ELSEVIER SCIENTIFIC PUBLISHERS IRELAND

BioSystems 31 (1993) 111–119

Bio Systems

Horizontal transfer of ATPase genes — the tree of life becomes a net of life

Elena Hilario, Johann Peter Gogarten*

Department of Molecular and Cell Biology, University of Connecticut, 75 North Eagleville Rd., Storrs, CT 06269-3044, USA

Abstract

An ancient gene duplication gave rise to the catalytic and non-catalytic subunits of each of the three types of proton pumping ATPases: vacuolar, archaebacterial and eubacterial. Previously, this gene duplication had been used to root the universal tree of life. However, recent findings of archaebacterial type ATPases in eubacteria and of eubacterial type in an archaebacterium suggested that both types of ATPases may have been already present in the last common ancestor. Here we show that a phylogenetic analysis of these ATPase subunits indicates that this conclusion is premature. We suggest that horizontal gene transfer can explain the data. In addition, we show that the analysis of glutamate dehydrogenases data neither affirm nor contradict any particular placement of the last common ancestor in the universal tree of life. The prevalence and the mode of horizontal gene transfer is discussed.

Key words: Archaebacteria; Origin of life; ATPase; Evolution; Gene duplication; Glutamate dehydrogenase

1. Introduction

Schwartz and Dayhoff (1978) proposed to use ancient gene duplications to root the universal tree of life by means of an out-group. This concept was endorsed by many scientists studying the evolution of life (e.g., Woese, 1987; Space Studies Board, 1990). In 1989 two groups (Gogarten et al., 1989; Iwabe et al., 1989) independently employed this idea using either genes encoding ATPase subunits dehydrogenases, tRNAs, or elongation factors. Although using different molecular markers and algorithms the two studies came to the same conclusion: the root of the universal tree of life appeared to be located between the Bacteria on one side and the Archaea and Eucarya on the other. The location of this placement of the universal ancestor within the tree of life has been widely endorsed (e.g., Wheelis et al., 1992; Rivera and Lake, 1992). (This paper uses the terminology of Woese et al., (1990) for the three domains of life (formerly Ur-Kingdoms): Archaea (= Archaebacteria), Bacteria (= Eubacteria) and Eucarya (Eukaryota)).

The vacuolar (V-), bacterial (F-), and archaeal (A-) ATPases exhibit a number of structural and functional similarities. All are large multi-subunit enzymes composed of a water soluble and a

* Corresponding author.

SSDI 0303-2647(93)01410-U

proton-conducting membrane spanning complex. Each hydrophilic sector contains three copies of a catalytic (β-subunit in F-ATPases, *A*-subunit in A/V-ATPases) and three copies of a non-catalytic subunit (α-subunit in F-ATPases, *B*-subunit in A/V-ATPases). The catalytic subunits bind and hydrolyze ATP. The non-catalytic or regulatory subunits also bind ATP; however, they do not hydrolyze it. V-ATPases normally are found in eukaryotic endomembranes. A very similar enzyme is present in the Archaea. The archaeal ATPase had also been classified as a V-ATPase; however, because of structural and functional differences, the term A-ATPase for the archaeal enzyme seems appropriate (Gogarten et al., 1992; Ihara and Mukohata, 1992; Schäfer and Meyering-Vos, 1992). The A-, V-, and F-ATPases are homologous enzymes, i.e. they evolved from the same ancestral enzyme (Zimniak et al., 1988, Gogarten et al., 1989). In addition, the catalytic and the non-catalytic subunits were also shown to have evolved from the same ancestral gene after an ancient gene duplication (Gogarten et al., 1989). For recent reviews on the structure and function of V-, F-, and A-ATPases see Harvey and Nelson (1992).

Based on rRNAs and other biochemical characters (cell wall, lipids, antibiotic resistance) the genus *Thermus* is classified as a Bacterium, with closest similarities to the genus *Deinococcus* (Hensel et al., 1986; Woese, 1987). However, *Thermus thermophilus* (= *T. aquaticus* HB8) does not have an F-ATPase like other Bacteria, but an A/V-ATPase (Yokoyama et al., 1990; Tsutsumi et al., 1991). Both major subunits of the *Thermus* ATPase are clearly archaeal in character (Gogarten et al., 1992). Furthermore, a sodium pumping ATPase from another Bacterium, *Enterococcus hirae* was determined to be an A/V ATPase (Kakinuma et al., 1991), and a partial DNA sequence encoding an F-ATPase catalytic subunit was obtained from the Archaeum *Methanosarcina barkeri* (Sumi et al., 1992). These findings were considered as evidence that both types of ATPases (A/V- and F-ATPases) were already present in the last common ancestor (Tsutsumi et al., 1991; Forterre et al., 1993). Thus, the phylogenetic relations between F- and A/V-ATPases would not reflect the relations between the three domains. Rather, the separation of the F- from the A/V-ATPases would reflect another ancient gene duplication event. The proposed rooting of the universal tree of life was also questioned because the analyses of other ancient gene duplications came to different results. In the case of glutamine synthase (Kumada et al., 1993), homologues of the heat shock protein family (Gupta and Singh, 1992) and glutamate synthase (Benachenhou-Lafha et al., 1993) at least some of the archaeal sequences group with the bacterial domain.

This paper shows that the assumption of both A/V- and F-type ATPases being present in the last common ancestor leads to untenable conclusions. Our analysis of the evolution of ATPase subunits suggests horizontal gene transfer as a more likely explanation.

2. Material and methods

The sequence encoding the *Thermotoga maritima* β-subunit was kindly provided by Dr W. Ludwig, University of Munich. All other sequences used were retrieved from GenBank. The sequence alignment used for the maximum likelihood analysis was calculated using Jotun Hein's progressive alignment program (Hein, 1990). Twenty-five different ATPase subunit amino acid sequences (catalytic and non-catalytic), and the ATPase-like peptide encoded on the flagellar operon (Galizzi et al., 1991; Vogler et al., 1991) were aligned. The non-conserved amino and carboxy terminal regions, the non-homologous region in the *A*-subunits of the V- and A- ATPases (Zimniak et al., 1988), and the 'protein intron' in the yeast sequence (Kane et al., 1990) were deleted by hand after an initial alignment. The resultant sequences were aligned using a gap penalty of $16 + 3k$, where k denotes the size of the gap. Different gap penalties (between $4 + 3k$ and $40 + 3k$) only slightly affected the alignment, and did not change the topology of the calculated tree. All positions that corresponded to a gap in the refined alignment were deleted, resulting in 396 aligned positions per sequence. Alignments of the partial ATPase subunit sequences used for protein parsimony and distance analyses were performed by hand using Eric Cabot's Eye Ball Sequence Editor

(Department of Biological Sciences, Simon Fraser University, Burnaby, B.C., V5A 1S6, Canada). Only those regions homologous to the fragments of the *Enterococcus hirae* Na^+ pumping ATPase (Kakinuma et al., 1991) (Figs. 2 and 3, 349 amino acid positions including gaps) or the *Methanosarcina barkeri* β-subunit (Sumi et al., 1992) (Figs. 4 and 5, 125 amino acid positions) were used. All three alignments are available from the authors upon request. For protein parsimony and protein distance matrix analyses gaps were encoded as missing values.

The topology and branch lengths of the reported maximum likelihood tree were calculated with the protein maximum likelihood method as implemented by Adachi and Hasegawa (1992) using the star decomposition, the semi-automatic mode and various user-defined trees. Protein parsimony, protein distance matrix analysis, and the generation of bootstrapped samples were performed using the algorithms as implemented by Felsenstein (1993). Protein parsimony was used with the global rearrangement option and with different random input orders. Protein distance matrices were calculated using Dayhoff's PAM matrix (1979). The least squares error trees were determined using the method of Fitch and Margoliash (1967) as implemented by Felsenstein (1993). The identity of the tree with the least squares error was verified using different input orders and tree topologies calculated using the neighbour joining procedure (Saitou and Nei, 1987). The branch lengths are scaled with respect to expected amino acid substitution events per site. For a comparative discussion of the different algorithms see Felsenstein (1988) and references therein.

For each bootstrap analysis 100 bootstrapped samples were generated. Numbers give the percentage of parsimony analyses of those samples that contained the indicated groups (non-integer numbers are due to samples that gave multiple equally parsimonious trees).

3. Results and discussion

3.1. The A/V-type ATPase from Thermus thermophilus

When the *Thermus* ATPase subunits (Tsutsumi et al., 1991) are included in a phylogenetic analy-

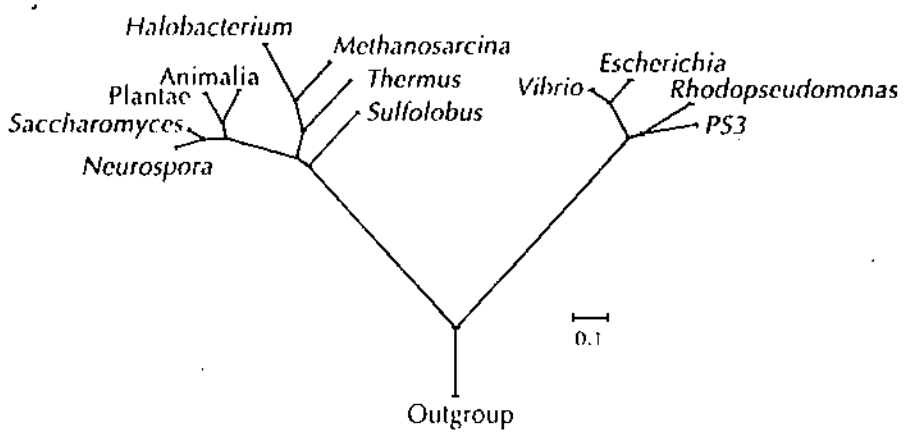

Fig. 1 Maximum likelihood phylogenetic tree depicting the evolution of V-, F-, and A-ATPases. A total of 792 positions from both the catalytic and the non-catalytic subunit were used for each terminal taxon. For Plantae the A subunit from carrot (*Daucus carota*) and the B-subunit from *Arabidopsis*, and for Animalia the bovine A (M831 30) and the human B subunit (M25809) were used. For *Sulfolobus acidocaldarius*, *Neurospora crassa*, *Methanosarcina barkeri*, *Halobacterium salinarium*, *Thermus thermophilus* (= *Thermus aquaticus* HB8), *Vibrio alginolyticus*, *Rhodopseudomonas blastica*, *Escherichia coli* and the thermophilic Bacterium PS3, both the catalytic (*A* respective β) and the non-catalytic subunits (*B* respective α) were used. The outgroup consisted of the aligned positions from the other subunit type of {[[Plantae *B/A*, *Halobacterium B/A*], *Sulfolobus B/A*], [*Vibrio* α/β, *Rhodopseudomonas* α β]]; [] indicate the branching order that was determined among the outgroup taxa. The same location for the outgroup was found using either catalytic or non-catalytic subunits separately, or by including the flagellar ATPase-like peptides in the outgroup, or by protein parsimony or protein distance matrix analyses. The length of each depicted branch was at least 3 times larger than the standard error calculated for that branch.

sis, they group within the archaeal domain together with the A-ATPases from *Halobacterium* and *Methanosarcina* (Fig. 1). How could *Thermus* end up with an archaeal type ATPase? Tsutsumi et al. (1991) and later Forterre et al. (1993) concluded that the presence of an A/V-ATPase in a Bacterium would show that the last common ancestor of Bacteria and Archaea would have already possessed both, an F-ATPase and an A/V-ATPase. Two sets of gene duplications would already have taken place in the universal ancestor. A first gene duplication gave rise to the catalytic and non-catalytic ATPase subunits. A second series of gene duplications would have led to the A/V- and to the F-ATPases (compare Fig. 4 in Forterre et al., 1993). However, this interpretation does not hold up to a closer examination. If this hypothesis would be correct, the A/V-ATPases alone could be used to trace the organismal evolution; i.e., the placement of the *Thermus* A/V-ATPase subunits relative to the archaeal and eucaryal A/V-ATPases should reflect the relations among the three domains, Bacteria, Archaea, and Eucarya. The root could be either represented by the non-catalytic subunits or by the F-ATPases. According to the current views on bacterial systematics (e.g., Woese, 1987), one should expect the bacterial A/V-ATPases to constitute a deep branch with respect to the other A/V-ATPases. However, as shown in Fig. 1, this is not the case for the *Thermus* ATPase. When the data set used for Fig. 1 was bootstrapped and the resulting 100 samples analysed by protein parsimony, in none of the most parsimonious trees did the *Thermus* sequences branch off closer to the outgroup than the *Sulfolobus* sequences.

3.2. The Na^+ pumping ATPase from *Enterococcus hirae*

The same result is obtained with the partial sequence from the sodium pumping A/V-type ATPase from *Enterococcus hirae* (Fig. 2). The two bacterial sequences group together with the mesophilic Archaea. The consensus tree of a bootstrap analysis of the data set is given in Fig. 3. Because only positions corresponding to the 1000-bp fragment encoding the *Enterococcus hirae* A/V-type ATPase catalytic subunit (Kakinuma et al., 1991) were used, the analysis of this data set (Figs. 2 and 3) is less reliable than the one depicted in Fig. 1. The single most parsimonious tree differs from the distance matrix tree depicted in Fig. 2 in that *Sulfolobus* branches off closer to the outgroup (*Escherichia* α and *Salmonella* fl) and separate from the other archaeal, eucaryal and the bacterial A/V-ATPases. Based on the bootstrap analysis, the possibility that the A/V-ATPases from *Ther-*

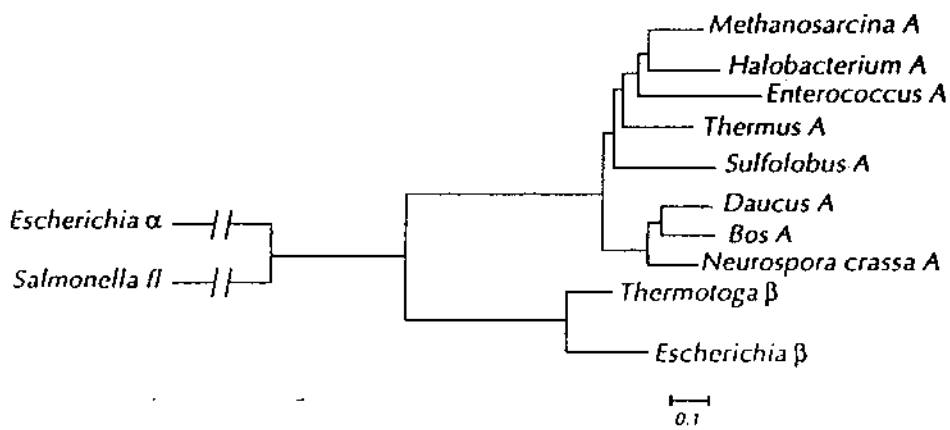

Fig. 2. Tree calculated by distance matrix analysis depicting the phylogenetic positions of bacterial A-ATPase subunits. 'fl' denotes the ATPase-like peptide that is encoded in the flagellar operon (Vogler et al., 1991); β and α denote the catalytic and non-catalytic F-ATPase subunits, respectively; *A* denotes the catalytic A/V-ATPase subunits. *Bos* and *Daucus* denotes the bovine and carrot V-ATPase *A*-subunits, respectively. Sequence fragments from *Enterococcus hirae* and *Salmonella typhimurium* were used in addition to the species listed in the legend to Fig. 1.

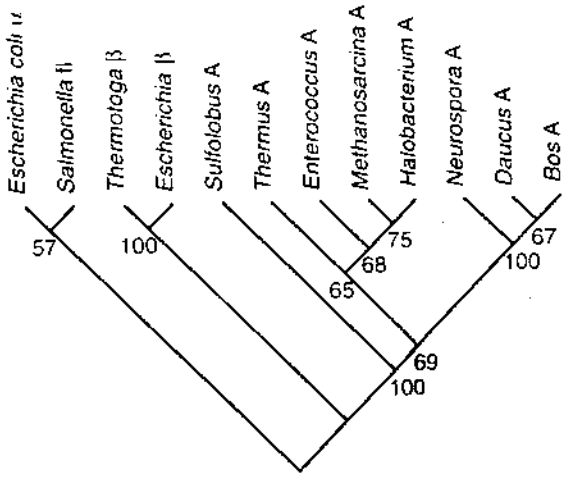

Fig. 3. Consensus tree from protein parsimony analyses of bootstrapped samples using the same sequence data as in Fig. 2.

mus and *Enterococcus* group together appears unlikely (found in 6.8% of the bootstrapped samples) but cannot be excluded. Nonetheless, there is no support for the hypothesis that the two bacterial A/V-ATPases form the deepest branch within the archaeal domain (found in none of the bootstrapped samples).

If the bacterial A/V-type ATPases are remnants of an A/V-ATPase already present in the last common ancestor, and if the A/V-ATPases catalytic subunits are considered as markers for the organismal evolution, then the root would be given by the non-catalytic subunits and by the catalytic F-ATPase. Therefore, Fig. 2 would lead one to conclude that Bacteria evolved from a mesophilic Archaeum. Clearly, this conclusion runs counter to traditional bacterial systematics (e.g., Woese, 1987; Woese et al., 1990), and, as discussed below, is also at odds with the conclusions obtained from the analysis of the archaeal F-ATPase.

3.3. An archaeal F-ATPase

The same argument given above for the bacterial A/V-ATPases can also be applied to the archaeal F-type ATPase. The partial sequence of an F-ATPase β subunit was obtained from *Methanosarcina barkeri* genomic DNA (Sumi et al., 1992). If this archaeal F-ATPase would have evolved from an ancestral F-ATPase already present in the last common ancestor, then the archaeal F-ATPase should represent the deepest branch in an F-ATPase phylogeny. However, the *Methanosarcina barkeri* β subunit groups together with *Anabaena* and branches off after the β-subunits from *Thermotoga maritima* and PS3 (a

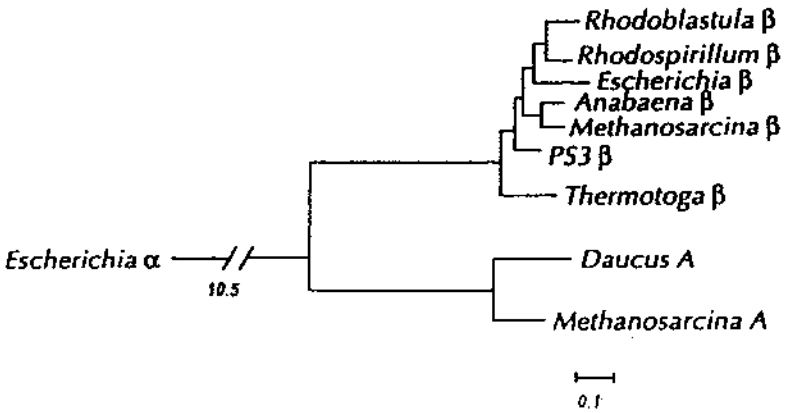

Fig. 4. Tree calculated by distance matrix analysis depicting the phylogenetic positions of the putative archaeal F-ATPase. Only those positions which correspond to the published sequence of the β subunit of the F-ATPase from *Methanosarcina barkeri* (Sumi et al., 1992) were used. Except for *Rhodospirillum rubrum* and *Anabaena* sp. the species and abbreviations are the same as in Figs. 1 and 2. Note that in the single-most parsimonious tree calculated from the same data set the sequence from PS3 branches off together with *Anabaena* and *Methanosarcina* (compare Fig. 5).

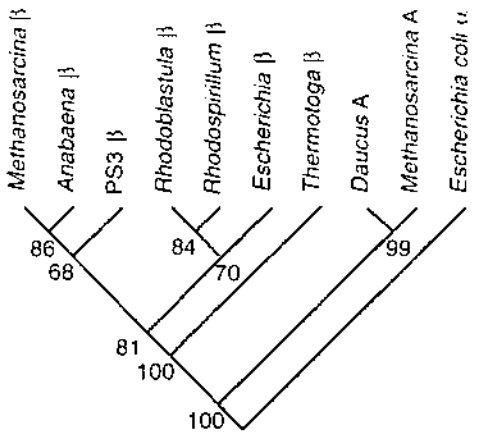

Fig. 5. Consensus tree from protein parsimony analyses of bootstrapped samples using the same sequence data as in Fig. 4. In none of the bootstrapped samples was the β subunit from *Methanosarcina* determined to form the deepest branch among the β-subunits.

thermophilic Bacterium) (Fig. 4). The bootstrap probabilities for the groupings within the bacterial domain are quite low (only 125 positions, including gaps in some of the sequences were used) (Fig. 5); however, in none of the bootstrapped samples did the *Methanosarcina* sequence form the deepest branch within the F-ATPase β-subunits.

3.4. A case of horizontal gene transfer?

The A/V-ATPase phylogeny (Figs. 1 and 2) suggests that Bacteria have evolved from Archaea. The F-ATPase phylogeny (Fig. 4) would suggest that the Archaea have evolved from a bacterial ancestor. Furthermore, both phylogenies are at odds with many other characters (for a listing of several discrete characters see Zillig et al., 1992) which, taken together, argue convincingly that Archaea and Bacteria constitute two separate domains. These contradictions compel us to reconsider the step from gene trees to species trees. As the assumption of both A/V- and F-ATPases being present in the last common ancestor leads to untenable conclusions, the possibility of horizontal gene transfers across domain boundaries remains the most likely alternative.

The assumption of horizontal gene transfer between species and even the merger between species to form new organisms is not without precedent. Well-studied examples are: resistance genes in Bacteria (e.g., Kaur and Rosen, 1992), transposable elements (Daniels et al., 1990), and between the nucleus and eukaryotic cell organelles (Margulis, 1981). In addition, horizontal gene transfer has been considered as an explanation for discrepancies between gene and species trees in the following cases: 16S rRNA of plant mitochondria (Gray et al., 1989), glyceraldehyde phosphate dehydrogenase from *Escherichia coli* (Doolittle et al., 1990), glucose-6-phosphate isomerase (Smith and Doolittle, 1992) and glutamine synthases (Smith et al., 1992; Tiboni et al., 1993). It appears that horizontal transfer of genes between species is not restricted to modern-day bacteria, but occurred throughout evolution.

3.5. Other ancient duplicated genes

The placement of the root in the universal tree of life is the same for a variety of different markers (Gogarten et al., 1989; Iwabe et al., 1989; Cammarano et al., 1992). Forterre et al. (1993) cited the study of glutamate dehydrogenases (GDH) (Benachenhou-Lafha et al., 1993) to arrive at a different conclusion. To assess the reliability of the interior branches in the GDH-phylogeny (i.e., the reliability with which the location of the last common ancestor can be inferred), we performed a bootstrap analysis on the aligned GDH data set that was kindly provided by B. Labedan (Fig. 6). The separation of the GDHs into two families, and the group comprising the eucaryal GDH family II are highly supported. In contrast, the grouping of the archaeal GDHs together with their bacterial counterparts was found only in 48.5% of the bootstrapped samples (Fig. 6). In 39% of the bootstrapped samples the GDH from *Halobacterium* grouped with its eucaryal counterparts, and in 10% both archaeal species grouped with the Eucarya. A similar result was obtained when only the most conserved portion of the GDHs (position 111–214) is used for analysis (data not shown). The GDH data set does not allow for a reliable placement of the root in the universal tree of life. In contrast to this, the grouping of the archaeal A-ATPase subunits together with the eucaryal V-ATPases (separate from the bacterial F-ATPases and the outgroup) is found in 100% of the

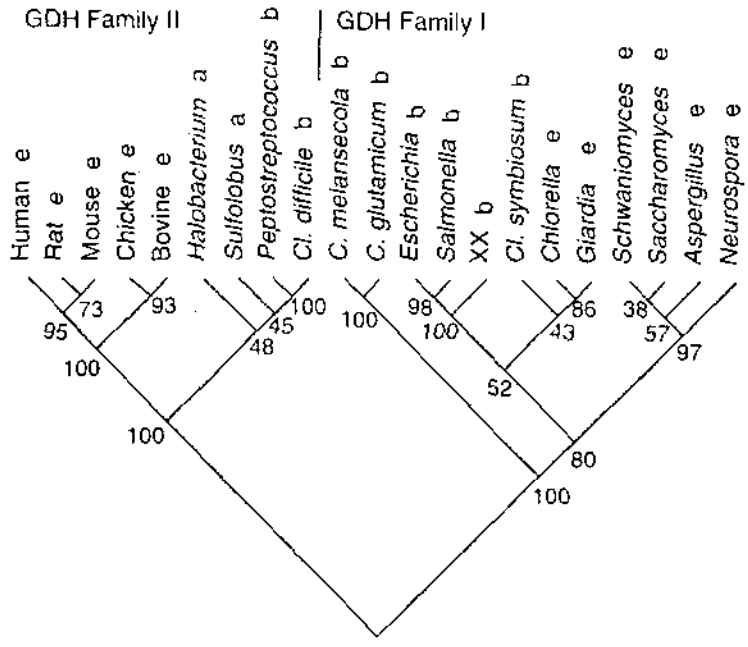

Fig. 6. Consensus tree of protein parsimony analyses of samples bootstrapped from the glutamate dehydrogenase (GDH) data set (Benachenhou-Lafha et al., 1993). The aligned sequences were kindly provided by B. Labedan. Only the conserved central region (position 111–494) was utilized. Note that the group in Family II that comprises the archaeal and bacterial GDHs has low bootstrap values. Abbreviations: C, *Corynebacterium*; Cl, *Clostridium*; XX, unidentified Bacterium contaminating *Chlorella* cultures; e b, and a denote the classification of the species as eucaryal, bacterial, and archaeal, respectively.

bootstrapped samples (Fig. 3). It seems that in the case of the GDHs, not enough information survived from the early evolutionary stage. We conclude that the analysis of the GDH sequences neither confirms nor contradicts the placement of the root that has been derived from other ancient duplicated genes.

The phylogeny of the glutamine synthases (Smith et al., 1992; Kumada et al., 1993; Tiboni et al., 1993) is similar to the phylogeny of the F-ATPases presented here. The glutamine synthase I from *Methanococcus voltae* groups with Gram-positive bacteria; furthermore, not only Archaea, but also Bacteria appear as polyphyletic. Tiboni et al. (1993) concluded that the glutamine synthase I data indicate at least one case of horizontal gene transfer across domain boundaries.

3.6. How many genes were transferred?

Additional molecular phylogenies might yield information whether horizontal gene transfers have occurred infrequently involving several functionally unrelated genes, or whether the transfer of single genes occurred more often. The latter scenario resembles the transfer of resistance genes in modern Bacteria, some of which (e.g., the resistance towards antimony and arsenic) are plasmid encoded multisubunit transport ATPases (Kaur and Rosen, 1992). This mode of horizontal gene transfer might have led to the Na^+ pumping A/V-type ATPase in *Enterococcus hirae*. The former case — i.e., the horizontal transfer of a larger portion of the genome — might have resulted from an endosymbiosis-like association of the involved organisms. This scenario applied to an ancestor of the genus *Thermus* (as recipient) and a mesophilic Archaeum (as donor) could explain why the coupling factor ATPase in the Bacterium *T. thermophilus* and the malate dehydrogenase in *T. flavus* (Iwabe et al., 1989) are both of the archaeal/

eucaryal type, whereas other characters clearly denote the genus *Thermus* as bacterial (Woese, 1987).

3.7. The origins of the eukaryotic cell

The finding that genes were exchanged between distantly related species implies that a single gene phylogeny can no longer be readily interpreted as a species tree. To determine the evolution of species more than one gene tree should be considered. The congruence of several marker molecules (Gogarten et al., 1989; Iwabe et al., 1989; Cammarano et al., 1992) suggests that the presence of archaeal type macromolecules in Eucarya is not due to a horizontal transfer. The different molecular phylogenies taken together indicate that an organism that branches close to the base of the archaeal domain gave rise to most of the eucaryal nucleocytoplasm.

Clearly, these findings do not exclude the possibility that other organisms made contributions to the eukaryotic cell, e.g. other Bacteria besides the mitochondria and plastids (Margulis, 1981, Zillig et al., 1992). Possibly an even more ancient independent lineage also could have been involved in the genesis of the eukaryotic cell (Sogin, 1992; Hartman, 1984). However, more data are needed if we want to trace the origins of the eukaryotic cell in more detail.

4. Acknowledgments

We thank Henrik Kibak, Lincoln Taiz, Patrick Forterre and Antonio Lazcano for the stimulating exchange of ideas, Richard Mercier, Maria Gogarten-Boekels and Scot Magnotta for critically reading the manuscript, Jeff Lary for setting up the PROTML program, and Bernard Labedan for providing us with the aligned GDH sequences. The research was funded by NSF grant BSR-9020868, E.H. was supported in part by a fellowship from the DGAPA-Universidad Nacional Autónoma de México.

5. References

Adachi, J. and Hasegawa, M., 1992, Computer Science Monographs, No 27, MOLPHY: Programs for Molecular Phylogenetics I-PROTML: Maximum Likelihood Inference of Protein Phylogeny (Institute of Statistical Mathematics, Tokyo, Japan).

Benachenhou-Lafha, N., Forterre, P. and Labedan, B., 1993, Evolution of glutamate dehydrogenase genes: evidence for two paralogous protein families and unusual branching patterns of the archaebacteria in the universal tree of life. J. Mol. Evol. 36, 335–346.

Cammarano, P., Palm, P., Creti, R., Ceccarelli, E., Sanangelantoni, A.M. and Tiboni, O., 1992, Early evolutionary relationships among known life forms inferred from elongation factor ef-2 (ef-g) sequences. Phylogenetic coherence and structure of the archaeal domain. J. Mol. Evol. 34, 396–405.

Daniels, S.P., Peterson, K.R., Strausbaugh, L.D., Kidwell, M.G. and Chovnick, A., 1990, Evidence for horizontal transmission of the P transposable element between Drosophila species. Genetics 124, 339–355.

Dayhoff, M.O., 1979, Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 3 (National Biomedical Research Foundation, Washington, DC).

Doolittle, R.F., Feng, D.F., Anderson, K.L. and Alberro, M.R., 1990, A naturally occurring horizontal gene transfer from a eukaryote to a prokaryote. J. Mol. Evol. 31, 383–388.

Felsenstein, J., 1988, Phylogenies from molecular sequences: interference and reliability. Annu. Rev. Genet. 22, 521–565.

Felsenstein, J., 1993, Phylogeny Inference Package, version 3.5c. (Distributed by the author; Department of Genetics, University of Washington, Seattle).

Fitch, W.M. and Margoliash, E., 1967, Construction of phylogenetic trees. A method based on mutation distances as estimated from cytochrome *c* sequences is of general applicability. Science 155, 279–284.

Forterre, P., Benachenhou-Lafha, N., Confalonieri, F., Duguet, M., Elie, C. and Labedan, B., 1993, The nature of the last universal ancestor and the root of the tree of life, still open questions. BioSystems 28, 15–32.

Galizzi, A., Scoffone, F., Crabb, W.D., Caramoni, T. and Albertini, A.M., 1991, The flaA locus of *Bacillus subtilis* is part of a large operon coding for flagellar structures, motility functions, and an ATPase like polypeptide. J. Bacteriol. 173, 3573–3579.

Gogarten, J.P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E.J., Bowman, B.J., Manolson, M.F., Poole, R.J., Date, T., Oshima, T., Konishi, J., Denda, K. and Yoshida, M., 1989, Evolution of the vacuolar H^+-ATPase: Implications for the origin of eukaryotes. Proc. Natl. Acad. Sci. USA 86, 6661–6665.

Gogarten, J.P., Starke, T., Kibak, H., Fichmann, J. and Taiz, L. 1992, Evolution and isoforms of V-ATPase subunits. J. Exp. Biol. 172, 137–147.

Gray, M.W., Cedergren, R., Abel, Y. and Sankoff, D., 1989, On the evolutionary origin of the plant mitochondrion and its genome. Proc. Natl. Acad. Sci. USA 86, 2267–2271.

Gupta, R.S. and Singh, B., 1992, Cloning of the HSP70 gene from *Halobacterium marismortui*: relatedness of ar-

chaebacterial HSP70 to its eubacterial homologs and a model of the evolution of the HSP70 gene. J. Bacteriol. 174, 4594–4605.

Hartman, H., 1984, The origin of the eukaryotic cell. Speculation Sci. Technol. 7, 77–81.

Harvey, W.R., Nelson, N., (eds.) 1992, V-ATPases. J. Exp. Biol. 172.

Hein, J., 1990, Unified approach to alignment and phylogenies. Methods Enzymol. 183, 626–645.

Hensel, R., Demharter, W., Kandler, O., Kroppenstedt, M. and Stackebrandt, E., 1986, Chemotaxonomic and molecular-genetic studies of the genus *Thermus*: evidence for a phylogenetic relationship of *Thermus aquaticus* and *Thermus ruber* to the genus *Deinococcus*. Int. J. Syst. Bacteriol. 36, 444–453

Ihara, K. and Mukohata, Y., 1992, The ATP synthase of *Halobacterium salinarium* is an archaebacteria type as revealed from the amino acid sequences of its two major subunits. Arch. Biochem. Biophys. 286, 111–116.

Iwabe, N., Kuma, K-I., Hasegawa, M., Osawa, S. and Miyata, T. 1989, Evolutionary relationships of archaebacteria, eubacteria and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. USA 86, 9355–9359.

Kakinuma, Y., Igarishi, K., Konishi, K. and Yamato, I., 1991, Primary structure of the alpha-subunit of vacuolar-type Na^+-ATPase in *Enterococcus hirae*, amplification of a 1000bp fragment by polymerase chain reaction. FEBS Lett. 292, 64–68.

Kane, P.M., Yamashiro, C.T., Wolczyk, D.F., Neff, N., Goebl, M. and Stevens, T.H., 1990, Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar adenosine triphbsphatase. Science 250, 651–657.

Kaur, P. and Rosen, B.P., 1992, Plasmid encoded resistance to arsenic and antimony. Plasmid 27, 29–40.

Kumada, Y., Benson, D.R., Hillemann, D., Hosted, T.J., Rochford, D.A., Thompson, C.J., Wohlleben, W. and Tateno, Y., 1993, Evolution of the glutamine synthase gene, one of the oldest existing and functioning genes. Proc. Nat. Acad. Sci. USA 90, 3009–3013.

Margulis, L., 1981, Symbiosis in Cell Evolution (W.H. Freeman and Company, New York).

Rivera, M.C. and Lake, J.A., 1992, Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. Science 257, 74–76.

Saitou, N. and Nei, M., 1987, The neighbor joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Schäfer, G. and Meyering-Vos, M., 1992, F-type or V-type? The chimeric nature of the archaebacteria ATPsynthase. Biochim. Biophys. Acta 1101, 232–235.

Schwartz, R.M. and Dayhoff, M.O., 1978, Origins of prokaryotes, mitochondria and chloroplasts. Science 199, 395–403.

Smith, M.W. and Doolittle, R.F., 1992, Anomalous phylogeny involving the enzyme glucose-6-phosphate isomerase. J. Mol. Evol. 34, 544–545.

Smith, M.W., Feng, D.F., Doolittle, R.F., 1992, Evolution by acquisition: the case for horizontal gene transfer. Trends Biochem. Sci. 17, 489–493

Sogin, M.L., 1992, Early Evolution and the origin of eukaryotes. Curr. Opinion Genet. Dev. 1, 457–463.

Space Studies Board 1990, The Search for Life's Origins (National Academy Press, Washington DC).

Sumi, M., Sato, M.H., Denda, K., Date, T. and Yoshida, M., 1992, A DNA fragment homologous to F_1-ATPase β subunit was amplified from genomic DNA of *Methanosarcina barkeri*. Indication of an archaebacterial F-type ATPase. FEBS Lett. 314, 207–210.

Tiboni, O., Cammarano, P. and Sanangelantoni, A.M., 1993, Cloning and sequencing of the gene encoding glutamine synthase I from the Archaeum *Pyrococcus woesei*: anomalous phylogenies inferred from analysis of archaeal and bacterial glutamine synthase I sequences. J. Bacteriol. 175, 2961–2969.

Tsutsumi, S., Denda, K., Yokoyama, K., Oshima, T., Date, T. and Yoshida, M., 1991, Molecular cloning of genes encoding major two subunits of a eubacterial V-Type ATPase from *Thermus thermophilus*. Biochim. Biophys. Acta 1098, 13–20.

Vogler, P.A., Homma, M., Irikura, V.M. and MacNab, R.M., 1991, *Salmonella typhimurium* mutants defective in flagellar filament regrowth and sequence similarity of FliI to F_0F_1, vacuolar and archaebacterial ATPase subunits. J. Bacteriol. 173, 3564–3572.

Wheelis, M.L., Kandler, O. and Woese, C.R., 1992, On the nature of global classification. Proc. Natl. Acad. Sci. USA 89, 2930–2934.

Woese, C.R., 1987, Bacterial evolution. Microbiol. Rev. 51, 221–271.

Woese, C.R., Kandler, O. and Wheelis, M.L., 1990, Towards a natural system of organisms: proposal for the domains Archaea, Bacteria and Eucarya. Proc. Natl. Acad. Sci. USA 87, 4576–4579.

Yokoyama, K., Oshima, T. and Yoshida, M., 1990, *Thermus thermophilus* membrane-associated ATPase: indication of a eubacterial V-type ATPase. J. Biol. Chem. 265 21946–21950.

Zillig, W., Palm, P. and Klenk, H-P., 1992. A model of the early evolution of organisms: the arisal of the three domains of life from the common ancestor. in: The Origin and Evolution of the Cell, H. Hartman and K. Matsuno (eds.) (World Scientific, Singapore, New Jersey, London, Hong Kong) pp. 163–182.

Zimniak, L., Dittrich, P., Gogarten, J.P., Kibak, H. and Taiz, L., 1988, The cDNA sequence of the 69 kDa subunit of the carrot vacuolar H^+-ATPase, homology to the beta-chain of F_0F_1-ATPases. J. Biol. Chem. 263, 9102–9112.