



HAL
open science

Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea

Guillaume Borrel, Panagiotis S Adam, Luke J Mckay, Lin-Xing Chen, Isabel Natalia Sierra-García, Christian M K Sieber, Quentin Letourneur, Amine Ghozlane, Gary Andersen, Wen-Jun Li, et al.

► **To cite this version:**

Guillaume Borrel, Panagiotis S Adam, Luke J Mckay, Lin-Xing Chen, Isabel Natalia Sierra-García, et al.. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nature Microbiology*, 2019, 10.1038/s41564-019-0363-3. pasteur-02059013

HAL Id: pasteur-02059013

<https://pasteur.hal.science/pasteur-02059013v1>

Submitted on 6 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

1 **Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea**

2 Guillaume Borrel^{1*}, Panagiotis S. Adam^{1,2}, Luke J. McKay³, Lin-Xing Chen⁴, Isabel Natalia
3 Sierra-García⁵, Christian M.K. Sieber^{4,6}, Quentin Letourneur⁷, Amine Ghozlane⁷, Gary L.
4 Andersen⁸, Wen-Jun Li⁹, Steven J. Hallam¹⁰, Gerard Muyzer¹¹, Valéria Maia de Oliveira⁵,
5 William P. Inskeep¹², Jillian F. Banfield⁴, and Simonetta Gribaldo^{1*}

6 ¹Department of Microbiology, Unit Evolutionary Biology of the Microbial Cell, Institut
7 Pasteur, Paris, France

8 ²Université Paris Diderot, Sorbonne Paris Cité, Paris, France

9 ³Center for Biofilm Engineering, Montana State University, Bozeman, MT 59717, USA

10 ⁴Department of Earth and Planetary Sciences, University of California, Berkeley, California
11 94720, USA.

12 ⁵Microbial Resources Division, Research Center for Chemistry, Biology and Agriculture
13 (CPQBA), Campinas University - UNICAMP, P.O. Box 6171, 13081-970, Campinas, SP, Brazil

14 ⁶Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA

15 ⁷Bioinformatics and Biostatistics Hub, C3BI, USR 3756 IP CNRS, Institut Pasteur, Paris, France

16 ⁸Department of Environmental Science, Policy, and Management, University of California,
17 Berkeley, California 94720, USA

18 ⁹State Key Laboratory of Biocontrol and Guangdong Provincial Key Laboratory of Plant
19 Resources, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, PR China

20 ¹⁰Department of Microbiology & Immunology, ECOSCOPE Training Program, Graduate
21 Program in Bioinformatics, and Genome Sciences and Technology Training Program,
22 University of British Columbia, Vancouver, BC, V6T 1Z3, Canada

23 ¹¹Microbial Systems Ecology, Department of Freshwater and Marine Ecology, Institute for
24 Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The
25 Netherlands

26 ¹²Department of Land Resources and Environmental Sciences, Montana State University,
27 Bozeman, MT 59717, USA

28

29 *Correspondence to:

30 simonetta.gribaldo@pasteur.fr

31 guillaume.borrel@pasteur.fr

32

33

34 **Abstract**

35 Methanogenesis is an ancient metabolism of key ecological relevance, with direct impact on
36 the evolution of Earth's climate. Recent results suggest that the diversity of methane
37 metabolisms and their derivations have probably been vastly underestimated. Here, by
38 probing thousands of publicly available metagenomes for homologues of methyl-coenzyme
39 M reductase complex (MCR), we have obtained ten metagenome-assembled genomes

40 (MAGs) belonging to potential methanogenic, anaerobic methanotrophic and short-chain
41 alkane oxidizing archaea. Five of these MAGs represent under-sampled (e.g.,
42 Verstraetearchaeota, Methanonatronarchaeia, ANME-1) or previously genomically
43 undescribed (ANME-2c) archaeal lineages. The remaining five MAGs correspond to lineages
44 that are only distantly related to previously known methanogens and span the entire
45 archaeal phylogeny. Comprehensive comparative annotation significantly expands the
46 metabolic diversity and energy conservation systems of MCR-bearing archaea. It also
47 suggests the potential existence of a yet uncharacterized type of methanogenesis linked to
48 short-chain alkane/fatty acid oxidation in a previously undescribed class of archaea (*'Ca.*
49 *Methanoliparia'*). We redefine a common core of marker genes specific to methanogenic,
50 anaerobic methanotrophic and short-chain alkane-oxidizing archaea, and propose a possible
51 scenario for the evolutionary and functional transitions that led to the emergence of such
52 metabolic diversity.

53 Methanogenesis is an archaeal-specific metabolism of key relevance in the anaerobic
54 degradation of organic matter and biogas production^{1,2}. It is considered one of the most
55 ancient energetic metabolisms^{3,4} with direct impact on the evolution of the Earth's climate
56 system⁵. Methanogens have been detected in virtually all types of anaerobic environments.
57 Until recently, all methanogens were thought to belong to two euryarchaeal clades, named
58 Class I and Class II methanogens⁶. The majority of Class I/II methanogens can grow by
59 reducing CO₂ into methane using H₂ as an electron donor⁷. Several representatives of the
60 Methanosarcinales (Class II methanogens) use additional energetic substrates, including
61 acetate and methylated compounds⁸. *Methanosphaera* spp. (Class I methanogens) are
62 restricted to the reduction of methanol with H₂⁹. Regardless of the encoded methanogenic
63 pathway, all members of Class I/II methanogens possess the H₄MPT methyl-branch of the
64 Wood-Ljungdahl pathway (m-WL), the N⁵-Methyltetrahydromethanopterin:coenzyme M
65 methyltransferase complex (MtrABCDEFGH or MTR), and the methyl-coenzyme M reductase
66 complex (McrABG or MCR)¹⁰. The same enzymes are present in ANAerobic MEthanotrophic
67 archaea (ANME) and are used in reverse to oxidize methane¹¹⁻¹³.

68 Our understanding of the diversity and metabolic versatility of methanogenic
69 archaea is undergoing a rapid transformation with the availability of additional isolates and
70 metagenome-assembled genomes (MAGs). This has revealed additional lineages only
71 distantly related to Class I/II methanogens¹⁴⁻¹⁶, including Methanomassiliicoccales¹⁰,
72 Methanofastidiosales¹⁷, Methanonatronarchaeia¹⁸ and Verstraetearchaeota¹⁹. A striking
73 characteristic of these recently described methanogens is the absence of the MTR complex,
74 a partial or missing m-WL pathway, and the presence of specific methyltransferases for the
75 utilization of methylated compounds. Accordingly, they are predicted to be limited to reduce
76 methylated compounds with H₂ for methanogenesis, which was experimentally validated for
77 Methanomassiliicoccales²⁰ and Methanonatronarchaeia¹⁸. Interestingly, the implication of a
78 divergent McrABG-like complex in the oxidation of short-chain alkanes (butane, propane)
79 has been demonstrated in two representatives of a recently described euryarchaeal order,
80 the 'Ca. Syntrophoarchaeales'²¹. Divergent MCR sequences were also found in two members
81 of the Bathyarchaeota²² (TACK superphylum), in the GoM-Arc1 (a lineage within the
82 Methanosarcinales)²³, and from environmental samples²⁴. Altogether, this suggests that
83 methanogens, anaerobic methanotrophs and short-chain alkane oxidizers may have an even
84 wider phylogenetic and environmental distribution than previously anticipated, provoking
85 new questions on the diversity and evolution of these metabolisms.

86

87

88 **Results and Discussion**

89

90 ***Additional lineages of archaea with an MCR or MCR-like complex***

91 To identify previously undescribed lineages of potential methanogens, anaerobic
92 methanotrophs and short-chain alkane oxidizers, we probed available metagenomes from
93 the JGI/IMG database for McrA homologues and identified sequences distantly related to
94 well characterized lineages (Methods). Ten MAGs were reconstructed from the
95 corresponding metagenomes sourced from a wide range of anoxic environments including
96 an inland petroleum reservoir from Brazil, oil seeps from USA²⁵, soda lake sediments from
97 Russia, and hot-springs from China and USA (Table 1). Nine of the ten MAGs had an
98 estimated completeness ranging from 78.4 to 94.4%, and one was only 51.5% complete.
99 Estimated contamination (without strain heterogeneity) ranged from 0 to 3.3%.

100 Four MAGs represent three previously undescribed lineages only distantly related to
101 known methanogenic/methanotrophic archaea (Fig. 1): NM1 (NM1a and NM1b MAGs)
102 branches within the Methanotecta superclass¹⁵, between Archaeoglobales and the clade
103 formed by ‘*Ca. Syntrophoarchaeales*’ and ‘*Ca. Methanophagales*’¹⁵ (ANME-1); NM3 branches
104 within the Acherontia superclass¹⁵, at the base of the clade formed by the non-
105 methanogenic Theionarchaea²⁶ and ‘*Ca. Methanofastidiosa*’ (former WSA2/Arc1); NM4
106 branches within the TACK superphylum and is related to *Korarchaeum cryptofilum*. In the
107 NM4 MAG, the markers for methane metabolism are present on two contigs with a lower
108 coverage than the other contigs and contain few genes related to *Korarchaeum*. However,
109 an independent study supports that these contigs belong to the same organism,
110 provisionally named “*Candidatus Methanodesulfokores washburnensis*” for its methane-
111 and sulfur-cycling capacities (McKay et al., submitted). One additional MAG (NM2) was too
112 partial to assess its phylogenetic placement.

113 Finally, five MAGs correspond to currently under-sampled archaeal lineages: a deep-
114 branching Verstraetearchaeota, a third Methanonatronarchaeia, a second ‘*Ca.*
115 *Methanophagales*’ (ANME-1), a second GoM-Arc1 (showing a close relationship with the
116 methanotrophic *Methanoperedens*), and the first representative of ANME-2c (Fig. 1).

117 **Phylogeny and functional inference of MCR/MCR-like complexes**

118 To investigate the evolutionary relationships and characteristics of the MCR complexes
119 identified in this study, we built a phylogeny based on a concatenated alignment of McrABG
120 subunits. This phylogeny is in overall agreement with recently published ones^{16,18}. Three
121 MAGs (NM1a, NM1b and GoM-Arc1-GOS) encode alternative McrABG-like complexes that
122 cluster with those of ‘*Ca. Syntrophoarchaeum*’ and Bathyarchaeota (Fig. 2A, in blue). The
123 presence of an MCR-like complex and the absence of a canonical MCR in GoM-Arc1-GOS are
124 consistent with the recent description of a MAG of this lineage²³, and represent so far a
125 unique feature within the Methanosarcinales.

126 The remaining MAGs harbour canonical MCR complexes (Fig. 2A), branching next to
127 their closest MCR-bearing neighbors in the reference archaeal phylogeny (Fig. 1), suggesting
128 no recent horizontal gene transfers (HGT). The clustering of NM3 with *Methanofastidiosa*
129 supports an early presence of methanogenesis in the Acherontia. The clustering of NM4 with
130 *Verstraetearchaeota*, support that it is a genuine methanogenic/methanotrophic
131 representative of the TACK. The separate branching of ANME-2c from the other ANME-2
132 lineages suggests that anaerobic methane oxidation in Methanosarcinales emerged multiple
133 times independently from methanogenic ancestors. Interestingly, both NM1a and NM1b
134 encode, in addition to the McrABG-like complex, a canonical MCR complex branching at the
135 base of Class II methanogens, consistently with the reference phylogeny. The coexistence of
136 MCR and MCR-like complexes in the same archaeon has never been observed before and
137 brings into question the metabolism of this lineage (see below).

138 It is striking to observe that most of the predicted or experimentally proven methyl-
139 dependent hydrogenotrophic methanogens are closely related in the MCR tree (Fig. 2A, in
140 red), irrespective of their placement in the reference phylogeny (Fig. 1). This might be the
141 consequence of ancient exchanges of the MCR complex among these lineages, whose
142 direction is hard to define. Nevertheless, some more recent transfers may be identified. For
143 example, ‘*Ca. Methanophagales*’ (ANME-1) MCRs branch far from their *Methanotecta*
144 relatives, and might have acquired their MCR complex from a methanogenic member of the
145 Acherontia.
146

147 The clustering of MCR-like homologues belonging to distantly related lineages (Fig.
148 2A, in blue) is also puzzling. This might be due to HGT and/or tree reconstruction artefacts
149 linked to their high sequence divergence with respect to canonical MCRs, exemplified by
150 their longer-than-average branches. Such divergence is probably related to a change in
151 function, as MCR-like complexes are involved in activating short-chain alkanes (butane and
152 propane) in '*Ca. Syntrophoarchaeum*'²¹. Accordingly, several residues playing an important
153 role in canonical MCR, either by interacting with cofactors, forming the catalytic site cavity
154 wall or being post-translationally modified, are not conserved in '*Ca. Syntrophoarchaeum*'
155 sequences (Fig 2B). The replacement of large aromatic residues (e.g. Phe330, Tyr333,
156 Tyr444, Tyr446) present in the cavity wall of canonical MCR²⁷ by smaller ones in '*Ca.*
157 *Syntrophoarchaeum*' MCR-like complexes could have occurred to accommodate
158 butane/propane (larger substrate than methane) in the catalytic site (Fig. 2B). The presence
159 of smaller amino acids at these positions in NM1 and Bathyarchaeota MCR-like complex
160 suggest a similar function in short-chain alkane oxidation. Finally, the MCR-like sequences of
161 GoM-Arc1 show fewer modifications at these sites, suggesting the utilization of a smaller
162 alkane, possibly ethane or methane.

163

164 ***Expanded diversity of methyl-dependent hydrogenotrophic methanogens***

165 The NM3 and NM4 MAGs share several similarities with the recently discovered order-level
166 lineages of methanogens that were proposed or experimentally proven to perform methyl-
167 dependent hydrogenotrophic methanogenesis^{10,17-19} (Fig. 3, Supplementary Table 1). First,
168 these relatively complete MAGs (85,5% completeness) lack at least 24 genes coding for the
169 MTR complex, H₄MPT biosynthesis, and the H₄MPT methyl-branch of the WL pathway,
170 otherwise present in all Class I/II methanogens (Supplementary Table 1). Second, they
171 encode [Ni-Fe] hydrogenases and methyltransferases with the potential to support
172 methanogenesis from methanol (MtaABC) in NM3 and NM4 and methanethiol (MtsAB) in
173 NM3 (Fig. 3; Supplementary Table 1). Interestingly, energy conservation complexes of NM3
174 are mostly similar to Methanofastidiosales¹⁷ (Supplementary Fig. 1), their closest related
175 methanogens in the reference phylogeny (Fig. 1). Altogether, this suggests that NM3 and
176 NM4 rely on methyl-dependent hydrogenotrophic methanogenesis (Fig. 3; Supplementary
177 Discussion for details on energy conservation in NM3 and NM4).

178 The predicted methanogenesis pathway in Verst-YHS (Verstraetearchaeota) and
179 Mnatro-ASL (Methanonatronarchaeia) MAGs also supports methyl-dependent
180 hydrogenotrophic methanogenesis (Fig. 3, Supplementary Table 1), as described in the first
181 genomic assemblies for these lineages^{18,19}. However, comparison of the energy conservation
182 enzymes in the seven currently available Verstraetearchaeota (order Methanomethyliales)
183 suggests an alternative model than previously described¹⁹ (Fig. 3, Supplementary Table 1).
184 Indeed, we found that all Methanomethyliales MAGs (95% average completeness) lack the
185 HdrA/MvhD and possibly MvhAG subunits of the electron-bifurcating complex
186 HdrABC/MvhADG, suggesting that this complex is absent in these archaea. In contrast, we
187 identified in these genomes a gene cluster encoding a potential complex composed of a
188 membrane-bound hydrogenase and of HdrBC (tentatively named Energy-converting
189 Hydrogenase D or Ehd; Supplementary Fig. 2). We propose that this complex could be
190 involved in a previously unreported mode of energy conservation associated with
191 methanogenesis (Fig. 3; Supplementary Discussion).

192

193 ***Insights into methane and short-chain alkane oxidizers***

194 GoM-Arc1-GOS, ANME-1-THS and ANME-2c MAGs possess a WL pathway and lack the
195 methyltransferases and [Ni-Fe] hydrogenases required for methylotrophic and
196 hydrogenotrophic methanogenesis, respectively (Fig. 3; Supplementary Table 1), similar to
197 all available MAGs of methanotrophs and short-chain alkane oxidizer (Supplementary Fig. 3).
198 Although they encode an AMP-producing acetyl-CoA synthetase (Acs) which is used for
199 acetoclastic methanogenesis in *Methanosaeta* spp., they could rather use it for acetate
200 assimilation¹¹. Comparison with methanotrophs and short-chain alkane oxidizers also reveals
201 a common core of enzymes for energy conservation, comprising the F₄₂₀H₂:quinone (or
202 phenazine) oxidoreductase (Fqo/Fpo) and a potential electron confurcating complex
203 (HdrABC/MvhD/FdhB²⁸) coded by a conserved gene cluster (Supplementary Fig. 4). ANME-2c
204 and GoM-Arc1-GOS encode 17 and 10 multiheme c-type cytochromes respectively,
205 supporting the importance of direct electron transfer to syntrophic partners in anaerobic
206 methane^{29,30} and short-chain alkane oxidation²¹ metabolisms (Supplementary Fig. 3;
207 Supplementary Table 1).

208 ANME-1-THS MAG is the first sequenced representative of a “Land clade” within the
209 ‘*Ca. Methanophagales*’ (Supplementary Fig. 5), suggesting different adaptations to
210 environmental conditions than members of the ANME-1b clade, which are mainly from
211 marine methane seeps. ANME-1-THS differs from the ANME-1b MAG³¹ by the presence of a
212 bacterial-like Rnf complex that could couple the NAD:ferredoxin oxidoreduction with
213 chemiosmotic gradient generation/utilisation (Fig. 3; Supplementary Fig. 6; Supplementary
214 Discussion). If these genes are not in the missing region of this MAG, ANME-1-THS might also
215 differ from the other ANMEs by the lack of multiheme c-type cytochromes to transfer
216 electrons from methane oxidation to a syntrophic partner (Fig. 3; Supplementary Fig. 3).
217 Alternatively, two PsrABC-like complexes, including a molybdenum/selenocysteine-
218 containing dehydrogenase subunit, could be involved in the reduction of inorganic
219 compounds such as polysulfide/elemental sulfur^{32,33} (Fig. 3). This contrasts with ANME-1b
220 MAG which misses the membrane integral (PsrC-like) subunit needed to transfer electrons
221 from membrane-associated electron transporters (Supplementary Fig. 3). These
222 characteristics might indicate growth of ANME-1-THS without bacterial syntrophs.

223 The gene content of GoM-Arc1-GOS is consistent with the recent description of the
224 first member of the GoM-Arc1 lineage²³. While GoM-Arc1 members encode an MCR-like
225 complex possibly involved in short-chain alkane oxidation (Fig. 3), they lack the beta-
226 oxidation pathway proposed to be involved in butane/propane utilization in ‘*Ca.*
227 *Syntrophoarchaeales*’²¹ (Supplementary Fig. 3). If GoM-Arc1 members are capable of
228 oxidizing ethane (CH₃CH₃), as suggested by the fewer modifications observed in the catalytic
229 site of its MCR-like complex relative to canonical MCRs (Fig. 2), the oxidation of the ethyl-
230 group would lead to an acetyl- group that could directly enter the oxidative WL pathway,
231 making the beta-oxidation pathway unnecessary (Fig. 3; Supplementary Discussion). With
232 the presence of Fqo, HdrABC/MvhD/FdhB, multiheme c-type cytochromes, and HdrDE
233 (Supplementary Table 1), the energy conservation system associated with this potential
234 ethane-oxidation metabolism in GoM-Arc1 would mostly resemble that associated with
235 methanotrophy in their closely related ANME-2 lineages (Supplementary Fig. 3). The
236 question remains whether the MCR-like homologs of GoM-Arc1 could also be capable of
237 methane oxidation.

238

239 ***A previously uncharacterised type of methanogenesis?***

240 The two NM1 MAGs represent the first archaea predicted to encode both an MCR and an
241 MCR-like complex (Fig. 2), suggesting that they might be potentially capable of both
242 methane and short-chain alkane metabolisms (Fig. 4; Supplementary Table 1). Interestingly,
243 while both NM1 MAGs encode the MTR and the m-WL pathway similarly to Class I/II
244 methanogens, they lack the [Ni-Fe] hydrogenases (MvhA and FrhA) and methyltransferases
245 needed for hydrogenotrophic and methylotrophic methanogenesis, respectively. They also
246 diverge from Class I/II methanogens by the replacement of the F₄₂₀ dependent methylene-
247 tetrahydromethanopterin dehydrogenase (Mtd) by MtdB, which relies on NAD(P) redox
248 cofactor in *Methylobacterium extorquens*³⁴.

249 Beyond the presence of an MCR-like complex, the potential ability of NM1 for short-
250 chain alkane oxidation is also suggested by the presence of a complete beta-oxidation
251 pathway with several gene copies per step, and a complete WL pathway (including
252 CODH/ACS) as in '*Ca. Syntrophoarchaeales*'²¹. In addition, NM1 encode multiple long-chain
253 fatty acid acyl-CoA synthases (FadD-like), not present in '*Ca. Syntrophoarchaeales*'. Long
254 chain fatty acids (LCFA) activated with these enzymes can enter the beta-oxidation pathway.
255 NM1a and NM1b also encode multiple AMP-forming acetyl-CoA synthetase (Acd) to
256 generate ATP from LCFA degradation. These enzymatic redundancies suggest a versatility
257 toward substrates, as previously proposed for *Syntrophus aciditrophicus*³⁵ and
258 *Archaeoglobus fulgidus*³⁶. Consistently, analysis of the environmental distribution of NM1
259 (Supplementary Fig. 7) reveals their common association with anoxic hydrocarbon-rich
260 environments including methane seeps and oil-rich environments, where short-chain
261 alkanes and long-chain carboxylic acids can be present in substantial concentrations^{37,38}. In
262 particular, NM1a and NM1b originate from an enrichment culture based on petroleum fluids
263 and from a natural oil seep.

264 In addition to this potential wide substrate range, NM1 also contrast with '*Ca.*
265 *Syntrophoarchaeales*' in terms of energy conservation by lacking homologues of the
266 NADH/F₄₂₀H₂:quinone oxidoreductase (Nuo/Fqo) and multiheme c-type cytochromes (Fig. 4;
267 Supplementary Table 1). Also, NM1 contain an Rnf complex potentially using NAD instead of
268 menaquinone for ferredoxin oxidoreduction, similarly to ANME-1-THS (Supplementary Fig. 6;
269 Supplementary Discussion). In the absence of membrane-bound enzymes involved in
270 oxidoreduction of lipid-soluble electron carriers, of multiheme c-type cytochromes for direct
271 interspecies electron transfer, of confurcating [Fe]-hydrogenase for interspecies H₂
272 transfer³⁹, and of enzymes involved in dissimilatory reduction of inorganic compounds, the
273 nature of the terminal electron acceptor coupled to alkane/LCFA oxidation remains elusive.
274 Although both MAGs are mostly complete (~90%), it cannot be excluded that some of these
275 enzymes are coded in their missing regions, or that an alternative way to transfer electrons
276 to a terminal acceptor exists (e.g. utilisation of the assimilatory-type sulfite reductase
277 present in both MAGs for dissimilatory reduction of sulfite, direct electron transfer not
278 relying on cytochromes, or utilization of cytochromes produced by a syntrophic partner).
279 Alternatively, we speculate that in NM1, methanogenesis involving the canonical MCR
280 complex could act as a sink for the electrons produced during alkane and LCFA oxidation.
281 Several electron-bifurcating/confurcating complexes encoded in the two NM1 MAGs
282 (Supplementary Fig. 8) together with the Rnf complex could be involved in this metabolism.
283 The conversion of alkane and LCFA into CH₄ and acetate is thermodynamically feasible but
284 was only reported to occur through syntrophic partnerships between a bacterium
285 (performing the beta-oxidation) and a H₂-consuming methanogen^{40,41}, and it thus remains to
286 be proven experimentally whether this can occur in a single organism.

287 Based on the presence of methane and short-chain alkane/fatty acid-related
288 enzymes and the preferential association with hydrocarbon-rich environments, we propose
289 the provisional class ‘*Candidatus Methanoliparia*’, with ‘*Candidatus Methanoliparum*
290 *thermophilum*’ for NM1a and ‘*Candidatus Methanolliviera hydrocarbonicum*’ for NM1b (see
291 Supplementary Discussion for full taxonomy and nomenclature).

292

293 ***A core of markers related to methane and short chain-alkane metabolisms***

294 A group of 38 genes present in most methanogens and absent from most other organisms,
295 generally referred to as “methanogenesis core markers”, was previously defined from Class
296 I/II methanogen genomes^{42,43} (Supplementary Table 2). Half of them have an unknown
297 function. The others correspond to MCR and MTR subunits, enzymes for biosynthesis and
298 activation of the F₄₃₀ prosthetic group of MCR, and post-translational modifications in the
299 McrA catalytic site^{44,45}. We reassessed the occurrence of these markers in the ten assembled
300 MAGs as well as reference genomes covering all recently discovered lineages of
301 methanogens, methanotrophs and short-chain alkane oxidizers (Table 2).

302 Our analysis shows that some markers are no longer universal in Class I/II
303 methanogens (e.g. m37, 38). Also, many marker genes shared by all or most Class I/II
304 methanogens were predicted to be nonessential in *Methanococcus maripaludis* S2⁴⁶ (Table
305 2). These non-universal and nonessential genes could possibly be involved in fine-tuning of
306 methanogenesis (e.g. post-translational modification of MCR⁴⁷) or in its regulation under
307 specific environmental conditions that are not encountered by all methanogens. For
308 example, m21 and m24 are missing in several methanogens from nutrient-rich
309 environments, such as *Methanobrevibacter*, *Methanosphaera* and *Methanocorpusculum*,
310 and could be involved in regulatory processes related to changes in substrate/nutrient
311 availability.

312 All the lineages of predicted and experimentally proven methyl-dependent
313 hydrogenotrophic methanogens¹⁷⁻¹⁹ lack numerous markers (Table 2), similarly to what was
314 previously noted in *Methanomassiliicoccales*⁴⁸. These markers correspond to MTR complex
315 subunits (m27-31), an McrA post-translational modification enzyme (m33)⁴⁵ and several
316 uncharacterized markers that are mostly nonessential in *M. maripaludis*⁴⁶ (Table 2). The
317 existence of the same pattern in NM3 and NM4 supports our inference of a potential
318 methyl-dependent hydrogenotrophic methanogenesis. Finally, Bathyarchaeota BA1 and
319 BA2²² which were described as methyl-dependent hydrogenotrophic methanogens²² but
320 possess an MCR-like complex instead of the canonical MCR (Fig. 2), lack almost all
321 methanogenesis markers (Table 2), questioning their actual metabolism.

322 Several homologues of the methanogenesis markers are also known to be present in
323 non-methanogenic archaea. This is the case of the MCR/MCR-like (m1-3) and MTR (m27-31)
324 complexes in archaeal methanotrophs⁷ and GoM-Arc1²³, as well as the MCR-like complex in
325 ‘*Ca. Syntrophoarchaeales*’²¹. Based on our analysis, archaeal methanotrophs and short-chain
326 alkane oxidizers also appear to possess numerous markers previously exclusively associated
327 with methanogenesis (Table 2), supporting the common origin and functional links of these
328 metabolisms.

329 In addition to the MCR/MCR-like complex subunits, the most specific and conserved
330 markers in all lineages of methanogens, methanotrophs and short-chain alkane oxidizers
331 appear to be the genes involved in the biosynthesis (*nfdD/cfbD*, *murD/cfbE* and possibly
332 *mcrD*⁴⁹) and activation (*atwA* and possibly *mcrC*⁵⁰) of the F₄₃₀ prosthetic group of MCR, along
333 with six genes encoding uncharacterized proteins (m4 to m9) (Table 2). These six genes are

334 co-localized in most genomes (Supplementary Fig. 9) and are among those that were
335 predicted to be co-transcribed in *Methanobrevibacter smithii* R15⁵¹, suggesting they
336 operate in a common process. These six marker enzymes do not co-purify with MCR⁵⁰.
337 However, their phylogeny (Supplementary Fig. 10) and their restriction to archaea having
338 MCR or MCR-like complexes strongly suggest they are involved in essential aspects of the
339 regulation, folding and/or function of the respective holoenzymes (Supplementary
340 Discussion).

341 Finally, several markers are present in archaeal lineages without MCR/MCR-like
342 complexes (Supplementary Table 3) and are possibly remnants of an ancestral methane-
343 metabolism (Supplementary Fig. 11-13; Supplementary Discussion).

344 Taken together, these observations indicate that none of the previously defined
345 methanogenesis markers are unique to methanogens but are rather more generally
346 indicative of metabolisms involving MCR or MCR-like complexes, including methanogenesis,
347 methanotrophy, and short-chain alkane oxidation. Elucidating the roles of these markers
348 (MCR-Associated Markers or MAM) will be essential not only for understanding
349 methanogenesis, but also anaerobic methanotrophy and short-chain alkane oxidation in
350 archaea.

351
352

353 ***Evolution of methane and short-chain alkane metabolisms***

354 Our results significantly extend recent data by highlighting the overwhelming presence of
355 lineages with an MCR or MCR-like complex in the Archaea (Fig. 1). This supports an early
356 origin of methanogenesis in this domain of life, and multiple losses of this metabolism during
357 archaeal diversification^{4,14,18,52}.

358 The sharing of a common set of genes (Table 2) clearly indicates that methanogens,
359 anaerobic methanotrophs and short-chain alkane oxidizers are evolutionarily linked.
360 However, it remains unclear which type of metabolism is the most ancient, and what
361 evolutionary and functional transitions led to such diversity¹⁴. The antiquity of the WL
362 pathway^{53,54}, and the recent proposal that the root of the archaeal tree might lie in between
363 Class I and II methanogens⁵², would suggest that CO₂-dependent hydrogenotrophic
364 methanogenesis is the ancestral type of methanogenesis. Nevertheless, the growing
365 diversity of methyl-dependent hydrogenotrophic methanogens, including this work (Fig. 1 in
366 red), indicates that this metabolism has been largely overlooked. Its origin and evolutionary
367 relationship with CO₂-dependent hydrogenotrophic methanogenesis remain unclear. The
368 fact that it is a simpler metabolism, requiring fewer genes than CO₂-dependent
369 hydrogenotrophic methanogenesis might suggest its earliest origin. However, it may also
370 signify that it could have emerged later through loss of the WL pathway and/or HGT, as
371 suggested by the grouping of most archaea sharing this metabolism in the phylogenies of
372 MCR (Fig. 2A) and of m4 to m9 markers (Supplementary Fig. 10). Also, the clustering of NM4
373 with Verstraetearchaeota on a separate and well-supported clade in the MCR tree (Fig. 2A) is
374 compatible with a possible inheritance of this metabolism from the last archaeal common
375 ancestor, even under the classical root in between Euryarchaeota and the TACK. However,
376 the possibility of an acquisition through ancient HGT cannot be excluded at present. More
377 insights into the ancestral type of methanogenesis might also be gained from re-examination
378 of the root of the archaeal tree⁵² including all recently discovered archaeal lineages.

379 The phylogenetic placement of the ANME lineages (Fig. 1), strongly suggests that the
380 capabilities for anaerobic methanotrophy emerged multiple times independently during

381 archaeal diversification. In the Methanosarcinales this could have occurred relatively
382 recently and repeatedly by reversal of methanogenesis, possibly through switch of function
383 of a resident canonical MCR, leading to the different ANME-2 (Fig. 2A) and possibly ANME-3
384 lineages. The pool of genes associated with energy conservation in methanogenic and
385 methanotrophic Methanosarcinales is in fact relatively similar⁵⁵ (Fig. 3 and Supplementary
386 Fig. 3) and some methanogenic Methanosarcinales encode c-type multiheme cytochromes¹¹
387 providing the necessary background for electron transfer in AOM archaea.

388 The identification and experimental demonstration of the capacity for oxidation of
389 short-chain fatty-acids (butane, propane) by a divergent MCR-like complex in the
390 Syntrophoarchaeales²¹ is among the most interesting findings of the recent years in the
391 field of environmental microbiology. Our results extend the distribution of these MCR-like
392 complexes in the archaea (Fig. 2A), and therefore also of potential short-chain alkane
393 oxidation capabilities (Figs. 3 and 4). The rapid evolutionary rates of MCR-like homologues
394 coupled to the change of key residues (Fig. 2B) suggest that these complexes might have
395 arisen from canonical MCRs through modifications in the catalytic site to accommodate
396 larger hydrocarbons than methane. Transitions between anaerobic methanotrophy and
397 short-chain alkane utilisation could have occurred in both directions as suggested by i) the
398 close phylogenetic relationships between ‘*Ca. Methanophagales*’ and ‘*Ca.*
399 *Syntrophoarchaeales*’ and the position of GoM-Arc1 within a clade comprising ANME-
400 2a/ANME-2d (Fig. 1), ii) the proposed mechanism of alkane activation in their MCR/MCR-like
401 complexes²¹, iii) their very similar modes of energy conservation (Supplementary Fig. 3), and
402 iv) their numerous shared markers (Table 2). If GoM-Arc1 is a short-chain alkane oxidizer, as
403 suggested by its MCR-like complex, this capacity could have emerged from methanotrophy.
404 Conversely, ‘*Ca. Methanophagales*’ (ANME-1) might have shifted from short-chain alkane
405 oxidation to methanotrophy after acquisition of their MCR through HGT (Fig. 2). Finally, the
406 first report of co-existence of an MCR and an MCR-like complex in members of the “*Ca.*
407 *Methanoliparia*” class opens up the possibility of an additional type of methanogenesis
408 associated with alkane and/or LCFA oxidation.

409 Further exploration of archaeal lineages with an MCR/MCR-like complex and their
410 experimental characterization will lead to a more complete understanding of methane
411 metabolisms and their derivations, as well as their environmental impact.

412

413

414 **Methods**

415

416 ***Metagenomic database probing and contig binning***

417 Contigs of 6108 metagenomes publicly available on the IMG/JGI database in April 2017 were
418 screened for the presence of COG4058, corresponding to McrA, using search tools of the
419 database. 819 contigs containing an McrA sequence with a minimal length of 750 bp were
420 downloaded. The McrA sequences present on these contigs were aligned on those of 188
421 published genomes by using Mafft⁵⁶ (mafft-linsi) and were trimmed with BMGE⁵⁷
422 (BLOSUM30). A maximum likelihood (ML) phylogeny was calculated in IQTree⁵⁸ with the
423 TEST option for best model selection and 100 bootstrap replicates. Metagenomes containing
424 one or several contigs coding for an McrA homologue that was only distantly related to
425 know lineages or belonged to undersampled lineages were downloaded from the IMG
426 database. These metagenomes were assembled with MetaSPAdes⁵⁹, IDBA-UD⁶⁰ and Newbler
427 (Roche) (see Supplementary Table 4 for details). The contigs were binned with ESOM,

428 MetaBAT⁶¹, ABAWACA 1.07 (<http://ggkbase.berkeley.edu/>), MaxBin 2.0⁶² and CONCO⁶³
429 (Supplementary Table 4). An in-house pipeline (Let-it-bin,
430 <https://github.com/QuentinLetourneur/Let-it-bin>) was used for read trimming, assembly,
431 and contig binning. Two of the MAGs were refined using DAS_Tool⁶⁴. Completeness and
432 contamination of the assembled MAGs were estimated with CheckM⁶⁵.

433

434 ***Phylogenomic analyses***

435 A reference archaeal phylogeny was built from a concatenation of 40 phylogenetic markers
436 corresponding to the 36 proteins of the PhyloSift dataset⁶⁶, plus the alpha and beta subunits
437 of the RNA polymerase and two universal ribosomal proteins (L30, S4) (Supplementary Table
438 5). We used a subset of the genomes available for each order/class/phylum level lineages
439 (Supplementary Table 6) to minimize biased associated with uneven distribution of taxa
440 among them (e.g. >100 taxa in Halobacteriales vs. 3 taxa in Methanocellales). The 147
441 genomes were chosen because they were the most complete and the most distant to each
442 other within each lineage. Two phylogenies were built from a concatenation of McrABG and
443 of six co-localized markers (m4 to m9) specific to genomes encoding an MCR/MCR-like
444 complex. Sequences used for these trees were searched by HMM in the ten MAGs obtained
445 in this study and in genomes present in the NCBI or IMG-databases, aligned with Mafft⁵⁶
446 (mafft-linsi), trimmed with BMGE⁵⁷ (BLOSUM30) and concatenated with an in-house script.
447 Before concatenation, the genes of the two datasets (McrABG and m4 to m9) were tested
448 for congruence using the Internode Certainty (IC)⁶⁷ test in RaxML⁶⁸. Maximum likelihood
449 phylogenies for each gene and blind concatenations were calculated in IQTree⁵⁸ with the
450 TEST option for best model selection and 100 bootstrap replicates. Sequences causing strong
451 incongruences (with a bootstrap>=80%) at high taxonomic ranks (order to phylum as
452 applicable) were removed, and the procedure was repeated until no further incongruence
453 was found. Bayesian phylogenies were constructed in PhyloBayes⁶⁹ under the CAT+GTR+Γ4
454 model. Four independent Markov chain Monte Carlo chains were run until convergence and
455 checked by sampling ever two cycles with a 25% burn-in. Support at nodes was evaluated by
456 posterior probability values. ML phylogenies were constructed in IQ-TREE⁵⁸ under the
457 LG+C60 model.

458

459 ***Metabolic prediction***

460 Gene prediction was performed using Prodigal⁷⁰. All metabolic genes were identified using
461 hidden Markov models (HMMs) searches with PFAM, TIGR and custom HMM profiles.
462 Annotation of proteins displayed in Supplementary Table 1 were improved by inspecting the
463 genomic context of the metabolic genes using RAST⁷¹ and SyntTax⁷² ([http://archaea.u-
464 psud.fr/synttax/](http://archaea.u-psud.fr/synttax/)), by phylogenetic analyses including the sequences of characterized
465 enzymes and by identifying their conserved domains using CD-search batch⁷³
466 (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>). Genomic context was also
467 inspected to identify conserved patterns among multiple genomes. Identification of energy
468 conservation systems was performed with MacSyFinder⁷⁴, by defining specific and sensitive
469 models for each system followed by manual curation.

470

471

472 **Data availability**

473 MAG sequences are available in the BioProject PRJNA472146 and Biosamples
474 SAMN10387997, SAMN10390728, SAMN10390732, SAMN10390733, SAMN10390735,

475 SAMN10390736, SAMN10390737, SAMN10390738, SAMN10390739. NM2 sequences
476 corresponding to markers reported in Table 2 are deposited under MK202738 to MK202758.
477 The data that support the findings of this study are available from the corresponding author
478 upon request.
479

480

481 References

482

- 483 1. Demirel, B. & Scherer, P. The roles of acetotrophic and hydrogenotrophic
484 methanogens during anaerobic conversion of biomass to methane: A review. *Rev.*
485 *Environ. Sci. Biotechnol.* **7**, 173–190 (2008).
- 486 2. Schink, B. Energetics of syntrophic cooperation in methanogenic degradation.
487 *Microbiol. Mol. Biol. Rev.* **61**, 262–280 (1997).
- 488 3. Ueno, Y., Yamada, K., Yoshida, N., Maruyama, S. & Isozaki, Y. Evidence from fluid
489 inclusions for microbial methanogenesis in the early Archaean era. *Nature* **440**, 516–
490 519 (2006).
- 491 4. Sousa, F. L. *et al.* Early bioenergetic evolution. *Philosophical Transactions of the Royal*
492 *Society B: Biological Sciences* **368**, (2013).
- 493 5. Kasting, J. F. & Siefert, J. L. Life and the evolution of Earth’s atmosphere. *Science* **296**,
494 1066–1068 (2002).
- 495 6. Bapteste, E., Brochier, C. & Boucher, Y. Higher-level classification of the Archaea:
496 evolution of methanogenesis and methanogens. *Archaea* **1**, 353–363 (2005).
- 497 7. Whitman, W. B., Bowen, T. L. & Boone, D. R. *The Methanogenic Bacteria. Prokaryotes*
498 *Vol. 3* (2006). doi:10.1007/0-387-30743-5_9
- 499 8. Kendall, M. M. & Boone, D. R. in *The Prokaryotes* 244–256 (2006). doi:10.1007/0-387-
500 30743-5_12
- 501 9. Oren, A. in *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*
502 **9783642389**, 165–193 (2014).
- 503 10. Borrel, G. *et al.* Phylogenomic data support a seventh order of methylotrophic
504 methanogens and provide insights into the evolution of methanogenesis. *Genome*
505 *Biol. Evol.* **5**, 1769–1780 (2013).
- 506 11. Timmers, P. H. A. *et al.* Reverse Methanogenesis and Respiration in Methanotrophic
507 Archaea. *Archaea* **2017**, (2017).
- 508 12. Scheller, S., Ermler, U. & Shima, S. in *Anaerobic Utilization of Hydrocarbons, Oils, and*
509 *Lipids* 1–29 (2017). doi:10.1007/978-3-319-33598-8_3-1
- 510 13. Hallam, S. J. *et al.* Reverse methanogenesis: Testing the hypothesis with
511 environmental genomics. *Science*. **305**, 1457–1462 (2004).
- 512 14. Borrel, G., Adam, P. S. & Gribaldo, S. Methanogenesis and the Wood-Ljungdahl
513 Pathway: An Ancient, Versatile, and Fragile Association. *Genome Biol. Evol.* **8**, 1706–
514 1711 (2016).
- 515 15. Adam, P. S., Borrel, G., Brochier-Armanet, C. & Gribaldo, S. The growing tree of
516 Archaea: New perspectives on their diversity, evolution and ecology. *ISME J.* **11**,
517 2407–2425 (2017).
- 518 16. Spang, A., Caceres, E. F. & Ettema, T. J. G. Genomic exploration of the diversity,
519 ecology, and evolution of the archaeal domain of life. *Science* **357**, (2017).
- 520 17. Nobu, M. K., Narihiro, T., Kuroda, K., Mei, R. & Liu, W.-T. Chasing the elusive
521 Euryarchaeota class WSA2: genomes reveal a uniquely fastidious methyl-reducing
522 methanogen. *ISME J.* **10**, 2478 – 2487 (2016).
- 523 18. Sorokin, D. Y. *et al.* Discovery of extremely halophilic, methyl-reducing euryarchaea
524 provides insights into the evolutionary origin of methanogenesis. *Nat. Microbiol.* **2**,
525 17081 (2017).
- 526 19. Vanwonterghem, I. *et al.* Methylotrophic methanogenesis discovered in the novel

- 527 archaeal phylum Verstraetearchaeota. *Nat. Microbiol.* **1**, 16170 (2016).
- 528 20. Brugère, J. F. *et al.* Archaeobiotics: Proposed therapeutic use of archaea to prevent
529 trimethylaminuria and cardiovascular disease. *Gut Microbes* **5**, 5–10 (2014).
- 530 21. Laso-Pérez, R. *et al.* Thermophilic archaea activate butane via alkyl-coenzyme M
531 formation. *Nature* **539**, 396–401 (2016).
- 532 22. Evans, P. N. *et al.* Methane metabolism in the archaeal phylum Bathyarchaeota
533 revealed by genome-centric metagenomics. *Science* **350**, 434–438 (2015).
- 534 23. Dombrowski, N., Seitz, K. W., Teske, A. P. & Baker, B. J. Genomic insights into potential
535 interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal
536 sediments. *Microbiome* **5**, (2017).
- 537 24. McKay, L. J., Hatzenpichler, R., Inskeep, W. P. & Fields, M. W. Occurrence and
538 expression of novel methyl-coenzyme M reductase gene (*mcrA*) variants in hot spring
539 sediments. *Sci. Rep.* **7**, (2017).
- 540 25. Hawley, E. R. *et al.* Metagenomes from two microbial consortia associated with Santa
541 Barbara seep oil. *Mar. Genomics* **18**, 97–99 (2014).
- 542 26. Seitz, K. W., Lazar, C. S., Hinrichs, K.-U., Teske, A. P. & Baker, B. J. Genomic
543 reconstruction of a novel, deeply branched sediment archaeal phylum with pathways
544 for acetogenesis and sulfur reduction. *ISME J.* **10**, 1696–1705 (2016).
- 545 27. Ermler, U., Grabarse, W., Shima, S., Goubeaud, M. & Thauer, R. K. Crystal structure of
546 methyl-coenzyme M reductase: the key enzyme of biological methane formation.
547 *Science* **278**, 1457–1462 (1997).
- 548 28. Arshad, A. *et al.* A metagenomics-based metabolic model of nitrate-dependent
549 anaerobic oxidation of methane by Methanoperedens-like archaea. *Front. Microbiol.*
550 **6**, (2015).
- 551 29. Wegener, G., Krukenberg, V., Riedel, D., Tegetmeyer, H. E. & Boetius, A. Intercellular
552 wiring enables electron transfer between methanotrophic archaea and bacteria.
553 *Nature* **526**, 587–590 (2015).
- 554 30. McGlynn, S. E., Chadwick, G. L., Kempes, C. P. & Orphan, V. J. Single cell activity
555 reveals direct electron transfer in methanotrophic consortia. *Nature* **526**, 531–535
556 (2015).
- 557 31. Meyerdierks, A. *et al.* Metagenome and mRNA expression analyses of anaerobic
558 methanotrophic archaea of the ANME-1 group. *Environ. Microbiol.* **12**, 422–439
559 (2010).
- 560 32. Grimaldi, S., Schoepp-Cothenet, B., Ceccaldi, P., Guigliarelli, B. & Magalon, A. The
561 prokaryotic Mo/W-bisPGD enzymes family: A catalytic workhorse in bioenergetic.
562 *Biochim. Biophys. Acta - Bioenerg.* **1827**, 1048–1085 (2013).
- 563 33. Jormakka, M. *et al.* Molecular mechanism of energy conservation in polysulfide
564 respiration. *Nat. Struct. Mol. Biol.* **15**, 730–737 (2008).
- 565 34. Hagemeyer, C. H., Chistoserdova, L., Lidstrom, M. E., Thauer, R. K. & Vorholt, J. A.
566 Characterization of a second methylene tetrahydromethanopterin dehydrogenase
567 from *Methylobacterium extorquens* AM1. *Eur. J. Biochem.* **267**, 3762–3769 (2000).
- 568 35. McInerney, M. J. *et al.* The genome of *Syntrophus aciditrophicus*: Life at the
569 thermodynamic limit of microbial growth. *Proc. Natl. Acad. Sci.* (2007).
570 doi:10.1073/pnas.0610456104
- 571 36. Klenk, H.-P. *et al.* The complete genome sequence of the hyperthermophilic, sulphate-
572 reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364–370 (1997).
- 573 37. Milkov, A. V. Molecular and stable isotope compositions of natural gas hydrates: A

- 574 revised global dataset and basic interpretations in the context of geological settings.
575 *Org. Geochem.* **36**, 681–702 (2005).
- 576 38. Meredith, W., Kelland, S. J. & Jones, D. M. Influence of biodegradation on crude oil
577 acidity and carboxylic acid composition. *Org. Geochem.* **31**, 1059–1073 (2000).
- 578 39. Sieber, J. R., McInerney, M. J. & Gunsalus, R. P. Genomic Insights into Syntrophy: The
579 Paradigm for Anaerobic Metabolic Cooperation. *Annu. Rev. Microbiol.* (2012).
580 doi:10.1146/annurev-micro-090110-102844
- 581 40. Zengler, K., Richnow, H. H., Rosselló-Mora, R., Michaelis, W. & Widdel, F. Methane
582 formation from long-chain alkanes by anaerobic microorganisms. *Nature* (1999).
583 doi:10.1038/45777
- 584 41. Stams, A. J. M., Sousa, D. Z., Kleerebezem, R. & Plugge, C. M. Role of syntrophic
585 microbial communities in high-rate methanogenic bioreactors. *Water Sci. Technol.*
586 (2012). doi:10.2166/wst.2012.192
- 587 42. Gao, B. & Gupta, R. S. Phylogenomic analysis of proteins that are distinctive of
588 Archaea and its main subgroups and the origin of methanogenesis. *BMC Genomics* **8**,
589 86 (2007).
- 590 43. Kaster, A. K. *et al.* More than 200 genes required for methane formation from H₂ and
591 CO₂ and energy conservation are present in *Methanothermobacter marburgensis* and
592 *Methanothermobacter thermoautotrophicus*. *Archaea* **2011**, (2011).
- 593 44. Nayak, D. D., Mahanta, N., Mitchell, D. A. & Metcalf, W. W. Post-translational
594 thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and
595 methanotrophic archaea. *Elife* **6**, e29218 (2017).
- 596 45. Lyu, Z. *et al.* Mmp10 is required for post-translational methylation of arginine at the
597 active site of methyl-coenzyme M reductase. *BioRxiv* 211441 (2017).
- 598 46. Sarmiento, F., Mrázek, J. & Whitman, W. B. Genome-scale analysis of gene function in
599 the hydrogenotrophic methanogenic archaeon *Methanococcus maripaludis*. *Proc.*
600 *Natl. Acad. Sci. U. S. A.* **110**, 4726–4731 (2013).
- 601 47. Wagner, T., Kahnt, J., Ermler, U. & Shima, S. Didehydroaspartate Modification in
602 Methyl-Coenzyme M Reductase Catalyzing Methane Formation. *Angew. Chemie - Int.*
603 *Ed.* **55**, 10630–10633 (2016).
- 604 48. Borrel, G. *et al.* Comparative genomics highlights the unique biology of
605 *Methanomassiliicoccales*, a Thermoplasmatales-related seventh order of
606 methanogenic archaea that encodes pyrrolysine. *BMC Genomics* **15**, 679 (2014).
- 607 49. Zheng, K., Ngo, P. D., Owens, V. L., Yang, X. & Mansoorabadi, S. O. The biosynthetic
608 pathway of coenzyme F₄₃₀ in methanogenic and methanotrophic archaea. *Science.*
609 **354**, 339–342 (2016).
- 610 50. Prakash, D., Wu, Y., Suh, S. J. & Duin, E. C. Elucidating the process of activation of
611 methyl-coenzyme M reductase. *J. Bacteriol.* **196**, 2491–2498 (2014).
- 612 51. Li, J. *et al.* Global mapping transcriptional start sites revealed both transcriptional and
613 post-transcriptional regulation of cold adaptation in the methanogenic archaeon
614 *Methanobrevibacter smithii*. *Sci. Rep.* **5**, 9209 (2015).
- 615 52. Raymann, K., Brochier-Armanet, C. & Gribaldo, S. The two-domain tree of life is linked
616 to a new root for the Archaea. *Proc. Natl. Acad. Sci.* **112**, 6670–6675 (2015).
- 617 53. Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor.
618 *Nat. Microbiol.* **1**, (2016).
- 619 54. Adam, P. S., Borrel, G. & Gribaldo, S. Evolutionary history of carbon monoxide
620 dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc.*

- 621 *Natl. Acad. Sci.* **115**, E5837 (2018).
- 622 55. McGlynn, S. E. Energy Metabolism during Anaerobic Methane Oxidation in ANME
623 Archaea. *Microbes Environ.* **32**, 5–13 (2017).
- 624 56. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
625 Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
- 626 57. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): A new
627 software for selection of phylogenetic informative regions from multiple sequence
628 alignments. *BMC Evol. Biol.* **10**, 210 (2010).
- 629 58. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and
630 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol.*
631 *Biol. Evol.* **32**, 268–274 (2015).
- 632 59. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. MetaSPAdes: A new versatile
633 metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
- 634 60. Peng, Y., Leung, H. C. M., Yiu, S.-M. & Chin, F. Y. L. IDBA-UD: a de novo assembler for
635 single-cell and metagenomic sequencing data with highly uneven depth.
636 *Bioinformatics* **28**, 1420–1428 (2012).
- 637 61. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately
638 reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165
639 (2015).
- 640 62. Wu, Y. W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: An automated binning
641 algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*
642 **32**, 605–607 (2015).
- 643 63. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat.*
644 *Methods* **11**, 1144–1146 (2014).
- 645 64. Sieber, C. M. K. *et al.* Recovery of genomes from metagenomes via a dereplication,
646 aggregation, and scoring strategy. *BioRxiv* 107789 (2017).
647 doi:<http://dx.doi.org/10.1101/107789>
- 648 65. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM:
649 Assessing the quality of microbial genomes recovered from isolates, single cells, and
650 metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
- 651 66. Darling, A. E. *et al.* PhyloSift: phylogenetic analysis of genomes and metagenomes.
652 *PeerJ* **2**, e243 (2014).
- 653 67. Kobert, K., Salichos, L., Rokas, A. & Stamatakis, A. Computing the Internode Certainty
654 and Related Measures from Partial Gene Trees. *Mol. Biol. Evol.* **33**, 1606–1617 (2016).
- 655 68. Stamatakis, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with
656 thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
- 657 69. Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: A Bayesian software package for
658 phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286–2288
659 (2009).
- 660 70. Hyatt, D. *et al.* Prodigal: Prokaryotic gene recognition and translation initiation site
661 identification. *BMC Bioinformatics* **11**, 119 (2010).
- 662 71. Aziz, R. K. *et al.* The RAST Server: Rapid annotations using subsystems technology.
663 *BMC Genomics* **9**, 75 (2008).
- 664 72. Oberto, J. SyntTax: A web server linking synteny to prokaryotic taxonomy. *BMC*
665 *Bioinformatics* **14**, 4 (2013).
- 666 73. Marchler-Bauer, A. *et al.* CDD: NCBI’s conserved domain database. *Nucleic Acids Res.*
667 **43**, D222–D226 (2015).

- 668 74. Abby, S. S., Néron, B., Ménager, H., Touchon, M. & Rocha, E. P. C. MacSyFinder: A
669 program to mine genomes for molecular systems with an application to CRISPR-Cas
670 systems. *PLoS One* **9**, (2014).
671 75. Kahnt, J. *et al.* Post-translational modifications in the active site region of methyl-
672 coenzyme M reductase from methanogenic and methanotrophic archaea. *FEBS J.* **274**,
673 4913–4921 (2007).
674
675

676 Corresponding authors

677
678 Correspondence and requests for data could be addressed to guillaume.borrel@pasteur.fr
679 and simonetta.gribaldo@pasteur.fr
680
681
682

683 Acknowledgements

684
685 We thank Rudolf Thauer for feedback on an earlier version of the manuscript. G.B.
686 acknowledges support from the Institut Pasteur through a Roux-Cantarini fellowship. P.S.A.
687 is supported by a PhD fellowship from Paris Diderot University and by funds from the PhD
688 Program “Frontières du Vivant (FdV)-Programme Bettencourt”. S.G. acknowledges funding
689 from the French National Agency for Research Grant ArchEvol (ANR-16-CE02-0005-01). This
690 work used the computational and storage services (TARS cluster) provided by the IT
691 department at Institut Pasteur, Paris. S.J.H. acknowledges support from the US Department
692 of Energy (DOE) JGI supported by the Office of Science of US DOE Contract DE-AC02-
693 05CH11231, the Natural Sciences and Engineering Research Council (NSERC) of Canada,
694 Genome British Columbia, Genome Canada, Canada Foundation for Innovation (CFI), and the
695 Tula Foundation. I.N.S.G. and V.M.O. are grateful to São Paulo Research Foundation - FAPESP
696 (process numbers 2011/14501-6 and 2013/20436-8) and Petrobras for financial support and
697 to Dr. Neil Gray and Dr. Ian Head from the School of Civil Engineering and Geosciences at
698 Newcastle University for lab facilities. W-J.L. was supported by Key Projects of Ministry of
699 Science and Technology (MOST) (Nos. 2013DFA31980, 2015FY110100). G.M. was supported
700 by the ERC Advanced Grant PARASOL (No. 322551). L.J.M. appreciates funding from the
701 NASA Postdoctoral Program through the NASA Astrobiology Institute and W.P.I. was
702 supported by the Montana Agricultural Experiment Station (Project 911300).
703
704

705 Author contributions

706
707 G.B. and S.G. conceived the study. L.J.M., L-X.C., I.N.S.-G., C.M.K.S., G.L.A., W.-J.L., S.J. H.,
708 G.M., V.M.d.O., W.P.I., and J.F.B. sequenced and assembled the metagenomes. G.B.
709 screened the IMG database for McrA and identified these metagenomes. Q. L. , A. G. and G.
710 B. developed the pipeline Let-it-bin. G.B. performed the contig binning of NM1a, NM1b,
711 NM2, NM3, NM4, Verst-YHS, and Mnatro-ASL MAGs. L-X.C. carried out the contig binning of
712 ANME-1-THS MAG and C.M.K.S. those of GoM-Arc1-GOS and ANME-2c MAGs. G.B. inferred
713 the metabolism associated to each MAG and performed all phylogenetic analyses. P.A.

714 performed the congruence tests. G.B. and S.G. wrote the manuscript. All authors read and
715 commented on the manuscript.

716

717

718 **Competing interests**

719

720 The authors declare no competing interests.

721 **Table legends**

722

723 **Table 1:** General information on the ten MAGs obtained in this study.

724

725 **Table 2:** Occurrence in methanogens, methanotrophs and short-chain alkane users of 38
726 genes previously suggested as methanogenesis markers.

727

728

729 **Figure legends**

730

731 **Figure 1:** Placement of nine MAGs described in this study in the reference phylogeny of
732 Archaea. NM2 was not included due to low completeness. Bayesian phylogeny (PhyloBayes,
733 CAT+GTR+ Γ 4) based on concatenation of 40 conserved phylogenetic markers (8,564 amino
734 acid positions) and 156 genomes/MAGs (see Supplementary Table 5 and Supplementary
735 Table 6 for detail). Node supports refer to posterior probabilities, and for reasons of
736 readability only values above 0.8 are shown. The tree is rooted according to Raymann et
737 al.⁵². The scale bar represents the average number of substitutions per site. Black arrows
738 point to the 9 obtained MAGs and accolated pie charts indicate their estimated
739 completeness. Colors indicate that genomes of these lineages encode an MCR/MCR-like
740 complex, Class I/II methanogens are in green, methyl-dependent hydrogenotrophic lineages
741 are in red, methanotrophs are in orange (some being within Class II), potential or validated
742 shot-chain alkane users are in in blue. NM1 could also have a methane metabolism (see text
743 for discussion).

744

745 **Figure 2:** Phylogeny of the MCR/MCR-like complex and conservation of important positions
746 in the catalytic site. A) Unrooted Bayesian phylogeny (CAT+GTR+ Γ 4) based on a
747 concatenation of McrABG/McrABG-like subunits (1,187 amino acid positions) from 109
748 genomes/MAGs (see Supplementary Table 6 for details). Node supports refer to posterior
749 probabilities, and for reasons of readability only values above 0.8 are shown. The scale bar
750 represents the average number of substitutions per site. The color code is similar to that in
751 Fig. 1 with the exception of NM1 which have both an MCR-like (in blue) and a canonical MCR
752 (in purple) (see text for discussion). B) Conservation of 17 residues previously described to
753 interact with CoM, CoB, F₄₃₀ cofactors, making part of the substrate cavity wall, or having
754 post-translational modifications^{27,47,75}. Replacement of conserved amino acids associated to
755 a negative value in the Blosum45 matrix are indicated by white on black background, those
756 with a null or positive value in the Blosum45 matrix are in bold, "." indicate conserved
757 positions and "-" indicate missing positions in the sequence due to sequencing
758 incompleteness.

759

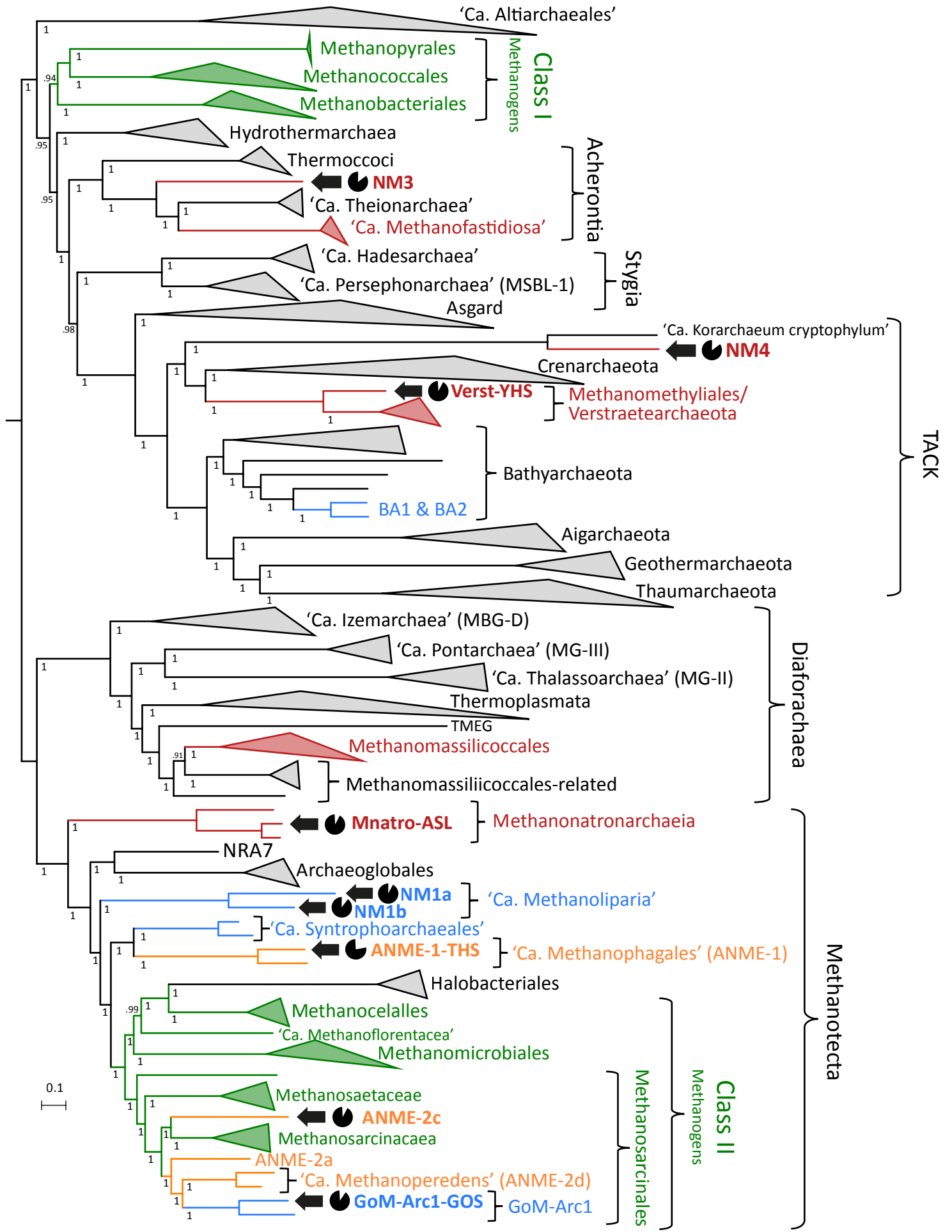
760 **Figure 3:** Predicted methane and short-chain alkane metabolism of the MAGs described in
761 this study, with the exception of NM1, which is presented in Fig. 4, and NM2, which has a
762 low completeness. Colored arrows correspond to reactions modifying or transferring the C1
763 carbon group of the substrate. Details on the annotation of the enzymes are presented in
764 Supplementary Table 1. MFR, methanofuran; H₄MPT, tetrahydromethanopterin; Fd,
765 ferredoxin; F₄₂₀, coenzyme F₄₂₀; LCFA, Long Chain Fatty Acids; MQ, menaquinone; Mp,
766 methanophenazine; Mhc, c-type multiheme cytochromes; DIET, Direct Interspecies Electron
767 Transfer. Grey color indicates the absence of the enzyme, complex, reaction or compound.

768 Comparisons of with other methane-cycling or short-chain alkane oxidizers, which are
769 discussed in the text, are presented in Supplementary Figs 1 and 3. The percentages
770 between brackets indicate the estimated completeness of the corresponding MAGs.

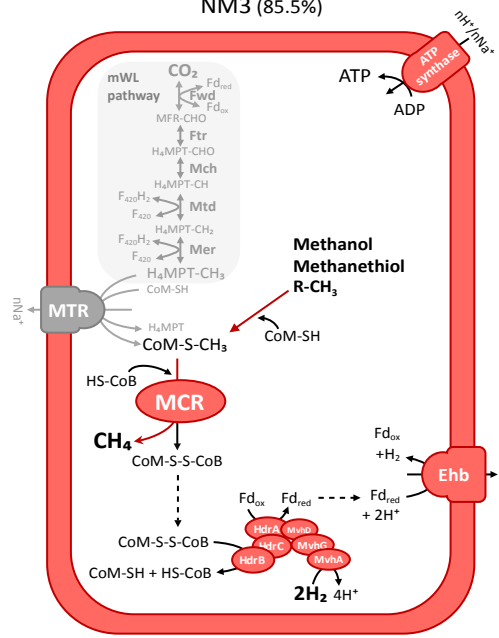
771

772 **Figure 4:** Predicted methane, short-chain alkane and long/medium-chain fatty acid
773 metabolism of the two MAGs NM1a ("*Ca. Methanoliparum thermophilum*" completeness of
774 92.5%) and NM1b ("*Ca. Methanolliviera hydrocarbonicum*" completeness of 90.2%)
775 belonging to the candidate class "*Ca. Methanoliparia*". Colored arrows correspond to
776 reactions present in both MAGs which modify or transfer the carbon group(s) of the
777 substrate. Predicted possible metabolisms are the utilisation of short-chain alkanes and
778 long/medium-chain fatty acids (L/MCFA) (blue), methanotrophy (orange) and
779 methanogenesis from short-chain alkanes or L/MCFA (purple). Details on the annotation of
780 the enzymes are provided in Supplementary Table 1. MFR, methanofuran; H₄MPT,
781 tetrahydromethanopterin; Fd, ferredoxin; F₄₂₀, coenzyme F₄₂₀; L/MCFA, Long/medium chain
782 fatty acids; MQ, menaquinone; Mhc, c-type multiheme cytochromes; DIET, Direct
783 Interspecies Electron Transfer. Grey color indicates the absence of the corresponding
784 enzyme, complex, reaction or compound in both MAGs.

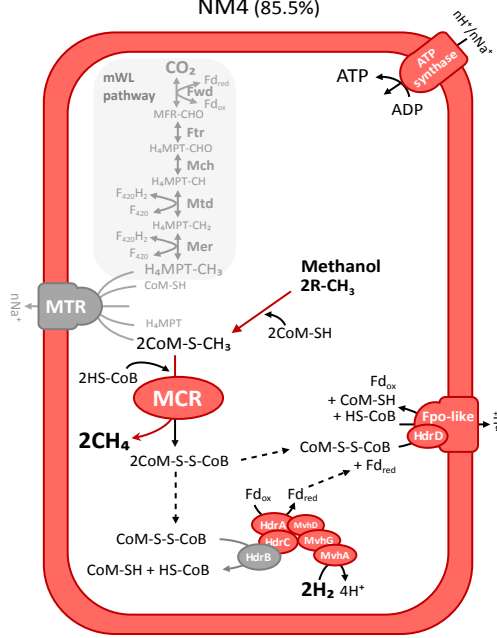
785



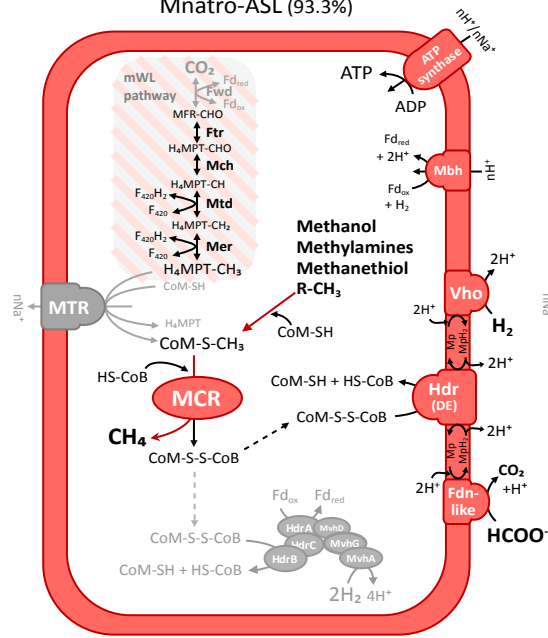
NM3 (85.5%)



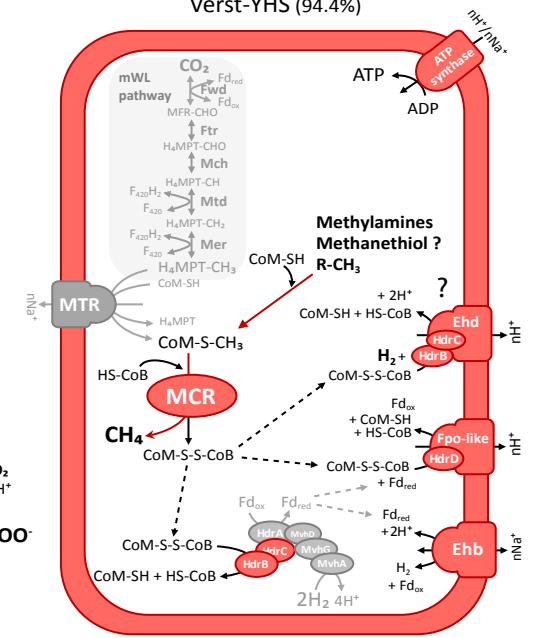
NM4 (85.5%)



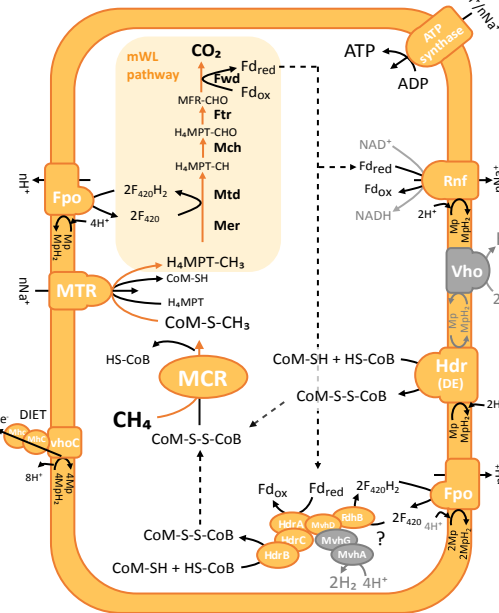
Mnatro-ASL (93.3%)



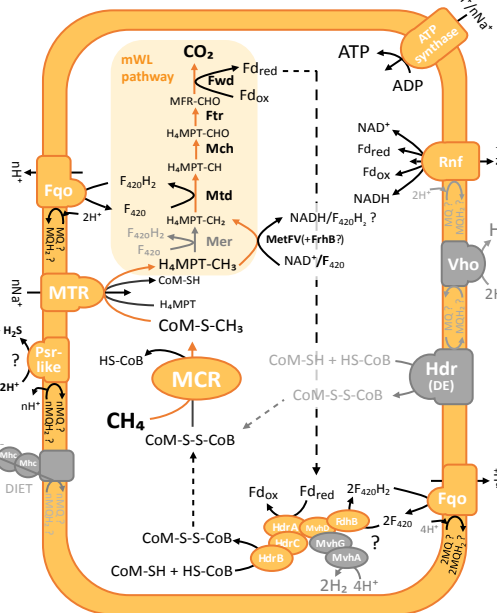
Verst-YHS (94.4%)



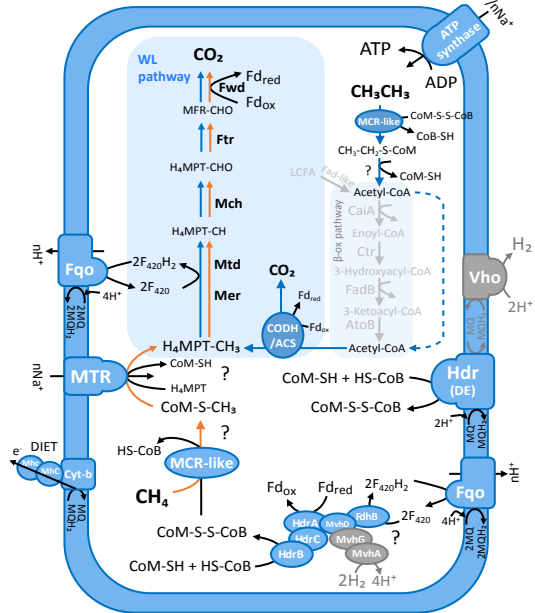
ANME-2c (92%)



ANME-1-THS (78.4%)



GoM-Arc1-GOS (91.2%)



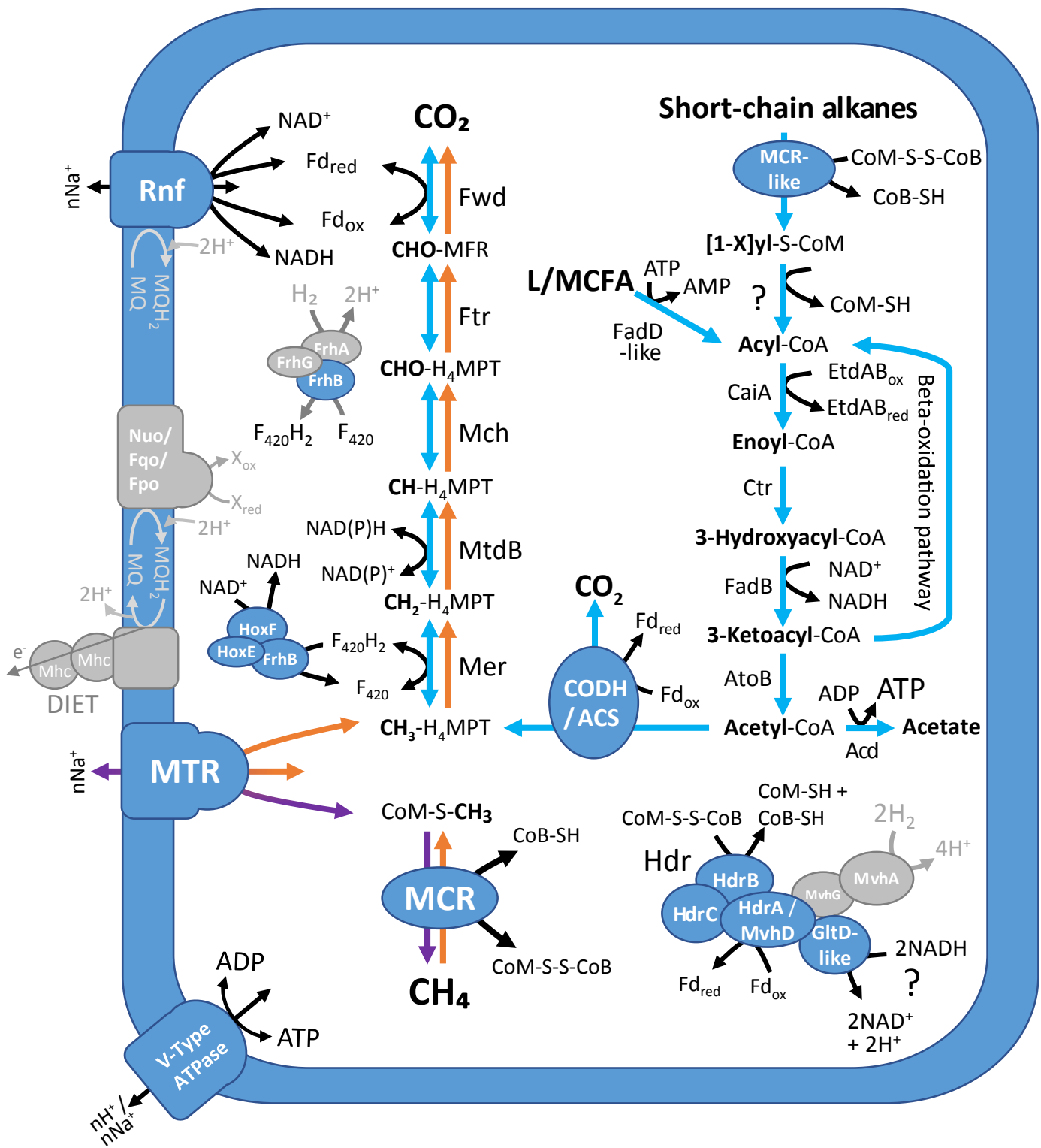


Table 1: General information on the ten MAGs obtained in this study.

Genome	Origin	Scaffold (nbr)	Size (Mb)	Genes (nbr)	GC (%)	Compl. (%)	Cont. (%)	Strain hetero.*	Cont. excl. strain hetero. (%)
NM1a	Enrichment culture (50°C) from petroleum sample, Brazil	12	1.26	1388	35.7	92.5	1.3	0	1.3
NM1b	Santa Barbara Channel oil seeps, USA	183	1.66	1860	43.8	90.2	3.6	66.7	1.2
NM2	Santa Barbara sediments, USA	210	1.03	1254	41.8	51.5	2	0	2
NM3	Enrichment culture (40°C) from petroleum sample, Brazil	26	1.49	1578	55.2	85.5	2.9	0	2.9
NM4	Yellowstone sulfidic hot spring, USA	122	1.42	1603	43.4	85.5	1.8	66.7	0.6
Verst-YHS	Yellowstone sulfidic Hot Spring, USA	46	1.05	1220	28.4	94.4	0	0	0
Mnatro-ASL	Altai Soda Lake sediments, Russia	73	1.34	1451	42.2	93.3	1.3	0	1.3
ANME-1-THS	Tibetan Hot Spring sediment, China	181	2.03	2155	48.7	78.4	4.9	33	3.3
GoM-Arc1-GOS	Gulf of Mexico natural Oil Seep, USA	119	1.46	1623	41.0	91.2	0	0	0
ANME-2c	Gulf of Mexico natural oil seep, USA	249	2.66	2867	48.5	92	2	0	2

Genome ID, origin, number of scaffolds, number of protein-coding genes, guanine-cytosine (GC) content, estimated completeness (Compl.), estimated contamination (Cont.), strain heterogeneity (Strain) and contamination excluding strain heterogeneity (Cont. excl. strain hetero.) are shown. *percentage of contamination that can be due to binning of contigs from closely related strains.

