A kingdom’s progress: Archezoa and the origin of eukaryotes

Patrick J. Keeling*

Summary
The taxon Archezoa was proposed to unite a group of very odd eukaryotes that lack many of the characteristics classically associated with nucleated cells, in particular the mitochondrion. The hypothesis was that these cells diverged from other eukaryotes before these characters ever evolved, and therefore they represent ancient and primitive eukaryotic lineages. The kingdom comprised four groups: Metamonada, Microsporidia, Parabasalia, and Archamoebae. Until recently, molecular work supported their primitive status, as they consistently branched deeply in eukaryotic phylogenetic trees. However, evidence has now emerged that many Archezoa contain genes derived from the mitochondrial symbiont, revealing that they actually evolved after the mitochondrial symbiosis. In addition, some Archezoa have now been shown to have evolved more recently than previously believed, especially the Microsporidia for which considerable evidence now indicates a relationship with fungi. In summary, the mitochondrial symbiosis now appears to predate all Archezoa and perhaps all presently known eukaryotes.

INTRODUCTION
Prior to the popularization of the endosymbiotic theory, it was widely believed that the evolutionary link between prokaryotes and eukaryotes was the presence of photosynthesis in cyanobacteria and algae. The biochemistry of oxygenic photosynthesis was considered too complicated and too similar in detail to have arisen twice independently. Therefore, it was reasoned that all photosynthetic organisms were related, and by extension that cyanobacteria had evolved into photosynthetic eukaryotes. This ancestral eukaryote was thought to be like red algae because their pigments and light-harvesting antennae most closely resemble those of cyanobacteria and they also lack flagella and basal bodies (for discussion see Ref. 1). However, according to the endosymbiotic theory, the reason photosynthesis is so similar in cyanobacteria and photosynthetic eukaryotes is that the plastids of plant and algal cells are derived from a cyanobacterial symbiont. With increasing acceptance of the origin of plastids from cyanobacteria, links between cyanobacteria and the nucleus dissolved, and with it our explanation for the origin of eukaryotes.

An alternative to the cyanobacterial origin of eukaryotes arose from what seemed like an unlikely source when Carl Woese and his colleagues discovered an unexpected division in prokaryotes in 1977. Woese’s group showed that prokaryotes are composed of two very distantly related groups, which they named Eubacteria and Archaeabacteria, now synonymous with Bacteria and Archaea. Archaeabacteriologists soon began to find molecular links between archaeabacteria and eukaryotes, and these observations were brought into focus by the demonstration that archaeabacteria and eukaryotes are one another’s closest relatives in rooted universal trees. This means that archaeabacteria share a recent common ancestor with eukaryotes, so it makes...
perfect sense that many aspects of archaeobacterial molecular biology also should be found to resemble their eukaryotic counterparts. Each of these shared similarities clarifies the prokaryote-eukaryote transition in a small way by showing that molecular traits previously considered “eukaryotic” actually predate nucleated cells. However, these shared characters tie the archaebacteria to all eukaryotes. Unlike the cyanobacteria-to-alga hypothesis, no single group of eukaryotes specifically resembles the archaebacteria. So what was the nature of the first eukaryote?

ARCHEZOA: ARCHETYPICAL EUKARYOTES

When the endosymbiotic theory became fashionable and the photosynthetic origin of eukaryotes less so, the absence of mitochondria in certain eukaryotic cells started to attract attention. If the mitochondrion, like the plastid, originated by an endosymbiotic event, then it was also possible that some amitochondrial eukaryotes diverged prior to this event. This was first suggested for the hypermastigotes, a group of Parabasalia, and was expanded and refined by Cavalier-Smith in 1983. Cavalier-Smith proposed the Archezoa to contain the descendants of ancient premitochondrial eukaryotes: the Metamonads, Parabasalia, Microsporidia, and a new group, the Archamoebae. Each of these groups comprised anaerobic, amitochondrial cells that are morphologically very simple. Not only do they lack mitochondria but also peroxisomes or microbodies, in most cases Golgi dictyosomes, and in some also flagella (see Table 1). What’s more, the ribosomes of Metamonads, Parabasalia, and Microsporidia also were known to be about the same size as those of prokaryotes. Eukaryotic ribosomes are typically 80S in size, whereas those of archezoa and prokaryotes are 70S. Similarly, eukaryotic rRNA molecules are for the most part 18S and 28S in contrast to the smaller rRNAs found in prokaryotes and archezoa.

Almost immediately after the Archezoa was proposed, evidence from molecular phylogeny was produced to support its validity. The small subunit rRNA genes from a microsporidian, a diplomonad (the most studied subdivision of Metamonada), and a parabasalian were all published within a short space of time, and these three genes branched deeper in the rRNA tree than any previously known eukaryotic sequence. Although the order of the three groups remains contentious in rRNA phylogeny, trees of other molecules involved in gene expression, such as EF-1a, EF-2, RNA polymerase subunits, isoleucyl-tRNA synthetase, and large subunit rRNA, corroborated the deep branching position of these organisms.

This phylogenetic evidence fulfilled the most basic prediction of the Archezoa hypothesis for these groups: if they predate the origin of the mitochondrion, then they must branch earlier than mitochondrion-containing eukaryotes in phylogenetic trees (see Fig. 1). However, this was the high-water mark for the Archezoa. In recent years evidence has emerged that several of these organisms contain a genetic residue of the mitochondrion, which means that they could not have evolved before the endosymbiosis. The Archezoa is fast losing membership, and it now appears that the mitochondrial endosymbiosis may have taken place before the evolution of any of the presently known eukaryotic lineages.

### TABLE 1. Some Common Names and Characteristics of Archezoan Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Some common genera</th>
<th>Mitochondria</th>
<th>Peroxisomes</th>
<th>Golgi</th>
<th>Flagella</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metamonads</td>
<td>Giardia, Hexamita, Trepomonas, Retortamonas, Pyrsonympha</td>
<td>Not recognized</td>
<td>Not recognized</td>
<td>Not recognized</td>
<td>Yes</td>
</tr>
<tr>
<td>Microsporidia</td>
<td>Encephalitozoon Nosema, Spraguea Vairimorpha</td>
<td>Not recognized</td>
<td>Not recognized</td>
<td>Not recognized</td>
<td>No</td>
</tr>
<tr>
<td>Parabasalla</td>
<td>Trichomonas, Tritrichomonas, Monocercomonas</td>
<td>Hydrogenosome?</td>
<td>Not recognized</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Archamoebae</td>
<td>Trichonympha, Entamoeba, Pelomyxa, Mastigamoeba, Phreatamoeba</td>
<td>Not recognized</td>
<td>Not recognized</td>
<td>In some</td>
<td>In some</td>
</tr>
</tbody>
</table>
ARCHAMOEBAE: FIRST TROUBLE FOR ARCHEZOA

Little can be said about the Archamoebae as a group because they share almost no unifying characters other than the fact that they are all amitochondriate amoebae. Otherwise, they are a very diverse group inhabiting a wide variety of environments and making a living in many diverse ways. The diversity of the Archamoebae has been reason to doubt the phylogenetic unity of the group,17 but at the same time the unparalleled simplicity of the members also has led to the suggestion that of all Archezoa, Archamoebae are the most primitive18 (Fig. 2).

The first hint that any Archamoeba may be secondarily amitochondrial came from the enteric pathogen, Entamoeba histolytica. When the small subunit rRNA gene from Entamoeba was characterized, it was found to branch later than some mitochondria-containing protists (herolobosea and euglenozoa), which suggested that the ancestor of Entamoeba had a mitochondrion.12 Since then, the small subunit rRNA from Phreatamoeba balamuthi and large subunit rRNA from a species of Pelomyxa also have been sequenced. In phylogenetic trees these also branch within the mitochondria-containing eukaryotes,19,20 so these too appear to have arisen from mitochondria-containing ancestors. Furthermore, in the small subunit rRNA tree, Entamoeba and Phreatamoeba sometimes branch together,19 and sometimes do not,19 raising doubts as to the validity of the Archamoebae as an evolutionary lineage, let alone a primitively amitochondrial one. These phylogenetic arguments are not beyond reasonable doubt, however, because the Archamoebae often branch very near the origin of mitochondria, and other molecular trees disagree with rRNA on the order of these deep branching taxa.22

More substantial evidence came from a direct search of the Entamoeba genome for molecular relics of the mitochondrial symbiont. Most of the many hundreds of mitochondrial protein-coding genes are encoded in the nucleus and targeted to the organelle after they are translated in the cytoplasm. These genes were transferred from the symbiont genome to the nucleus but are recognizable today both because of this targeting and because the genes themselves closely resemble homologues from the type of bacteria from which the mitochondrion evolved, the alpha-proteobacteria.23 Finding such genes in the nucleus of an organism shows that ancestors of that lineage contained a mitochondrion, even if the organelle cannot otherwise be recognized today. This concept proved to be a breakthrough for testing the Archezoa and was first applied to the nuclear genome of Entamoeba histolytica. Clark and Roger24 found two genes of mitochondrial origin in Entamoeba, one for pyridine nucleotide transhydrogenase and another for a 60-kilodalton chaperonin (cpn60). In phylogenetic analyses of cpn60, the Entamoeba protein branched very strongly with homologues from the mitochondrion of other eukaryotes, which in turn were related to cpn60 proteins from alpha-proteobacteria.

These observations could be interpreted as evidence for a lateral transfer or symbiosis involving some other eubacteria25,26 but to do so would require a number of lateral
transfers or symbioses involving alpha-proteobacteria that happened to be very closely related to the mitochondrial symbiont. The simplest explanation is that *Entamoeba* evolved from mitochondrial-containing ancestors, and this may be said with some confidence because the *Entamoeba* chaperonin not only conforms to the phylogenetic expectations of a mitochondrial-derived protein but also is specifically related to eukaryotic homologues that are targeted to the mitochondrion.

**PARABASALIA AND THE HYDROGENOSOME**

Although they lack classical mitochondria, the Parabasalia do contain a double membrane-bound metabolic organelle called the hydrogenosome whose origin has been the source of some debate. Parabasalian hydrogenosomes do not resemble mitochondria in morphology, they do not appear to contain a genome, and unlike oxidative phosphorylation in the mitochondrion, energy is released in the hydrogenosome from the conversion of pyruvate or malate into acetate, carbon dioxide, and hydrogen gas (see Ref. 27 for review). However, despite these differences, there are reasons to suspect that parabasalian hydrogenosomes may share common ancestry with mitochondria. First, hydrogenosomes are not restricted to Parabasalia but are found in isolated members of several unrelated eukaryotic lineages (percolozoa, ciliates, and chytrid fungi), and these organisms invariably lack mitochondria. Moreover, some of these hydrogenosomes do resemble mitochondria morphologically. The mutual exclusion of mitochondria and hydrogenosomes throughout eukaryotes led to the suggestion that mitochondria may have turned into hydrogenosomes in these organisms, and for this reason, the Parabasalia were removed from Archezoa. However, in some of these organisms the hydrogenosome also has been argued to have evolved from peroxisomes, and considering the differences between hydrogenosomes and mitochondria, their mutual exclusion is not in itself sufficient evidence that they are of common ancestry. Without a direct link between the mitochondrial and hydrogenosome, the question remains open (Fig. 3).

For the Parabasalia, such a link has now been clearly established by the application of the same strategy that was so successful in *Entamoeba* to the parabasalian pathogen, *Trichomonas vaginalis*. Chaperonin genes, cpn10, cpn60, and cpn70, were sought and found in the genome of *Trichomonas*. In each case these genes were found to be specifically related to mitochondrial homologues. In addition, *T. vaginalis* cpn60 and cpn70 antibodies were shown to cross-react specifically with the purified hydrogenosomes, and bacterial cpn60 antibodies have been localized specifically to the hydrogenosome by in situ immunoelectron microscopy. These chaperonins are apparently localized in the hydrogenosome and derived from the same endosymbiont as the mitochondrion, creating a compelling argument that the parabasalian hydrogenosome and the mitochondrion descended from a common ancestor.

How, then, are the hydrogenosome and mitochondrion related? In those ciliates and fungi where hydrogenosomes evolved from mitochondria, they must have evolved from a highly specialized organelle because mitochondria were well developed in the ancestors of these groups. But this is not necessarily the case in Parabasalia; mitochondria are not found in any lineage known to predate Parabasalia, so all that can be said is that their hydrogenosome evolved from the same symbiont. Whether that symbiont was anything like what we would call a mitochondrion is not certain. If the ancestral parabasalian had a “proper” mitochondrion, then the transformation to the hydrogenosome may have occurred as it did in ciliates and fungi. This process would entail the loss of the critical metabolic enzymes found in mitochondria and the conscription of the nonmitochondrial enzymes hydrogenase and pyruvate:ferredoxin oxidoreductase from some other source. The loss of the mitochondrial genome would be expected to occur when the enzymes that it encoded were no longer used in the organelle. Alternatively, if the hydrogenosome evolved from a not-yet-specialized symbiont, then it merely followed a different evolutionary trajectory than did the mitochondrion, and this would reflect the likelihood that the symbiont had a greater metabolic diversity than contemporary mitochondria. Indeed, this also would bring into question the widely held assumption that the original reason the symbiont was retained was for the reactions that the mitochondrion now performs.

**MICROSPORIDIA: ARCHEZOA, PROTISTS, OR FUNGI?**

Microsporidia are obligate intracellular parasites that share a complex and unique infection strategy. Outside their host
cells, Microsporidia can only survive as spores with a tough double coat of chitin and protein. Inside the cytoplasm is a tightly wound projectile known as the polar tube. This organelle could be thought of as a hybrid between a harpoon and a hypodermic needle; when a spore encounters a susceptible host, it rapidly everts the polar tube, which then penetrates the host membrane. The infectious cytoplasm is squeezed through the polar tube (which is typically only approximately 0.1 µm in diameter) and is injected directly into the host cytoplasm where it lives as an amoeba, dividing and producing more spores (see Ref. 38 and references therein for details) (Fig. 4).

Microsporidia are obviously accomplished parasites, but they stand out among eukaryotes in others ways too. Microsporidia have the smallest known nuclear genomes, in some instances smaller even than many bacterial genomes. Moreover, microsporidian ribosomes resemble prokaryotic ribosomes in that the sequence homologous to the 5.8S rRNA molecule is covalently linked with the 23S rRNA; in other eukaryotes this rRNA is a separate molecule. Because the fused 5.8S-23S rRNA is unique to Microsporidia and prokaryotes, it has been cited to support the notion that Microsporidia are the most primitive of all eukaryotes. However, the 5.8S gene is located immediately upstream of the large subunit gene in the rDNA operon of eukaryotes. When this operon is transcribed, these two species of rRNA are cleaved at specific processing sites. Microsporidian rRNA sequences are very odd in general and contain numerous deletions. If one of these deletions affected a single processing site in the ancestor of Microsporidia, it could easily have led to the reformation of fused 5.8S-23S rRNA.

All of these characteristics can be interpreted as evidence that Microsporidia are descendants of ancient or primitive eukaryotes, but they also could be the result of the highly adapted, parasitic lifestyle characteristic of the group. This consideration has always figured in speculation of their evolutionary origin, but even with such doubts, current arguments regarding the nature of their evolutionary history still come as a surprise. Evidence is now emerging that suggests that Microsporidia actually evolved recently from so-called “crown” eukaryotes (the twigs, such as animals, plants, and fungi) and may share a close ancestry with fungi. The strongest evidence supports a general relationship between Microsporidia and crown-taxa. First, EF-1α proteins of microsporidia, animals, and fungi all contain an insertion that is unique to these taxa. Similarly, the dihydrofolate reductase and thymidylate synthase are two separate enzymes in microsporidia, animals, and fungi but are fused in plants and other protists. A specific relationship with fungi was first proposed on the basis of parallels found between the unusual meiotic cycle of Microsporidia and that of certain fungi and also has now been supported by the phylogeny of both alpha- and beta-tubulins, which place Microsporidia within the fungal radiation.

The balance between evidence for the ancient or crown status of the Microsporidia was tipped in favor of the latter by the recent discovery of mitochondrial cpn70 genes in the genomes of Nosema locustae and Vairimorpha necatrix. In phylogenetic trees these sequences branch convincingly with mitochondrial homologues, undermining the argument that they are primitive descendants of amitochondrial eukaryotes. Interestingly, the microsporidian cpn70 genes also branch, albeit weakly, with mitochondrial genes from fungi.

So why would the Microsporidia branch at the base of the eukaryotes in phylogenetic trees if they actually arose from the crown of the tree? Highly divergent sequences will often fall in the wrong place in phylogenetic trees, and the rate of substitution in microsporidian genes is exceptionally high. Such genes will sometimes branch preferentially with other highly divergent sequences, and in the case of the eukaryotic tree, this means branching deeply. In fact, the divergence rate of most of the deep branching eukaryotes is relatively high, although to a much lesser extent than Microsporidia. It may be that some of these other taxa also are misplaced due to substitution rate, but there is no reason to dismiss the rRNA tree without evidence for some other relationship, such as we now have for Microsporidia.

**MITOCHONDRIAL RELICS IN METAMONADS?**

Most of what we know about the Metamonads is from one type, the diplomons. This is perhaps unfortunate because diplomons are probably the most highly derived Metamonads and, therefore, poorly represent the ancestral state of the group. Other Metamonads are flagellates with a single nucleus and four kinetosomes, one of which is recurrent and often has a flagellum associated with a feeding
organelle or cytostome. Most diplomonads are essentially two such cells fused back-to-back in axial symmetry. There are two nuclei associated with four kinetosomes each, and in heterotrophic species there are also two symmetrical cytostomes (see Refs. 49,50 for review) (Fig. 5).

Historically, there has been little reason to suspect Metamonads of ever having harbored a mitochondrion. They contain no mysterious organelle comparable with the parasibasal hydrogenosome and, unlike the Microsporidia, their molecular sequences are not so unusual that they stand out as obvious candidates for misplacement in phylogenetic trees. Indeed, molecular data have provided little reason to doubt the early divergence of these organisms; genes for transcription and translation genes commonly used for phylogeny consistently place diplomonads at the very base of eukaryotes.14,16,21

So what reasons are there to doubt the primitively amitochondrial nature of the Metamonads? There is an intriguing report that the diplomonad Giardia lamblia contains a protein that cross-reacts with antibodies against mitochondrial cpn60.51 However, in situ immunofluorescent labeling with this antibody is scattered throughout the cytoplasm and not concentrated in a particular part of the cell as one would expect of an organelar protein. One possibility is that this chaperonin is derived from the mitochondrion but has ascribed a cytosolic role. This would contrast with other chaperonin proteins in eukaryotes may be derived from the mitochondrial symbiont. These proteins are more related to homologues from proteobacteria (the closest relatives of the mitochondrion) than they are to homologues from archaebacteria (the closest relatives of the nuclear-cytosolic lineage). This relationship has been seen in both glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and triosephosphate isomerase (TPI).54 In both cases the genes from diplomonads are known, and they do not differ remarkably from those of other eukaryotes. If the source of either of these proteins is the mitochondrial symbiont, then it implies that diplomonads descend from a lineage that also contained the symbiont. Unfortunately, neither GAPDH nor TPI is totally unambiguous. The phylogeny of GAPDH is complex, with numerous paralogous gene families distributed in a pattern whose interpretation is not yet clear from the available data.53,55 TPI, on the other hand, does not contain enough phylogenetically useful information to discriminate between a specifically alpha-proteobacterial origin of eukaryotic TPI or simply the proteobacteria in general.54

The idea that the mitochondrial symbiont could have contributed more to the nucleus and cytosol than just mitochondrially targeted genes is an interesting assertion and may partially explain a few nagging problems in the phylogeny of several other proteins. There are a number of eukaryotic proteins that appear to be closer to proteobacterial homologues than expected or at least closer to eubacteria than to archaebacteria. These have generally been interpreted as the result of lateral transfers or even an ancient cellular fusion event, or chimerism.26,56 Alternatively, some of these incongruencies may be the result of genes derived from the mitochondrial symbiont acquiring a role in cytoplasmic metabolism (variations on this theme are discussed in Ref. 57). Unfortunately, for many of these genes there is either insufficient information in the sequences or inadequate taxon sampling to make any believable conclusions as to their origin. Moreover, because there are no mitochondrially targeted homologues of any of the genes for which this has been proposed (including TPI or GAPDH), their evolutionary origin in eukaryotes will never be as clear-cut as proteins, such as cprn60, which have a functional link to the organelle in mitochondrion-containing eukaryotes. This also means that, even if these genes are derived from the mitochondrial symbiont, their presence is not evidence for the presence of the organelle. With no functional link between the protein and the organelle, the organelle could be

Figure 5. A metamonad, Retortamonas.
lost without affecting the protein, and similarly, the protein could be the result of a transfer that took place when the symbiosis was still transient.54

CONCLUSIONS: THE RIGHT ANCESTORS FOR THE WRONG REASONS

The purpose for creating the Archezoa was that it united primitively amitochondrial eukaryotes. This is not fulfilled for Parabasalia, Microsporidia, or Archamoebae, and there are now growing doubts for the Metamonads as well.

However, the kingdom Archezoa also was proposed explicitly as a “phylogenetic hypothesis”9 intended to draw attention to these organisms as putative descendants of early eukaryotes. In this regard it has been an outstanding success because two of the original Archezoa, Parabasalia, and Metamonads are shown still branching deeply among eukaryotes but after the mitochondrial symbiosis. This can be said with some confidence in the case of Parabasalia, but there is only preliminary evidence to support this conclusion for Metamonads (hence the lingering question mark).

Figure 6. The Archezoa today. The color scheme is as in Figure 1. The Archamoebae and Microsporidia are shown branching much later than previously proposed, and the Parabasalia and Metamonads are shown still branching deeply among eukaryotes but after the mitochondrial symbiosis. This can be said with some confidence in the case of Parabasalia, but there is only preliminary evidence to support this conclusion for Metamonads (hence the lingering question mark).

strongly allied with other Parabasalia, but these relationships have yet to be tested with molecular data. Moreover, some other relationships are not so obvious. Oxymonads, for instance, are only tenuously classified with Metamonads and have characteristics, such as meiosis and a sexual cycle60 that make them stand out among Metamonads. It is distinctly possible that Oxymonads are not closely related to diplomonads or retortamonads at all. These “details” are central to our understanding of the nature of these groups, and the nature of the ancestral eukaryote.

Of course, there is also hope that new groups and deeper branches will be revealed to us in the future as organisms known only from morphology or some never seen before are characterized or identified. The molecular approach to biological diversity has greatly increased our understanding of diversity and the ancestral state of eubacteria and especially archaeabacteria.25,61 The same has yet to be applied to eukaryotes on as grand a scale, but when it is, it will hopefully yield as many surprises and confirmations as it has in the prokaryotes. Even if none of the eukaryotes we know today evolved before the acquisition of the mitochondrion, we might still find an archezoan somewhere in the long branch between archaeabacteria and eukaryotes.

ACKNOWLEDGMENTS

I thank A.J. Roger, W.F. Doolittle, M. Müller, G.C. Clark, G. Brugerolle, and T. Cavalier-Smith for past discussions and lessons on this subject, and G.I. McFadden and N.M. Fast for critical reading of the manuscript. Retortamonas and Tritrichomonas line drawings were based on reconstructions and diagrams in References 58,59.

NOTE ADDED IN PROOF


REFERENCES

Problems and paradigms


