[Frontiers in Bioscience 13, 6886-6892, May 1, 2008]

The basal phylogenetic position of Nanoarchaeum equitans (Nanoarchaeota)

Sergio Branciamore[1], Enzo Gallori[1], Massimo Di Giulio[2]

[1]Department of Animal Biology and Genetics, University of Florence, Via Romana 17/19, 50125 Florence, Italy, [2]Laboratory of Molecular Evolution, Institute of Genetics and Biophysics 'Adriano Buzzati Traverso', CNR, Via P. Castellino, 111, 80131 Naples, Napoli, Italy

TABLE OF CONTENTS

1. ABSTRACT

Using sequences of ribosomal RNA from organisms belonging exclusively to the Archaea domain and by means of two methods to remove the phylogenetic noise, we investigate the phylogenetic position of *Nanoarchaeum equitans* The results obtained are compatible with the hypothesis that *N. equitans* represents a new phylum within the Archaea domain because the characteristic long branch of *N. equitans* in phylogenetic trees is conserved even after most of the phylogenetic noise has been removed, thus implying that its rRNA might indeed be singular. However, our analysis is unable to be equally as clear on the phylogenetic position of *Methanopyrus kandleri*.

2. INTRODUCTION

The identification of the first lines of divergence of the three domains of life, Archaea, Bacteria and Eukarya, might help to understand the mysterious nature of the Last Universal Common Ancestor. Therefore, the description of Nanoarchaeum equitans as a representative of a new phylum of Archaea might be particularly important and is based primarily on the unusual characteristics of this organism's ribosomal RNA (1). With the genome sequencing of N. equitans (2), this suggestion is strengthened because it seems to possess some truly singular characteristics, such as its high number of split genes and the absence of operons (2). In particular the presence of six tRNA split genes, whose 5' and 3' halves

are codified on two completely separate genes that are not contiguous in the genome of N. equitans (3-4) seems to be particularly interesting as these half genes are expected to be one of the evolutionary stages through which the evolution of the tRNA molecule might have passed (5). In other words, the tRNA molecule might have originated by direct duplication of a hairpin structure of RNA and the tRNA genes whose 5' and 3' halves are codified on two completely separate genes might therefore be molecular fossils bearing witness to this evolution (5-6). Therefore, the split genes of tRNA and, more generally, all the split genes of N. equitans should be considered plesiomorphic traits (2,5-6). Furthermore, as the split genes of tRNA in N. equitans have only been described in this organism, they seem to suggest that it was an ancient lineage, perhaps one of the earliest branches separating from the last universal common ancestor (7). All the above appears to justify the suggestion that N. equitans might effectively constitute a new phylum of Archaea (1).

Some phylogenetic analyses have pointed out that *N. equitans* is the first line of divergence in the Archaea domain (2, 8-11). In particular, Boussau and Gouy (8) make use of rRNA sequences and use a complex model of molecular evolution, noting that *N. equitans* behaves as if it were the first line of divergence in the Archaea domain. Whereas, Brochier *et al.* (12) make use of protein concatenation in their phylogenetic analysis and come to the conclusion that *N. equitans* might represent a fast-evolving euryarchaeal lineage (possibly related to Thermococcales) and is not the representative of a novel and early diverging archaeal phylum. This has been more recently suggested by Gribaldo and Brochier-Armanet (13) who build phylogenetic trees (by means of protein concatenation) which show *N. equitans* as the first line of divergence in the Archaea domain but they believe that this is attributable only to the high speed of its evolution and not to the possibility that *N. equitans* might actually represent one of the earliest lineages in the Archaea domain.

Here we make use of sequences of 16S ribosomal RNA and conduct an extensive phylogenetic analysis aiming to clarify the phylogenetic position of *N. equitans*. Moreover, given that the phylogenetic position of *Methanopyrus kandleri* is under discussion and is not at all clear (14), we have also followed this organism in our analysis.

3. MATERIALS AND METHODS

We used a total of 81 sequences of 16S ribosomal RNA only from Archaea (see Figure 1 and Appendix). All the sequences were taken from the site KEGG (www.genome.jp/kegg/kegg2.html). These were aligned using ClustalX (15) obtaining what Tables 1 and 2 call alignment A1. The construction of the other alignments is reported below and these are available on request.

3.1. Construction of the alignments

A first alignment (A1) was obtained by simply using ClustalX with the default options and no further intervention. The regions which, at first sight, did not seem to be well aligned were then removed from this first alignment, thus obtaining alignment A2. This alignment was then objectively 'cleansed' to remove phylogenetic noise using the two methods described below.

3.1.1. Elimination of the sites with more phylogenetic noise by means of PAUP

The 'cleansing' with the parsimony criterion was carried out using the phylogenetic tree topology reported in the appendix as a topological constraint, in which *N. equitans* is located between the Crenarchaeotes and the Euryarchaeotes while *Methanopyrus kandleri* is among the Crenarchaeotes. PAUP's 'describe tree' and 'character diagnostic' options (16) were used to obtain the maximum number of nucleotide substitutions for all the positions in the alignment. We then built 10 alignments progressively including the sites with a larger number of substitutions. The 10 alignments thus obtained include only the sites with at most one (P1), two (P2) … ten (P10) substitutions. Tab. 1 reports the characteristics of the obtained alignments.

3.1.2. Elimination of the sites with more phylogenetic noise by means of MrBayes

The MrBayes program makes it possible to estimate the speed of evolution of the single positions in an alignment, using the 'report siterates' option (17). The speed of evolution for every position in the alignment was calculated on the alignment A2. Twelve new alignments were then built, MB1 … MB12, in which the sites with greater evolutionary speed were progressively included. The alignment MB1 therefore only includes the invariant and more slowly evolving sites; the alignment MB2 includes the sites of MB1 plus some slightly faster evolving sites, and so on up to MB12 which includes 90% of the variable sites contained in the starting alignment A2. The characteristics of the alignments are reported in Tab. 2.

3.2. Modeltest

In order to assess the most appropriate evolutionary model for the various alignments, we used the Modeltest 3.7 software (18). For each alignment, all 56 different evolutionary models available were assessed. In order to compare the results of the different models, we used the Akaike information criterion (AIC). The model chosen was always the GTR.

3.3. PHYML

Part of the maximum likelihood analysis made use of the PHYML 2.4.4 software (19-20). In this case we conducted a non-parametric bootstrap analysis with at least 1000 replicates for each alignment. The evolutionary model chosen for all alignments, in compliance with the Modeltest analysis, was GTR and, when necessary, we considered the presence of invariant sites (+I option), a gamma distribution (+G), or a combination of the two. All the parameters, including nucleotide frequency, were used during the simulation.

3.4. MrBayes

Bayesian analysis was conducted using GTR as the substitution model (17). The rate of site evolution was

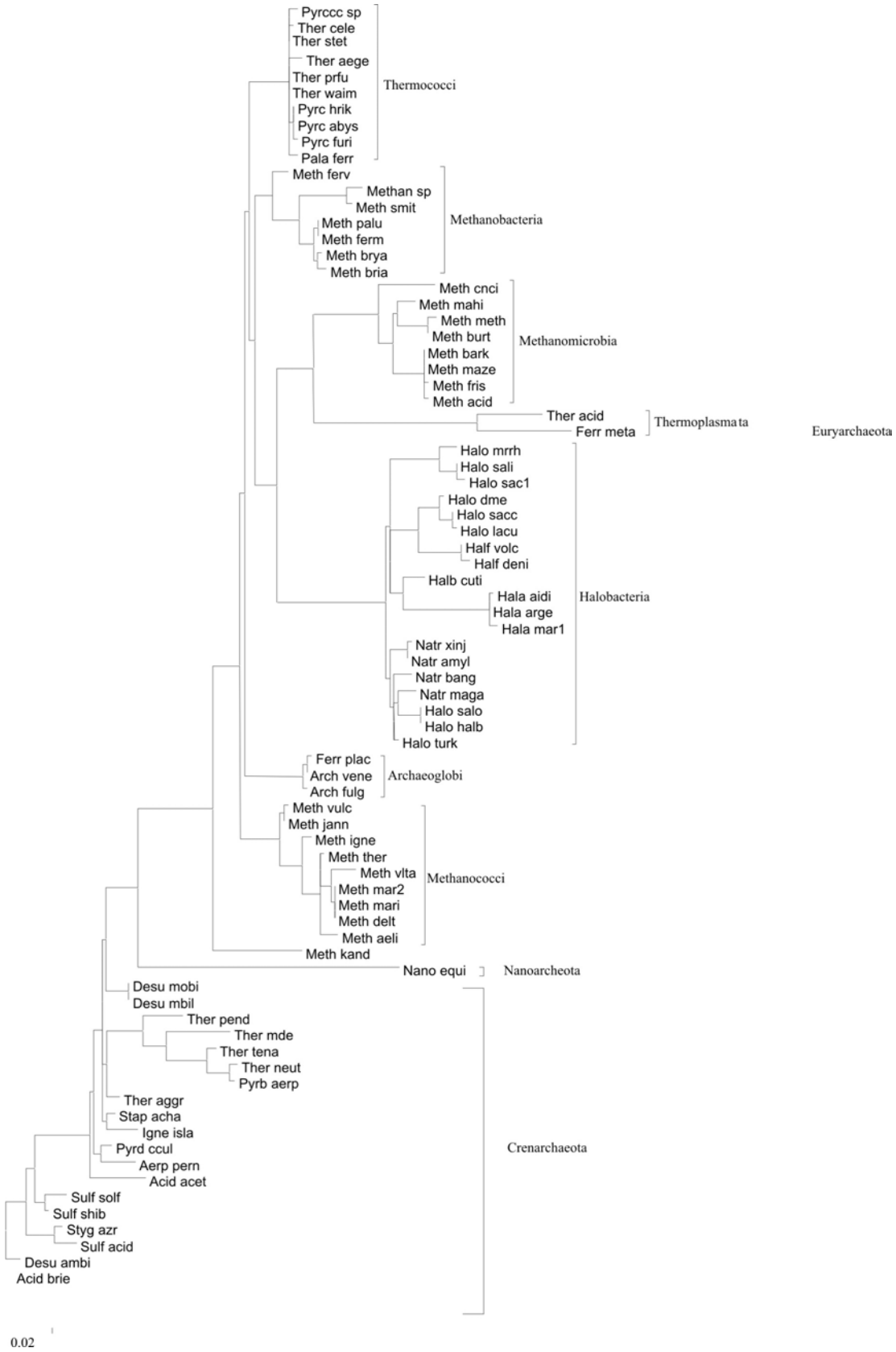

Figure 1. Phylogenetic tree derived from the analysis using PHYML on the alignment MB3 (Table 2).

Table 1. Alignments derived from the cleansing with PAUP

Alignment	Variable Sites	Total sites
P1	153	584
P2	269	700
P3	355	786
P4	413	844
P5	466	897
P6	504	935
P7	536	967
P8	561	992
P9	574	1005
P10	592	1023
A2	621	1052
A1	1115	1593

This reports the number of variable and total sites in the various alignments used in the phylogenetic analysis derived from the cleansing with PAUP. See Materials and Methods for further information.

Table 2. Alignment derived from the cleansing with MrBayes

Alignment	Variable Sites	Total sites
MB1	245	676
MB2	321	752
MB3	348	779
MB4	390	821
MB5	408	839
MB6	426	857
MB7	440	871
MB8	462	893
MB9	467	898
MB10	475	906
MB11	515	946
MB12	565	996
A2	621	1052
A1	1115	1593

This reports the number of variable and total sites in the various alignments used in the phylogenetic analysis derived from the cleansing with MrBayes. See Materials and Methods for further information.

considered by setting the corresponding options to +I (propinv) and +I+G (invgamma) when necessary and in compliance with modeltest. The analysis was conducted on one million replicates, recording one tree every hundred. For the final analysis, only the last 7500 trees acquired were used.

3.5. Distance criterion

3.5.1. Phylip: F84

The model used was F84 with one thousand bootstrap replicates. The programs Seqboot, DNAdist, Neighbor and Consense were used in sequence (21). When necessary and in compliance with Modeltest, we considered the presence of invariant sites (+I option) or a gamma distribution (+G) or a combination of the two was used. The values of G and I were calculated using Modeltest.

3.5.2. Mega: logdet

The evolutionary model used was logdet (22) and the method for tree resolution was neighbour-joining (23). When necessary and in compliance with Modeltest, we considered the presence of invariant sites (+I option) or a gamma distribution (+G) or a combination of the two was used. The values of G and I were calculated using Modeltest. The bootstrap analyses was carried out on at least one thousand replicates.

3.6. Parsimony criterion (PAUP)

For parsimony analysis, PAUP was used (16). TBR was used as the branch swapping algorithm and random sequence addition with ten replicates. The total number of replicates was at least one hundred.

4. RESULTS AND DISCUSSION

4.1. The phylogenetic position of *N. equitans*

Tables 3 and 4 report the results of the phylogenetic analysis conducted using five different methods and a ribosomal RNA alignment from which phylogenetic noise was progressively removed using two different methods (see Materials and Methods). In other words, given that it is thought (24-25) that the more slowly evolving sites are less disposed to confuse the phylogenetic signal than sites with multiple substitutions, we removed these sites from the alignment, thus selecting the positions with few substitutions and that are held to be less influenced than the artefacts deriving from the phylogenetic analysis (25). As can be seen, both methods used to remove the phylogenetic noise display an equivalent pattern (Tables 3 and 4). That is to say, *N. equitans* is a deep branch separate from both the Crenarchaeotes and the Euryarchaeotes (this result is substantially equivalent to the one obtained by the concatenation of many proteins (12-13,26). Only for the alignments P1 and P2 and MB1 and MB2 is it observed that N. equitans clusters mostly among the crenoarchaeotes (Tables 3 and 4). We interpret this as an artefact due to the small number of variable sites available in these alignments, even if it has been observed that *N. equitans* clusters with Crenarchaeotes as the first line of divergence in an analysis also using an rRNA (8). Whereas, in other analyses *N. equitans* groups among Euryarchaeotes and, in particular with Thermococcales (12-13,26). However, our analysis gives a different result from that of Boussau and Gouy (8) in that *N. equitans* is not the first line of divergence but groups among Crenoarchaeotes but with non-significant bootstrap percentages (Tables 3 and 4). Furthermore, the alignment P1 has been omitted from Tab. 3 because the relative phylogenetic trees are unresolved and present widespread polytomies (data not shown).

We interpret the results of Tables 3 and 4 as follows. We believe that in alignments with more phylogenetic noise, such as P10 and P9 or MB12 and MB11 (Tabs, 3 and 4), *N. equitans* apparently behaves as one of the first lines of divergence in the Archaea domain and this is due to the fact that in this case it is a fast evolving species which makes it behave like a long-branch attraction (12-13,26). While for alignments such as P3 and P4 or MB3 and MB4, which always see *N. equitans* as one of the first lines diverging in the Archaea domain, it should be concluded that this is the most likely hypothesis because it continues to be the first line of divergence in this domain when most of the phylogenetic noise is removed. In support of this is the observation that as 'cleansing' of alignments increases, for instance in the analysis conducted using

Table 3. Phylogenetic analysis for *N. equitans* on the alignments derived from the cleansing with PAUP

Alignment	Phylip (F84)		Mega (logdet)		Paup		Phyml		MrBayes	
	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota
P2	Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 52		Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 41		Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 65		Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 49		Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 87	
P3	99	94	95	79	90	96	83	97	100	100
P4	99	95	98	85	91	100	87	99	100	100
P5	96	89	99	88	92	98	98	90	100	100
P6	97	99	97	97	100	100	98	98	100	100
P7	97	100	97	98	100	94	99	94	100	100
P8	99	100	96	98	99	96	98	97	100	100
P9	99	100	97	98	100	95	98	97	100	100
P10	97	99	97	99	97	95	97	97	100	100
A2	96	100	97	98	96	97	97	98	100	100
A1	75	100	94	99	79	99	88	100	100	100

This reports the results of the phylogenetic analysis for *N. equitans* on the alignments obtained after removing the positions with more phylogenetic noise by means of PAUP (see Materials and Methods). The numbers represent the bootstrap percentages with which *N. equitans* is separate from both Euryarchaeotes and Crenarchaeotes, thus taking up a basal position in the Archaea domain.

Table 4. Phylogenetic analysis for *N. equitans* on the alignments derived from the cleansing with MrBayes

Alignment	Phylip (F84)		Mega (logdet)		Paup		Phyml (GTR)		MrBayes (GTR)	
	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota	Euryarchaeota	Crenarchaeota
MB1	98	64	97	41	97	51	Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 42		Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 92	
MB2	99	61	Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 35		Groups inside crenarchaeota with a polyphyletic topology		Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 45		Groups inside crenarchaeota,the first divergent line of crenoacheota are thermoproteales, 90	
MB3	100	83	99	66	100	78	64	100	96	100
MB4	100	96	99	84	99	93	90	100	100	100
MB5	99	73	99	81	100	94	86	100	100	100
MB6	80	65	99	89	100	97	88	100	100	100
MB7	99	85	99	90	98	97	87	100	100	100
MB8	99	89	98	91	100	97	96	90	100	100
MB9	99	91	97	94	98	98	98	93	100	100
MB10	98	99	98	95	98	98	97	94	100	100
MB11	98	100	98	96	98	100	98	98	100	100
MB12	98	100	98	98	99	99	99	98	100	100
A2	96	100	97	98	96	97	97	98	100	100
A1	75	100	94	99	79	99	88	100	100	100

This reports the results of the phylogenetic analysis for N. equitans on the alignments obtained after removing the positions with more phylogenetic noise by means of MrBayes (see Materials and Methods). The numbers represent the bootstrap percentages with which *N. equitans* is separate from both Euryarchaeotes and Crenarchaeotes, thus taking up a basal position in the Archaea domain.

PHYML, we observe a general reduction in the length of branches for all species. For instance, passing from alignment A1 to P3 or MB3, the branch length of *N. equitans* falls from 0.35 units to 0.085 units (for P3) or 0.077 units (for MB3), while for *M. kandleri* it falls from 0.12 units to 0,024 units (for P3) or 0.023 (for MB3). For the branch of *P. furiosus* taken as the overall length at the first common ancestor which is not a Thermococcales, it falls from 0.12 to 0.015 units. Finally, for *Ferroplasma acidiphilum* which presents the longest branch in the alignment A1, the length falls from 0.5 units to 0.086 units (for P3) and 0.076 (for MB3) (see Figure 1). One particularly interesting observation is that this variation is not uniform for all organisms. If we consider, for instance, the ratios between the branches of *N. equitans* and *P. furiosus*, we observe that initially the branch of *N. equitans* is three times longer, while after the removal of phylogenetic noise, it is five times longer, with a proportional increase in length of approximately 65% compared to Thermococcales. In the same way, the branch of *Ferroplasma acidiphilum* , which is initially 40% longer than that of *N. equitans*, tends to have the same length after the phylogenetic noise is removed. According to the hypothesis supported by Brochier *et al.* (12-13,26) the basal position of *N. equitans* in the Archaea domain is a typical example of long branch attraction (LBA) due to the fact that this species has a high evolutionary speed. The observation that the reduction of the branch length is not the same between the various species under examination, and in particular that the branch of *N. equitans* tends to increase proportionally as the cleansing of the alignment increases, would seem to contradict this hypothesis, suggesting a real, early divergence for this organism. In other words, alignments with less phylogenetic noise conserve only sites with few nucleotide substitutions and the fact that *N. equitans* still presents an even longer branch would seem to imply that there is a singular diversity of the rRNA of this organism. Therefore, the greater length of the branch of *N. equitans* should, in the cleaner alignments, be considered not as an artefact due to the LBA but as a real manifestation of an ancient event of evolutionary divergence.

4.2. The phylogenetic position of *Methanopyrus kandlerii*

A phylogenetic situation which is very similar to that of N. equitans has been described for M. kandleri. For instance, Brochier *et al.* (14) maintain that the basal

Table 5. Phylogenetic analysis for *Methanopyrus kandleri* on the alignments derived from the cleansing with PAUP

Alignment	Phylip (F84)	Mega (logdet)	Paup	Phyml (GTR)	MrBayes (GTR)
P2	70	83	Grups inside euryarchaeota	74	100
P3	48	42	Grups inside euryarchaeota	49	82
P4	59	43	Grups inside euryarchaeota	Groups near methanobacteria	Groups together with methanobactria, and the group of methanococci and archaeoglobi.
P5	63	Groups near methanococci	Grups inside euryarchaeota	Groups near methanococci	Groups near methanococci
P6	54	39	Grups inside euryarchaeota	47	52
P7	48	46	Grups inside euryarchaeota	55	95
P8	49	44	Grups inside euryarchaeota	59	93
P9	57	54	Grups inside euryarchaeota	63	96
P10	45	Groups near methanobacteria	Grups inside euryarchaeota	54	83
A2	54	Groups near methanobacteria	Grups inside euryarchaeota	47	97
A1	84	51	92	38	100

This reports the results of the phylogenetic analysis for *Methanopyrus kandleri* on the alignments obtained after removing the positions with more phylogenetic noise by means of PAUP (see Materials and Methods). The numbers represent the bootstrap percentages with which *M. kandleri* is separate from Euryarchaeotes, thus taking up a basal position in the Archaea domain.

Table 6. Phylogenetic analysis for *Methanopyrus kandleri* on the alignments derived from the cleansing with MrBayes

Alignment	Phylip (F84)	Mega (logdet)	Paup	Phyml (GTR)	MrBayes (GTR)
MB1	55	53	75	98	100
MB2	46	35	57	59	100
MB3	57	39	68		100
MB4	48	36	65	48	Groups together with methanobactria, and the group of methanococci and archaeoglobi.
MB5	44	24	Groups inside Euryarchaeota	43	Groups near archaeoglobi
MB6	47	Groups near methanococci	52	48	Groups near archaeoglobi
MB7	43	39	Groups inside Euryarchaeota	57	79
MB8	43	36	52	54	84
MB9	41	39	52	58	88
MB10	58	43	56	61	99
MB11	56	Groups near methanococci	Groups inside Euryarchaeota	51	74
MB12	52	38	Groups inside Euryarchaeota	52	93
A2	54	Groups near methanobacteria	Groups inside Euryarchaeota	47	97
A1	84	51	92	38	100

This reports the results of the phylogenetic analysis for *Methanopyrus kandleri* on the alignments obtained after removing the positions with more phylogenetic noise by means of MrBayes (see Materials and Methods). The numbers represent the bootstrap percentages with which *M. kandleri* is separate from Euryarchaeotes, thus taking up a basal position in the Archaea domain.

position observed for M. kandleri in the Archaea domain reflects the high evolutionary rate of this organism, which in their trees has a very long branch. On the other hand, in an original analysis Bucknam *et al* (11) report that M. kandleri might be close to N. equitans, implying that M. kandleri might be a very deep branch of the Archaea domain. Unfortunately, in our analysis the behaviour of M. kandleri is extremely difficult to interpret as the relative bootstrap percentages are mostly low and non-significant (Tables 5 and 6) but our analysis is such as not to exclude Buckman *et al.*'s conclusion.

5. APPENDIX

Topological constraint used to remove the phylogenetic noise using PAUP (see Materials and Methods):

(Meta_sedu,Meta_prun,(Acid_brie,(Desu_ambi,((Sulf_acid,Styg_azr),(Sulf_shib,Sulf_solf),((((Acid_acet,Aerp_pern),Pyrd_ccul),((((Desu_mbil,Desu_mobi),Ther_aggr),Stap_acha),Igne_isla)),(((((Pyrb_aerp,Ther_neut),Ther_tena),Ther_mde),Ther_pend),(Nano_equi,((((Meth_kand,(((((Ferr_meta,Ther_acid),((((((Hala_arge,Hala_mar1),Hala_aidi),((Half_deni,Half_volc),((Halo_lacu,Halo_sacc),Halo_dme))),(((Halo_halb,Halo_salo),(Halo_turk,(Natr_bang,Natr_maga))),(Natr_amyl,Natr_xinj))),Halb_cuti),((Halo_sac1,Halo_sali),Halo_mrrh))),((((Meth_fris,Meth_maze),(Meth_acid,Meth_bark)),((Meth_burt,Meth_meth),Meth_mahi)),Meth_cnci)),(((Meth_bria,Meth_brya),(Meth_ferm,Meth_palu)),(Meth_smit,Methan_sp))),Meth_ferv)),((((((Meth_delt,Meth_mari,Meth_mar2),Meth_vlta),Meth_aeli),Meth_ther),Meth_igne),(Meth_jann,Meth_vulc))),((Pyrc_abys,Pyrc_hrik,Pyrc_furi),(((Ther_prfu,Ther_stet,Ther_cele,Pyrccc_sp),(Ther_waim,Pala_ferr)),Ther_aege))),((Arch_fulg,Arch_vene),Ferr_plac)))))))))

6. REFERENCES

1. Harald Huber, Michael J. Hohn, Reinhard Rachel, Tanja Fuchs, Verena C. Wimmer, Karl O. Stetter, A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 417, 63-67 (2002)

2. Elizabeth Waters, Michael J. Hohn, Ivan Ahel, David E. Graham, Mark D. Adams, Mary Barnstead, Karen Y. Beeson, Lisa Bibbs, Randall Bolanos, Martin Keller, Keith Kretz, Xiaoying Lin, Eric Mathur, Jingwei Ni, Mircea Podar, Toby Richardson, Granger G. Sutton, Melvin

Simon, Dieter Söll , Karl O. Stetter, Jay M. Short, Michiel Noordewier. The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc. Natl. Acad. Sci. U.S.A*. 100, 12984-12988 (2003)
3. Lennart Randau, Michael Pearson, Dieter Soll**,** D., *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their 5'- and 3'-halves. *Nature* 433, 537-541 (2005)
4. Lennart Randau, Michael Pearson,Dieter Soll,The complete set of tRNA species in *Nanoarchaeum equitans*. *FEBS* 579, 2945-2947 (2005)
5. Massimo Di Giulio, The non-monophyletic origin of the tRNA molecule and the origin of genes only after the evolutionary stage of the Last Universal Common Ancestor (LUCA). *J. Theor. Biol.* 240, 343-352 (2006)
6. Massimo Di Giulio, *Nanoarchaeum equitans* is a living fossil. *J. Theor. Biol*. 242, 257-260. (2006)
7. Massimo Di Giulio, The tree of life might be rooted in the branch leading to Nanoarchaeota. Gene, 401, 108-113 (2007)
8. Bastien Boussau, Manolo Gouy: Efficient likelihood computations with nonreversible models of evolution. Syst. Biol. 55, 756-768 (2006)
9. Christa Schleper, German Jurgens, Melanie Jonuscheit, Genomic studies of uncultivated Archaea. *Nature Rev. Microb.* 3, 479-488 (2005)
10. Minglei Wang, Gustavo Caetano-Anollés, Global phylogeny determined by the combination of protein domains in proteomes. Mol. Biol. Evol. 23, 2444-2454 (2006)
11. James Bucknam, Yan Boucher, Eric Bapteste:. Refuting phylogenetic relationships. *Biol. Direct* 1, 26, doi:10.1186/1745-6150-1-26 (2006)
12. Celine Brochier, Simonetta Gribaldo, Yvan Zivanovic, Fabrice Gonfalonieri, Patrick Forterre: Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol.* 6, R42 doi:10.1186/gb-2005-6-5-r42 (2005)
13. Simonetta Gribaldo, Celine Brochier-Armanet, The origin and evolution of Archaea: a state of art. *Phil. Trans. R. Soc. B* 361, 1007-1022 (2006)
14. Céline Brochier, Patrick Forterre, Simonetta, Gribaldo: Archaeal phylogeny based on proteins of the transcription translation machineries: tackling the Methanopyrus kandleri paradox. *Genome Biol.* 5, R17 (2004)
15. Julie D. Thompson, Toby J. Gibson, Frédéric Plewniak, François Jeanmougin, Desmond G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876-4882 (1997)
16. David L. Swofford, 1998. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). version 4.0b10 (PPC), Sinauer Associates, Sunderland MA.
17. Fredrik Ronquist, John P. Huelsenbeck**,** MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 9, 1572-4 (2003)
18. David Posada, Keith A. Crandall, Modeltest: testing the model of DNA substitution. *Bioinformatics* 14, 817-818 (1998)
19. Stéphane Guindon, Olivier Gascuel, A simple, fast, and accurate algorithm to estimate large phylogenies by maximun likelihood. *Syst. Biol.* 52, 696-704 (2003)
20. Stéphane Guindon, Franck Lethiec, Patrice Duroux, Olivier Gascuel, PHYML Online- a web server for fast maximum likelihood- based phylogenetic inference. Nucleic Acids Res. 33, W557-W559 (2005)
21. Joseph Felsenstein, program version 3.63: http://evolution.genetics.washington.edu/phylip/software.html
22. Koichiro Tamura, Sudhir Kumar, Evolutionary distance estimation under heterogeneous sunstitution pattern among lineages. *Mol. Biol. Evol.* 19, 1727-1736 (2002)
23. Sudhir Kumar, Koichiro Tamura, Masatoshi Nei, MEGA 3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.*, 150-163 (2004)
24. Joseph Felsenstein, Taking variation of evolutionary rates between sites into account in inferring phylogenies. *J. Mol. Evol*. 53, 447-555 (1996)
25. Celine Brochier, Hervé Philippe: A non-hyperthermophilic ancestor for Bacteria. *Nature* 41**7**, 244 (2002)
26. Céline Brochier, Patrick Forterre, Simonetta, Gribaldo: An emergenting phylogenetic core of Archaea: phylogenies of transcription and translation machineries converge following addition of new genome sequences. *BMC Evolut. Biol*, 5, 36 doi:10.1186/1471-2148-5-36 (2005)

Key Words: Deep Divergence, Removal Of Phylogenetic Noise, Phylogenetic Analyses, *Methanopyrus kandleri*; *Nanoarchaeum equitans*; LUCA

Send correspondence to: Massimo Di Giulio, Laboratory of Molecular Evolution, Institute of Genetics and Biophysics 'Adriano Buzzati Traverso', CNR, Via P. Castellino, 111, 80131 Naples, Napoli, Italy, Tel: 39-081-6132369, Fax: 39-081-6132706, E-mail: digiulio@igb.cnr.it

http://www.bioscience.org/current/vol13.htm