

Figure S1

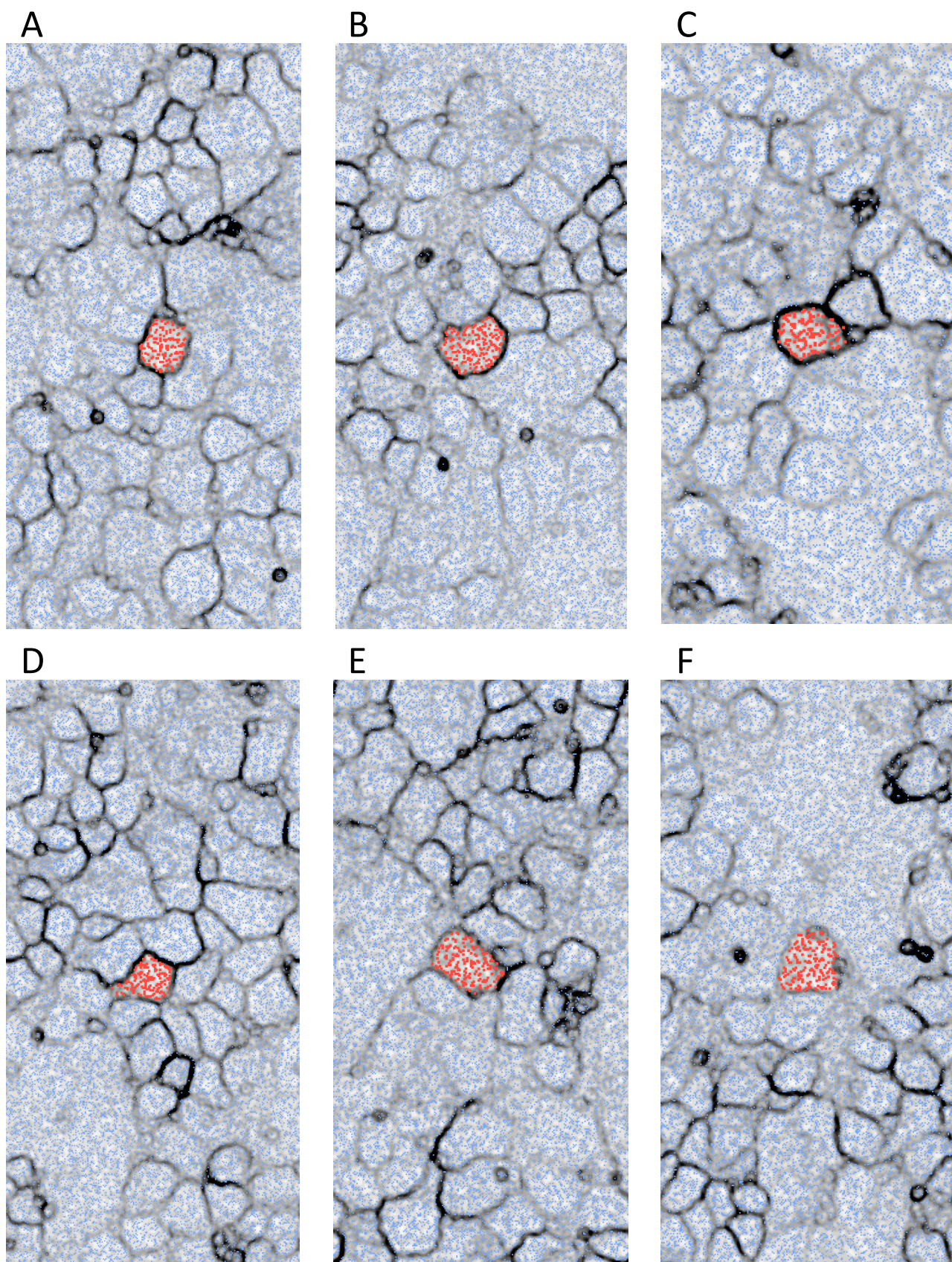
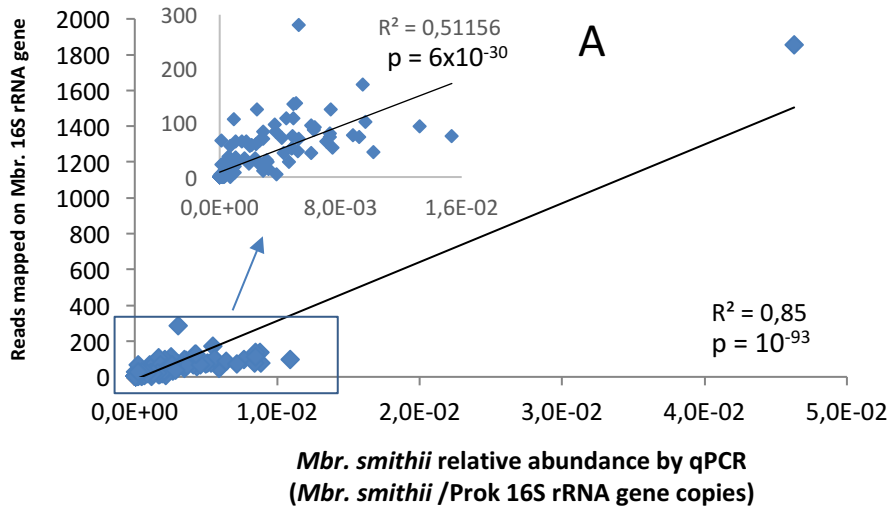


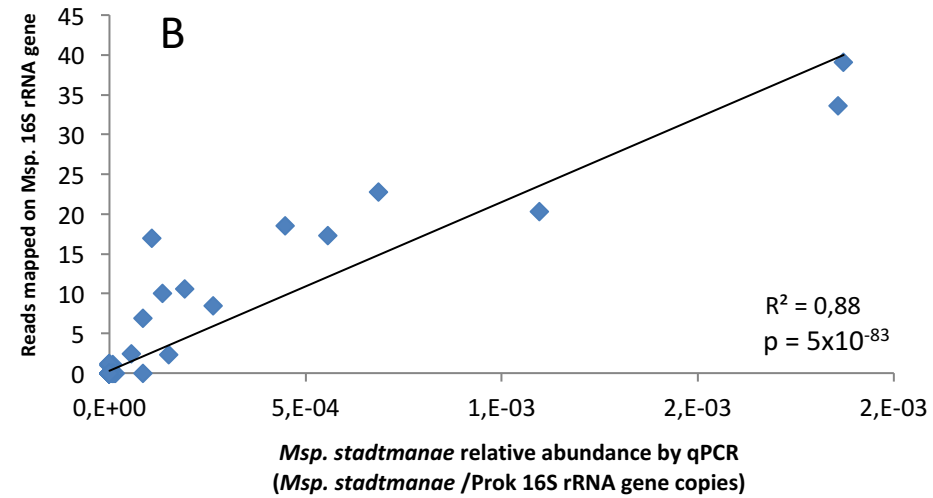
Figure S1: ESOM maps generated from 6 metagenomic assemblies (related to Figure 2). Dots correspond to contigs or contig fragments (5-10 kb). Contigs forming Methanomassiliicoccales bins are colored in red, others are in blue. A) Methanomassiliicoccales Mx-02, B) Methanomassiliicoccales Mx-03, C) Methanomassiliicoccales Mx-06, D) "Ca. Methanomethylophilus alvus 271", E) "Ca. Methanomassiliicoccus intestinalis 138" and F) "Ca. Methanomassiliicoccus intestinalis 183".

Figure S2

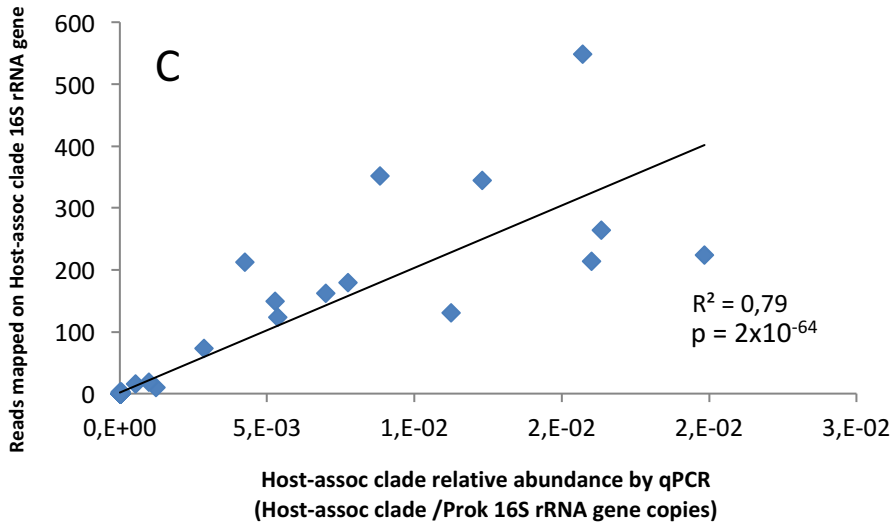
Mbr. smithii



Msp. stadtmanae



“Host-associated clade” Methanomassiliicoccales



“Free-living clade” Methanomassiliicoccales

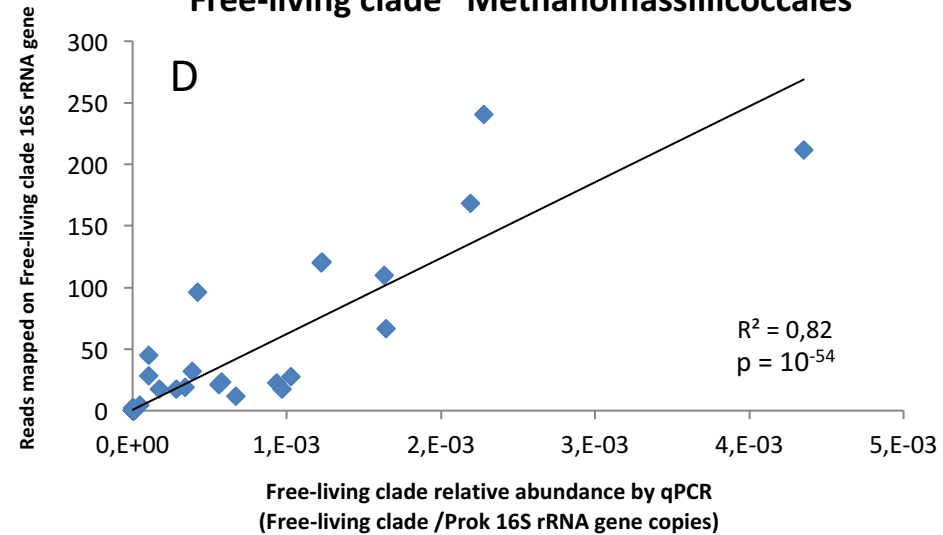


Figure S2: Cross-evaluation of read mapping and quantitative PCR approaches (related to Experimental Procedures). Correlations between the number of reads mapped on 16S rRNA genes and the relative abundance of corresponding 16S rRNA genes determined by qPCR for *Methanobrevibacter smithii* (A), *Methanosphaera stadtmanae* (B), the “Host-associated clade” Methanomassiliicoccales (C) and the “Free-living clade” Methanomassiliicoccales (D).

Figure S3

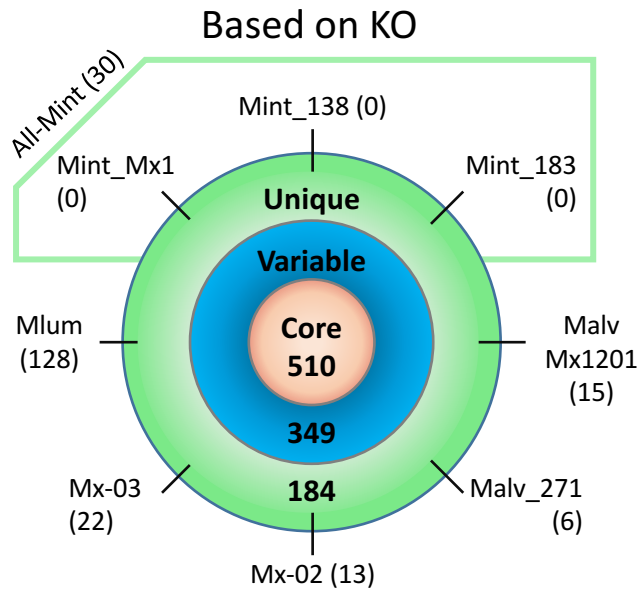


Figure S3 : Core, variable and unique gene families having a KO annotation in the pan-genome of the human-associated Methanomassiliococcales (related to Figure 2). The number of unique gene families within each genome is given between brackets. The number of unique gene families is also given for the pan-genome of "Ca. Mmc. intestinalis" as the three genomes of this species are very similar. Mint, "Ca. Mmc. intestinalis"; Malv, "Ca. Mmp. alvus"; Mlum, *Mmc. luminyensis*.

Figure S4

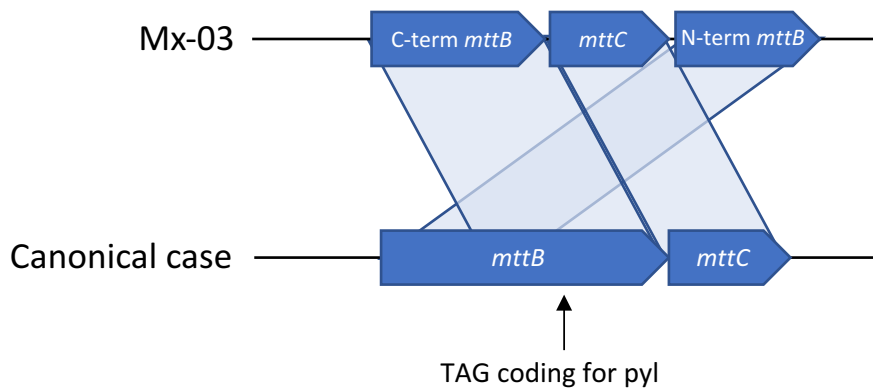


Figure S4: Map of the *mttB* and *mttC* genes in Mx-03 (upper) and the canonical *mttBC* gene cluster in Methanomassiliicoccales (below). The shaded connections between Mx-03 and canonical *mttBC* genes indicate the regions that aligned between both gene clusters. In Mx-03, a gene coding for the N-terminal part of MttB is situated upstream *mttC*. The position of the amber codon (TAG) coding for pyrrolysine (pyl) in canonical *mttB* genes is indicated by the arrow. This codon is not present in the gene coding for MttB C-terminal region of Mx-03.

Figure S5

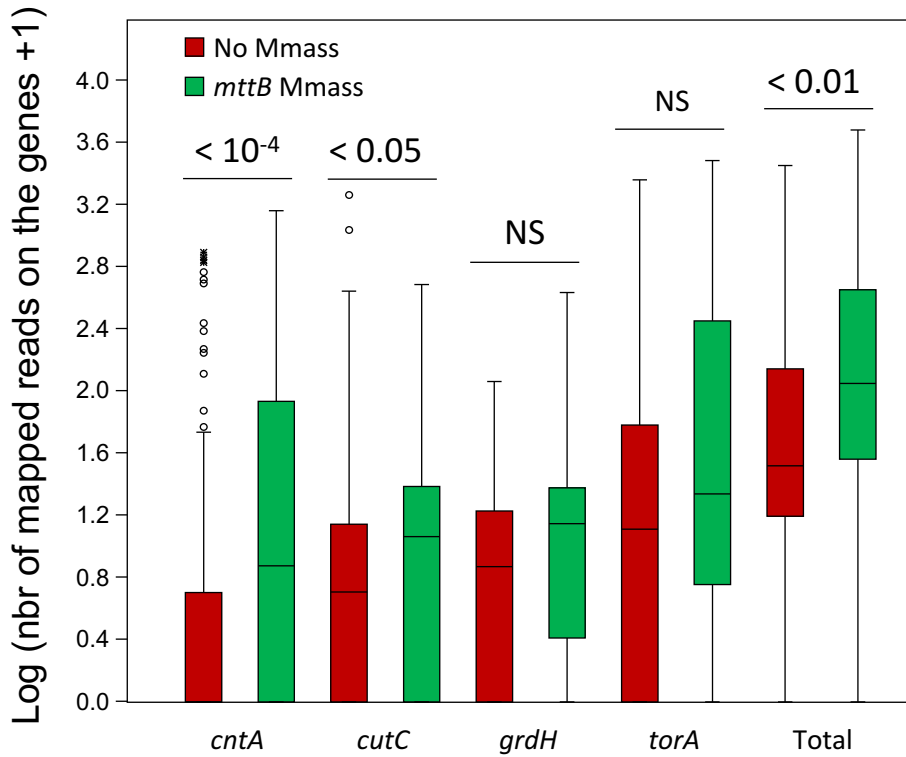


Figure S5: Level of four TMA-production marker genes in subjects without Methanomassiliicoccales (“No Mmass”, in red) and subjects harboring Methanomassiliicoccales with a predicted capacity to utilize TMA (“*mttB* Mmas”, in green) (related to Figure 3) . Level of each marker was determined by mapping reads of ELDERMET metagenomes on the marker genes present in the ELDERMET metagenomes and publically available genomes. “Total” refer to the sum of the reads mapped on all four marker genes. Mann-Whitney test was used to determine significant differences between “No Mmass” subjects and “*mttB* Mmass” subjects.

Figure S6

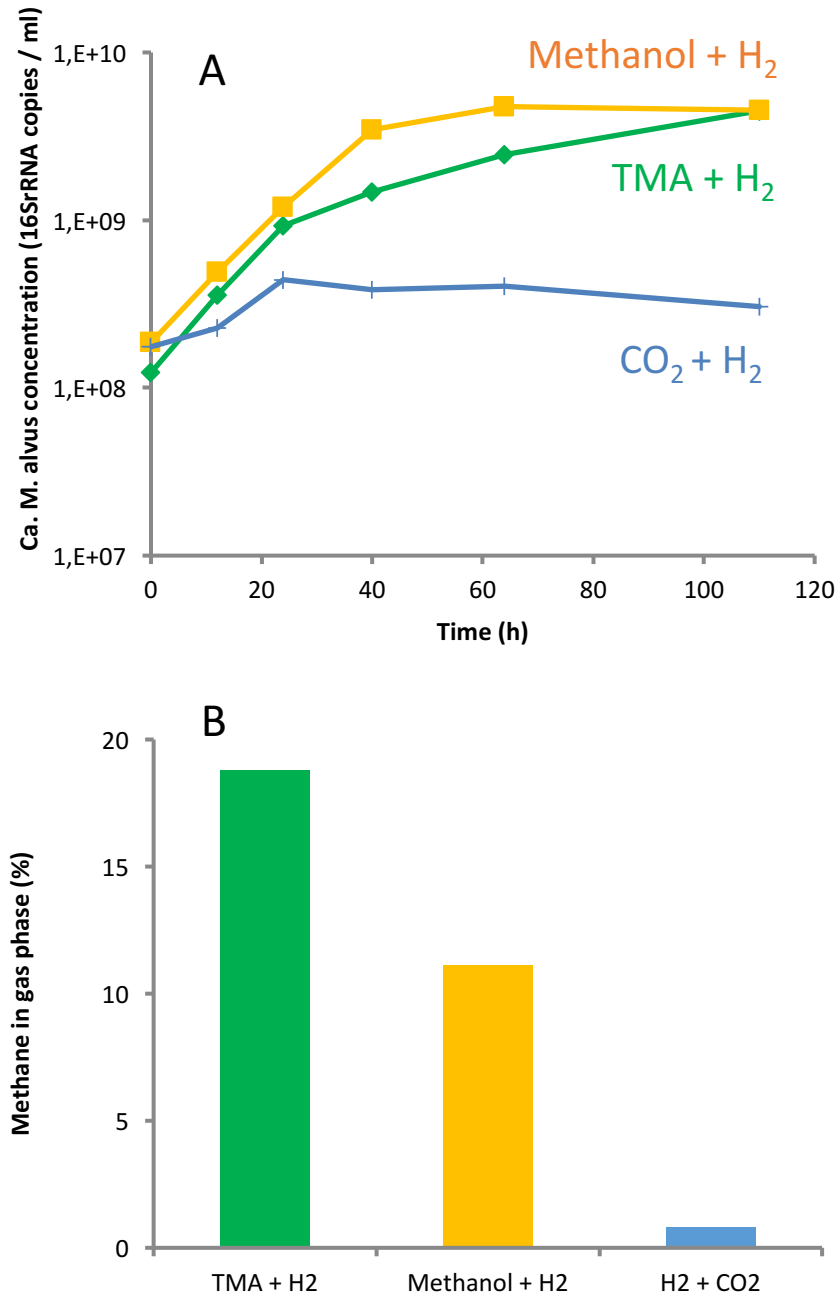


Figure S6: Growth (A) and methane production (B) of “*Candidatus Methanomethylophilus alvus*” in enrichment cultures supplemented with TMA+H₂, methanol+H₂ or H₂+CO₂ (related to Figure 3). “*Ca. Mmp. alvus*” is the only archaea detected in the culture and is mixed with a complex community of bacteria. Specific growth of “*Ca. Mmp. alvus*” was determined by qPCR using primers targeting the “Host-associated clade” Methanomassiliicoccales. The slight growth and methane production in absence of added methyl-compound but in presence of H₂ and CO₂ in the gas phase is attributed to the presence of methyl-compound remaining from the previous growth of “*Ca. Mmp. alvus*”. Values are averages of duplicate growth.

Figure S7

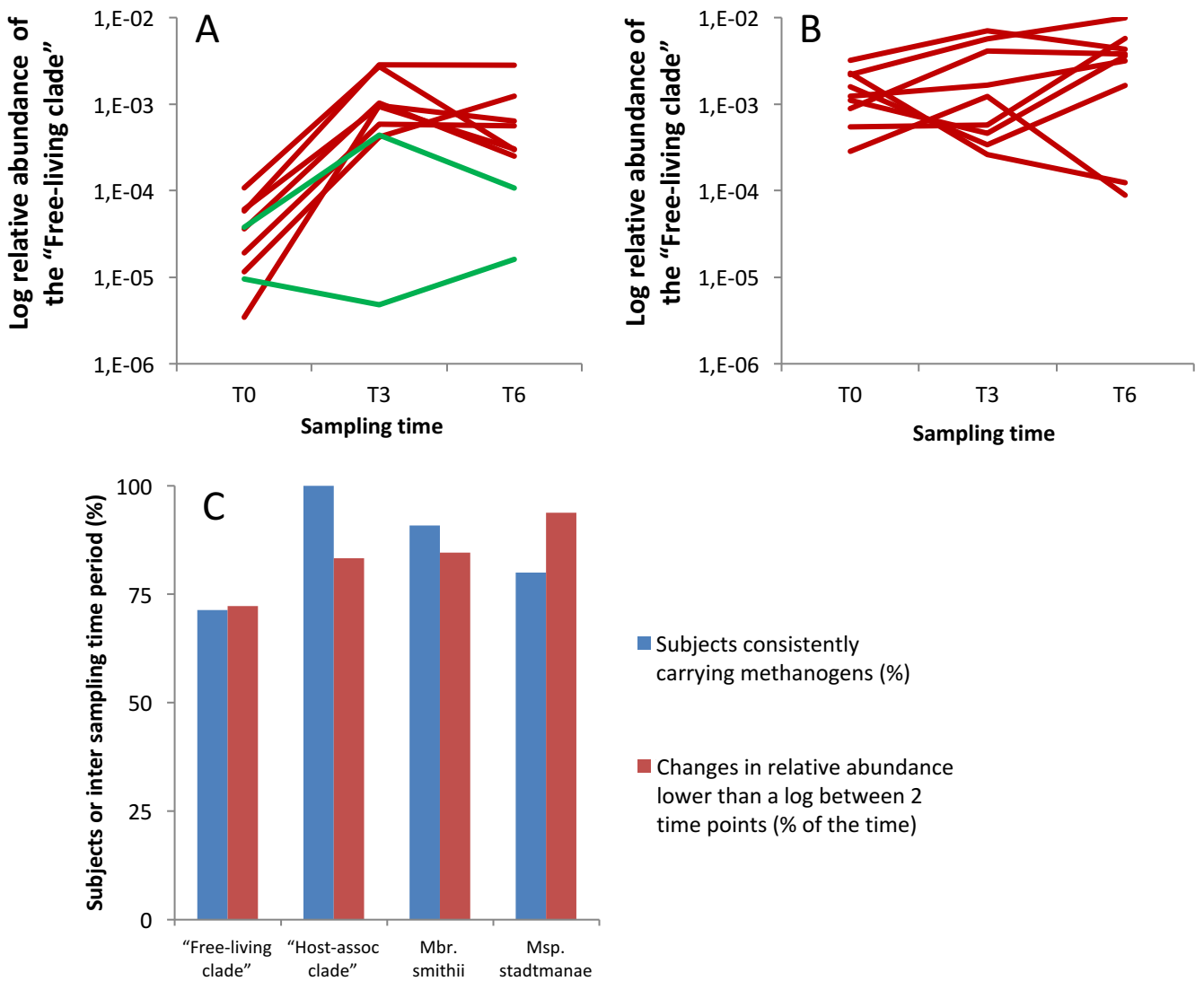


Figure S7: Time stability of Methanomassiliococcales and Methanobacteriales among subjects sampled at three time points over 6 months (related to Figure 4). A) Changes in relative abundance of "Free-living clade" Methanomassiliococcales in subjects having less than 0.01% of these methanogens as a proportion of the whole microbiota at T0 and B) more than 0.01% at T0. Red curves correspond to subjects in long-term residential care and green curves correspond to community-dwelling subjects. C) Stability of colonization in term of carriage (blue) or relative abundance (red).

Figure S8

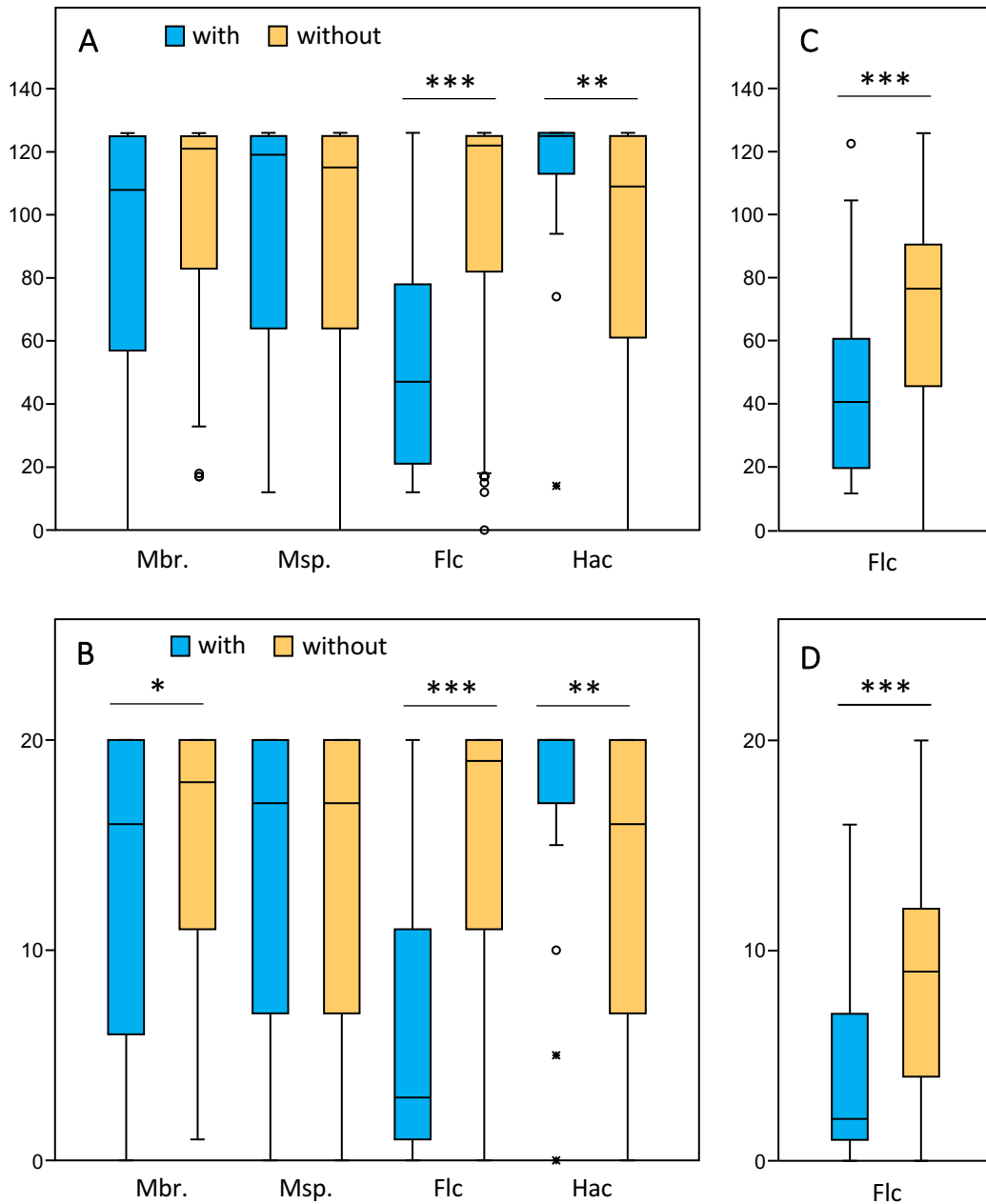


Figure S8: Frailty level measured with FIM index (A-C) and Barthel Score (B-D) among subjects with or without methanogens. A) and B) is a comparison among all subjects, and C and D is a comparison among subjects in residential care for the “Free-living clade”. Differences in FIM index and Barthel score were determined by a generalized linear model of standard gaussian family, in order to adjust for the age of the subjects. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Except for the “Free-living clade” among subjects in residential care (B-D), no differences were observed in the level of FIM index and Barthel score according to the presence/absence of methanogens when subjects were stratified for location. Mbr., *Methanobrevibacter smithii*; Msp., *Methanosphaera stadtmanae*; Flc, “Free-living clade”; Hac, “Host-associated clade”.

Figure S9

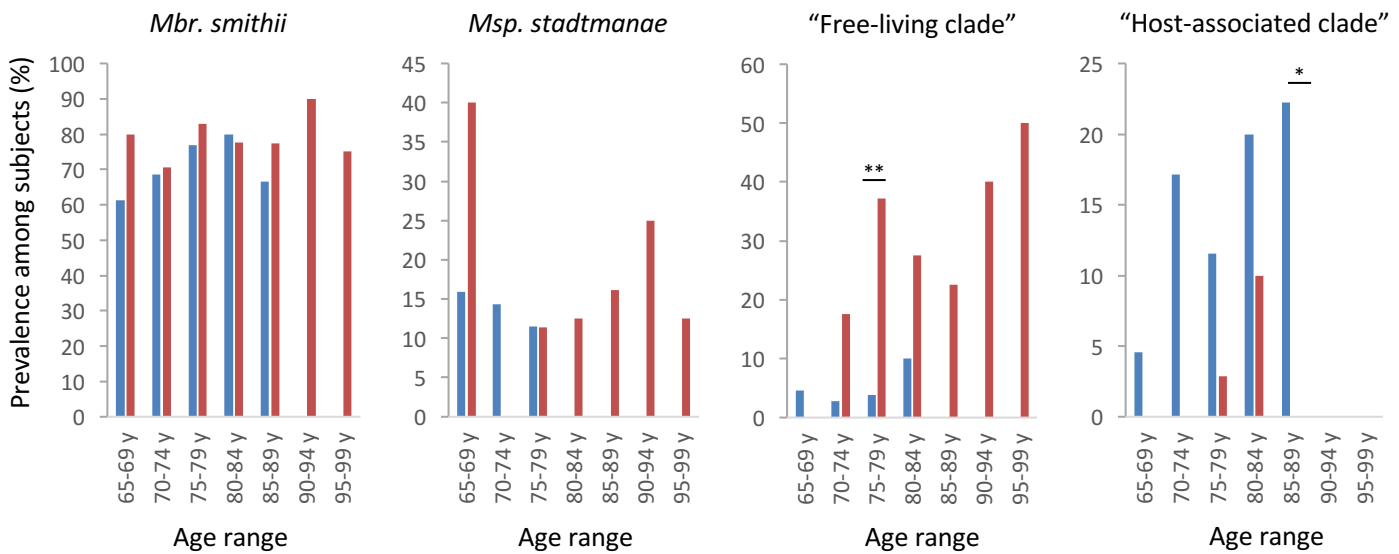


Figure S9: Prevalence of methanogens by class of age (5 year range) among community dwellers (blue) and residential care dwellers (red). 65-69y: Com n=44, LS n=5; 70-74y: Com n=17, LS n=35; 75-79y: Com n=35, LS n=26; 80-84y: Com n=40, LS n=10; 85-89y: Com n=31, LS n=9; 90-94y: LS n=20; 95-99y, Com n=8. There is no community dwelling subject being 90 to 99 year old. Differences in prevalence of the methanogen species/lineages between community and residential care dwellers of same age were determined with Fisher exact test. * p < 0.05, ** p < 0.01.

Figure S10

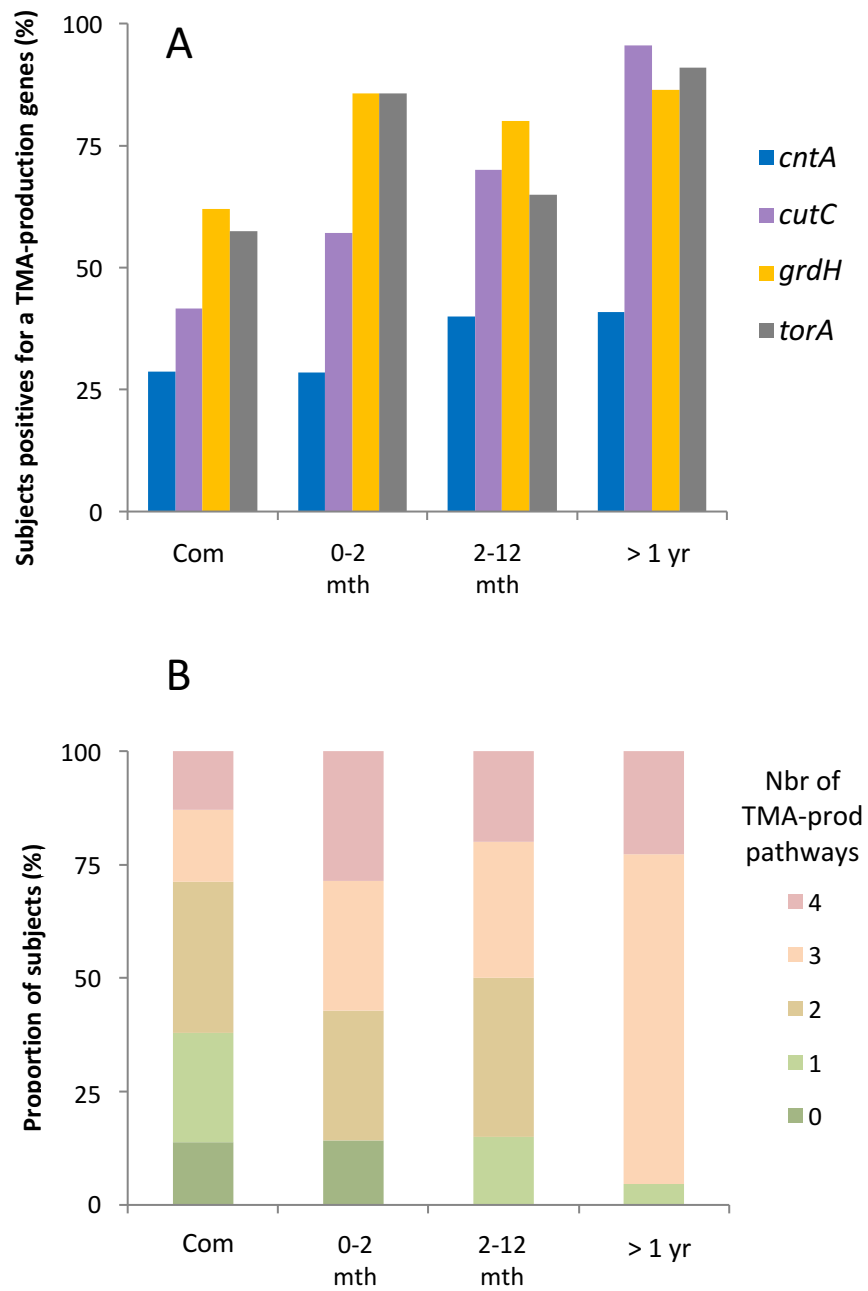


Figure S10: Increase in TMA-production potential of the gut microbiota with the time spent in long-term residential care (related to Figure 4). Change in the prevalence of the four TMA-production pathways (A) and the number of distinct TMA-production pathways per subject (B) according to the time subject spent in residential care. The presence of each of the four TMA-production pathways is predicted by the detection an associated marker genes (*cntA*, *cutC*, *grdH*, *torA*) in the gut microbiota metagenomes of the subjects.

Table S1: Genes present in single copy in all Diaforarchaea (related to Experimental Procedures).

Annotation	K number	Protein accession number in " <i>Ca. Mmc. alvus</i> "
nuoB; NADH-quinone oxidoreductase subunit B [EC:1.6.5.3]	K00331	YP_007712876.1
nuoD; NADH-quinone oxidoreductase subunit D [EC:1.6.5.3]	K00333	YP_007712878.1
nuoM; NADH-quinone oxidoreductase subunit M [EC:1.6.5.3]	K00342	YP_007712885.1
TRMT1; tRNA (guanine26-N2/guanine27-N2)-dimethyltransferase [EC:2.1.1.215 2.1.1.216]	K00555	YP_007713660.1
DPH5; diphthine synthase [EC:2.1.1.98]	K00586	YP_007713827.1
purH; phosphoribosylaminoimidazolecarboxamide formyltransferase / IMP cyclohydrolase [EC:2.1.2.3 3.5.4.10]	K00602	YP_007712853.1
pyrB; aspartate carbamoyltransferase catalytic subunit [EC:2.1.3.2]	K00609	YP_007713857.1
pyrI; aspartate carbamoyltransferase regulatory subunit	K00610	YP_007713858.1
purF; amidophosphoribosyltransferase [EC:2.4.2.14]	K00764	YP_007714008.1
ribE; riboflavin synthase [EC:2.5.1.9]	K00793	YP_007713679.1
ribH; 6,7-dimethyl-8-ribityllumazine synthase [EC:2.5.1.78]	K00794	YP_007713680.1
folP; dihydropteroate synthase [EC:2.5.1.15]	K00796	YP_007714135.1
ppnK; NAD+ kinase [EC:2.7.1.23]	K00858	YP_007714200.1
adk; adenylate kinase [EC:2.7.4.3]	K00939	YP_007714031.1
ndk; nucleoside-diphosphate kinase [EC:2.7.4.6]	K00940	YP_007712788.1
PRPS; ribose-phosphate pyrophosphokinase [EC:2.7.6.1]	K00948	YP_007712821.1
nadM; nicotinamide-nucleotide adenyltransferase [EC:2.7.7.1]	K00952	YP_007713238.1
E3.1.27.9; tRNA-intron endonuclease, archaea type [EC:3.1.27.9]	K01170	YP_007712711.1
E3.3.1.1; adenosylhomocysteinase [EC:3.3.1.1]	K01251	YP_007712738.1
URA4; dihydroorotase [EC:3.5.2.3]	K01465	YP_007714012.1
folD; methylenetetrahydrofolate dehydrogenase (NADP+) / methenyltetrahydrofolate cyclohydrolase [EC:1.5.1.5 3.5.4.9]	K01491	YP_007712870.1
comEB; dCMP deaminase [EC:3.5.4.12]	K01493	YP_007712685.1
purE; 5-(carboxyamino)imidazole ribonucleotide mutase [EC:5.4.99.18]	K01588	YP_007712717.1
E2.3.3.10; hydroxymethylglutaryl-CoA synthase [EC:2.3.3.10]	K01641	YP_007713226.1
ENO; enolase [EC:4.2.1.11]	K01689	YP_007713145.1
purB; adenylosuccinate lyase [EC:4.3.2.2]	K01756	YP_007713435.1
rpe; ribulose-phosphate 3-epimerase [EC:5.1.3.1]	K01783	YP_007713132.1
TPI; triosephosphate isomerase (TIM) [EC:5.3.1.1]	K01803	YP_007713652.1
YARS; tyrosyl-tRNA synthetase [EC:6.1.1.1]	K01866	YP_007712964.1
WARS; tryptophanyl-tRNA synthetase [EC:6.1.1.2]	K01867	YP_007714047.1
TARS; threonyl-tRNA synthetase [EC:6.1.1.3]	K01868	YP_007713905.1
IARS; isoleucyl-tRNA synthetase [EC:6.1.1.5]	K01870	YP_007713130.1
VARs; valyl-tRNA synthetase [EC:6.1.1.9]	K01873	YP_007713752.1
MARS; methionyl-tRNA synthetase [EC:6.1.1.10]	K01874	YP_007713323.1
SARS; seryl-tRNA synthetase [EC:6.1.1.11]	K01875	YP_007713036.1
GARS; glycyl-tRNA synthetase [EC:6.1.1.14]	K01880	YP_007713750.1
PARs; prolyl-tRNA synthetase [EC:6.1.1.15]	K01881	YP_007712822.1
CARS; cysteinyl-tRNA synthetase [EC:6.1.1.16]	K01883	YP_007713696.1
EARS; glutamyl-tRNA synthetase [EC:6.1.1.17]	K01885	YP_007713671.1
RARS; arginyl-tRNA synthetase [EC:6.1.1.19]	K01887	YP_007714040.1
FARSb; phenylalanyl-tRNA synthetase beta chain [EC:6.1.1.20]	K01890	YP_007713453.1
HARS; histidyl-tRNA synthetase [EC:6.1.1.21]	K01892	YP_007713906.1
purC; phosphoribosylaminoimidazole-succinocarboxamide synthase [EC:6.3.2.6]	K01923	YP_007713241.1
purM; phosphoribosylformylglycinamide cyclo-ligase [EC:6.3.3.1]	K01933	YP_007714302.1
fhs; formate--tetrahydrofolate ligase [EC:6.3.4.3]	K01938	YP_007713710.1
purA; adenylosuccinate synthase [EC:6.3.4.4]	K01939	YP_007713662.1
ATPVA; V/A-type H+-transporting ATPase subunit A [EC:3.6.3.14]	K02117	YP_007712692.1
ATPVB; V/A-type H+-transporting ATPase subunit B	K02118	YP_007712693.1
ATPVC; V/A-type H+-transporting ATPase subunit C	K02119	YP_007712690.1
ATPVD; V/A-type H+-transporting ATPase subunit D	K02120	YP_007712694.1
ATPVE; V/A-type H+-transporting ATPase subunit E	K02121	YP_007712689.1
ATPVF; V/A-type H+-transporting ATPase subunit F	K02122	YP_007712691.1
ATPVI; V/A-type H+-transporting ATPase subunit I	K02123	YP_007712687.1
DPB1; DNA polymerase II large subunit [EC:2.7.7.7]	K02322	YP_007713023.1
DPB2; DNA polymerase II small subunit [EC:2.7.7.7]	K02323	YP_007712696.1
rlmE; 23S rRNA (uridine2552-2'-O)-methyltransferase [EC:2.1.1.166]	K02427	YP_007714301.1
rdgB; XTP/dITP diphosphohydrolase [EC:3.6.1.66]	K02428	YP_007714103.1
gyrA; DNA gyrase subunit A [EC:5.99.1.3]	K02469	YP_007714305.1
gyrB; DNA gyrase subunit B [EC:5.99.1.3]	K02470	YP_007714306.1
ksgA; 16S rRNA (adenine1518-N6/adenine1519-N6)-dimethyltransferase [EC:2.1.1.182]	K02528	YP_007714186.1
nusG; transcriptional antiterminator NusG	K02601	YP_007712725.1
PRI; DNA primase [EC:2.7.7.-]	K02683	YP_007714324.1
ribB; 3,4-dihydroxy 2-butanone 4-phosphate synthase [EC:4.1.99.12]	K02858	YP_007713677.1

Table S1 (continued)

Annotation	K number	Protein accession number in "Ca. Mmc. alvus"
RP-L1; large subunit ribosomal protein L1	K02863	YP_007712730.1
RP-L10; large subunit ribosomal protein L10	K02864	YP_007712731.1
RP-L10e; large subunit ribosomal protein L10e	K02866	YP_007713039.1
RP-L11; large subunit ribosomal protein L11	K02867	YP_007712726.1
RP-L12; large subunit ribosomal protein L12	K02869	YP_007712732.1
RP-L13; large subunit ribosomal protein L13	K02871	YP_007714125.1
RP-L14; large subunit ribosomal protein L14	K02874	YP_007714066.1
RP-L15; large subunit ribosomal protein L15	K02876	YP_007714054.1
RP-L15e; large subunit ribosomal protein L15e	K02877	YP_007714189.1
RP-L18; large subunit ribosomal protein L18	K02881	YP_007714057.1
RP-L18e; large subunit ribosomal protein L18e	K02883	YP_007714126.1
RP-L19e; large subunit ribosomal protein L19e	K02885	YP_007714058.1
RP-L2; large subunit ribosomal protein L2	K02886	YP_007714073.1
RP-L21e; large subunit ribosomal protein L21e	K02889	YP_007714183.1
RP-L22; large subunit ribosomal protein L22	K02890	YP_007714071.1
RP-L23; large subunit ribosomal protein L23	K02892	YP_007714074.1
RP-L24; large subunit ribosomal protein L24	K02895	YP_007714065.1
RP-L3; large subunit ribosomal protein L3	K02906	YP_007714076.1
RP-L30; large subunit ribosomal protein L30	K02907	YP_007714055.1
RP-L32e; large subunit ribosomal protein L32e	K02912	YP_007714059.1
RP-L40e; large subunit ribosomal protein L40e	K02927	YP_007713143.1
RP-L44e; large subunit ribosomal protein L44e	K02929	YP_007712954.1
RP-L4e; large subunit ribosomal protein L4e	K02930	YP_007714075.1
RP-L5; large subunit ribosomal protein L5	K02931	YP_007714063.1
RP-L6; large subunit ribosomal protein L6	K02933	YP_007714060.1
RP-L7Ae; large subunit ribosomal protein L7Ae	K02936	YP_007712785.1
RP-S10; small subunit ribosomal protein S10	K02946	YP_007713000.1
RP-S11; small subunit ribosomal protein S11	K02948	YP_007713021.1
RP-S12; small subunit ribosomal protein S12	K02950	YP_007712996.1
RP-S15; small subunit ribosomal protein S15	K02956	YP_007714203.1
RP-S17e; small subunit ribosomal protein S17e	K02962	YP_007714261.1
RP-S19; small subunit ribosomal protein S19	K02965	YP_007714072.1
RP-S19e; small subunit ribosomal protein S19e	K02966	YP_007713460.1
RP-S24e; small subunit ribosomal protein S24e	K02974	YP_007713107.1
RP-S27Ae; small subunit ribosomal protein S27Ae	K02977	YP_007713108.1
RP-S27e; small subunit ribosomal protein S27e	K02978	YP_007712953.1
RP-S28e; small subunit ribosomal protein S28e	K02979	YP_007712786.1
RP-S3Ae; small subunit ribosomal protein S3Ae	K02984	YP_007714016.1
RP-S4; small subunit ribosomal protein S4	K02986	YP_007713020.1
RP-S4e; small subunit ribosomal protein S4e	K02987	YP_007714064.1
RP-S5; small subunit ribosomal protein S5	K02988	YP_007714056.1
RP-S6e; small subunit ribosomal protein S6e	K02991	YP_007712791.1
RP-S7; small subunit ribosomal protein S7	K02992	YP_007712997.1
RP-S8; small subunit ribosomal protein S8	K02994	YP_007714061.1
RP-S8e; small subunit ribosomal protein S8e	K02995	YP_007712985.1
RP-S9; small subunit ribosomal protein S9	K02996	YP_007714124.1
rpoA1; DNA-directed RNA polymerase subunit A' [EC:2.7.7.6]	K03041	YP_007714284.1
rpoA2; DNA-directed RNA polymerase subunit A" [EC:2.7.7.6]	K03042	YP_007714283.1
rpoD; DNA-directed RNA polymerase subunit D [EC:2.7.7.6]	K03047	YP_007713022.1
rpoE1; DNA-directed RNA polymerase subunit E' [EC:2.7.7.6]	K03049	YP_007713104.1
rpoE2; DNA-directed RNA polymerase subunit E" [EC:2.7.7.6]	K03050	YP_007713105.1
rpoF; DNA-directed RNA polymerase subunit F [EC:2.7.7.6]	K03051	YP_007714184.1
rpoL; DNA-directed RNA polymerase subunit L [EC:2.7.7.6]	K03056	YP_007713840.1
tfs; transcription elongation factor	K03057	YP_007712858.1
secY; preprotein translocase subunit SecY	K03076	YP_007714053.1
SRP19; signal recognition particle subunit SRP19	K03105	YP_007712984.1
SRP54; signal recognition particle subunit SRP54	K03106	YP_007713445.1
ftsY; fused signal recognition particle receptor	K03110	YP_007714327.1
EIF1; translation initiation factor 1	K03113	YP_007714068.1
topA; DNA topoisomerase I [EC:5.99.1.2]	K03168	YP_007713817.1
ubiX; 3-octaprenyl-4-hydroxybenzoate carboxy-lyase UbiX [EC:4.1.1.1-]	K03186	YP_007713582.1
EEF1A; elongation factor 1-alpha	K03231	YP_007712999.1
EEF1B; elongation factor 1-beta	K03232	YP_007712748.1
EEF2; elongation factor 2	K03234	YP_007712998.1
EIF1A; translation initiation factor 1A	K03236	YP_007714179.1
EIF2S1; translation initiation factor 2 subunit 1	K03237	YP_007712952.1
EIF2S2; translation initiation factor 2 subunit 2	K03238	YP_007712994.1
EIF2S3; translation initiation factor 2 subunit 3	K03242	YP_007712792.1

Table S1 (continued)

Annotation	K number	Protein accession number in "Ca. Mmc. alvus"
EIF5B; translation initiation factor 5B	K03243	YP_007712789.1
EIF5A; translation initiation factor 5A	K03263	YP_007712907.1
EIF6; translation initiation factor 6	K03264	YP_007713464.1
ETF1; peptide chain release factor subunit 1	K03265	YP_007714041.1
psmB; proteasome beta subunit [EC:3.4.25.1]	K03433	YP_007713580.1
rnhB; ribonuclease HII [EC:3.1.26.4]	K03470	YP_007714082.1
MBF1; putative transcription factor	K03627	YP_007712834.1
RRP4; exosome complex component RRP4	K03679	YP_007712974.1
dnaJ; molecular chaperone DnaJ	K03686	YP_007714242.1
GRPE; molecular chaperone GrpE	K03687	YP_007714244.1
helS; helicase [EC:3.6.4.-]	K03726	YP_007713235.1
lonB; Lon-like ATP-dependent protease [EC:3.4.21.-]	K04076	YP_007713237.1
radA; DNA repair protein RadA	K04483	YP_007712935.1
flpA; fibrillar-like pre-rRNA processing protein	K04795	YP_007714201.1
FEN1; flap endonuclease-1 [EC:3.-.-.]	K04799	YP_007713131.1
rfcL; replication factor C large subunit	K04800	YP_007713483.1
rfcS; replication factor C small subunit	K04801	YP_007713702.1
PCNA; proliferating cell nuclear antigen	K04802	YP_007712859.1
scpA; segregation and condensation protein A	K05896	YP_007712898.1
scpB; segregation and condensation protein B	K06024	YP_007712899.1
truA; tRNA pseudouridine38-40 synthase [EC:5.4.99.12]	K06173	YP_007714275.1
ABCE1; ATP-binding cassette, sub-family E, member 1	K06174	YP_007713319.1
truD; tRNA pseudouridine13 synthase [EC:5.4.99.27]	K06176	YP_007713841.1
PDCD5; programmed cell death protein 5	K06875	YP_007713461.1
NTPCR; nucleoside-triphosphatase [EC:3.6.1.15]	K06928	YP_007713659.1
tiaS; tRNA(Ile2)-agmatinylcytidine synthase [EC:6.3.4.22]	K06932	YP_007714331.1
Predicted GTPase, probable translation factor [Translation, ribosomal structure and biogenesis]	K06942	YP_007714187.1
NOG1; nucleolar GTP-binding protein	K06943	YP_007713024.1
Predicted GTPase [General function prediction only]	K06944	YP_007713913.1
PELO; protein pelota	K06965	YP_007713169.1
TatD-related deoxyribonuclease	K07049	YP_007714323.1
TRM12; tRNA wybutosine-synthesizing protein 2 [EC:2.5.1.114]	K07055	YP_007712864.1
UPF0271 protein	K07060	YP_007712890.1
Uncharacterized protein (ATP-grasp superfamily) [General function prediction only]	K07159	YP_007712951.1
atm56; tRNA (cytidine56'-O)-methyltransferase [EC:2.1.1.206]	K07254	YP_007714294.1
TRM61; tRNA (adenine57-N1/adenine58-N1)-methyltransferase [EC:2.1.1.219 2.1.1.220]	K07442	YP_007712824.1
trm-G10; tRNA (guanine10-N2)-dimethyltransferase [EC:2.1.1.213]	K07446	YP_007712934.1
tgtA2; archaeosine synthase [EC:2.6.1.97]	K07557	YP_007713856.1
dph2; 2-(3-amino-3-carboxypropyl)histidine synthase [EC:2.5.1.108]	K07561	YP_007713040.1
NMD3; nonsense-mediated mRNA decay protein 3	K07562	YP_007714314.1
tsaC; L-threonylcarbamoyladenylate synthase [EC:2.7.7.87]	K07566	YP_007713015.1
putative nucleotide binding protein	K07572	YP_007714185.1
PUA domain protein	K07575	YP_007713054.1
putative methylase	K07579	YP_007712852.1
rfk; riboflavin kinase, archaea type [EC:2.7.1.161]	K07732	YP_007712828.1
TSR3; pre-rRNA-processing protein TSR3	K09140	YP_007714318.1
gatD; glutamyl-tRNA(Gln) amidotransferase subunit D [EC:6.3.5.7]	K09482	YP_007714272.1
hypothetical protein	K09721	YP_007712722.1
pyrH; uridylate kinase [EC:2.7.4.22]	K09903	YP_007714042.1
NTH; endonuclease III [EC:4.2.99.18]	K10773	YP_007713808.1
DKC1; H/ACA ribonucleoprotein complex subunit 4 [EC:5.4.99.-]	K11131	YP_007714051.1
purN; phosphoribosylglycinamide formyltransferase 1 [EC:2.1.2.2]	K11175	YP_007714175.1
RRP41; exosome complex component RRP41	K11600	YP_007712973.1
folC; dihydrofolate synthase / folylpolyglutamate synthase [EC:6.3.2.12 6.3.2.17]	K11754	YP_007713809.1
mtaD; 5-methylthioadenosine/S-adenosylhomocysteine deaminase [EC:3.5.4.31 3.5.4.28]	K12960	YP_007712939.1
coaBC; phosphopantothenoylecysteine decarboxylase / phosphopantothenate--cysteine ligase [EC:4.1.1.36 6.3.2.5]	K13038	YP_007712818.1
idsA; geranylgeranyl diphosphate synthase, type I [EC:2.5.1.1 2.5.1.10 2.5.1.29]	K13787	YP_007714100.1
rpoB; DNA-directed RNA polymerase subunit B [EC:2.7.7.6]	K13798	YP_007714285.1
RTCB; tRNA-splicing ligase RtcB [EC:6.5.1.3]	K14415	YP_007713056.1
NOP56; nucleolar protein 56	K14564	YP_007712909.1
SDO1; ribosome maturation protein SDO1	K14574	YP_007712975.1
ribL; FAD synthetase	K14656	YP_007713678.1
TRM5; tRNA (guanine37-N1)-methyltransferase [EC:2.1.1.228]	K15429	YP_007713826.1
TYW1; tRNA wybutosine-synthesizing protein 1 [EC:4.1.3.44]	K15449	YP_007713484.1
E2.5.1.89; tritans,polycis-undecaprenyl-diphosphate synthase [geranylgeranyl-diphosphate specific] [EC:2.5.1.89]	K15888	YP_007712873.1
E2.5.1.41; phosphoglycerol geranylgeranyltransferase [EC:2.5.1.41]	K17104	YP_007712932.1
E2.7.8.39; archaetidylinositol phosphate synthase [EC:2.7.8.39]	K17884	YP_007712743.1
AK6; adenylate kinase [EC:2.7.4.3]	K18532	YP_007712742.1

Table S2: Primers used in the study (related to Experimental Procedures).

Targeted taxon	Targeted gene	Primer name	Sequence	Amplicon length (bp)	T _m (°C)	Concentration (μM)	Reference
"Host-associated clade" Methanomassiliicoccales	16S rRNA	Mx-HAC-F	GATTCTGAGACACGAATCCAGG	155	60	0.35	This study
		Mx-HAC-R	CGTCTTACCCAGCCCTTATTC				
"Free-living clade" Methanomassiliicoccales	16S rRNA	Mx-FLC-F	CTGCTGCTTGGGGAGTATGT	135	60	0.35	This study
		Mx-FLC-R	GACCTTCATATATCCGTCGCTC				
<i>Methanomassiliicoccus luminyensis</i>	16S rRNA	Mlum-16S-F	GCGAGTCGAGAGACGTAAGG	131	60	0.2	This study
		Mlum-16S-R	GCCTAGTAGCATTCCAGCATG				
	<i>mcrA</i>	Mlum-mcrA-F	AGAGGAAGAGGGAGTGGGTC	291	63	0.2	This study
		Mlum-mcrA-R	AGTGGTTGATGGTCTCAGGG				
"Ca. Methanomassiliicoccus intestinalis"	16S rRNA	Mint-16S-F	GCGAGTCGAGAGACATTTGG	131	60	0.2	This study
		Mint-16S-R	GCCTAAAAGCATTCCAGCAGA				
	<i>mcrA</i>	Mint-mcrA-F	AGAGAAAAGAGGGAAATGGGTT	291	59	0.2	This study
		Mint-mcrA-R	AGTTGTTGATTGTTTCAGGT				
<i>Methanobrevibacter smithii</i>	16S rRNA	Mbs-955-F	GCCAGGTTGATGACTTTGCTTG	227	60	0.2	This study
		Mbs-1162-R	GCGTGTTGCCAGAGGATTC				
<i>Methanosphaera stadtmanae</i>	16S rRNA	Mss-1057-F	CGCCCTTAGTTACCAACACAATC	154	60	0.35	This study
		Mss-1193-R	CAGAAACCCATTGCCATAGCC				
Bacteria (Universal)	16S rRNA	F_Bact1369	CGGTGAATACGTTCCCGG	145	60	0.2	Sokol et al. 2009
		R_Prok1492	TACGGCTACCTTGTTACGACTT				

Table S3: Presence/absence of genes related to methanogenesis and energy-conservation in all genomes of the human-associated Methanomassiliicoccales

Genes	Notes	Genomes from this study				Previously published genomes				
		"Ca. Mmp. alvus 271"	Mx-03	Mx-02	Mx-06	"Ca. Mmc. Intestinalis 138"	"Ca. Mmc. Intestinalis 183"	"Ca. Mmp. alvus Mx1201"	"Ca. Mmc. intestinalis Mx1"	Mmc. luminyensis
Methanogenesis pathways										
20 methanogenic markers*	they include <i>mcrABGCD</i>	+	+	+	19/20	+	+	+	+	+
<i>fwd/fmd, ftr, mch, mtd/hmd, mer, mtr</i>	reduction of CO ₂ to a methyl-moiety	-	-	-	-	-	-	-	-	-
<i>mtaBC</i>	first step of methanol utilisation	+	+	<i>mtxC</i>	+	+	+	+	+	+
<i>mtmBC</i>	first step of monomethylamine utilization	+	+	<i>mtxC</i>	+	+	+	+	+	+
<i>mtbBC</i>	first step of dimethylamine utilization	+	-	<i>mtxC</i>	+	+	+	+	+	+
<i>mttBC</i>	first step of trimethylamine utilization	+	- †	<i>mtxC</i>	+	+	+	+	+	+
<i>mtsAB</i>	methanethiol/dimethyl sulfite utilization	+	-	-	-	-	-	+	-	+
<i>mtaA/mtbA</i>	second step of mono- di- trimethylamine and/or methanol**	+	+	+	+	+	+	+	+	+
<i>pylSBCD</i>	encode for pyrrolysine, an amino acid present in mono- di- trimethylamine methyltransferase	+	+	-	+	+	+	+	+	+
<i>comDE, serC, CS</i>	coenzyme M synthesis	<i>serC</i>	<i>serC</i>	<i>serC</i>	<i>serC</i>	+	+	+	+	+
<i>aksADEF</i>	coenzyme B synthesis	+	+	-	partial	+	+	+	+	+
Energy conservation										
<i>hdrABC</i>	heterodisulfite reductase/ferredoxin reductase	+	+	+	+	+	+	+	+	+
<i>hdrD</i>	heterodisulfite reductase	+	+	+	+	+	+	+	+	+
<i>hdrE</i>	membrane bound cytochrome	-	-	-	-	-	-	-	-	-
<i>mvhADG</i>	hydrogenase forming a complex with HdrABC	+	+	+	partial	+	+	+	+	+
<i>fpoABCDHIJKLMN</i>	truncated F ₄₂₀ H ₂ hydrogenase (membrane bound) generating a proton gradient	+	+	+	+	+	+	+	+	+
<i>echABCDEF</i>	energy converting hydrogenase (membrane bound) generating a with proton gradient	-	+	-	+	+	+	-	+	+
<i>hyfBCDEFGHIG-like</i>	membrane-bound Hydrogenase 4 (membrane bound) potentially generating a proton gradient	-	-	-	partial	-	-	-	-	+
<i>ahaABCDEFHIK</i>	ATP synthesis using the proton motive force	+	+	+	+	+	+	+	+	+

Presence of all genes is marked by a "+" and absence of all genes by a "-". *mtxC* means that one or several corrinoid proteins are present but their involvement in the utilization of a particular methyl-compound is unclear. The only gene implied in CoM biosynthesis found in Mx-02, Mx-03, Mx-06 and "Ca. Mmc. alvus 271" is an homologue of *serC*. SerC enzyme can catalyze the transamination of other amino-acids (e.g. aspartate, glutamate) than L-cysteine for CoM biosynthesis.

* As defined in Borrel et al., BMC genomics, 2014

** Discussed in Borrel et al., Genome Biology Evolution, 2013

† *mttB* is interrupted and is not coding for pyrrolysine (See Figure S4)

Table S4: Percentage of identity between the TMA-production genes from the ELDERMET metagenomes (at the amino acid level) and the reference sequences from publically available genomes (related to Figure 3). Several ELEREMET sequences could have the same closest reference sequence but with different level of identity, indicated by min and max here. The total number of reads mapped on these ELDERMET sequences is also displayed.

cntA

Closest species	Identity to the closest reference sequence (aa level, %)		number of reads mapped
	min	max	
	Citrobacter freundii	100	
Enterobacter aerogenes	99	99	89
Escherichia coli	94	100	9009
Klebsiella pneumoniae	99	100	2342

cutC

Closest species	Identity to the closest reference sequence (aa level, %)		number of reads mapped
	min	max	
	Anaerococcus hydrogenalis	100	
Clostridium asparagiforme	100	100	435
Clostridium baratii str. Sullivan	89	91	233
Clostridium chauvoei	84	84	18
Clostridium citroniae	100	100	158
Clostridium clostridioforme	82	89	167
Clostridium hathewayi	98	100	501
Clostridium senegalense	74	74	217
Clostridium sporogenes	100	100	10
Collinsella tanakaei	100	100	85
Desulfosporosinus youngiae	52	52	45
Desulfovibrio alaskensis	89	89	28
Desulfovibrio desulfuricans	89	89	24
Dorea sp. 5-2	57	57	168
Enterobacter aerogenes	99	100	1876
Enterococcus haemoperoxidus	56	56	56
Escherichia coli	100	100	1207
Escherichia fergusonii	100	100	74
Klebsiella oxytoca	100	100	94
Klebsiella pneumoniae	99	100	2782
Klebsiella sp. 07A044	100	100	329
Proteus mirabilis	100	100	76
Streptococcus dysgalactiae	90	90	40
Yokenella regensburgei	100	100	8

Table S4 (continued)**grdH**

Closest species	Identity to the closest reference sequence (aa level, %)		number of reads mapped
	min	max	
Brachyspira pilosicoli	94	94	134
Clostridium bolteae ATCC_BAA-613	100	100	501
Clostridium clostridioforme 2_1_49FAA	97	100	1822
Dorea formicigenerans 4_6_53FAA	100	100	1297
Lachnospiraceae bacterium 6_1_37FAA	100	100	1032
Leptotrichia sp.	73	73	4
Olsenella profusa	94	95	29
Parvimonas micra ATCC_33270	100	100	24
Synergistes sp. 3_1	97	100	310

torA

Closest species	Identity to the closest reference sequence (aa level, %)		number of reads mapped
	min	max	
Actinobacillus succinogenes 130Z	89	89	97
Aggregatibacter aphrophilus	91	91	334
Citrobacter freundii	94	100	308
Citrobacter sp. KTE32	97	100	363
Citrobacter werkmanii NBRC 105721	99	99	13
Escherichia albertii	98	100	6
Escherichia coli	83	100	33927*
Escherichia sp. KTE159	99	100	268
Haemophilus haemolyticus	92	94	330
Hafnia alvei	61	100	169
Ignatzschineria larvae	65	66	235
Providencia rettgeri	56	56	361
Shigella flexneri	95	100	2341

* 33847 reads (99.8%) are mapped on sequences with more than 95% identity to *Escherichia coli* TorA

Table S5: Organisms with the highest occurrence of genes coding for Flg_{new} (A) and Sel1-repeat domain (B), their environment of origin and/or their lifestyle (related to Figure 5). Origin and lifestyle is only presented for prokaryotes. Methanomassiliicoccales genomes are in red.

A) Flg_{new} domain

Genome	Gene Count	Origin / lifestyle	Domain
Bacteroidales bacterium PSC KfJoeBact1-1	52	endosymbionts of flagellates within the termite gut	Bacteria
Bacteroidales bacterium PSC KfJoeBact2-2	48	endosymbionts of flagellates within the termite gut	Bacteria
Bacteroidales bacterium PSC KfJoeBact2-3	45	endosymbionts of flagellates within the termite gut	Bacteria
Bacteroidales bacterium PSC KfJoeBact1-2	41	endosymbionts of flagellates within the termite gut	Bacteria
Candidatus Methanomassiliicoccus intestinalis Issoire-Mx1	38	human gut	Archaea
Butyrivibrio sp. MC2021	33	host-associated	Bacteria
Candidatus Methanomassiliicoccus intestinalis 183	31	human gut	Archaea
Anaerostipes caccae DSM 14662	29	host-associated	Bacteria
Candidatus Methanomassiliicoccus intestinalis 138	29	human gut	Archaea
Robinsoniella sp. KNHs210	26		Bacteria
Bifidobacterium asteroides Bin2	25	honey bees hindgut	Bacteria
Bifidobacterium sp. A11	24	honey bees hindgut	Bacteria
Butyrivibrio sp. AE3009	21	rumen	Bacteria
Bifidobacterium asteroides PRL2011	21	honey bees hindgut	Bacteria
Sphaerochaeta coccoides SPN1, DSM 17374	21	termite	Bacteria
Bifidobacterium asteroides ATCC 25910	20	honey bees hindgut	Bacteria
Butyrivibrio sp. VCD2006	20	rumen	Bacteria
Treponema sp. SAG7	19	termite	Bacteria
Acholeplasma axanthum ATCC 25176	18	murine leukemia tissue culture	Bacteria
Listeria monocytogenes sv. 3b SLCC 2540	18	host-associated	Bacteria
Anaerostipes sp. 3_2_56FAA	17	host-associated	Bacteria
Bifidobacterium sp. 7101	16	honey bees hindgut	Bacteria
Listeria monocytogenes sv. 4b CLIP 80459	15	host-associated	Bacteria
Listeria monocytogenes J1816	15	host-associated	Bacteria
Acholeplasma palmae J233, ATCC 49389	15		Bacteria
Listeria monocytogenes L312	15	host-associated	Bacteria
Eubacterium saphenum ATCC 49989	15	human periodontal pocket	Bacteria
Methanomassiliicoccales Mx-03	15	human gut	Archaea
Listeria monocytogenes sv. 1/2b SLCC 2755	14	host-associated	Bacteria
Lachnospiraceae bacterium V9D3004	14	rumen	Bacteria

Table S5 (continued)

B) Sel1-repeat domain

Genome	Gene Count	Origin / lifestyle	Domain
Aureococcus anophagefferens CCMP 1984	124		Eukaryota
Micromonas pusilla CCMP1545	89		Eukaryota
Micromonas pusilla CCMP 490(RCC 114)	72		Eukaryota
Naegleria gruberi NEG-M	70		Eukaryota
Oxalobacter formigenes HOxBLS	51	host-associated/human	Bacteria
Trichomonas vaginalis G3	45		Eukaryota
Candidatus Amoebophilus asiaticus 5a2	42	obligate endosymbiont of Amoeba	Bacteria
Snodgrassella alvi SCGC AB-598-O11	42	honey-bees gut symbiont	Bacteria
Succinimonas amylolytica DSM 2873	40	bovine host	Bacteria
Oxalobacter formigenes OXCC13	40	host-associated/human	Bacteria
Candidatus Odysseella thessalonicensis L13	40	obligate intracellular parasite	Bacteria
Candidatus Paracaedibacter symbiosus	37	obligate intracellular symbiont of Acanthamoeba	Bacteria
Trichoplax adhaerens Grell-BS-1999	32		Eukaryota
Orpinomyces sp. C1A	30		Eukaryota
Allomyces macrogynus ATCC 38327	29		Eukaryota
Candidatus Methanomethylophilus alvus Mx1201	28	human gut	Archaea
Emiliana huxleyi CCMP 1516	27		Eukaryota
Neocallimastix sp. SC4 (gene modules)	25		Eukaryota
Helicobacter felis CS1, ATCC 49179	25	gastric mucosa of cat	Bacteria
Epulopiscium sp. N.t. morphotype B	24	fish gut symbiont	Bacteria
Candidatus Methanomethylophilus alvus 271	23	human gut	Archaea
Cardiobacterium valvarum F0432	22	human dental plaque	Bacteria
Thiomargarita sp. Thio36	22	sediment	Bacteria
Magnetospira sp. QH-2	22	marine	Bacteria
Asticcacaulis sp. AC460	22		Bacteria
Commensalibacter intestini A911	20	gut epithelia	Bacteria
Asticcacaulis biprosthecum C19, ATCC 27554	20	freshwater	Bacteria
Ruminobacter sp. RM87	20	rumen	Bacteria
Thermoplasmatales archaeon BRNA1	19	rumen	Archaea
Helicobacter bizzozeronii CCUG 35545	19	gastric mucosa of dog	Bacteria

Supplemental Experimental Procedures

Cross evaluation of new group specific primers and mapping of metagenomic reads on corresponding 16S rRNA genes

ELDERMET metagenome reads were mapped on 16S rRNA genes of *Methanobrevibacter smithii*, *Methanosphaera stadtmanae* and human-associated Methanomassiliicoccales, including the 6 novel genomes from ELDERMET, *Mmc. luminyensis*, “*Ca. Mmc. intestinalis Mx1*”, and “*Ca. Mmp. alvus Mx1201*”. Read mapping was performed using Bowtie 2 (Langmead and Salzberg, 2012) with the alignment parameters -D 15 -R 2 -L 22 -N 1 -i S,1,1.15 in “end-to-end” mode. The output was normalized according to the total number of reads in each metagenome and compared to the relative abundance of “Host-associated clade” and “Free-living clade” Methanomassiliicoccales, as well as *Mbr. smithii*, *Msp. stadtmanae* determined by quantitative PCR. The relative abundance of the four archaeal groups was significantly correlated with number of reads mapped on the 16S rRNA genes corresponding to these groups (Figure S7), with all $r^2 > 0.79$ and $p < 10^{-54}$ (Pearson linear correlation). This strong correlation highlights the high specificity and coverage of the newly designed primers and the accuracy of read mapping. Discrepancies in detected presence/absence of the groups between the two techniques were only observed for low values. The quantitative PCR approach is complementary to metagenomic approach by providing a good estimation of abundance/relative abundance and allowed the investigation of a larger number of samples. Detection of the presence of Methanomassiliicoccales was thus primarily based on quantitative PCR approach results (probing of “Host-associated clade” and “Free-living clade” Methanomassiliicoccales) rather than metagenome read mapping on Methanomassiliicoccales 16S rRNA genes.

Identification of reference TMA-production marker genes from publically available sources

Four distinct microbial pathways generating TMA from four different substrates (L-carnitine, choline, betaine and trimethylamine-N-oxide) have been previously identified and characterized (Craciun and Balskus, 2012; Koeth et al., 2014; Méjean et al., 1994; Meyer et al., 1995). It should be noted that in the human gut, L-carnitine can be directly converted into TMA or first get converted into γ -butyrobetaine and then into TMA (Koeth et al., 2014). From these pathways, we chose four enzymes expected to be markers of TMA generation: CntA (also referred as YeaW) involved in L-carnitine/ γ -butyrobetaine conversion (Koeth et al., 2014; Zhu et al., 2014), CutC involved in choline conversion (Craciun & Balskus 2012), GrdH involved in betaine conversion (Meyer et al. 1995) and TorA involved in trimethylamine-N-oxide conversion (Méjean et al. 1994). The following sequences of biochemically characterized marker enzymes were identified using MetaCyc (www.metacyc.org) and used as references to query NCBI for further homologs (BlastP): CntA from *Acinetobacter baumannii* ATCC1960 (WP_001277086.1), CutC from *Desulfovibrio alaskensis* G20 (WP_011369019.1), GrdH from *Eubacterium acidaminophilum* DSM3953 (O69406) and TorA from *Escherichia coli* K-12 (NP_415517.1). Homologs of each marker protein were identified in the NCBI database using blastp and in human-associated bacterial genomes sequenced by the Human Microbiome Project (HMP). In order to eliminate more distant homologs which have a different function manual curation was carried out. The presence of conserved residues in the region aligned to the catalytic sites of the reference sequences was used to validate hits. Further validation was made possible for CutC and CntA genes using results from Craciun & Balskus (2012) and Zhu et al. (2014).

Additionally, phylogenetic analyses were carried out to separate the sequences clustering with the reference sequence of each marker and the sequences clustering with more distant homologs which have a characterized different function than the marker. The lowest percent identity between predicted homologs (amino acid level) of a given marker was 65% for CntA/YeaW, 60% for CutC, 56% for GrdH and 50% for TorA.

Identification of TMA-production marker genes in ELDERMET metagenomes

A two-step approach was adopted to identify sequences coding for marker proteins within the 190 ELDERMET metagenomes. Firstly, the reference sequence of each marker gene identified above was blasted (tblastn) against all ELDERMET metagenomes. An identity cut-off 10% below the minimal identity displayed between sequences of each marker (displayed above) was applied. In a second time, in order to ensure there were no false positives, protein sequences with a lower identity to reference marker proteins for TMA-production than to homologous protein having a different function were identified and eliminated.

Abundance of genes involved in TMA modulation

ELDERMET metagenome reads were mapped with Bowtie2 (Langmead & Salzberg 2012) on the marker genes identified in the ELDERMET cohort and in genomes for HMP using the same parameters as described above for 16S rRNA genes. Additionally, reads were mapped on *mttB* genes identified in all genomes of human associated Methanomassiliicoccales. In order to minimize false positive in presence/absence analysis, a gene was considered present when more than 3 reads mapped to it. This threshold was refined based on comparisons of reads mapped on Methanomassiliicoccales 16S rRNA genes and their presence/absence based on quantitative PCR method. The number of reads mapped on marker genes was normalized according to the total number of reads of their metagenome.

Determination of different categories of samples depending on presence/absence of Methanomassiliicoccales, “*mttB* lacking Methanomassiliicoccales” and “*mttB* Methanomassiliicoccales”, and their concentration per gram of stool

Presence of Methanomassiliicoccales was determined by qPCR, as mentioned above. Also, the Methanomassiliicoccales genomes reconstructed from ELDERMET metagenomes revealed that some predominant human-associated Methanomassiliicoccales are likely not able to deplete TMA. In order to evaluate the influence of Methanomassiliicoccales on faecal TMA concentration, we separated samples that contain Methanomassiliicoccales potentially not having the *mttB* gene from samples containing Methanomassiliicoccales having the *mttB* gene. Samples having more than 3 reads mapped on a Methanomassiliicoccales *mttB* were considered as containing “*mttB* Methanomassiliicoccales”. Additionally, because there was a high correlation between reads mapped on “*Ca. Mmc. intestinalis*” *mttB* and their detection by qPCR, we considered that samples without metagenomic sequencing data but having a “*Ca. Mmc. intestinalis*” detected by qPCR as containing “*mttB* Methanomassiliicoccales”. A similar assumption was not extended to samples without metagenomic sequencing data and containing “Host-associated clade” Methanomassiliicoccales as all methanogens of this clade do not consistently have the *mttB* gene. This led to the removal of one sample from the analysis of the impact of Methanomassiliicoccales on faecal TMA concentration. We also considered the possibility that low concentration of “*mttB* Methanomassiliicoccales” could have a lower impact on TMA removal than high

concentration of “*mttB* Methanomassiliicoccales”. This hypothesis was tested using concentration (copies/g of stool) determined by qPCR in samples containing “*mttB* Methanomassiliicoccales”.

Determination of fecal trimethylamine concentration by $^1\text{H-NMR}$

A 250 μL aliquot of fecal water was placed in a 1.5 mL eppendorf tube followed by the addition of 35 μL of D_2O and 65 μL of a standard buffer solution (585 mM NaHPO_4 [pH 7.0], 11.667 mM DSS [disodium-2,2-dimethyl-2-silapentane-5-sulphonate], and 0.47% NaN_3 in H_2O). The samples (350 μL) were then transferred to a shigemi NMR tube for subsequent NMR spectral analysis.

All $^1\text{H-NMR}$ spectra were collected on a Varian 500 MHz Inova spectrometer equipped with a 5 mm HCN Z-gradient pulsed-field gradient (PFG) cryogenic probe. $^1\text{H-NMR}$ spectra were acquired at 25 °C using the first transient of the Varian tnoesy pulse sequence, which was chosen for its high degree of selective water suppression and quantitative accuracy of resonances around the solvent. Water suppression pulses were calibrated to achieve a bandwidth of 80 gauss. Spectra were collected with 128 transient and 8 steady-state scans using a 4 second acquisition time (48,000 complex points) and a 1 second recycle delay.

Prior to spectral analysis, all FIDs were zero-filled to 64,000 data points and line broadened 0.5 Hz. The methyl singlet produced by a known quantity of DSS was used as an internal standard for chemical shift referencing (set to 0 ppm) and for quantification. All $^1\text{H-NMR}$ spectra were processed and analyzed using the Chenomx NMR Suite Professional software package version 7.7 (Chenomx Inc., Edmonton, AB). The Chenomx NMR Suite software allows for qualitative and quantitative analysis of an NMR spectrum by manually fitting spectral signatures from an internal database to the spectrum.

Supplemental References

- Craciun, S., and Balskus, E.P. (2012). Microbial conversion of choline to trimethylamine requires a glyceryl radical enzyme. *Proceedings of the National Academy of Sciences* *109*, 21307-21312.
- Koeth, R.A., Levison, B.S., Culley, M.K., Buffa, J.A., Wang, Z., Gregory, J.C., Org, E., Wu, Y., Li, L., and Smith, J.D. (2014). γ -Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell metabolism* *20*, 799-812.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods* *9*, 357-359.
- Méjean, V., Lobbi-Nivol, C., Lepelletier, M., Giordano, G., Chippaux, M., and Pascal, M.C. (1994). TMAO anaerobic respiration in *Escherichia coli*: involvement of the tor operon. *Molecular microbiology* *11*, 1169-1179.
- Meyer, M., Granderath, K., and Andreesen, J.R. (1995). Purification and characterization of protein PB of betaine reductase and its relationship to the corresponding proteins glycine reductase and sarcosine reductase from *Eubacterium acidaminophilum*. *European Journal of Biochemistry* *234*, 184-191.
- Zhu, Y., Jameson, E., Crosatti, M., Schäfer, H., Rajakumar, K., Bugg, T.D., and Chen, Y. (2014). Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota. *Proceedings of the National Academy of Sciences* *111*, 4268-4273.

>Candidatus_Methanomassiliicoccus_intestinalis 16S rRNA gene sequence from NC_021353.1

ATTCGGTTGATCCTGCCGGCGGCCACCGCTATTGGAATGCGATTAAGACATGCGAGTCGAGAGACATTTGGTCTCGGCC
GACCGCTCAGTAACACGTGGACAACATGCCCTAAGTGGGGGATAATCTCGGGAAACTGAGGATAATACCCCATAGGTCT
AATCTGCTGGAATGCTTTTAGGCCGAAAACCTCCGGTGCCTTAGGATTGGTCTGCGGTCTATCAGGTTGTAGTGGGTGTAA
TGGACCCACTAGCCCGTGACGGATATGGGCCTTGAGAGAGGGAGCCAGAGATGGACTCTGAGACATGAGTCCAGGCCCT
ACGGGGCGCAGCAGTCGCGAAAACCTTCGCAATGGGGGAAAACCCCGACGAGGGAATTCCAAGTGCCAGCATTTAATGTTGG
CTGTTCTTTTGTCTAAAAACATGAGAAGTAAGGGGCGGGTAAGACGGGTGCCAGCCGCCGCGGTAATACCCGCGCCCCA
AGTGGTGGTTCGATATTATTGAGCCTAAAACGTCCGTAGCCGGTTTTCTAAATCCTTGGGTAAATCGGGCAGCTTAACGT
CCGAATTCGGGGGAGACTGGGAGACTTGGGACCGGGAGAGGTGAGAGTACTTCTGGGGTAGGGGTAATAATCCTGTAATC
CTGGAAGGACCACCGGTGGCGAAGGCGTCTGACTAGAACGGATCCGACGGTGAGGGACGAAGCCCTGGGGCGCAAACGGG
ATTAGATACCCCGTAGTCCAGGGTGTAAACGCTGCGGGCTTGGTGTGGGGTCTTAGTGGGCGCCAGTCCGGGAGA
GAAGTTGTTAAGCTGCTGCTTGGGAGTATGTCCGAAGGATGAACTTAAAGGAATTGGCGGGGAGCACCGCAACGG
GAGGAGCGTGCAGTCAATTGGATTCAACGCCGAAAACCTCACCAGGAGCGACGGATATATGAAGGTCAAGCCGATGACT
TTACCTGATTTTCCGAGAGGTGGTGCATGGCCGTCGTAGTTCGTACCGTAAGGCGTCTCTTAAGTGAATAACGAACG
AGACCTCACCATTAATTGCTAATCCTTTCTCCGGAGGGATGCACATTAATGGGACCCTGGCGCTAAGTCAGAGGAAG
GAGAGTCAACGGTAGGTCGTATGCCCGAATCTCTGGGCAACACGCGGCTACAAAGGACGGGACAATGGGCTTCAA
CACGAAAGGTGAAGACAATCCTGAAACCCGTCATAGTTCGGATTGGGGCTGTAACTCGCCCCATGAAGTGGATTTC
CGTAGTAATCGCTTGTCAACATCAAGCGGTGAATATGCCCTGCTCCTTGCACACACCGCCCGTCAAACACCCGAGTTG
GGTTTCAGTGAGGATGTTTCTAATTGGGGCTTTTCGAACTGAGATTTCAGCAAGGAGGTTAAGTCGTAACAAGGTATCTGT
AGGGGAACCTGCAGATGGATCACCTCCGTAGGGGAACCTGCAGATGGATCACCTCCTA

>Candidatus_Methanomethylophilus_alvus 16S rRNA gene sequence from NC_020913.1

ACTCCGGTTGATCCTGCCGGCGGCCACCGCTATAGGAATTCGATTAAGACATGCGAGTCGAGAGCGCAAGCTCGGCGGAC
TGCTCAGTAACACGTGGACAACGTGCCCTAAGTGGGGGATAATCAGGGGAAACTCTGGATAATACCCCATAGATCATGA
GATCTGGAATGACTTATGTTCAAAGTTCCGGCGCTTTAGGATCGGTCTGCGGCCTATCAGGTAGTAGTGGGTGTAATGT
ACCCCTAGCCTATTACGGGTATGGCCTTGAGAGAGGGAGCCAGAGTTGGATTCTGAGACACGAATCCAGGCCCTACG
GGGCGCAGCAGTCGCGAAAACCTTCAATGGGCGAAAGCCGATGAGGGAATTCCTAGTGCTAGCACTTTTGTGTTAGCT
TTTCTTTAGCGTAGATAACTAGAGGAATAAGGGCTGGGTAAGACGGGTGCCAGCCGCCGCGGTAATACCCGAGCCCGAG
TGGTGGTTCGATTTTATTGAGTCTAAAACGTTTCGTAGCCGGTCTGATAGATCCTTGGGTAAATCGGGGGGCTTAACCTTCC
GAATTCGAGGAGACCGTTCAGGCTTGGGATCGGGAGAGGTAAGAGGTACTTCAGGGGTAGGGGTAATAATCCTGTAATCCT
TGGAGGACCACCGGTGGCGAAGGCGTCTTACTAGAACAATCCGACGGTGAGGGACGAAGCCCTAGGTCGCAAACGGGAT
TAGATACCCCGGTAGTCTAGGGTGTAAACGCTGCAGACTTGGTGTGGAGGCCCTTCGGGGGCATTTCAGTCCGGGAGAGA
AGTTGTTAAGTCTGCTACTTGGGGAGTACGTCCGCAAGGATGAACTTAAAGGAATTGGCGGGGAGCACCGCAACGGGA
GGAGCGTGCAGTTAATTGGATTCAACACCGGAAAACCTCACCAGGGAGACCATCACATGAAAGCCAGGCTAATGACTTT
GCTTGATTCTTGGAGAAGTGGTGCATGGCCATCGTCAGTTCGTACTGTAAGCGTTCCTTAAGTGAATAACGAACAAG
ACCCTCACTTATAATTGCTAACTAGATCTCCGGATCTGGTGCACATTATCGGGACCCTGGCGCTAAGTCAGAGGAAGGA
GAGGTCAACGGTAGGTCAGTATGCCCTGAATCTCTTGGGCTACACGCGGCTACAAAGGGCGGGACAATGAGTTCGACA
CCGAAAGGTGAAGTAATCTCAAACCCGTCCTAGTTCGGACTGAGGGTTGTAACCTACCCTCACGAAGCTGGATTCCG
TAGTAATCGCAATCAAAAACCTCGCGGTGAATATGCCCTGCTCCTTGCACACACCGCCCGTCAAACACCCGAGTCCGG
TTTTCAGTGAGGCTGCCTTTATTGGGGTATTTCGAGCTGAGATTTCAGCAAGGAGGTTAAGTCGTAACAAGGTATCTGTAG
GGGAACCTGCAGATGGATCACCTCCT

>Methanomassiliicoccales_Mx-02 16S rRNA gene sequence from LVVQ01000000.1

TCCGGTTGATCCTGCCGGCGGCCACCGCTATAGGAATTCGATTAAGACATGCGAGTCGAGAGTCGTAATGGACTCGGCGG
ACTGCTCAGTAACACGTGGATAACATGCCCTAGAGTGGAGGATAATCTCGGGAAATGAGAATAATACTCCATAGATCAT
GAAATCTGGAATGATTCATGGTTCAAAGTTCCGGCGCTCTAGGATTGGTCTGCGGCTTATCAGGTTGTAGTGGGTGTAAC
GGACCCACTAGCCAGTGACGAGTATGGGCCTTGAGAGAGGGAGCCAGAGATGGATTCTGAGACACGAATCCAGGCCCTA
CGGGGCGCAGCAGTCGCGAAAACCTTCAATGGAGGCAACTCCGATGAGGGAATTCCTAGTGCTAGTACATTAGTATTAG
CTTTCTTTAGTATAGAACGCTAAAGGAATAAGGGCTGGGTAAGACGGGTGCCAGCCGCCGCGGTAATACCCGAGCCCA
AGTGGTGGTTCGATTTTATTGAGTCTAAAACGTTTCGTAGCCGGTTCATTAATCCTTGGGTAAATCGGGAAGCTTAACCTT
CCGACTTCCGAGGAGACTGGTGAACCTTGGGACCGGGAGAGGCAAGAGGTACTTCTGGGGTAGGGGTAATAATCCTGTAATC
CTAGAAGGACCACCGGTGGCGAAGGCGTCTTGTAGAACGGATCCGACGGTGAGGGACGAAGCCCTGGGTGCGAAACGGG
ATTAGATACCCCGTAGTCCAGGGTGTAAACGCTGCAGACTTGGTGTGGGGACCCCTCGGGGTGTTTCAGTCCGGGAGA
GAAGTTGTTAAGTCTGCTACTTGGGGAGTACGTCCGCAAGGATGAACTTAAAGGAATTGGCGGGGAGCACCGCAACGG
GAGGAACGTGCGGTTAATTGGATTCAACACCGGAAAACCTCACCAGCAGCGACTGTTACATGAAAGTCAAGCCGATGACC
TTACTAGATTTTTCAGAGAGGTGGTGCATGGCCGTCGTAGTTCGTACCGTAAGGCGTTCCTTAAGTGAATAACGAACA
AGACCTCACCAATAATTGCTACTATATCGTCCCGATGTAGGCACATTAATGGGACCCTGGCGCTAAGTCAGAGGAAG
GAGAGGTCAACGGTAGGTCAGTATGCCCGAATCTGCTGGGCTACACGCGGTTACAAAGGGCGGGACAATGGGCTCCGA

CGCCGAAAGGCGAAGCAATCTCGAAACCCGTCGGTAGTTTCGGATTGTGGGTTGTAACTCACCCACATGAAGCTGGATTCCGTAGTAATCGCGAATCAACAACCTCGCGGTGAATATGCCCTGCTCCTTGCACACACCCGCCCGTCAAACCACCTGAGTTGGTTTCAGTGAGGCTATCTCTAATTGGGGTATTCGAACTGAGATTTAGCAAGGAGGGTTAAGTCGTAACAAGGTA

>Methanomassiliicoccales_Mx-03 16S rRNA gene sequence from LVVR01000000.1
ACTCCGGTTGATCCTGCCGGCGCCACCGCTATAGGAATTCGATTAAGACATGCGAGTCGAGAGTCTTCGGGACTCGGCGGACTGCTCAGTAACACGTGGAGAACGTGCCCTTAAGTGGAGGATAATCTCGGGAAATTGAGGATAATACTCCATAGATCATGGGATCTGGAATGACCCATGGTTGAAAGTTCGGCGCTTAAGGATCGCTCTGCGGCCTATCAGGTAGTAGTGGGTGTAACTACCCCTAGCCGACGACGGGTATGGGCCTTGAGAGAGGGAGCCAGAGTTGGATTCTGAGACACGAATCCAGGCCCTACGGGGCGCAGCAGTCGCGAAAACCTCACACTGGGGCAACCCCGATGAGGGAATTCCTAGTGCTAGGACATTTGTTCTAGCTTTTCTCAAGCGTAGATAACTTGAGGAATAAGGGCTGGGTAAGACGGGTGCCAGCCGCCGCGGTAATACCTGCAGCCC AAGTGGTGGTTCGATTTTATTGAGTCTAAAACGTTCCGTAGCCGGTCTGGTAAATCCTTGGGTAAATCGGGAGCTTAACCTTCCGACTTCCGAGGAGACTGCCAGACTTGGGACCGGGAGAGGCAAGAGGTACTTCTGGGGTAGGGGTAAAATCCTGTAAT CCTAGAAGGACCACCGGTGGCGAAGGCGTCTTGCTAGAACGGATCCGACGGTGAGGGACGAAGCCCTGGGTGCGAAACGG GATTAGATACCCCGTAGTCCAGGGTGTAAACGCTGCAGACTTGGTGTGGAGGCCCTTCGGGGCATTAGTCCGGGAG AGAAGTTGTTAAGTCTGCTACTTGGGGAGTACGTCGCAAGGATGAAACTTAAAGGAATTGGCGGGGAGCACCAGCGG GAGGAGCGTGCAGTTAATTGGATTCAACACCGGAAAACCTACCAAGGGCGACTGTTACATGAAAGCCAGCTAATGAC CTTGCTAGATTTTTAGAGAGGTGGTGCATGGCCGTCAGTTTCGTACCGTAAGGCGTCTCTTAAAGTGAAGATAACGAAC GAGACCTCACTGATAATTGCTACTCCATCCTCCGGGTGGAGGCACATTATTGGGACCGCTGGCGCTAAGTCAGAGGAA GGAGAGGTCAACGGTAGGTGAGTATGCCCCGAATCTCTGGGCTACACGCGCGCTACAAGGGCGGGACAATGGGCTCCG ACGCGAAAAGGCGAAGGTAATCTCGAAACCCGTCCTATAGTTCGGATTGAGGGTTGTAACCTCACCTCATGAAGCTGGATT CCGTAGTAATCGCGAATCAACAACCTCGCGGTGAATATGCCCTGCTCCTTGCACACACCCGCCCGTCAAACCATCCGAGTT GGGTTTCAGTGAGGTTGCCTCTAATTAGGGTATTGCAACTGAGATTTAGCAAGGAAGGTTAAGTCGTAACAAGGTATCTG TAGGGGAACCTGCAGATGGATCACCTCCTA

>Methanomassiliicoccales_Mx-06 16S rRNA gene sequence from LVVU01000000.1
ACTCCGGTTGATCCTGCCGGCGCCACCGCTATAGGAATTCGATTAAGACATGCGAGTCGAGAGGCATAATGGCCTCGGC GAAGTCTCAGTAACACGTGGAGAACATGCCCTAGAGTGGGGATAGTCAGGGGAACTCTGCATAATACCCCATAGGTC AGGAGATCTGGAATGACTCCTGGCTCAAAGTTCGGCGCTCTAGGATTGCTCTGCGGCCTATCAGGTAGTAGTGGGTGTA ATGTACCCCTAGCCGACGACGGGTATGGGCCTTGAGAGAGGGAGCCAGAGTTGGATTCTGAGACACGAATCCAGGCC TACGGGGCGCAGCAGTCGCGAAAACCTCACACTGGGGCAACCCCGATGAGGGAATTCCTAGTGCTAGGACATTTGTTCT AGCTTTTCTCAAGCGTAGATAACTTGAGGAATAAGGGCTGGGTAAGACGGGTGCCAGCCGCCGCGGTAATACCTGCAGCC CAAGTGGTGGTTCGATTTTATTGAGTCTAAAACGTTCCGTAGCCGGTCTGGTAAATCCTTGGGTAAATCGGGGAGCTCAACT TTCCGAATTCGAGGAGACTGCCAGACTTGGGACCGGGAGAGGCTAGAGGTACTTCTGGGGTAGGGTAAAATCCTGTAA TCCTAGAAGGACCACCGGTGGCGAAGGCGTCTAGCTAGAACGGATCCGACGGTGAGGGACGAAGCCCTGGGTGCGAAACG GGATTAGATACCCCGTAGTCCAGGGTGTAAACGCTGCAGACTTGGTGTGGAGGCCCTTCGGGGCATTAGTCCGGGA GAGAAGTTGTTAAGTCTGCTACTTGGGGAGTACGTCGCAAGGATGAAACTTAAAGGAATTGGCGGGGAGCACCAGCAAC GGGAGGAGCGTGCAGTTAATTGGATTCAACACCGGAAAACCTACCAAGGGCGACCATTACATGAAAGCCAGGCTAATGA CCTTGCTAGATTTTTGGAGAGGTGGTGCATGGCCGTCAGTTTCGTACCGTAAGGCGTCTCTTAAAGTGAAGATAACGAA CGAGACCTCACTGATAATTGCTAATCTATTCTCCGAGTAGATGCACATTATTGGGACCGCTGACGCTAAGTCAGAGGA AGGAGAGGTCAACGGTAGGTGAGTATGCCCCGAATCTCTGGGCTACACGCGCGCTACAAGGGCGGGACAATGGGCTCC GACGCCGAGAGGCGAAGGTAATCTCGAAACCCGTCCTATAGTTCGGACTGAGGGTTGTAACCTCACCTCACGAAGCTGGAT TCCGTAGTAATCGCGAATCAACAACCTCGCGGTGAATATGCCCTGCTCCTTGCACACACCCGCCCGTCAAACCATCCGAGT TGGGTTTCAGTGAGGTGGCCTTTAATTGGGGTCTTCAAACCTGAGATTTAGCAAGGAAGGTTAAGTCGTAACAAGGTATCT

>Methanomassiliicoccales_Mx-04 FJ752572
CGGCCACCGCTATAGGAATTCGATTAAGACATGCGAGTCGAGAGGCATAATGGCCTCGGCGGACTGCTCAGTAACACGTG GAGAACGTGCCCTAAAGCCACTAATACAATCGGGAACTGATAATAATGAGTGATAGGTGATAGGTCAGGAGATCTGGAATGACTCC TGGCTCAAAGTTCGGCGCTTTAGGATCGCTCTGCGGCCTATCAGGTAGTAGTGGGTGTAACGTACCCCTAGCCGACGA CGGGTATGGGCCTTGAGAGAGGGAGCCAGAGTTGGATTCTGAGACACGAATCCAGGCCCTACGGGGCGCAGCAGTCGCG AAAACTTCACACTGGGGCAACCCCGATGAGGGAATTCCTAGTGCTAGGACATTTGTTCTAGCTTTTCTCATGCGTAGAT AACATGAGGAATAAGGGCTGGGTAAGACGGGTGCCAGCCGCCGCGGTAATACCTGCAGCCAAAGTGGTGGTTCGATTTTAT TGAGTCTAAAACGTTTCGTAGCCGGTCCGGTAAATCCTTGGGTAAATCGGGGTGCTTAACATTCCGACTTCCGAGGACACT GCCGACTTGGGACCGGGAGAGGCTAGAGGTACTTCTGGGGTAGGGGTAAAATCCTGTAATCCTGGAAGGACCACCGGTG CGAAGGCGTCTAGCTAGAACGGATCCGACGGTGAGGGACGAAGCCCTGGGTGCGAAACGGGATTAGATACCCCGGTAGT CCAGGGTGTAAACGCTGCAGACTTGGTGTGGAGGCCCTTCGGGGCATTAGTCCGGAGAGAAGTTGTTAAGTCTGCT ACTTGGGGAGTACGTCCGCAAGGATGAACTTAAAG

>Methanomassiliicoccales_Mx-07 SRR1763675.40242.1 40242
TACCTGCAGCCCAAGTGGTGGTCGATTTTATTGAGTCTAAAGCGTTCGTAGCCGGTTTGATAAATCTTTGGGTAAATCGG
GAAGCTTAACTTTCCGACTTCCGGAGAGACTGTCAAACCTTGGGACCGGGGAGGCTAGAGGTACTTCTGGGGTAGGGGTA
AAATCCTGTACTCCTAGAAGGACCACCGGTGGCGAAGGCGTCTAGCTAGAACGGATCCGACGGTGAGGGACGAAGCCCTG
GGTCGCAAACGGG

>Methanomassiliicoccales_Mx-08 10052.MG.F5_7016 orig_bc=AGCGACTGTGCA
new_bc=AGCGACTGTGCA bc_diffs=0
TACCTGCAGCCCAAGTGGTGGTCGATTTTATTGAGTCTAAAACGTTTCGTAGCCGGTCTGGTAAATCCTTGGGTAAATCGC
GGAGCTCAACTTTCCGAATTCCGTGGAGACTGCCAGACTTGGGTCCGGGAGAGGCTAGTGGAACCTTCTGGG

>Methanomassiliicoccales_Mx-09 10052.MG.F48_126 orig_bc=GCTTACATCCAG
new_bc=GCTTACATCGAG bc_diffs=1
TACGTGCAGCCCAAGTGGTGGTCGATTTTATTGAGTCTAAAACGTTTCGTAGGCGGTCTGGTAAATCTTTGGGTAAATCGG
GGAGCTCAACTTTACGAATTCTTAGGAGACTGTCAAACCTTGGGGCCGGGAGAGGCTAGAGGAACTCCTGGG