The Phylogeny of Methanopyrus kandleri

M aria C. Rivera* and James A. Lake

Molecular Biology Institute and Department of Biology, University of California, Los Angeles, Los Angeles, California 90024

The phylogenetic position of Methanopyrus kandleri has been difficult to determine because reconstructions of phylogenetic trees based on rRNA sequences have been ambiguous. The most probable trees determined by most algorithms place the genus Methanopyrus at the base of a group that includes the halobacteria and the methanogens and their relatives, although occasionally some algorithms place this genus near the eocytes (the hyperthermophilic, sulfur-metabolizing prokaryotes), suggesting that it may belong to this lineage. In order to resolve the phylogeny of the genus Methanopyrus, we determined the sequence of an informative region of elongation factor 1-alpha that contains an 11-amino-acid insertion in eocytes and eukaryotes which is replaced by a 4-amino-acid insertion in methanogens, halobacteria, and eubacteria. On the basis of the results of our elongation factor 1-alpha gene analysis, we concluded that the genus Methanopyrus diverged from the eocyte branch before the eukaryotic and eocyte lineages separated and therefore is not an eocyte.

M ethanopyrus kandleri is a rod-shaped, gram-positive, methanogenic bacterium that was isolated from a deep thermal vent in the Gulf of California (16). This organism is the first known hyperthermophilic methanogen that grows at temperatures up to 110°C (16). Although extreme thermophily has been found in three phylogenetically different groups of prokaryotes (the eubacteria, the methanogens and their relatives, and the eocytes [hyperthermophilic, principally sulfur-metabolizing prokaryotes]), members of the genus Pyrodictium are the only other prokaryotes that are known to grow at a temperature of 110°C or above (37, 38). On the basis of the results of a 16S rRNA and protein synthesis elongation factor 1-alpha (EF-1α) sequence analysis, the genus Pyrodictium was classified as an eocyte taxon (24, 32, 33, 37); therefore, we wanted to determine whether the genus Methanopyrus is also an eocyte taxon.

M ethanopyrus kandleri appears to represent a unique lineage within the phylogenetically diverse methanogens (18, 30), and, consistent with this, the results of a Jukes-Cantor distance and parsimony analysis of the 16S rRNA sequence of this organism have shown that it is only distantly related to the other methanogens (8). A suggestion that M. kandleri may be related to the eukaryotes resulted from the discovery that a eukaryote-like topoisomerase I molecule is present in Methanopyrus cells (35).

The most probable tree determined in our analysis of small-subunit rRNA sequences in which the Jukes-Cantor (19), Kimura (21) and paralinear distance (27) algorithms were used is shown in Fig. 1; the probabilities of this eocyte tree are 72.5, 60.5, and 65.5%, respectively, and the probabilities of the archaeabacterial tree (data not shown) are 20.5, 29.5, and 20.0%, respectively. The most probable location of the genus Methanopyrus on the trees obtained with these three algorithms is at the branch labelled A; branch B is less likely, and branch C is least likely. The trees that have been constructed, the similarity of the maximal growth temperatures of Methanopyrus and Pyrodictium strains, and the presence of a eukaryotic topoisomerase I-like molecule all raised the possibility that the genus Methanopyrus may be more closely related to eukaryotes than previously thought. Given that sequence analysis studies can be biased by unequal rate effects, by site-to-site variation, and by the order of alignment, we sought additional, more reliable evidence for the phylogenetic position of the genus Methanopyrus.

In a previous study, the phyletic distribution of the two variants of EF-Tu and EF-1α was used to infer an immediate relationship between the eocytes and the eukaryotes (32). Previously, this relationship had been difficult to analyze by direct sequence comparisons because of the extremely long divergence times between the prokaryotes and the eukaryotes. In that study, we showed that the members of the genus Pyrodictium and the other eocytes, as well as all of the eukaryotes, contain an 11-amino-acid segment in EF-1α that is not present in other prokaryotes. In the methanogens, the halobacteria, and the eubacteria this segment is replaced by a 4-amino-acid segment. In light of the questionable phylogenetic position of the genus Methanopyrus, determining the sequence of the EF-1α gene of a member of this taxon could reveal if it is also an immediate eukaryotic relative (containing the 11-amino-acid segment and corresponding to tree branch C) or if it is more closely related to the methanogens and halobacteria (containing the 4-amino-acid segment and corresponding to either tree branch A or tree branch B).

As previously noted (22, 28, 31), EF-1α appears to be one of the best molecules available for studying deep divergences. This molecule is a slowly evolving protein found in all cells, where it binds, transports, and participates in the selection of the correct aminoacyl-tRNA. In addition, EF-1α interacts with cellular components encoded by genes dispersed throughout the genome, including aminoadyl- tRNAs, ribosomal proteins, 16S and 18S rRNAs, and elongation factor EF-Ts; therefore, it is unlikely that EF-1α is transferred laterally between organisms.

To determine whether the genus Methanopyrus is an eocyte genus, we isolated by PCR, cloned, sequenced, and analyzed a fragment in the GDP-binding domain of the gene coding for EF-1α from several taxa whose sequences had not been determined previously. Total genomic DNAs were isolated from frozen cell pastes of Methanopyrus kandleri, Archaeoglobus fulgidus (36), and Thermococcus marinus by using the sodium dodecyl sulfate-proteinase K lysis method (14). Total nucleic acids were extracted from Desulfoviridans salina (40) and purified by CsCl centrifugation (1). Degenerate oligonucleotide primers were designed on the basis of conserved amino acid motifs in all known EF-Tu and EF-1α sequences (Fig. 2) and used for PCR amplification (34). The primers spanned the

* Corresponding author.
Methanococcus vannielii; Archae., Archaeoglobus fulgidus.

The PCR fragments were cloned into the pT7 Blue vector and the insertion were determined by using a Sequenase 2.0 kit (U.S. Biochemicals) and M13 universal and reverse primers for reading both strands of the amplified fragments.

The nucleotide sequences of the methanogen Methanosarcinalesander, Archaeoglobus fulgidus, the halophilic cyanobacterium Dactylococcopsis salina, and the eocyte Thermococcus maritimus are shown in Fig. 3. The deduced amino acid sequence of the Methanosarca kandleri EF-1α fragment was aligned with the sequences of the homologous fragments of eocytes, methanogens and their relatives, a halobacterium, eubacteria, and eukaryotes (Fig. 2). The Methanosarcina kandleri sequence contained the 4-amino-acid segment (GVMP) characteristic of eubacteria, halobacteria, and methanogens, and it lacked the 11-amino-acid segment found in eocyte and eukaryotic sequences (Fig. 2).

The absence of the 11-amino-acid segment in the genus Methanosarcina is most parsimoniously consistent with branches A and B on Fig. 1. On either branch A or branch B, only a single change from the 4-amino-acid segment to the 11-amino-acid segment (Fig. 1, solid rectangle) is needed to explain the observed distribution of insertions. In comparison, one additional change from the 11-amino-acid segment to the 4-amino-acid segment on branch C would be required to explain the branch C alternative. Therefore, these results strongly indicate that the genus Methanosarcina is not an eocyte genus and is less closely related to the eukaryotes. In addition, it should be noted that all of the new sequences described in this paper are consistent with and provide additional support for the eocyte tree, on which the eocytes are the closest relatives of the eukaryotes.

![FIG. 1. Rooted tree showing the three possible positions of the Methanosarcina lineage. This tree is the most probable tree derived from an analysis of 16S and 18S rRNA sequences in which the Kimura (21), Jukes-Cantor (19), and panc-
linear distance (25, 27) algorithms were used. The tree is rooted in the branch leading to the eubacteria, as proposed by other workers (12, 17). The three possible locations of the genus Methanosarcina that have finite probabilities are designated branches A, B, and C. The eubacterium used was Thermotoga maritima. Abbreviations: H.vol., Halobacterium volcanii; Mc.vann., Methanospirillum hungatii; Mc.vann., Methanococcus vannielii; Archae., Archaeoglobus fulgidus.

The EF-la fragment was aligned with the sequences of methanogens, halobacteria, eubacteria, eocytes, and eukaryotes. The 4- and 11-amino-acid segments are underlined. The lowercase letters represent sequences used as the PCR primers. The methanogens and their relatives which we used were Methanosarcina kandleri (Mp.kand.) (this study), Thermococcus celer (T.celer) (3), Pyrococcus woesei (P.woes.) (11), Archaeoglobus fulgidus (A.fulg.) (this study), Methanococcus vannielii (Mc.vann.) (26), and Thermoplasma acidophilum (T.acido.) (39). The halobacterium used was Halobacterium marismortui (H.mar.) (6). The eubacterium used were the thermophilic eubacterium Thermotoga maritima (Th.mar.) (5), the halophilic cyanobacterium Dactylococcopsis salina (D.sal.) (this study), and the enterobacterium E. coli (9). The eocytes used were Thermococcus maritimus (Th.mar.) (this study), Pyrococcus occultum (P.occu.) (32), Desulfo bacterium mucosus (D.muco.) (32), Acidianus infernus (A.infe.) (32), and Sulfolobus acidocaldarius (Su.acid.) (4). The eukaryotes used were Giardia lamblia (Giardia) (13), Tetrahymena pyriformis (Tetrah.) (23), Saccharomyces cerevisiae (Yeast) (10), Lycopersicon esculentum (Tomo.) (29), Drosophila melanogaster (Dros.) (15), Rattus norvegicus (Rat) (2), and Homo sapiens (Human) (7).
Support for the hypothesis that 11-amino-acid segments occur in eukaryotes and eocyte and 4-amino-acid segments or their variants occur in all other prokaryotes was provided by the new sequences described in this paper. These sequences give rise to the eocytes and eukaryotes (branch B). We thank Karl Stetter for providing *Methanopyrus, Archaeoglobus*, and *Thermodesmus maritimus* cells and thank Anthony Walsby for providing *Dactylococcopsis* cells.

We thank the National Science Foundation to J.A.L.

**REFERENCES**

15. Hovemn, B., S. Richter, U. Wulfdorf, and C. Crippfel. 1988. Two genes encode related cytoplasmic elongation factors 1 α (EF-1α) in *Drosophila melanogaster* and *Pyrococcus woesei*, which are located before the node which gives rise to the eocytes and eukaryotes (branch B).

![FIG. 3. Nucleotide sequences of the KNMITG,5-QTREH, fragments from *Methanopyrus kandleri*, *Archaeoglobus fulgidus*, *Dactylococcopsis salina*, and *Thermodesmus maritimus*. Triplet corresponding to amino acids are separated by spaces. Gaps are represented by solid lines. Dashes represent nucleotides identical to the ones in the *M. kandleri* reference sequence (top row).](image-url)


