Correspondence

The root of the eukaryote tree pinpointed

Alexandra Stechmann and Thomas Cavalier-Smith

Determining the precise position of the root of the eukaryote evolutionary tree is very important for understanding cell evolution [1], but has been a major challenge, because of systematic biases in gene sequence trees [2,3]. However, one can deduce the position indirectly, by using derived genic properties to exclude the possibility that the root lies within clades that share them [4]. A derived gene fusion between dihydrofolate reductase (DHFR) and thymidylate synthase (TS) genes strongly indicated that the root is not amongst the eukaryote groups ancestrally with two cilia (bikonts [1]). Bikonts share a fused DHFR-TS gene that is clearly derived [4]. However, the precise position of the root of the eukaryote tree remained uncertain, because of the unclear status of the Amoebozoa [4]. We have sequenced the region upstream of the TS gene in the amoebozoan *Hartmannella cantabrigiensis* and find that the TS and DHFR genes are separate. As the two genes are encoded on opposite strands they must be translated separately and thus be unfused (inset in Figure 1). Also the choanozoan *Corallochytrium limacisporum* contains only non-coding DNA upstream of the TS gene, but not a DHFR gene. Thus, all major opisthokont groups lack fused DHFR-TS genes. Until now, this had only been inferred from the fact that we were not able to detect the fusion gene in *Corallochytrium*, the fact that Choanozoa are sisters to animals [5] and from the absence of the fusion gene in both animals and Fungi.

Given the improbability of reversal of the fusion in the bikont ancestor [4], this provides the best evidence to date that Amoebozoa are not secondarily derived from the bikonts, and that bikonts are a clade, not a paraphyletic group. Do other Amoebozoa also have separate DHFR and TS genes? Using our new sequences from *Hartmannella,* we searched ongoing genome sequencing projects of different *Entamoeba* species and *Dictyostelium discoideum* without success [4]. Our inability to amplify the bifunctional gene from two other Amoebozoa, *Phalansterium* and *Phreatamoeba,* suggests that all Amoebozoa lack the gene fusion.

Another derived gene fusion, of three of the six enzymes in the pyrimidine synthesis pathway [7], provides the first really compelling support for Amoebozoa being sisters to opisthokonts rather than to bikonts. In eubacteria and archaebacteria all six enzymes of this pathway are separately translated, as are the two subunits of the first of these, carbamoyl-phosphate synthetase II (CPSII) [7]. CPSII had clearly undergone a fusion between its two subunits in the ancestral eukaryote, as they are fused in the trypanosomatids and Sporozoa. In plants (at least angiosperms), this fused enzyme has been replaced by unfused paralogues from the cyanobacterial ancestor of chloroplasts. In animals, Fungi and *Dictyostelium*, however, the first three enzymes are fused into a multienzyme protein [7]. No bikonts are known to have this three-gene fusion, which is hence a shared derived character that unites opisthokonts and Amoebozoa into a single clade and thus implies that the root of the tree does not lie within them. We previously designated this clade unikonts [6], because we inferred that its common ancestor was uniciliate and probably unicentriolar (unikont [1]).

As the DHFR-TS fusion can be used to exclude the root of the tree from the bikonts and the pyrimidine pathway three-gene fusion excludes it from the unikonts, the root of the tree lies precisely between these two clades. Strictly speaking, we cannot yet firmly rule out the possibility that Amoebozoa are paraphyletic and that only some of them are sisters of unikonts. However, we can already use a uniquely derived fusion of the mitochondrial cytochrome oxidase 1 and 2 genes shared by *Dictyostelium* and *Acanthamoeba* [8] to show that they are part of a broader amoebozoan clade and that neither the eukaryote nor the opisthokont root can lie within this amoebozoan subclade.

An internal gene duplication to form an N-terminal catalytic part and C-terminal regulatory part of phosphofructokinase (PFK) also cleanly partitions eukaryotes into unikonts and bikonts. This duplication/fusion is a shared derived state of unikonts, so far found only in animals, fungi and *Dictyostelium* [9,10]. If it originated in their last common ancestor it would be a unikont synapomorphy like the triple gene fusion. Bikonts have very different PFK paralogues, lacking both this duplicated PFK and its unduplicated orthologue [11], widespread in bacteria but apparently absent in actinobacteria or archaebacteria. As the ancestor of eukaryotes [1] is likely to be a transient intermediate between these two groups, this paralogue probably entered eukaryotes either by lateral gene transfer into unikonts, which would make it a unikont synapomorphy, or, more likely, via the ancestral mitochondrial symbiosis, which would make its loss a bikont synapomorphy. Thus, this unikont gene duplication implies that unikonts or bikonts (or both) are holophyletic. If the PFK duplication continues to prove absent from all bikonts with more extensive taxon sampling, we would have two shared derived characters uniting the unikonts, as well as two independent shared derived characters uniting the bikonts, all together supporting a position of the root precisely between unikonts and bikonts.

Our pinpointing the root of the eukaryote tree between the unikonts and bikonts has numerous important evolutionary implications. It shows, contrary to earlier ideas [3], that there are no extant eukaryotes that branched off from the tree prior to the divergence of the ancestors of animals and plants. For the first

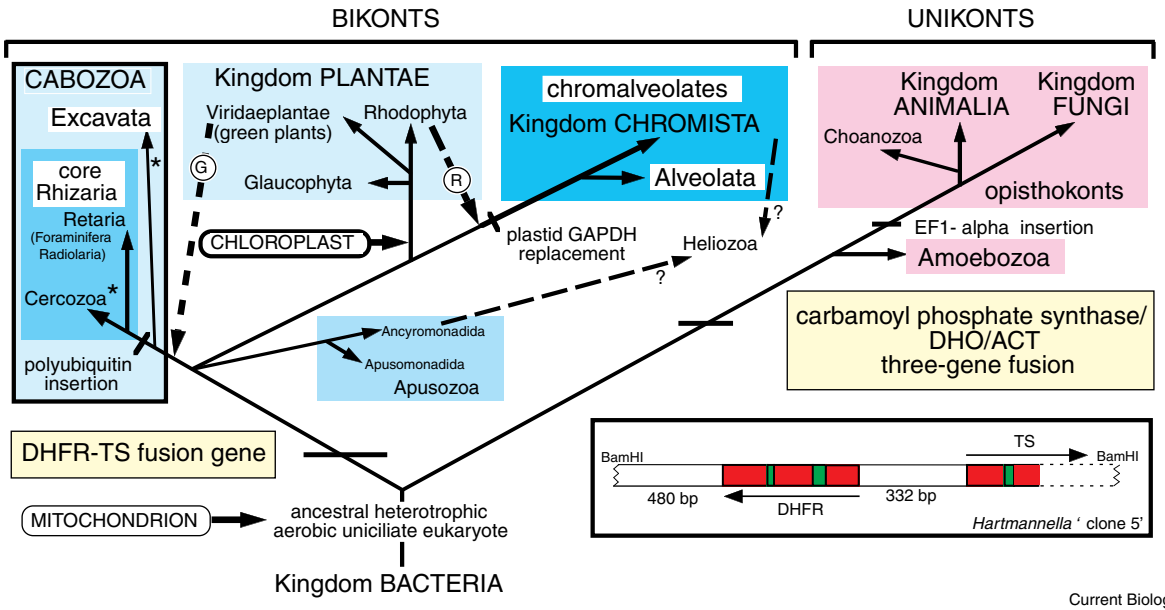

Figure 1. Eukaryote phylogeny integrating ultrastructure, sequence trees, gene fusions and molecular cladistic markers.
The unikont topology is established, but the branching order of the six bikont groups remains uncertain. The single enslavement [12] of a red alga (R) to create chromalveolates is supported by a plastid glyceraldehyde phosphate dehydrogenase (GAPDH) replacement [13]. Whether there was a single enslavement of a green alga (G) to form cabozoa or two separate enslavements (asterisks) to form Cercozoa and Excavata is uncertain [12], as is the position of Heliozoa [14]. Polyubiquitin [15] and EF-1α [16] insertions strongly support the clades core Rhizaria and opisthokonts. The inset shows the *Bam*HI restriction fragment from *H. cantabrigiensis* that was sequenced and analysed in this study, spanning the DHFR and the amino terminus of the TS gene (red, introns are green). The length of the noncoding regions upstream and downstream of the DHFR gene from one of the clones is indicated.

time it enables us to reconstruct the phenotype of the last common ancestor (cenancestor) of eukaryotes [1]. Any homologous character present in two taxa on opposite sides of the fundamental unikont/bikont divide must have been in the cenancestor, in the absence of lateral transfer. Nevertheless, it is highly desirable that other genic characters of comparable phylogenetic decisiveness are found and used to further test our conclusion.

Supplemental Data

Supplemental data are available at http://images.cellpress.com/supmat/supmatin.htm.

References

1. Cavalier-Smith, T. (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. Int. J. Syst. Evol. Microbiol. *52*, 297–354.
2. Philippe, H., Lopez, P., Brinkmann, H., Budin, K., Germot, A., Laurent, J., Moreira, D., Müller, M. and Le Guyader, H. (2000). Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. Proc. R. Soc. Lond. B Biol. Sci. *267*, 1213–1221.
3. Roger, A.J. (1999). Reconstructing early events in eukaryotic evolution. Am. Nat. *154*, S146–S163.
4. Stechmann, A. and Cavalier-Smith, T. (2002). Rooting the eukaryote tree by using a derived gene fusion. Science *297*, 89–91.
5. Cavalier-Smith, T. and Chao, E.E.-Y. (2003). Phylogeny of Choanozoa, Apusozoa and other protozoa and early eukaryote megaevolution. J. Mol. Evol. *56*, 540–563.
6. Stechmann, A. and Cavalier-Smith, T. (2003). Phylogenetic analysis of eukaryotes using heat-shock protein Hsp90. J. Mol. Evol., in press.
7. Nara, T., Hashimoto, T. and Aoki, T. (2000). Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. Gene *257*, 209–222.
8. Gray, M.W., Lang, B.F., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T.G., *et al.* (1998). Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res. *26*, 865–878.
9. Poorman, R.A., Randolph, A., Kemp, R.G. and Heinrikson, R.L. (1984). Evolution of phosphofructokinase – gene duplication and creation of new effector sites. Nature *309*, 467–469.
10. Estevez, A.M., Martinez-Costa, O.H., Sanchez, V. and Aragon, J.J. (1997). Cloning, sequencing and developmental expression of phosphofructokinase from *Dictyostelium discoideum.* Eur. J. Biochem. *243*, 442–451.
11. Bapteste, E., Moreira, D. and Philippe, H. (2003). Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. Gene, in press.
12. Cavalier-Smith, T. (2003). Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). Philos. Trans. R. Soc. London Ser. B *358*, 109–134.
13. Fast, N.M., Kissinger, J.C., Roos, D.S. and Keeling, P.J. (2001). Nuclear encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. Mol. Biol. Evol. *18*, 418–426.
14. Cavalier-Smith, T. and Chao, E.E.-Y. (2003). Molecular phylogeny of centrohelid Heliozoa, a novel lineage of bikont eukaryotes that arose by ciliary loss. J. Mol. Evol. *56*, 387–396.
15. Archibald, J.M., Longet, D., Pawlowski, J. and Keeling, P.J. (2003). A novel polyubiquitin structure in Cercozoa and Foraminifera: evidence for a new eukaryotic supergroup. Mol. Biol. Evol. *20*, 62–66.
16. Baldauf, S.L. and Palmer, J.D. (1993). Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. Proc. Natl. Acad. Sci. U.S.A. *90*, 11558–11562.

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.