Reply to 'Evolutionary placement of Methanonatronarchaeia'

Sorokin, Dimitry Y.; Makarova, Kira S.; Abbas, Ben; Ferrer, Manuel; Golyshin, Peter N.; Galinski, Erwin A.; Ciorda, Sergio; Mena, María Carmen; van Loosdrecht, Mark C.M.; More Authors

DOI
10.1038/s41564-019-0358-0

Publication date
2019

Document Version
Accepted author manuscript

Published in
Nature Reviews Microbiology

Citation (APA)
Sorokin, D. Y., Makarova, K. S., Abbas, B., Ferrer, M., Golyshin, P. N., Galinski, E. A., Ciorda, S., Mena, M. C., van Loosdrecht, M. C. M., & More Authors (2019). Reply to 'Evolutionary placement of Methanonatronarchaeia'. *Nature Reviews Microbiology*, *4*(4), 560-561. https://doi.org/10.1038/s41564-019-0358-0

More genomes needed to resolve archaeal phylogeny

Dimitry Y. Sorokin[1,2*], Kira S. Makarova[3], Ben Abbas[2], Manuel Ferrer[4], Peter N. Golyshin[5], Erwin A. Galinski[6], Sergio Ciorda[7], María Carmen Mena[7], Alexander Y. Merkel[1], Yuri I. Wolf[3], Mark C.M. van Loosdrecht[2], Eugene V. Koonin[3*]

Response to Monique Aouad, Guillaume Borrel, Céline Brochier-Armanet, and Simonetta Gribaldo

"Methanonatronarchaeia are not evolutionary intermediates on the path from methanogens to extreme halophiles"

[1]*Winogradsky Institute of Microbiology, Centre for Biotechnology, Russian Academy of Sciences, Moscow, Russia;*
[2]*Department of Biotechnology, Delft University of Technology, Delft, The Netherlands;*
[3]*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA;*
[4]*Institute of Catalysis, CSIC, Madrid, Spain;*
[5]*School of Biological Sciences, Bangor University, Gwynedd, UK*
[6]*Institute of Microbiology and Biotechnology, Rheinische Friedrich-Wilhelms University, Bonn, Germany*
[7]*Proteomics Facility, Centro Nacional de Biotecnología, CSIC, Madrid, Spain*

*Corresponding authors:
Dimitry Y. Sorokin: soroc@inmi.ru; d.sorokin@tudelft.nl
Eugene V. Koonin: koonin@ncbi.nlm.nih.gov

Standfirst

Different phylogenetic methods applied to different gene sets yield alternative positions for the proposed archaeal class "*Methanonatronoarchaeia*" in the archaeal tree. A more representative sampling of archaeal genomes is essential to resolve this phylogenetic impasse.

We appreciate the interest of Aouad and colleagues in our work on the proposed archaeal class "*Methanonatronoarchaeia"* [1,2] and their effort to clarify the phylogenetic position of this unique group of extremely halophilic, methyl-reducing methanogens. In our analysis, *Methanonatronoarchaeia* formed a clade with the class *Halobacteria*, the non-methanogenic euryarchaeal extreme halophiles. Notably, this phylogenetic placement is 100% bootstrap-supported in maximum-likelihood (ML) phylogenetic trees for both 16S rRNA and concatenated alignments of ribosomal proteins [1]. Given the congruence of the two trees, the strong support for the *Methanonatronoarchaeia-Halobacteria* clade, the biological plausibility of this affinity and the fact that these trees conformed with the currently favored solutions for difficult problems in archaeal phylogeny (such as the monophyly of the DPANN superphylum and the euryarchaeal assemblage including Class I methanogens and Thermococci), we did not perform a more thorough phylogenetic analysis. Such an in-depth analysis was undertaken by Aouad and colleagues [3]. Their results suggest a different position for *Methanonatronoarchaeia*, much deeper in the archaeal tree, outside the branch that consists of Methanomicrobia (formerly, Class II Methanogens), including *Halobacteria* (denoted "Stenosarchaea" by Aouad et al.) and the class *Archaeoglobi*, and at the base of the group which Aouad et al. denote the "superclass Methanotecta". This difference between the results of the two phylogenetic analyses stems primarily from the increasingly stringent removal of fast-evolving sites from the alignment prior to the phylogenetic tree construction that was applied by Aouad and colleagues. After a certain fraction of the fastest sites was removed, the tree topology abruptly transitioned to the deep placement of *Methanonatronoarchaei*a. This procedure is supposed to eliminate the false signal produced by sites with multiple substitutions, and therefore, Aouad et al. conclude that the affinity of *Methanonatronoarchaeia* with *Halobacteria* was an artifact caused by such sites. Aouad et al. also obtained the "deeper" placement of *Methanonatronoarchaeia* with extended sets of conserved protein families and expanded taxon sampling, in these cases, even without removing the fast-evolving sites.

In our view, the position of *Methanonatronoarchaeia* in the archaeal phylogeny remains an open question. Removal of fast-evolving sites is a double-edged sword: it reduces the noise introduced by multiple substitutions but phylogenetic information that is contained in comparatively variable positions is lost as well [4]. The most highly conserved sites are phylogenetically uninformative and so are the most variable ones, whereas those with intermediate variability carry the bulk of the phylogenetic signal [5]. The loss of phylogenetic signal can result in exactly what is observed for *Methanonatronoarchaeia*, namely, losing the information on a specific affinity, in this case, with *Halobacteria*, and pushing a branch down the tree, closer to the root. Inclusion of additional protein families, although potentially enhancing the phylogenetic signal, also has its own caveats. Many of these families are less strongly conserved during evolution than ribosomal proteins are, which leads to less reliable alignments, and many are prone to horizontal gene transfer (HGT), which can dilute the signal. Also, the observations on protein phylogenies cannotexplain away the affinity between *Methanonatronoarchaeia* and *Halobacteria* in the 16S RNA tree.

The highly conserved ribosomal-based phylogeny is not the only line of evidence that links *Methanonatronoarchaeia* with *Halobacteria*. The two groups share a variety of genes that are not commonly found in other archaea, in particular, those encoding multiple membrane ion transport systems involved in halophily and uncharacterized membrane proteins (see Supplementary Table 3 in Ref. 1). Especially conspicuous is the UspA family of stress response proteins [6] that is dramatically expanded in both *Methanonatronoarchaeia* and *Halobacteria* (see Supplementary Figure 8 in Ref. 1). It appears most likely that these proteins contribute to the extreme salt tolerance. Phylogenetic analysis of the UspA family shows a complex picture, but for a number of branches, inheritance of the respective genes from a common ancestor of *Methanonatronoarchaeia* and *Halobacteria* appears to be the most likely scenario (Supplementary File 1).The two sequenced genomes of *Methanonatronoarchaeia* encompass integrated virus-like elements (His2-like proviruses) that closely resemble viruses of *Halobacteria* (see Table 1 in Ref. 1). Given the generally narrow host range of archaeal viruses [7], the presence of these elements in *Methanonatronoarchaeia* seems to suggest a common evolutionary history with *Halobacteria*. Together, these observations appear to be compatible with a common ancestor of *Methanonatronoarchaeia* and *Halobacteria* that was already adapted to hypersalinity including the expansion of the UspA family. Admittedly, none of this is incontrovertible

evidence, and in particular, HGT always offers an alternative. However, in cases like the UspA family and His2-like elements, the HGT scenario seems less parsimonious than common ancestry.

As Aouad and colleagues point out [3], repositioning *Methanonatronoarchaeia* in the archaeal phylogenetic tree would have distinct biological implications, in particular, indicating independent origins of the adaptations to hypersalinity in *Methanonatronoarchaeia* and *Halobacteria*. The problem runs even deeper because another recent study by Aouad and colleagues [8] suggests also the relocation of the candidate division Nanohaloarchaea from the DPANN superphylum to "Stenosarachaea", suggesting two independent origins of non-methanogenic extreme halophiles from different lineages of Methanomicrobia and putting into question the monophyly of DPANN. A recent comprehensive phylogenetic modeling study has yielded a clear support for a monophyletic DPANN [9]. These phylogenetic travails also resemble the long debate on the position of Nanoarchaea [10-12] that, with the discovery of many other archaea with miniature genomes, seemed to have been settled on the DPANN superphylum. The impending changes to the archaeal phylogeny and taxonomy could be quite profound. A phylogenetic tree of archaea generated from a set of 122 marker proteins using a recently developed methodology for genome phylogenies [13] has led to the proposal of the phylum *Halobacterota* that is placed outside the Euryarchaeota and unites *Archaeoglobi*, *Halobacteria*, *Methanomicrobia*, *Methanonatronoarchaeia*, *Methanosarcini*, and NRA6, with deeply placed *Methanonatronoarchaeia* (http://gtdb.ecogenomic.org/tree).

Deep phylogenies are fraught with uncertainty, so that definitive solutions might be out of reach. However, one remedy seems to be consistently efficient, namely, improved taxon sampling [14,15] which, indeed, has been attempted by Aouad and colleagues [3]. However, the representation of *Methanonatronoarchaeia* remains obviously insufficient to reach compelling conclusions, with the current sample including only two genomes (but, notably, two additional sequences clustering with *Methanonatronoarchaeia* in the 16S RNA tree). Further progress in microbial genome sequencing, in particular, by methods of metagenomics and single-cell genomics, will substantially expand the diversity of archaea available for phylogenomic analysis, providing for more robust phylogenies in the near future. Indeed, a high quality draft single-cell genome corresponding to one of these additional 16S RNA sequences (SA1) has recently become available [16]. There is no doubt that, within a few years,

more genomes will follow, likely, providing for the resolution of the current phylogenetic impasse.

References

1 Sorokin, D. Y. *et al.* Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nat Microbiol* **2**, 17081, doi:10.1038/nmicrobiol.2017.81 nmicrobiol201781 [pii] (2017).

2 Sorokin, D. Y. *et al.* Methanonatronarchaeum thermophilum gen. nov., sp. nov. and 'Candidatus Methanohalarchaeum thermophilum', extremely halo(natrono)philic methyl-reducing methanogens from hypersaline lakes comprising a new euryarchaeal class Methanonatronarchaeia classis nov. *Int J Syst Evol Microbiol* **68**, 2199-2208, doi:10.1099/ijsem.0.002810 (2018).

3 Aouad, M., Borrel, G., Brochier-Armanet, C. & Gribaldo, S. Methanonatronarchaeia are not evolutionary intermediates on the path from methanogens to extreme halophiles. *Nature Microbiology* (2018).

4 Felsenstein, J. *Inferring Phylogenies*. (Sinauer Associates, 2004).

5 Wagele, J. W. & Rodding, F. A priori estimation of phylogenetic information conserved in aligned sequences. *Mol Phylogenet Evol* **9**, 358-365, doi:S1055-7903(98)90501-2 [pii] 10.1006/mpev.1998.0501 (1998).

6 Vollmer, A. C. & Bark, S. J. Twenty-Five Years of Investigating the Universal Stress Protein: Function, Structure, and Applications. *Adv Appl Microbiol* **102**, 1-36, doi:S0065-2164(17)30042-4 [pii] 10.1016/bs.aambs.2017.10.001 (2018).

7 Prangishvili, D. *et al.* The enigmatic archaeal virosphere. *Nat Rev Microbiol* **15**, 724-739, doi:10.1038/nrmicro.2017.125 nrmicro.2017.125 [pii] (2017).

8 Aouad, M. *et al.* Extreme halophilic archaea derive from two distinct methanogen Class II lineages. *Mol Phylogenet Evol* **127**, 46-54, doi:S1055-7903(17)30697-8 [pii] 10.1016/j.ympev.2018.04.011 (2018).

9 Williams, T. A. *et al.* Integrative modeling of gene and genome evolution roots the archaeal tree of life. *Proc Natl Acad Sci U S A* **114**, E4602-E4611, doi:10.1073/pnas.1618463114 1618463114 [pii] (2017).

10 Waters, E. *et al.* The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci U S A* **100**, 12984-12988, doi:10.1073/pnas.1735403100 1735403100 [pii] (2003).

11 Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F. & Forterre, P. Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol* **6**, R42, doi:gb-2005-6-5-r42 [pii] 10.1186/gb-2005-6-5-r42 (2005).

12 Petitjean, C., Deschamps, P., Lopez-Garcia, P. & Moreira, D. Rooting the domain archaea by phylogenomic analysis supports the foundation of the new kingdom Proteoarchaeota. *Genome Biol Evol* **7**, 191-204, doi:10.1093/gbe/evu274 evu274 [pii] (2014).

13 Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol*, doi:10.1038/nbt.4229 nbt.4229 [pii] (2018).

14 Townsend, J. P. & Lopez-Giraldez, F. Optimal selection of gene and ingroup taxon sampling for resolving phylogenetic relationships. *Syst Biol* **59**, 446-457, doi:10.1093/sysbio/syq025 syq025 [pii] (2010).

15 Pick, K. S. *et al.* Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Mol Biol Evol* **27**, 1983-1987, doi:10.1093/molbev/msq089 msq089 [pii] (2010).

16 Ngugi, D. K. & Stingl, U. High-Quality Draft Single-Cell Genome Sequence Belonging to the Archaeal Candidate Division SA1, Isolated from Nereus Deep in the Red Sea. *Genome Announc* **6**, doi:e00383-18 [pii] 10.1128/genomeA.00383-18

6/19/e00383-18 [pii] (2018).
17 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490, doi:10.1371/journal.pone.0009490 (2010).

Supplementary File 1. Phylogenetic tree of the UspA family.

The tree (Newick format) is constructed from an alignment of 4,550 UspA domain sequences from 4,184 distinct loci in 427 archaeal genomes using the FastTree program (WAG evolutionary model, gamma-distributed site rates) [17]. Sites with more than 50% of gap characters and homogeneity less than 0.1 were removed; both the raw (https://ftp.ncbi.nlm.nih.gov/pub/wolf/_suppl/archtre/UspA.raw.afa) and the filtered (https://ftp.ncbi.nlm.nih.gov/pub/wolf/_suppl/archtre/UspA.tre.afa) alignments are available.

The sequences of the following genes of Methanonatronarchaeia are included in the tree: BTN85_0146, BTN85_0312, BTN85_1038, BTN85_1108 (two UspA domains), BTN85_1119, BTN85_1447 (two domains), BTN85_1704, BTN85_1755. BTN85_1868, BTN85_1870 from Candidatus Methanohalarchaeum thermophilum and AMET1_RS00685, AMET1_RS00685, AMET1_RS01465 (2 domains), AMET1_RS02155 (2 domains), AMET1_RS02155, AMET1_RS03320, AMET1_RS03675, AMET1_RS03980 (2 domains), AMET1_RS05120, AMET1_RS06595 from Methanonatronarchaeum thermophilum AMET1.

1 Sorokin, D. Y. *et al.* Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nat Microbiol* **2**, 17081, doi:10.1038/nmicrobiol.2017.81 nmicrobiol201781 [pii] (2017).