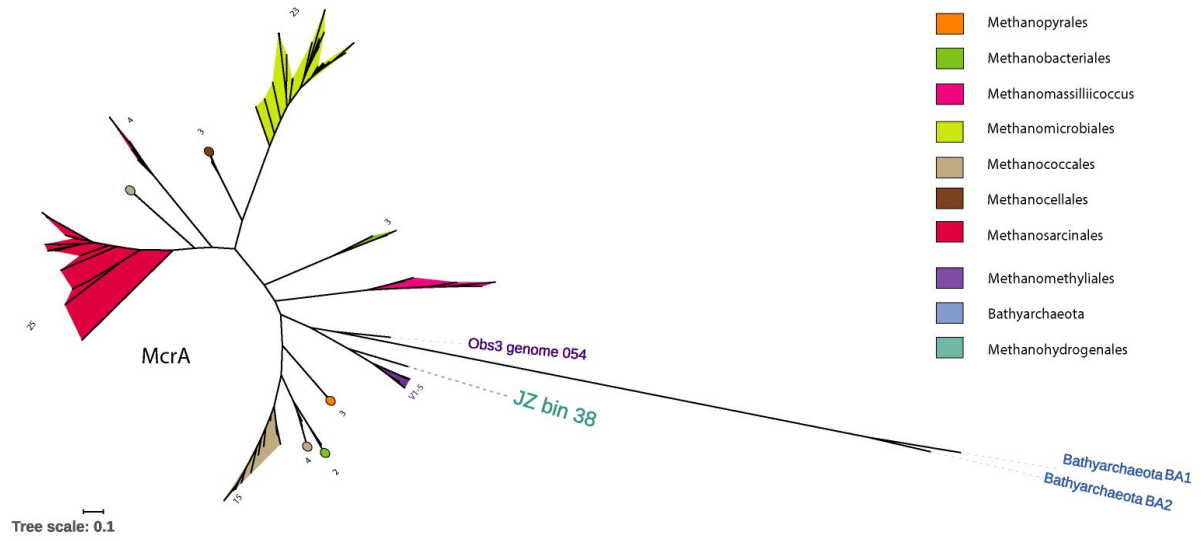


Fig S11. Phylogeny of McrA

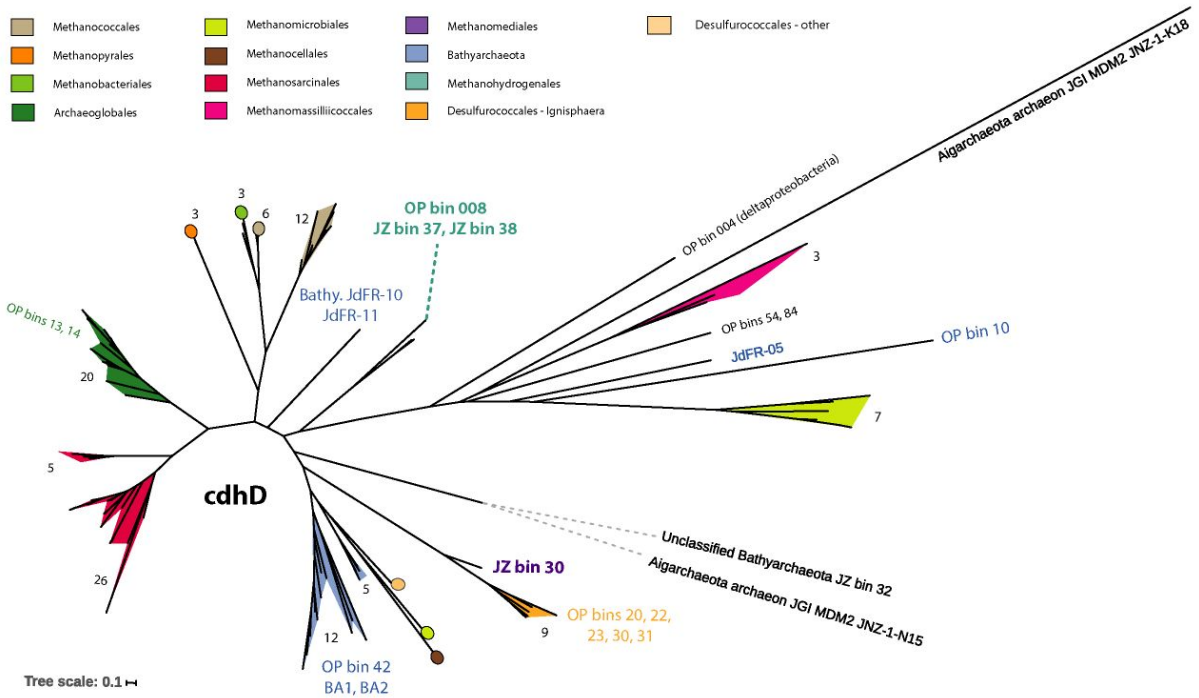


Fig. S12. Tree of CdhD.

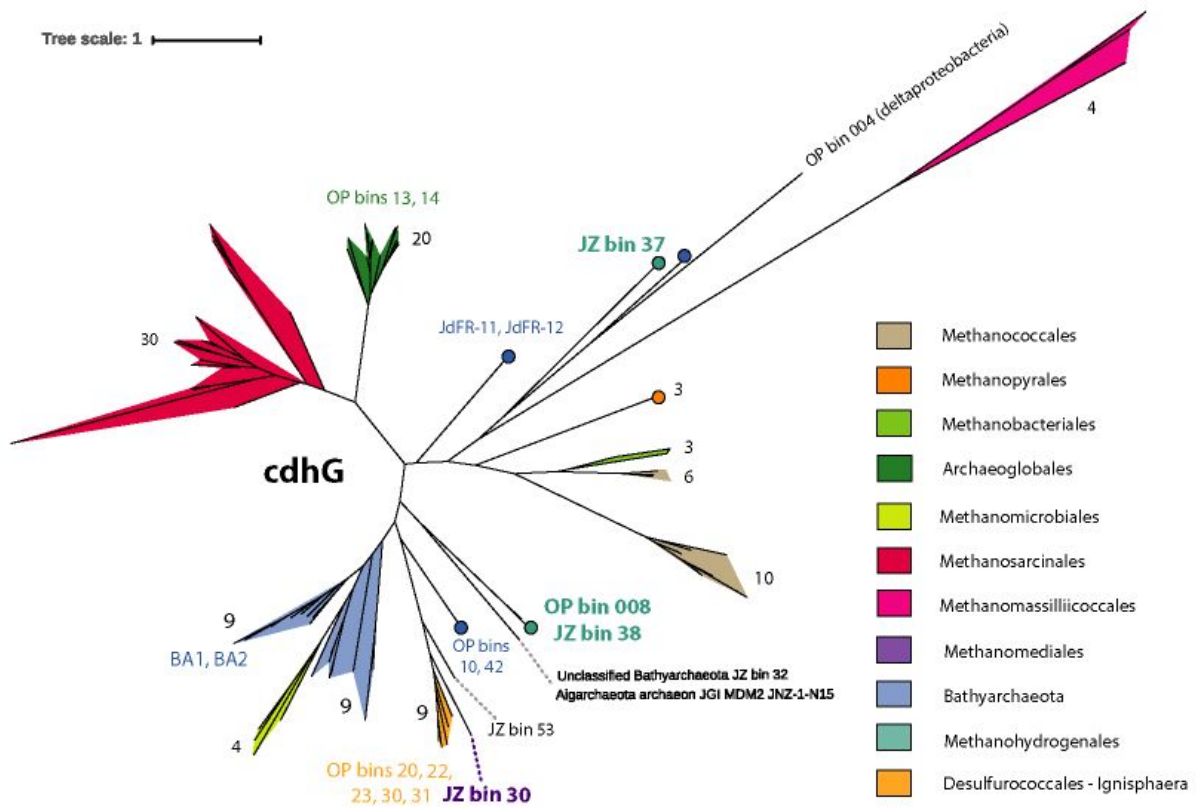


Fig. S13. Tree of CdhG.

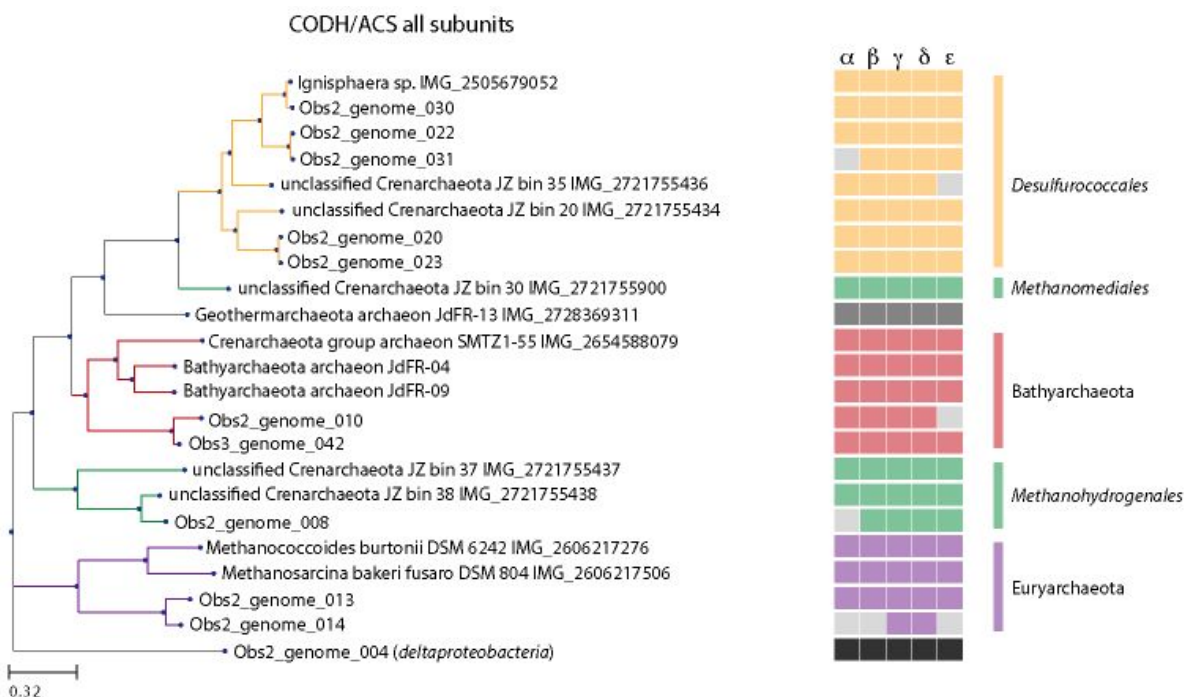


Figure S14. Phylogeny of Obsidian Pool genomes containing Cdh, as calculated from the concatenated alignment of the subunits.

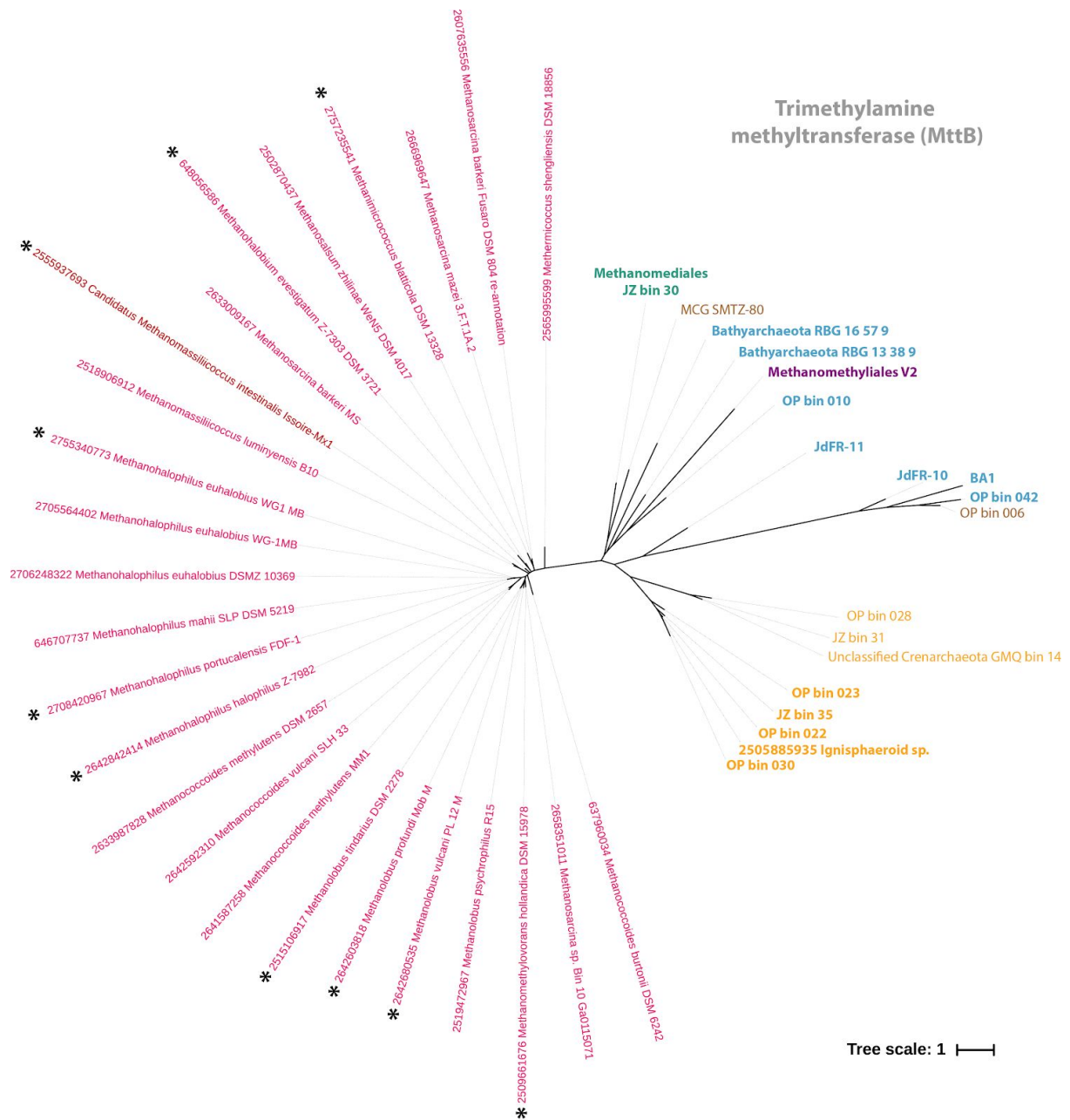


Figure S15. Gene tree of trimethylamine methyltransferase (MttB). The Euryarchaeota (*Methanosarcinales* (pink) and *Methanomassiliicoccales* (red)) form a tight clade, whereas members of the TACK superphylum are more divergent. Shown are *Methanomediales* (green), *Methanomethyliales* (purple), Bathyarchaeota (blue), *Ignisphaera* (yellow, bold), other *Desulfurococcales* (yellow) and other Crenarchaeota (brown). Amino acid sequences containing pyrrolysine (O) are indicated with a *.