

The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa

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

T. Cavalier-Smith

Tel: +44 1865 281065. Fax: +44 1865 281310. e-mail: tom.cavalier-smith@zoo.ox.ac.uk

Eukaryotes and archaeobacteria form the clade neomura and are sisters, as shown decisively by genes fragmented only in archaeobacteria and by many sequence trees. This sisterhood refutes all theories that eukaryotes originated by merging an archaeobacterium and an α -proteobacterium, which also fail to account for numerous features shared specifically by eukaryotes and actinobacteria. I revise the phagotrophy theory of eukaryote origins by arguing that the essentially autogenous origins of most eukaryotic cell properties (phagotrophy, endomembrane system including peroxisomes, cytoskeleton, nucleus, mitosis and sex) partially overlapped and were synergistic with the symbiogenetic origin of mitochondria from an α -proteobacterium. These radical innovations occurred in a derivative of the neomuran common ancestor, which itself had evolved immediately prior to the divergence of eukaryotes and archaeobacteria by drastic alterations to its eubacterial ancestor, an actinobacterial posibacterium able to make sterols, by replacing murein peptidoglycan by *N*-linked glycoproteins and a multitude of other shared neomuran novelties. The conversion of the rigid neomuran wall into a flexible surface coat and the associated origin of phagotrophy were instrumental in the evolution of the endomembrane system, cytoskeleton, nuclear organization and division and sexual life-cycles. Cilia evolved not by symbiogenesis but by autogenous specialization of the cytoskeleton. I argue that the ancestral eukaryote was uniciliate with a single centriole (unikont) and a simple centrosomal cone of microtubules, as in the aerobic amoebozoan zooflagellate *Phalansterium*. I infer the root of the eukaryote tree at the divergence between opisthokonts (animals, Choanozoa, fungi) with a single posterior cilium and all other eukaryotes, designated 'anterokonts' because of the ancestral presence of an anterior cilium. Anterokonts comprise the Amoebozoa, which may be ancestrally unikont, and a vast ancestrally biciliate clade, named 'bikonts'. The apparently conflicting rRNA and protein trees can be reconciled with each other and this ultrastructural interpretation if long-branch distortions, some mechanistically explicable, are allowed for. Bikonts comprise two groups: corticoflagellates, with a younger anterior cilium, no centrosomal cone and ancestrally a semi-rigid cell cortex with a microtubular band on either side of the posterior mature centriole; and Rhizaria [a new infrakingdom comprising Cercozoa (now including Ascetosporea classis nov.), Retaria phylum nov., Heliozoa and Apusozoa phylum nov.], having a centrosomal cone or radiating microtubules and two microtubular roots and a soft surface, frequently with reticulopodia. Corticoflagellates comprise photokaryotes (Plantae and chromalveolates, both ancestrally with cortical alveoli) and Excavata (a new protozoan infrakingdom comprising Loukozoa, Discicristata and Archezoa, ancestrally with three microtubular roots). All basal

This paper is an elaboration of part of an invited presentation to the XIIIth meeting of the International Society for Evolutionary Protistology in České Budějovice, Czech Republic, 31 July–4 August 2000.

Two notes added in proof are available as supplementary materials in IJSEM Online (<http://ijs.sgmjournals.org/>).

Abbreviations: snoRNP, small nucleolar ribonucleoprotein; SRP, signal recognition particle.

eukaryotic radiations were of mitochondrial aerobes; hydrogenosomes evolved polyphyletically from mitochondria long afterwards, the persistence of their double envelope long after their genomes disappeared being a striking instance of membrane heredity. I discuss the relationship between the 13 protozoan phyla recognized here and revise higher protozoan classification by updating as subkingdoms Lankester's 1878 division of Protozoa into Corticata (Excavata, Alveolata; with prominent cortical microtubules and ancestrally localized cytostome – the Parabasalia probably secondarily internalized the cytoskeleton) and Gymnomyxa [infrakingdoms Sarcomastigota (Choanozoa, Amoebozoa) and Rhizaria; both ancestrally with a non-cortical cytoskeleton of radiating singlet microtubules and a relatively soft cell surface with diffused feeding]. As the eukaryote root almost certainly lies within Gymnomyxa, probably among the Sarcomastigota, Corticata are derived. Following the single symbiogenetic origin of chloroplasts in a corticoflagellate host with cortical alveoli, this ancestral plant radiated rapidly into glaucophytes, green plants and red algae. Secondary symbiogeneses subsequently transferred plastids laterally into different hosts, making yet more complex cell chimaeras – probably only thrice: from a red alga to the corticoflagellate ancestor of chromalveolates (Chromista plus Alveolata), from green algae to a secondarily uniciliate cercozoan to form chlorarachneans and independently to a biciliate excavate to yield photosynthetic euglenoids. Tertiary symbiogenesis involving eukaryotic algal symbionts replaced peridinin-containing plastids in two or three dinoflagellate lineages, but yielded no major novel groups. The origin and well-resolved primary bifurcation of eukaryotes probably occurred in the Cryogenian Period, about 850 million years ago, much more recently than suggested by unwarranted backward extrapolations of molecular 'clocks' or dubious interpretations as 'eukaryotic' of earlier large microbial fossils or still more ancient steranes. The origin of chloroplasts and the symbiogenetic incorporation of a red alga into a corticoflagellate to create chromalveolates may both have occurred in a big bang after the Varangerian snowball Earth melted about 580 million years ago, thereby stimulating the ensuing Cambrian explosion of animals and protists in the form of simultaneous, poorly resolved opisthokont and anterokont radiations.

Keywords: Corticata, Rhizaria, Excavata, centriolar roots of bikonts, Amoebozoa and opisthokonts, symbiogenetic origin of mitochondria

Introduction: revising the neomuran theory of the origin of eukaryotic cells

In 1987, I published seven papers that together developed an integrated view of cell evolution, ranging from the origin of the first bacterium, a Gram-negative eubacterium (negibacterium; Cavalier-Smith, 1987a), and the nature of bacterial DNA segregation (Cavalier-Smith, 1987b), through the origins of archaeobacteria (Cavalier-Smith, 1987c) and eukaryotes (Cavalier-Smith, 1987c, d) and the symbiogenetic origins of mitochondria and chloroplasts and their secondary lateral transfers (Cavalier-Smith, 1987e) to the origins and diversification of plant (Cavalier-Smith, 1987f) and animal and fungal cells (Cavalier-Smith, 1987g). Central to those publications was the then novel view that eukaryotes were sisters to archaeobacteria and that both diverged from a common ancestor that itself arose by the drastic evolutionary transformation of a

Gram-positive eubacterium. I argued that the most important change in this radical transformation of a eubacterium into the common ancestor of eukaryotes and archaeobacteria was the replacement of the eubacterial peptidoglycan murein by *N*-linked glycoprotein and that the concomitant changes in the replication, transcription and translation machinery were of comparatively trivial evolutionary significance. I therefore called the postulated clade comprising archaeobacteria and eukaryotes 'neomura' (meaning 'new walls') and will continue this usage here. I further argued that the archaeobacteria, though becoming adapted to hot, acid environments by replacing the eubacterial acyl ester lipids by prenyl ether lipids, used the new glycoproteins as a rigid cell wall (exoskeleton) and therefore retained the ancestral bacterial cell division and generally bacterial cell and chromosomal organization. Though chemically radically changed, they remained prokaryotic because, like other bacteria, they retained an

exoskeleton, which prevented more radical innovation. Their pre-eukaryote sisters, in striking contrast, by using the new glycoproteins to develop a more flexible surface coat, were able to evolve phagocytosis for the first time in the history of life, which necessarily led to the rapid origin of the eukaryotic cytoskeleton, mitosis and endomembrane system and ultimately the nucleus, cilia and sex. Phagotrophy was also the prerequisite for the uptake of the symbiotic eubacterial ancestors of mitochondria and chloroplasts. I also interpreted the fossil record as showing that eukaryotes were less than half as old as eubacteria and emphasized that, according to the neomuran theory, archaeobacteria must be equally young and not a primordial group as has often been supposed to be the case. A shared neomuran character that I particularly stressed was the co-translational *N*-linked glycosylation of cell-surface proteins, which offered special insights into the origin of the eukaryotic endomembrane system (Cavalier-Smith, 1987c).

The present paper revises this neomuran theory of the origin of the eukaryotic cell, re-emphasizing the role of phagotrophy in the origin of eukaryotes (Cavalier-Smith, 1975, 1987c), in the light of recent substantial phylogenetic advances, notably the evidence that mitochondria were already present in the common ancestor of all extant eukaryotes (Cavalier-Smith, 1998a, 2000a; Embley & Hirt, 1998; Gupta, 1998a; Keeling, 1998; Keeling & McFadden, 1998; Roger, 1999) and that all anaerobic eukaryotes have evolved by the loss of mitochondria or their conversion into hydrogenosomes (Cavalier-Smith, 1987e; Müller & Martin, 1999). My detailed explanations of the origins of the 18 suites of shared neomuran characters (many more than were apparent in 1987) and of the many fewer unique archaeobacterial characters have been presented in a separate paper (Cavalier-Smith, 2002a). As that paper also discusses the palaeontological evidence for the origin of neomura – especially its remarkable recency – and the widespread misinterpretations of evolutionary artefacts in molecular trees, which have hindered our understanding of the proper positions of their roots, it provides an essential background and broader context to the present one, which focuses specifically on the origin and early diversification of the eukaryotic cell. Note that I always use ‘bacterium’ in its proper historical sense as a synonym for ‘prokaryote’, never as a synonym for eubacteria alone (Woese *et al.*, 1990), a thoroughly confusing, highly undesirable and entirely unnecessary change to established usage (Cavalier-Smith, 1992a), which I urge others also to eschew.

If there really are no primitively amitochondrial eukaryotes (Cavalier-Smith, 1998a, 2000a), the simplest explanation of the great mixture of genes of archaeobacterial and negibacterial character in eukaryotes (Golding & Gupta, 1995; Brown & Doolittle, 1997; Ribeiro & Golding, 1998) is that the negibacterial genes originated from the α -proteobacterium that evolved into the first mitochondrion and that the

archaeobacterial-like genes were derived from the host (Cavalier-Smith & Chao, 1996; Cavalier-Smith, 2002a). Postulating a fusion or symbiogenesis between an archaeobacterium and a negibacterium prior to the origin of mitochondria (Zillig *et al.*, 1989; Golding & Gupta, 1995; Gupta & Golding, 1996; Lake & Rivera, 1994; Margulis *et al.*, 2000; Moreira & López-García, 1998) is entirely unnecessary if the establishment of mitochondria, the endomembrane system and the eukaryotic cytoskeleton were virtually contemporaneous, as I argue here. I maintain that the origin of phagotrophy was the essential prerequisite for all three, for the reasons given in my first discussion of the origin of the nucleus (Cavalier-Smith, 1975). What changed markedly between 1975 and 1987, and less radically since, is the phylogenetic context of these fundamental mechanistic changes.

Prior to my proposal that eukaryotes and archaeobacteria are sisters (Cavalier-Smith, 1987c), it had been argued in three seminal papers that archaeobacteria were ancestral to eukaryotes (Van Valen & Maiorana, 1980; Searcy *et al.*, 1981; Zillig *et al.*, 1985). I opposed that view primarily because it entailed an independent origin of acyl ester lipids in eubacteria and eukaryotes. One could readily explain the changeover from eubacterial acyl esters to archaeobacterial isoprenoid ether lipids as a secondary adaptation for hyperthermophily (Cavalier-Smith, 1987a, c), an explanation that remains valid today (Cavalier-Smith, 2002a). Together with palaeontological evidence for very early eubacterial photosynthesis, it was and is a primary reason for arguing that eubacteria are the paraphyletic ancestors of archaeobacteria, not the reverse (Cavalier-Smith, 1987a, 1991a, b). But, if eukaryotes had evolved substantially before mitochondria, as suggested earlier (Cavalier-Smith, 1983a, b) and rRNA (Vossbrinck & Woese, 1986; Vossbrinck *et al.*, 1987) tempted us to believe (Cavalier-Smith, 1987d), there was no obvious reason why the first eukaryote should have switched from archaeobacterial lipids to acyl esters and no obvious means of doing so; postulating that archaeobacteria were sisters rather than ancestors of eukaryotes seemed the obvious solution, since one could argue that their remarkable shared characters all evolved in their immediate common ancestor. If, as now appears to be the case, the ancestral eukaryote was aerobic and had acquired mitochondria prior to its diversification into any extant lineages, this part of my 1987 argument becomes somewhat less compelling. In principle, it would have been possible for the host to have been an archaeobacterium and for the prenyl ether lipids to have been replaced by acyl ester lipids derived from the α -proteobacterial ancestor of mitochondria, as Martin (1999) suggested. Such wholesale lipid replacement, if it had occurred, would have been a remarkably complex and improbable evolutionary phenomenon, but it could not be dismissed as altogether impossible. However, recent molecular-phylogenetic evidence (Cavalier-Smith, 2002a), also briefly discussed below, now clearly refutes the idea that

archaeobacteria are the paraphyletic ancestors of eukaryotes and firmly establishes their holophyly. Thus, lipid replacement did not occur during the origin of eukaryotes: essentially all their lipids, including both acyl ester phospholipids and sterols, were probably already present in their actinobacterial ancestor.

The present phagotrophy theory of the essentially simultaneous origin of eukaryotes and mitochondria leaves unchanged most details of the actual transition postulated by the earlier phagotrophy theory of the serial origin of eukaryotes and mitochondria (Cavalier-Smith, 1987c, e) including their mechanisms and selective advantages, but telescopes them into a single geological period, thus allowing synergy between them and thereby strengthening the overall thesis. In addition to discussing the origin of eukaryotes, I evaluate new evidence bearing on the position of the root of the eukaryote tree and conclude that it lies among aerobic amoeboflagellates having mitochondria, not among the anaerobic amitochondrial flagellates as was thought until recently. Thus, what is most significantly changed in the revised phagotrophy theory is the nature of the first eukaryote – a chimaeric aerobic unikont flagellate resembling the zooflagellate *Phalansterium*, instead of a non-chimaeric anaerobic unikont flagellate like the pelobiont amoebozoan *Mastigamoeba*, as was proposed earlier (Cavalier-Smith, 1991c). Actually, in cell structure, *Phalansterium* and *Mastigamoeba* have a lot in common and are probably not cladistically widely separated (Cavalier-Smith, 2000a); indeed, as my own unpublished gamma-corrected distance trees actually place *Phalansterium* within the Amoebozoa, as sister to the Acanthamoebidae, I here transfer *Phalansterium* into a revised Amoebozoa. Thus, rooting a morphologically based eukaryote tree on either *Phalansterium* or mastigamoebids would give a very similar tree, broadly consistent with many protein trees but contradicting the rooting (but little of the topology) of rRNA trees. Now, however, with more balanced views of the strengths and weaknesses of molecular trees (Philippe & Adoutte, 1998; Embley & Hirt, 1998; Philippe *et al.*, 2000; Roger, 1999; Stiller & Hall, 1999) in the ascendant, there should be less pressure on us to accept every detail of every rRNA tree as gospel truth. If we also develop a healthier balance between the molecular and the cell-biological or ultrastructural evidence, we can use the latter to help us decide which of the conflicting molecular trees are more reliable (Cavalier-Smith, 1981a, 1995a; Taylor, 1999). I shall argue that, although the Amoebozoa are probably very close to the base of the eukaryotic tree, extant or ‘crown’ Amoebozoa may actually be holophyletic and that there are probably no extant eukaryotic groups that diverged prior to the fundamental bifurcation between the protozoan ancestors of animals and plants.

New phylogenetic insights, especially those concerning ciliary root evolution discussed here, lead me to revise the higher classification of the kingdom Protozoa in four main ways. Firstly, I place the secondarily

amitochondrial Archezoa (phyla Metamonada and Parabasalia), from which I now exclude oxymonads, as a superphylum within a new infrakingdom Excavata. Excavata also includes Discicristata (phyla Euglenozoa and Percolozoa) and the recently established phylum Loukozoa (Cavalier-Smith, 1999; here augmented by the Oxymonadida and Diphylleida) and is almost certainly a derived group, not an early branching one, as was previously widely believed. Secondly, I group infrakingdoms Excavata and Alveolata together as subkingdom Corticata (from which the kingdoms Plantae and Chromista are derived). Thirdly, I establish a new subkingdom Gymnomyxa to embrace the majority of the former sarcodine protozoa (i.e. virtually all except the Heterolobosea, which are percolozoa of obvious corticate ancestry) and a variety of soft-surfaced zooflagellates with pronounced pseudopodial tendencies within the phyla Cercozoa (into which I transfer the parasitic Ascetosporea and the pseudociliate *Stephanopogon*), Choanozoa (the closest protozoan relatives of kingdoms Fungi and Animalia; Cavalier-Smith, 1998b), both of which also contain a few former sarcodines, and Apusozoa. Finally, the classification of Gymnomyxa is rationalized by establishing a new infrakingdom Rhizaria that groups the phyla (Cercozoa and Retaria) in which reticulopodia are widespread with the axopodial Heliozoa and the Apusozoa, zooflagellates that often have ventral branched pseudopods. I shall present evidence that the Corticata are derived from Gymnomyxa and are cladistically closer to the Rhizaria than to the other gymnomyxan infrakingdom, Sarcomastigota (Amoebozoa, Choanozoa), within which the eukaryote tree is probably rooted.

Intracellular co-evolution: the key to understanding the origin of the eukaryote cell

To understand cell evolution, we must consider evenly the evolution of three things: genomes, membranes and cytoskeletons (Cavalier-Smith, 2001, 2002a). Though I shall discuss aspects of genome and metabolic evolution, I will emphasize the primary importance of membranes and the cell skeleton as providing the environment within which genomes evolve (thus profoundly determining the selective forces acting on them) and their very *raison d'être*.

Recent genome sequencing has fostered a simplistic view of organisms as essentially random aggregates of genes and proteins, a molecular-genetic bias worse than the earlier biochemists' oversimplification of them as a ‘mere bag of enzymes’; it forgets both the bag and the skeleton that gives it form and the ability to divide and evolve complexity! Molecular cell biology has taught us that organisms arise only by the co-operation of genes, catalysts, membranes and a cell skeleton (Cavalier-Smith, 1987a, 1991a, b, 2001). The most fundamental events in converting a bacterium into a eukaryote were not generalized changes in the genome or modes of gene expression, but two pro-

found cellular changes: (i) a radical change in membrane topology associated with the origin of coated-vesicle budding and fusion and nuclear pore complexes and (ii) a changeover from a relatively passive exoskeleton (the bacterial cell wall) to an endoskeleton of microtubules and microfilaments associated with the molecular motors dynein, kinesin and myosin. Paradoxically, these innovations fed back onto the genome itself, endowing eukaryotic DNA with a novel function as a nuclear skeleton; viral and bacterial chromosomes are indeed essentially aggregates of genes, but eukaryotic DNA (most of the DNA in the biosphere) is primarily a skeletal polyelectrolyte gel in which genes are only sparsely embedded (Cavalier-Smith & Beaton, 1999).

But I must not oversimplify. The origin of the eukaryote cell was the most complex transformation and elaborate example of quantum evolution (Simpson, 1944) in the history of life. Thousands of DNA mutations caused ten major suites of innovations: (i) origin of the endomembrane system (ER, Golgi and lysosomes) and coated-vesicle budding and fusion, including endocytosis and exocytosis; (ii) origin of the cytoskeleton, centrioles, cilia and associated molecular motors; (iii) origin of the nucleus, nuclear pore complex and trans-envelope protein and RNA transport; (iv) origin of linear chromosomes with plural replicons, centromeres and telomeres; (v) origin of novel cell-cycle controls and mitotic segregation; (vi) origin of sex (syngamy, nuclear fusion and meiosis); (vii) origin of peroxisomes; (viii) novel patterns of rRNA processing using small nucleolar ribonucleoproteins (snoRNPs); (ix) origin of mitochondria; and (x) origin of spliceosomal introns. Each of these ten novelties is so complex that it needs its own long review for discussion in the depth that present knowledge of cell biology requires. Here, I concentrate instead on placing them in phylogenetic context and re-emphasizing the co-evolutionary interconnections between them. I argued previously that the first six innovations arose simultaneously in response to the loss of the bacterial cell wall and the origin of phagotrophy (Cavalier-Smith, 1987c). I argued that each affected the others and that such intraorganismic molecular co-evolution made it counterproductive to attempt to understand the evolution of any one of them in isolation.

The present paper extends the thesis of intracellular co-evolution during the origin of eukaryotes by arguing that the last four innovations also accompanied the others and fed back on some of them (the eighth innovation at least partially preceded the origin of eukaryotes, at least one of the two types of snoRNPs having evolved in the ancestral neomuran; Cavalier-Smith, 2002a). It also uses advances in understanding ciliary transformation (Moestrup, 2000) to resolve conflicts between molecular trees and establish with reasonable confidence the approximate location of the eukaryote root and thus the properties of their cenancestor (latest common ancestor; Fitch & Upper,

1987). The structural evolution of cilia, centrioles and ciliary roots is more central to, and reveals much more about, eukaryote evolution than the sequence-dominated community generally realizes.

The unbacterial relatives of eukaryotes

Eukaryotes are sisters of archaeobacteria not derived from them

Deciding whether archaeobacteria are sisters of eukaryotes or actually ancestral to them is very important for knowing the origin of certain eukaryote genes. Over 40 genes are found in eukaryotes and eubacteria but not apparently in archaeobacteria, e.g. Hsp90 (Gupta, 1998a; note that not every gene listed there is truly absent from archaeobacteria – catalase is actually present). If eukaryotes and archaeobacteria are sisters (Cavalier-Smith, 1987c, 2002a; Pace *et al.*, 1996), these genes could all have been lost in the common ancestor of archaeobacteria, but retained by eukaryotes by vertical inheritance from their neomuran common ancestor. On the other hand, if eukaryotes branched within a paraphyletic archaeobacteria, as Gupta (1998a) and Martin & Müller (1998) assume, we should reasonably conclude instead that these genes came from a eubacterium by lateral transfer, possibly simply donated by the proteobacterial ancestor of mitochondria.

Van Valen & Maiorana (1980) proposed that eukaryotes (i.e. the host that enslaved a proteobacterium to make a mitochondrion) evolved from archaeobacteria, whereas, for the reasons stated above, I (Cavalier-Smith, 1987c) postulated instead that archaeobacteria are sisters of eukaryotes, but their common ancestor was neither: it was instead a radically changed derivative of a posibacterial mutant, which I shall refer to as a stem neomuran, i.e. one branching prior to the neomuran cenancestor. If eukaryotes evolved from a crown archaeobacterium (any descendant of the archaeobacterial cenancestor), then they should nest in trees (!) within archaeobacteria; but if they evolved from stem archaeobacteria (i.e. any before the archaeobacterial cenancestor), the sequence trees could not distinguish the two theories as they do not show the phenotype of their common ancestor. This unavoidable cladistic ambiguity of sequence trees, coupled with the fact that existing trees are contradictory, is why neither theory has been unambiguously refuted. Of 32 gene sequence trees showing archaeobacterial and eukaryotic genes as related in a recent analysis and including more than one archaeobacterium (Brown & Doolittle, 1997), 19 actually portray a sister relationship, as I predicted, and only 15 an ancestor–descendant one. However, to say that this slight numerical advantage supports my sister theory would be naïve, for two reasons. Firstly, nine of the trees did not include both euryarchaeotes and crenarchaeotes (also known as eocytes), so would be less likely to show archaeobacterial paraphyly, just because of poor taxon sampling. Perhaps more importantly, quantum evol-

ution during early eukaryote evolution would be likely to make their genes more distant from archaeobacterial ones than expected on a molecular clock. In practice, this means that genuine archaeobacterial paraphyly would often be converted by the consequent long-branch phylogenetic artefact into apparent holophyly. In most cases, bootstrap support for archaeobacterial holophyly was low: in one case (vacuolar ATPase), other authors have shown trees with eukaryotes within archaeobacteria (as sisters to euryarchaeotes; Gogarten & Kibak, 1992).

As an aside, I must point out that the cladistic terms 'stem' and 'crown' were invented and defined as in the preceding paragraph by the palaeontologist Jefferies (1979), but were used incorrectly by Knoll (1992): the phrase 'crown eukaryotes', therefore, properly includes all extant eukaryotes, not just those with short branches on rRNA trees (which are not a holophyletic group): the latter misuse, ignorantly adopted by GenBank and many others, should be discontinued so as not to destroy the utility of the distinction made by Jefferies, which is fundamentally important for phylogenetic discussion (Cavalier-Smith, 2002a).

The presence of an insertion of seven amino acids in EF-1 α shared by eukaryotes and crenarchaeotes alone has been used to argue that eukaryotes are more closely related to them than to euryarchaeotes (Rivera & Lake, 1992; Gupta, 1998a; Baldauf *et al.*, 1996). But this is very weak evidence, since this insertion could have been present in the ancestor of all archaeobacteria and deleted in euryarchaeotes – in fact, some have two or three amino acids here, showing that the region has undergone differential deletions or insertions within euryarchaeotes. Trees for the same gene sometimes do weakly depict eukaryotes within archaeobacteria (Baldauf *et al.*, 1996), as does a large-subunit rRNA tree from an unusually sophisticated maximum-likelihood analysis (Galtier *et al.*, 1999), which should be superior to the usual small-subunit trees that show them as sisters (Kyrpides & Olsen, 1999). However, of the 14 trees showing archaeobacterial paraphyly, only five (including EF-1 α) placed them as sisters to eocytes; only one did so as sisters to euryarchaeotes, but nine place them within euryarchaeotes.

If archaeobacteria were paraphyletic, the presence of genuine histones in euryarchaeotes but never crenarchaeotes or eubacteria (Reeve *et al.*, 1997) would favour a euryarchaeote not a crenarchaeote ancestor, as postulated by Moreira & López-García (1998), Martin & Müller (1998) and Sandman & Reeve (1998). However, this also is relatively weak evidence, as histones can be lost secondarily, as we know for dinoflagellates; indeed, I have argued elsewhere (Cavalier-Smith, 2002a) that they did evolve in the ancestral archaeobacterium. From their distribution among euryarchaeotes (Moreira & López-García, 1998), we can conclude that histones were present in the cenancestor of euryarchaeotes but were lost by *Thermoplasma*. The absence of histones in *Thermoplasma* rules out the idea

of Margulis *et al.* (2000) that the ancestor of eukaryotes was a *Thermoplasma*-like cell; *Thermoplasma* is highly derived and also too genomically and cytologically reduced in other ways to be a serious candidate. Margulis *et al.* (2000) were mistaken in calling the *Thermoplasma* basic protein 'histone-like' (it is actually more like the non-histone DNA-binding proteins of eubacteria) and in calling *Thermoplasma* an eocyte; it is not a crenarchaeote but is nested well within the euryarchaeotes. As its sister genus *Picrophilus* has a glycoprotein wall, *Thermoplasma* probably lost its wall hundreds of millions years after the origin of eukaryotes (Cavalier-Smith, 2002a). The posibacterial mycoplasmas, which Margulis (1970) originally favoured as a host, almost certainly lost their walls and suffered massive genomic reduction after their endobacterial ancestors (Cavalier-Smith, 2002a) became obligate parasites of pre-existing eukaryotic cells. Thus, no extant wall-free bacteria are suitable models for the ancestor of eukaryotes; all are too greatly reduced and none is specifically phylogenetically related to us. As *Thermoplasma* has lost histones, the crenarchaeote cenancestor might also have done so, in which case histones would have evolved in a stem neomuran, as I have argued (Cavalier-Smith, 2002a). A crenarchaeote (eocyte) ancestry for eukaryotes would only be possible if there had been multiple losses of histones within crenarchaeotes after the origin of eukaryotes and so is less plausible than a euryarchaeote ancestry. Some crenarchaeotes (*Sulfolobus*) have CCT chaperonins with nine rather than the usual eight subunits (Archibald *et al.*, 1999) so can be ruled out as potential eukaryotic ancestors.

The most decisive evidence for a sister relationship between eukaryotes and archaeobacteria (Cavalier-Smith, 1987c) is the fragmentation of two unrelated genes so as to encode more than one distinct protein in all archaeobacteria but in no eukaryotes or eubacteria: RNA polymerase RpoA became divided into two (Klenk *et al.*, 1999) and glutamate synthetase into three separate genes (Nesbø *et al.*, 2001). This clearly refutes all recent syntrophy theories (Martin & Müller, 1998; Moreira & López-García, 1998; Margulis *et al.*, 2000) and any other theory that, like them, assumes that an archaeobacterium was directly ancestral to eukaryotes (e.g. those of Gupta, 1998a; Lake & Rivera, 1994; Sogin, 1991). On such theories, the five fragmented RNA polymerase A and glutamate synthetase genes would together have had to undergo three refusion events to make two single genes in the common ancestor of eukaryotes, highly improbable and selectively dubious reversals. Such theories would also require the complete replacement of the host isoprenoid lipids by acyl ester lipids derived from the mitochondrion. The theory that archaeobacteria and eukaryotes are sisters diverging from a neomuran common ancestor (Cavalier-Smith, 1987c) avoids postulating this complex lipid replacement or the refusion of these genes and is thus much more parsimonious and almost certainly correct. The neomuran theory is

also much simpler, in that no special explanation is needed of how eukaryotes acquired the numerous genes such as *hsp90* that are absent from archaeobacteria (Gupta, 1998a); they were simply present in the neomuran ancestor, through vertical descent from its posibacterial ancestor, but lost by archaeobacteria following the eukaryote/archaeobacteria divergence. The theories that assume an archaeobacterial ancestor must suppose that these diverse genes were all re-acquired secondarily by symbiogenesis or massive lateral gene transfer. In addition to the general eubacterial genes listed by Gupta (1998a), there are many others shared specifically by eukaryotes and some or all actinobacteria that are absent from archaeobacteria, e.g. those for sterol synthesis; both these and the cell-biological similarities between eukaryotes and actinobacteria are unexplained by the theories of an archaeobacterial ancestry for eukaryotes.

Recency of the neomuran revolution

Elsewhere, I reviewed the extensive palaeontological evidence that eukaryotes are over four times younger than bacteria, evolving only about 850 million years (My) ago compared with ~ 3850 My for eubacteria (Cavalier-Smith, 2002a). Eukaryotes are emphatically not a 'primary line of descent', as Pace *et al.* (1986) so misleadingly called them. Archaeobacteria are also not a primary line of descent, and there is no evidence whatever that they are any older than eukaryotes. Since the evidence that archaeobacteria are sisters of eukaryotes is compelling (Cavalier-Smith, 2002a), it is highly probable that the neomuran common ancestor arose by the radical transformation of a posibacterium around 850 My ago. The widespread idea that both neomuran groups are primary lines of descent is based on the misrooting of some molecular trees (as shown by Cavalier-Smith, 2002a) and ignorance of the palaeontological evidence. Reconciling the palaeontological and molecular evidence is a complex matter that demands thorough discussion. It requires the recognition that the temporal pattern of molecular change is very different in different categories of molecules, which show the classical phenomenon of mosaic evolution: different molecules alter their rates of evolution to greatly differing degrees in the same lineage. As explained in great detail elsewhere (Cavalier-Smith, 2002a), the hundreds of molecules that were specifically involved in the drastic changes that created the ancestral neomuran (e.g. rRNA, protein-secretion molecules, vacuolar ATPase) underwent temporarily vastly accelerated evolution (quantum evolution) during those innovations in the stem neomuran, but thousands of other genes, notably most metabolic enzymes, were more clock-like. A subset of genes underwent similar drastic quantum evolution during the evolution of the stem archaeobacterium, as did an only partially overlapping set of genes during the evolution of the stem eukaryotes during the origin of phagotrophy, the cytoskeleton, endomembranes and other eukaryote-specific characters.

If a gene underwent quantum evolution during all three major transitions (the neomuran, archaeobacterial and eukaryote origins), then its molecular tree will show clear-cut separation into the three domains, as for rRNA, but if quantum evolution occurred in none of them (or in only one or two) for that particular molecule, the pattern will be different. The confusing effects of mosaic and quantum evolution and how they can be disentangled by making a proper synthesis with the direct evidence from the fossil record of the actual timing of historical events are explained thoroughly elsewhere (Cavalier-Smith, 2002a). These arguments are crucial for understanding the pattern of molecular evolution during the origin of eukaryotes, but too lengthy to repeat here. They do, however, help us to understand why many trees clearly support the sister relationship between eukaryotes and archaeobacteria, whereas a significant minority suggest instead that eukaryotes may branch within the archaeobacteria. The latter trees are usually for enzymes that did not undergo quantum evolution during the three transitions, so there are not enough changes in the archaeobacterial stem to show the holophyly of the archaeobacteria robustly, and they can appear paraphyletic because random noise or minor systematic biases make the eukaryotes branch misleadingly among them. The situation is made worse by the fact that, for many of these trees (Brown & Doolittle, 1997), the taxonomic sampling is so sparse that such artefacts will be relatively more likely. Trees with better sampling more often support archaeobacterial holophyly. But even they can be expected to do so only if a substantial number of synapomorphies evolved in the archaeobacterial stem. If the origin of archaeobacteria was relatively rapid after the neomuran revolution, as is likely (Cavalier-Smith, 2002a), and if the divergence of crenarchaeotes and euryarchaeotes occurred soon afterwards, one would expect many trees not to show archaeobacterial holophyly robustly – unless they had undergone marked quantum evolution in the archaeobacterial stem. Such quantum evolution can be very useful in accentuating the evidence for the holophyly of a particular group (Cavalier-Smith *et al.*, 1996a; Cavalier-Smith, 2002a), but it is highly misleading as to the relative temporal duration of different segments of the tree and, if extreme, can also give false topologies, so critical interpretation of a variety of trees for functionally unrelated molecules is essential for accurate phylogenetic reconstruction.

An actinobacterial ancestry for eukaryotes

According to the neomuran theory (Cavalier-Smith, 1987c, 2002a), eukaryotes evolved by the radical transformation of one particular posibacterial lineage to generate the cytoskeleton and endomembrane system and the associated or shortly following symbiogenetic implantation of an α -proteobacterial cell as a protomitochondrion. The most plausible ancestor for the host component of eukaryotes is a derivative of an aerobic, heterotrophic, Gram-positive bacterium, not

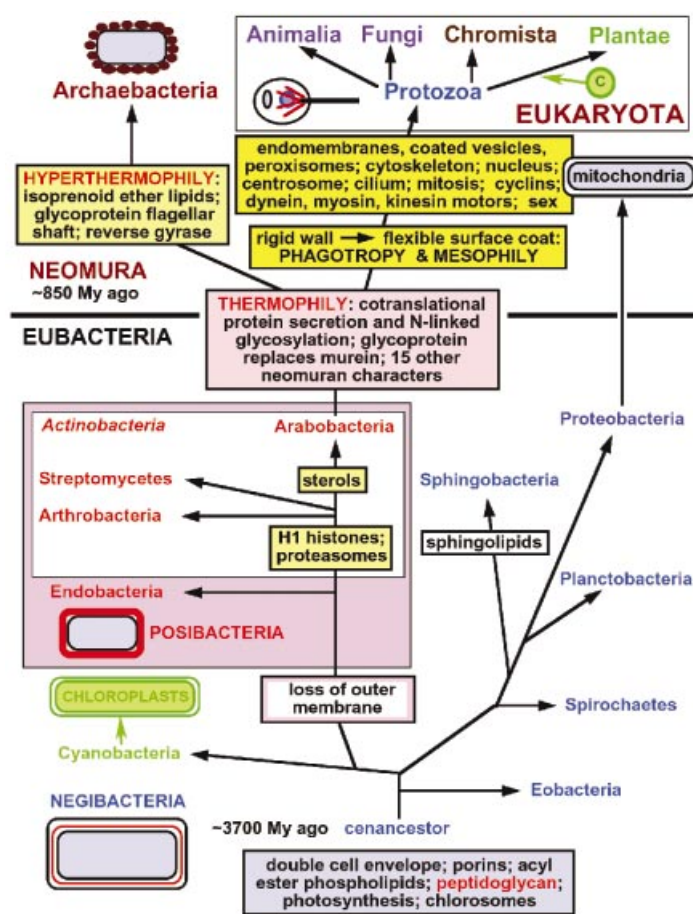


Fig. 1. The bacterial origins of eukaryotes as a two-stage process. The ancestors of eukaryotes, the stem neomura, are shared with archaeobacteria and evolved during the neomuran revolution, in which *N*-linked glycoproteins replaced murein peptidoglycan and 18 other suites of characters changed radically through adaptation of an ancestral actinobacterium to thermophily, as discussed in detail by Cavalier-Smith (2002a). In the next phase, archaeobacteria and eukaryotes diverged dramatically. Archaeobacteria retained the wall and therefore their general bacterial cell and genetic organization, but became adapted to even hotter and more acidic environments by substituting prenyl ether lipids for the ancestral acyl esters and making new acid-resistant flagellar shafts (Cavalier-Smith, 2002a). At the same time, eukaryotes converted the glycoprotein wall into a flexible surface coat and evolved rudimentary phagotrophy for the first time in the history of life. This triggered a massive reorganization of their cell and chromosomal structure and enabled an α -proteobacterium to be enslaved and converted into a protomitochondrion to form the first aerobic eukaryote and protozoan, around 850 My ago. Substantially later, a cyanobacterium (green) was enslaved by the common ancestor of the plant kingdom to form the first chloroplast (C).

an anaerobic methanogen (Martin & Müller, 1998; Moreira & López-García, 1998), which would need immensely more metabolic changes to make the eukaryote cenancestor, which, as explained below, was undoubtedly an aerobic heterotroph. As in my previous scenario (Cavalier-Smith, 1987c), I argue that this ancestor was pre-adapted for phagotrophy by secreting a number of digestive enzymes. *Bacillus subtilis* secretes about 300 proteins, the vast majority co-translationally as in eukaryotes, not post-translationally as in proteobacteria like *Escherichia coli* (Tjalsma *et al.*, 2000). Posibacteria are thus pre-adapted to have evolved the co-translational secretion mechanism used by the endomembrane system of eukaryotes. However, I previously (Cavalier-Smith, 1987c) gave several reasons for thinking that neomura

evolved, not from the posibacterial subphylum to which *Bacillus* belongs (Endobacteria; Cavalier-Smith, 2002a), but from the other posibacterial subphylum, Actinobacteria, which includes actinomycetes (e.g. *Streptomyces*) and their relatives such as mycobacteria and coryneforms. Fig. 1 emphasizes that the origin of eukaryotes from this actinobacterial ancestor occurred in two phases: the first phase was shared with the ancestors of archaeobacteria and involved the evolution of co-translational *N*-linked glycosylation and the substitution of the eubacterial peptidoglycan wall by one of glycoprotein.

The several reasons for favouring an actinobacterial origin for eukaryotes included the facts that *Streptomyces* was the only known bacterium to produce

Table 1. Neomuran characters shared by some or all actinobacteria but not other eubacteria

General neomuran characters
1. Proteasomes
2. 3'-Terminal CCA of tRNAs mostly (actinobacteria) or entirely (neomura) added post-transcriptionally
Characters shared by eukaryotes generally but not archaeobacteria
1. Sterols
2. Chitin
3. Numerous serine/threonine phosphotransferases and protein kinases related to cyclin-dependent kinases
4. Tyrosine kinases
5. Long H1 linker histone homologues related to eukaryote ones throughout
6. Calmodulin-like proteins
7. Phosphatidylinositol (in all actinobacteria)
8. Three-dimensional structure of serine proteases
9. Primary structure of alpha amylases
10. Fatty acid synthetase a complex assembly
11. Desiccation-resistant exospores
12. Double-stranded DNA repair Ku protein with C-terminal HEH domain (Aravind & Koonin, 2001)

chitin, that the actinobacterial fatty acid synthetase is a macromolecular aggregate as in fungi and animals (not separate soluble molecules as in other bacteria) and that the formation of exospores can be interpreted as a precursor for the evolution of eukaryote zygo-spores, probably the ancestral condition for sexual life-cycles (Cavalier-Smith, 1987c). Since then, it has been found that actinobacteria resemble eukaryotes more than do any other bacteria in five other key features (Table 1). Their histone H1 homologue is longer than in other eubacteria (absent from archaeobacteria) and related to eukaryotic H1 over more of its length (Kasinsky *et al.*, 2001). They have calmodulin-like proteins. They have a greater variety of serine/threonine kinases than any other bacteria (Av-Gay & Everett, 2000). *Mycobacterium* synthesizes sterols (Lamb *et al.*, 1998), like eukaryotes. They are rich in phosphatidylinositol lipids. Thus, in several very important respects, actinobacteria are more similar to eukaryotes than are any other bacteria. There are no other eubacteria with as many important similarities, so it is highly probable that neomura evolved from an actinobacterium having all these properties and that those that are not shared by archaeobacteria (e.g. sterols, histone, H1, chitin, spores) were lost after they diverged from their eukaryote sisters: as explained elsewhere, there is evidence for very extensive gene loss and genome reduction during the origin of archaeobacteria (Cavalier-Smith, 2002a).

Precisely which actinobacterial group is closest to eukaryotes is more problematic. The presence of sterols in *Mycobacterium* would favour the class Arabobacteria, in which it is now placed (Cavalier-Smith, 2002a), as shown in Fig. 1. However, the eukaryotic-like Ku double-strand repair protein with a fused downstream HEH domain (Aravind & Koonin, 2001) in *Streptomyces* would favour the class Streptomycetes instead. In view of the probable actinobacterial ancestry of eukaryotes, the suggestion that eukaryotes and *Streptomyces* independently fused a

similar HEH domain to Ku (Aravind & Koonin, 2001) is most unparsimonious; the absence of Ku proteins from archaeobacteria other than *Archaeoglobus* can be attributed to a single loss in the common archaeobacterial ancestor plus a single lateral transfer from a non-HEH-containing eubacterium into *Archaeoglobus*. Gene and character losses are more frequent than is often supposed, and complicate phylogenetic inference. A much better understanding of actinobacterial cell biology, a substantially improved knowledge of gene and character distribution within actinobacteria and a more robust molecular phylogeny for the group based on numerous genes are all needed to provide a sounder basis for understanding the origin of neomuran and eukaryotic characters.

The presence of cholesterol (Lamb *et al.*, 1998) is a particularly important pre-adaptation for the origin of phagotrophy and the endomembrane system. The fact that it is made by actinobacteria also means that to regard the presence of steranes as early as 2.7 billion years (Gy) ago as 'evidence' for eukaryotes (Brocks *et al.*, 1999) is incorrect; they are more likely to have been produced by actinobacteria or by the two groups of proteobacteria that make sterols (e.g. Kohl *et al.*, 1983; see Cavalier-Smith, 2002a). Moreira & López-García (1998) invoke a highly complex and cell-biologically unacceptable cell fusion between a δ -proteobacterium making sterols and an archaeobacterium to create the ancestor of eukaryotes prior to the symbiotic origin of mitochondria. They make this highly implausible suggestion mainly to explain how eukaryotes got sterols, serine/threonine kinases and calmodulin as well as the characters they share with archaeobacteria. But this elaborate hypothesis is entirely unnecessary, as all three characters would already have been present in the actinobacterial ancestors of neomura, which then evolved the shared neomuran characters prior to the divergence of eukaryotes and archaeobacteria (Fig. 1). Thus, their syntrophy hypothesis is phylogenetically unnecessary, as well

as being refuted by the three gene splits mentioned above and a fourth in methanogens discussed below and being mechanistically probably impossible. Actinobacteria should be studied carefully to see whether any also have other characters suggested by Moreira & López-García (1998) to be derived from δ -proteobacteria (e.g. the core structure of the lipid anchor).

Autogenous and exogenous mechanisms of eukaryogenesis

Low-trauma wall-to-coat transformation, actin and eukaryotic cytokinesis

Halobacteria are unusual among walled bacteria in that not all are rods or cocci, but some are pleomorphic. This indicates that their glycosylated exoskeletons are able to support a greater variety of shapes, as in eukaryotes. Their glycoprotein walls are aggregates of globular proteins constituting an 'S-layer', which I have argued was the ancestral state for all archaeobacteria (Cavalier-Smith, 1987c) and evolved first in the common ancestor of all neomura from an actinobacterial S-layer (Cavalier-Smith, 2002a). I suggested previously that the sudden and complete loss of the peptidoglycan wall created a naked ancestor of eukaryotes (Cavalier-Smith, 1975, 1987c). However, instead of losing the bacterial wall entirely, I now suggest that only the peptidoglycan and lipoproteins were lost and that the proteins of the S-layer were simply converted into surface coat/wall glycoproteins by evolving hydrophobic tails to anchor them in the membrane and co-translational N-linked glycosylation in the ancestral neomuran (Cavalier-Smith, 2002a). If stem neomura had a glycoprotein wall similar to that of halobacteria, the transition to archaeobacteria and to eukaryotes would have been less traumatic and thus evolutionarily somewhat easier than originally envisaged (Cavalier-Smith, 1987c). In particular, the neomuran ancestor would have been able to retain the eubacterial cell-division mechanism and divide satisfactorily during eukaryogenesis (Cavalier-Smith, 2002a). Relatively small genetic changes probably then sufficed to transform the early neomuran glycoprotein wall into a surface coat. I will call the hypothetical intermediate between the stem neomura and the first eukaryote a prekaryote for the probably short period between the first evolution of its surface coat and the origin of the nucleus.

I originally proposed that actomyosin was the key molecular innovation that created eukaryotes by enabling phagotrophy to evolve (Cavalier-Smith, 1975). We now know, as we then did not, that actomyosin does mediate phagocytosis. Actin was once thought to have evolved from the distantly related FtsA (Sánchez *et al.*, 1994), a protein that plays a role together with FtsZ in bacterial division. A much better candidate for an actin ancestor is MreB, a shape-determining protein of rod-shaped eubacteria that forms similar filaments that associate with membranes (van den Ent *et al.*, 2001). I speculate that actin polymerization, actin

membrane-anchoring proteins, actin cross-linking proteins and actin-severing proteins were the first elements of the eukaryotic cytoskeleton to evolve, in order to help stabilize the osmotically sensitive prekaryote against a varying ionic and osmotic environment while still allowing cell growth. Actin polymerization, if suitably anchored and oriented, could also be used to push the cell surface out and partially surround potential prey, even in the absence of myosin. Complete engulfment would be more sophisticated and would depend on fusogenic plasma-membrane proteins.

Actinobacterial homologues of MukB and Smc, large proteins with coiled-coil domains reminiscent of myosin, deserve study as potential precursors of eukaryotic mechanochemical motors. MukB is involved in active chromosome partitioning in negibacteria (Niki *et al.*, 1991), as is the more eukaryotic-like Smc in posibacteria. I suggested previously that a DNA helicase both moves bacterial chromosomes actively and was a precursor of eukaryotic molecular motors (Cavalier-Smith, 1987b). Active bacterial chromosome segregation is now well established (Sharpe & Errington, 1999; Møller-Jensen *et al.*, 2000) and remarkably similar to my prediction, but insufficiently understood for one to suggest exactly how it evolved into the eukaryotic system, which it probably did smoothly. The DNA translocase SpoIIIE that actively moves the chromosome terminus away from the division septum is conserved throughout eubacteria (Errington *et al.*, 2001). Its apparent absence from neomura suggests that it was lost at the same time as was peptidoglycan; but might it instead have been transformed radically beyond recognition into a novel neomuran protein?

I suggest that, after the evolution of the flexible surface coat instead of a rigid wall, MreB was converted to actin and initially functioned to hold the cell together passively in association with actin cross-linking proteins. Soon, a bacterial chromosomal motor was recruited to form an ancestral myosin to cause active sliding of actin filaments in the equator of dividing cells to supplement the activity of FtsZ, which I speculate could have become less efficient as the surface became less rigid. Once the actomyosin contractile ring became efficient, the FtsZ ring could be dispensed with, as also happened in mitochondria in a common ancestor of the opisthokonts (animals, fungi and choanozoa; see below) even though FtsZ was retained in plant and chromist mitochondria (Beech & Gilson, 2000). However, in the prekaryote, FtsZ was not lost as in opisthokont mitochondria; instead, after being relieved of its previous function, it was free to triplicate to form tubulins, centrosomes and microtubules. Thus, eukaryote cytokinesis by an actomyosin contractile ring replaced the bacterial FtsZ functionally, paving the way for the evolution of a microtubule-based mitosis.

It may not matter much whether actomyosin contraction originated for cytokinesis, as just suggested, and was then recruited to help with phagocytosis or the

reverse, or else was applied to both processes simultaneously. Both probably became essential for effective feeding and reproduction of the prekaryote.

Phagotrophy and mitochondrial symbiogenesis

My earlier analysis (Cavalier-Smith, 1987c) explained how an incipient ability to engulf prey led to the perfection of phagocytosis and to the formation of endomembranes and lysosomes and exocytosis to return membrane to the surface and allow it to grow; the reader is referred to Fig. 8 therein and its vast legend for details. The essential logic of the steps remains sound, but it is probable that, as suggested earlier (Cavalier-Smith, 1975) and outlined in a later section, peroxisomes may have evolved autogenously by differentiation from the early endomembrane system and need not have been a later symbiogenetic addition, as I then suggested (Cavalier-Smith, 1987e). Thus, even though in my present analysis the mitochondrial symbiogenesis occurred much earlier in eukaryote evolution than previously thought, the overall contributions of symbiogenesis to the evolution of aerobic eukaryotes are greatly reduced in comparison with my 1987 theory and the importance of autogenous transformation and innovation further increased: probably only two bacteria, not three, were symbiogenetically involved.

Since the seminal generalization of Stanier & Van Niel (1962) that prokaryotes never harbour cellular endosymbionts has only one probable exception (von Dohlen *et al.*, 2001), I persist in arguing that phagocytosis was essential for uptake of the α -proteobacterial symbiont (Cavalier-Smith, 1975, 1983a, 1987e) and that at least the beginnings of phagotrophy had evolved before mitochondria originated. This exceptional case where bacteria apparently harbour endosymbionts involves β -proteobacteria that are themselves obligate endosymbionts of mealy bugs and contain γ -proteobacteria within their inflated cytosol (von Dohlen *et al.*, 2001). This interesting example shows that it is not physiologically impossible for bacteria to harbour endosymbionts. The reason why free-living bacteria have never been found to do so is probably twofold: they are generally too small to be able to accommodate other cells and their usually rigid walls must impose a strong barrier to accidental uptake. The β -proteobacterial hosts are unusually large for proteobacteria (10–20 μm in diameter) and are vertically transmitted from mealy bug to mealy bug within a host vacuolar membrane, analogously to eukaryotic organelles. I suggest that they may also lack peptidoglycan walls and that the resulting greater flexibility of the surface may have enabled them to take up intimately associated γ -proteobacteria at some stage in the history of the mealy bug bacteriome. I consider that such large bacteria with relatively flexible surfaces able to take up other bacteria could have evolved only within the protective cytoplasm of pre-existing eukaryotic cells and are therefore irrelevant to the mechanical problems of the origin of mitochondria.

I continue to argue that loss of peptidoglycan and the origin of at least a crude form of phagocytosis and proto-endomembrane system were mechanical prerequisites for the origin of mitochondria.

Mitochondrial symbiogenesis, however, may temporally have overlapped the later stages of these more fundamental cellular transformations, as Rizzotti (2000) suggests. Recent versions of the symbiogenetic theory emphasizing syntrophy instead of phagocytosis (Martin & Müller, 1998; Moreira & López-García, 1998; Margulis *et al.*, 2000) are as mechanistically implausible as Margulis' original, now abandoned, symbiotic theory (Margulis, 1970, 1981), which asserted that the mitochondrial symbiogenesis was the prerequisite for phagocytosis, and which I previously refuted (Cavalier-Smith, 1983a). I continue to argue that replacement of a bacterial cell wall by a glycoprotein surface coat was the primary facilitating cause and that evolution of phagotrophy (De Duve & Wattiaux, 1964; Stanier, 1970) was the key secondary, but effective, cause of the transformation of a bacterium into a eukaryote. As soon as phagotrophy was adopted, however inefficient, the possibility immediately opened up for some phagocytosed cells to escape digestion and to become cellular endosymbionts, whether parasites, mutualisms or slaves (Cavalier-Smith, 1975). By inserting a host ATP/ADP exchange protein into the proteobacterial envelope (John & Whatley, 1975), the host made the bacterium an energy slave. This essential first step could have occurred quite early, before the endomembrane system, cytoskeleton and nucleus were fully developed, but need not necessarily have done so and might have slightly followed the origin of the nucleus. The transfer of proteobacterial genes to the host could have occurred more easily if the nucleus had not developed properly. As a result, the near-neutral substitution of some host soluble enzymes by ones from the symbiont (e.g. valyl-tRNA synthetase; Hashimoto *et al.*, 1998) could also have occurred very early. Such substitution is mechanistically easier than the next logical step in the evolution of mitochondria, developing a generalized mitochondrial import system, which would have been much more mutationally onerous (Cavalier-Smith, 1983a, 1987c). Likewise, the loss of the cell wall would have allowed sex to evolve at once, so it probably evolved very early and there may be no primitively asexual eukaryotes (Cavalier-Smith, 1975, 1980, 1987c).

The origin of phagotrophy created the most radically new adaptive zone in the history of life: living by engulfing other cells, which favoured the elaboration of the endomembrane system, internal cytoskeleton and associated molecular motors, yielding a new Empire of life: the Eukaryota. Phagotrophy and the consequent internalization of the membrane-attached chromosome necessarily entailed a profound change in the mechanisms of chromosome segregation and cell division; FtsZ evolved into tubulin and underwent triplication to yield γ -tubulin to make centrosomes and

α - and β -tubulins to make spindle microtubules to segregate chromosomes mitotically. MreB became actin, which was recruited for both cytokinesis and phagocytosis as well as a general osmotically protective cross-linked cytoskeleton. In turn, the new division mechanisms entailed quantum innovations in cell-cycle controls involving the origin of cyclin-dependent kinases by recruitment from amongst the very diverse serine/threonine protein kinases already present in actinobacteria.

The cytoskeleton and endomembranes were much more important causes than the neomuran changes in transcriptional control (Cavalier-Smith, 2002a) in allowing a vast increase in cellular complexity and the origin of highly complex multicellular organisms: significantly, no archaeobacteria ever became multicellular, though sharing the same gene-control mechanisms as eukaryotes. As soon as these changes had occurred, there would inevitably have been an explosive radiation of protists into every available niche, their functional diversification being accentuated by their entirely novel ability to engulf other cells and either digest them for food or maintain them instead as permanent slaves providing energy, fixed carbon or other useful metabolites. Thus, phagotrophy led to symbiogenesis, not the reverse. Once symbiogenesis occurred, it had far-reaching effects. The simplest explanation of the large number of eukaryotic genes of eubacterial character (Brown & Doolittle, 1997) is twofold: many simply reflect the retention of most of those actinobacterial genes that did not undergo quantum evolution during the origin of neomura (Cavalier-Smith, 1987c, 2002a); the significant minority that are negibacterial rather than posibacterial in affinity (Feng *et al.*, 1997; Ribiero & Golding, 1998; Rivera *et al.*, 1998) could have come mainly or entirely by the incidental substitution of the host actinobacterial genes by genes from the α -proteobacterial ancestor of mitochondria encoding functionally equivalent proteins (Cavalier-Smith & Chao, 1996; Roger, 1999; Cavalier-Smith, 2000a).

Such trivial gene replacement (Koonin *et al.*, 1996) is sometimes said to be unexpected (Doolittle, 1998a), but was predicted on general evolutionary grounds (Cavalier-Smith, 1990a) before evidence for it became widespread. As first stressed by Martin (1996, 1998) and Martin & Schnarrenberger (1997), similar substitution of host or symbiont soluble enzymes by functionally equivalent ones occurred during the symbiogenetic origin of chloroplasts, the only other primary symbiogenetic event in the history of life. It has also probably occurred in the three secondary symbiogenetic events that transferred pre-existing plastids to non-photosynthetic hosts: incorporation of a red alga to form the chromalveolates and of green algae into euglenoids and chlorarachneans (Cavalier-Smith, 1999, 2000b). However, the frequency of gene replacement in early eukaryote evolution may have been considerably overestimated by excessive faith in the reliability of single-gene trees:

for example, though I accept that cytosolic triose-phosphate isomerase almost certainly replaced the original plastid one in green plants and that the cyanobacterial/plastid one probably replaced their cytosolic one (Martin, 1998), the trees for both proteins are so dominated by long-branch problems that it is much more likely that they are misrooted and partially topologically incorrect than that the eukaryote cytosolic enzyme actually came from the mitochondrion or other eubacterial symbiont, as Keeling & Doolittle (1997) and Martin (1998) have postulated.

The fact that many eubacterial-like genes resemble those of negibacteria more than posibacteria (Golding & Gupta, 1995; Feng *et al.*, 1997) does not fit my original assumption that all of them came directly from the neomuran ancestor (Cavalier-Smith, 1987c), but favours descent for many of them from the α -proteobacterium. Two such key genes are the Hsp70 and Hsp90 chaperones. Both underwent duplication to create ER luminal as well as cytosolic versions. A third, mitochondrial version was retained for Hsp70, but not Hsp90. All are more negibacterial than posibacterial in character and group with proteobacteria on trees, though the cytosolic and ER versions do not do so specifically with α -proteobacteria (Gupta, 1998a), but this may be because they have evolved more rapidly than proteobacterial sequences; the cytosolic and ER Hsp70 are more divergent than the mitochondrial one, presumably because their function changed more significantly. Both have signature sequences that support a relationship with proteobacteria, in preference to most other negibacteria or posibacteria. Gupta (1998a) suggested that these and other proteobacterial-like genes were contributed by an additional symbiotic merger between a negibacterium and an archaeobacterium, but, if we accept that mitochondria arose at about the same time as the nucleus, these and other similar assumptions of an earlier symbiogenesis (e.g. Sogin, 1991; Lake & Rivera, 1994; Moreira & López-García, 1998) are entirely unnecessary (Cavalier-Smith & Chao, 1996; Cavalier-Smith, 1998a), as well as mechanistically implausible.

Mitochondrial Hsp70 is generally accepted as deriving from the α -proteobacterial ancestor of mitochondria, as it groups specifically with α -proteobacteria on trees. However, it typically appears as their sister (Roger, 1999), whereas, if it were evolving chronometrically, it ought to be nested relatively shallowly within them, since mitochondria are probably over three times younger than α -proteobacteria (Cavalier-Smith, 2002a). Mitochondrial Hsp60 is nested within α -proteobacteria, but not as shallowly as clock dogma would expect. The excessive depth of the mitochondrial versions of both chaperone molecules on trees is a typical artefact of accelerated evolution. Chloroplast genes, whether rRNA or protein, show similar several-fold accelerated evolution compared with their cyanobacterial ancestors and therefore branch much more deeply within the cyanobacterial tree (Turner *et al.*, 1999) than expected from the clear palaeontological

evidence that they are about five times younger (Cavalier-Smith, 2002a) or, in some cases, can even appear as sisters of cyanobacteria as a whole (Zhang *et al.*, 2000). Most nucleomorph genes show similar severalfold increases in rate, sometimes sufficiently great to prevent them nesting correctly within their ancestral groups (Archibald *et al.*, 2001; Douglas *et al.*, 2001). Thus, severalfold accelerated evolution is a general phenomenon for all symbiogenetically enslaved genomes, just as it appears to be for the obligately parasitic mycoplasmas. Not only does this provide yet another refutation of the dogma of the molecular clock, now devoid of any secure theoretical or empirical justification (Ayala, 1999), but it also urges extreme caution in interpreting the fact that the cytosolic and ER versions of Hsp70 do not group specifically (at least not reproducibly) with the mitochondrial versions, even though they appear to be of proteobacterial origin (Gupta, 1998a). Although we cannot exclude the possibility that they were acquired by lateral gene transfer independently of the mitochondrial symbiogenesis (Doolittle, 1998a), the indubitable marked tendency of genes of symbiogenetic origin to suffer accelerated evolution that drags them too deeply down trees is a more parsimonious explanation. I therefore consider it most likely that the host Hsp70 gene was lost accidentally after one or more copies of the protomitochondrial gene was transferred into the nucleus. The transferred protein may have taken over the function of the cytosolic protein before one copy of it acquired a mitochondrial pre-sequence for targeting it to the mitochondrion. Only after the latter happened could the mitochondrial copy of the gene have been lost.

Analogous considerations apply to eukaryotic Hsp90 genes, probably of negibacterial origin as they lack the conserved five amino acid insertion found uniquely in all posibacteria (Gupta, 1998b). I suggest that their ancestor moved from the protomitochondrial genome into the nucleus, but Hsp90 was never retargeted back into the mitochondrion, and that the actinobacterial gene was lost. No archaeobacteria have Hsp90. Fungi secondarily lost the ER Hsp90, possibly because they abandoned phagotrophy. It would be interesting to know if this occurred in the fungal ancestor or later within the Archaeomycota when the Golgi became secondarily unstacked in the ancestor of the Alcomycetes and Neomycota (Cavalier-Smith, 2000c).

The trivial replacement of a host gene for a cytosolic protein by an equivalent one from a symbiont therefore requires fewer steps than the effective transfer of a symbiont gene to the nucleus and the retargeting of its protein to the symbiont. Analogous trivial replacement instead of a symbiont gene by a host gene can also occur by the addition of appropriate targeting sequences (Cavalier-Smith, 1987e, 1990a), as exemplified by the fact that most of the soluble enzymes of mitochondria (unlike those of the cristae membranes) are probably of actinobacterial or archaeobacterial, rather than α -proteobacterial, affinity and by the

retargeting of a duplicated host glyceraldehyde-phosphate dehydrogenase to plastids of the ancestral chromalveolates (Fast *et al.*, 2001). Another nuclear duplicate of the α -proteobacterial Hsp70 acquired signal sequences for targeting into the ER and became the main ER chaperone. If the ancestors of the cytosolic and ER versions of Hsp70 (Cavalier-Smith, 2000a) and Hsp90 were both acquired from the protomitochondrion, then we must conclude that the establishment of the protomitochondrion overlapped with the final stages of evolution of the ER, when it was being made more efficient by the acquisition of both chaperones. Since the cytosolic versions of both proteins are involved in centrosome function, this implies that it also overlapped with the perfection of mitosis. If this interpretation is correct, symbiont genes played a role in the otherwise autogenous origins of the endomembrane system and cytoskeleton. However, unlike Margulis (1970), I do regard the symbiogenetic origin of mitochondria as making an essential, qualitative contribution to the origin of the eukaryotic cell, which was fundamentally driven by the selective advantages of phagotrophy (Stanier, 1970; Cavalier-Smith, 1975) once the neomuran revolution in wall chemistry and protein secretion mechanisms (Cavalier-Smith, 2002a) made this mutationally possible. If, purely by chance, the host version of the Hsp70 gene had been retargeted to the mitochondrion before the mitochondrial one, I argue that the mitochondrial version instead would have been lost and the ER version would have evolved from a duplicate of the original actinobacterial gene, rather than the replacement α -proteobacterial gene. If that had occurred, the mitochondrion, ER and cytosol would probably have worked equally well. If this is correct, and also true for all other cases of symbiogenetic replacement of host genes [e.g. glyceraldehyde-phosphate dehydrogenase (Henze *et al.*, 1995); valyl-tRNA synthetase (Hashimoto *et al.*, 1998)] by α -proteobacterial ones, then these contributions of the protomitochondrion to the origin of eukaryotes were trivial and purely incidental historical accidents, in no way essential for the tremendous qualitative changes in cell structure that occurred as a result of the origin of phagotrophy – with the sole exception of the origin of the structure of the mitochondrion itself, which is dispensable for eukaryotic life, as its multiple independent losses attest (Roger, 1999). As I argue in a later section, the major contribution of the protomitochondrion to the evolution of the eukaryotic cell was a purely quantitative one: greatly increasing the efficiency of use of the spoils of phagotrophy.

Though it is widely assumed that proteobacterial genes that replaced stem neomuran ones are functionally equivalent, on the neomuran theory, replacement need not have been entirely neutral. Most shared neomuran characters are explicable as adaptations to thermophily (see Cavalier-Smith, 2002a). Thus, it is highly probable that the prekaryote was initially a thermophile. However, it could enter the biological main-

stream as a phagotroph only by colonizing ordinary seawater, soil and freshwater under mesophilic conditions. If this reversion to mesophily coincided with the origin of the protomitochondrion, there might therefore have been a significant selective advantage in losing host genes rather than normal symbiont genes. Perhaps this explains why so many host enzymes were replaced, far more, apparently, than by the later plastid symbiogenesis (an alternative explanation for this large difference might be that the mitochondrial symbiogenesis took place before the origin of the nuclear envelope or the loss of its early free permeability, whereas that of chloroplasts undoubtedly took place afterwards). By replacing many heat-adapted enzymes by more mesophilic versions, symbiogenesis might have helped convert a specialized thermophilic pre-karyote with narrow ecological range into a hugely adaptable phagotrophic eukaryote that would spawn descendants able to live anywhere except very hot places, which their sister archaeobacteria were then colonizing for the first time (Cavalier-Smith, 2002a). In principle, multiple point mutations could easily have modified each gene for mesophily, but gene replacement might have been faster and thus more likely.

Contributions of lateral gene transfer to eukaryogenesis?

Doolittle (1998a) recently stressed another important consequence of the evolution of phagotrophy. It should be much easier to acquire novel genes by lateral transfer from phagocytosed prey than in old-fashioned bacterial ways. Some of the genes that Moreira & López-García (1998) suggest entered pre-karyotes in their over-elaborate pre-phagotrophy fusion event from myxobacteria or other δ -proteobacteria might have done so instead by phagocytosed myxobacterial prey contributing genes but no genetic membranes. The most significant contributions might have been two categories of gene suggested by Moreira & López-García (1998): small G proteins, essential for cytosin and, therefore, endomembrane compartmentation, and retroelements and reverse transcriptase, which could have played a key role in the explosive spread of spliceosomal introns after they evolved from group II introns immediately after the origin of the nucleus (Cavalier-Smith, 1991d). Unless future studies reveal that these genes are also present in the actinobacteria or in some α -proteobacteria, it is possible that such lateral gene transfer from myxobacteria played a minor but significant role in eukaryogenesis. Mutants of the MglA protein of *Myxococcus* can be complemented by eukaryotic Sar1p (Hartzell, 1997), but the assumption that this gene family entered eukaryotes via δ -proteobacteria (Moreira & López-García, 1998) need not be correct, as homologues are also known from *Aquifex* (ϵ -proteobacterium; Cavalier-Smith, 2002a) and the Hadobacteria (*Thermus* and *Deinococcus*). Their suggestion that phosphatidylinositol signalling proteins, which form the basis for eukaryotic cell signalling, also

came from δ -proteobacteria is less plausible, since phosphatidylinositol lipids are particularly well developed in actinobacteria and lateral gene transfer may have been unnecessary. When a myxobacterial genome sequence is available, it will be particularly interesting to see whether there is evidence for any gene contributions to eukaryotes by lateral gene transfer or whether all the key eukaryotic genes came either vertically from an actinobacterium or laterally by the mitochondrial symbiogenesis, as is possible.

As myxobacteria have a typical negibacterial envelope, I do not see how their outer membrane could ever have been lost or even fuse, as Moreira & López-García (1998) assume. Their scenario for the origin of the nucleus is totally implausible compared with the classical vesicle-fusion hypothesis (Cavalier-Smith, 1987c, 1988a). However, none of these cytological absurdities is necessary if any myxobacterial genes that may eventually be shown to have entered the pre-karyote were simply got from food. Likewise, if it can be proven that myxobacterial enzymes making the glycosylated *myo*-inositol phosphate lipid headgroup that is identical to the core of the eukaryotic lipid anchor for external globular proteins are homologues of the eukaryotic ones, but found in no other bacteria, their genes could also have come in the food. This is far preferable to invoking a mechanistically unsound pre-phagotrophy fusion (Moreira & López-García, 1998). Given the eukaryote phylogeny advocated here, that membrane-anchor machinery must have been present in the eukaryote cenancestor.

None of the above possible lateral transfers is yet established. Perhaps some or even all of them will turn out, on closer investigation, to be convergent red herrings. The possible contributions from myxobacteria suggested by Moreira & López-García (1998) are potentially so important that they should be pursued in depth. However, if we reject their syntrophy hypothesis because of its unparsimonious and mechanistically unsound fusions and membrane losses, there is no reason to single out sulphate-reducing myxobacteria as potential donors. Those most deserving of a major genome sequencing effort would be the aerobic predators, because of both their possible donation of genes to eukaryotes and their developmental complexity, remarkable for bacteria.

As primitive phagotrophy was a prerequisite for uptake of the ancestor of the mitochondrion, protomitochondria could not have originated before a stem neomuran began to be converted into a eukaryote and the rudiments of endomembranes and cytoskeleton had arisen. However, if my arguments about the protomitochondrial origins of the cytosolic and ER Hsp70s and Hsp90s are correct, mitochondria must have originated so early during the evolution of the eukaryotic cell that the prospect of our finding a primitively amitochondrial protozoan is effectively zero. If, however, these and other proteobacterial genes that seem to have replaced vital host genes came

instead from the food of the transitional prekaryote, and mitochondria were implanted substantially later, then primitively amitochondrial eukaryotes might exist and remain to be discovered. On present evidence, I think this very unlikely, but not impossible.

Spliceosomal intron origin, spread and purging

According to the mitochondrial seed theory of spliceosomal introns (Cavalier-Smith, 1991d; Roger *et al.*, 1994), they originated from group II introns in genes transferred from the protomitochondrion to the nucleus after the evolution of the nuclear envelope (Cavalier-Smith, 1987c, 1988a) allowed a slower splicing *in trans* by the spliceosome to evolve. Spread was by reverse splicing followed by reverse transcription. If neither the actinobacterial host nor the proteobacterial symbiont had reverse transcriptase, addition of reverse transcriptase from myxobacterial food would have created an explosive cocktail that allowed the introns to insert rapidly into genes throughout the genome. The new eukaryote phylogeny presented below greatly strengthens this version of the 'introns-late' scenario. For a period, when I accepted that the Archezoa were all secondarily amitochondrial but still thought that their basal branching on rRNA trees (Cavalier-Smith & Chao, 1996) might be correct, why they seemed to be virtually free of introns (Logsdon, 1998) was a puzzle. Why had introns not invaded immediately following the acquisition of mitochondria? Now that scepticism of the rRNA tree's root and the rerooting on putatively unikont eukaryotes, not Archezoa, is effected (see below), it is clear that spliceosomal introns did invade rapidly in the cenancestor. Their virtual absence in archezoa must be a secondary loss, like their almost total absence in kinetoplastids, which belong in Euglenozoa, the putative sister group to Archezoa. These selfish genetic elements must somehow have been largely, or perhaps entirely in *Giardia*, purged from the genome.

Sterols, cell-cycle controls and the nuclear skeleton

Sterols provided the rigidifying properties and acyl esters the fluid properties needed for highly flexible membrane functions. Since eukaryotes have a gradient with an increased sterol/phospholipid ratio running from rough endoplasmic reticulum (RER) to plasma membrane, it can hardly be doubted that present functions are optimized by a proper balance between them. Having both might have been a prerequisite for the origin of coated-vesicle budding and fusion. The absence of sterols in α -proteobacteria and the demonstration that actinobacteria make sterols (Lamb *et al.*, 1998) provide the final refutation of the suggestion (Margulis, 1970) that sterol biosynthesis came into a pre-eukaryote via the mitochondrion and that a claimed membrane 'fluidizing' function for them was a prerequisite for the origin of phagotrophy and the endomembrane system. As I have long argued, the reverse is true: phagotrophy was a prerequisite for the symbiogenetic origin of mitochondria (Cavalier-

Smith, 1983a). The discovery of actinobacterial sterols renders acquisition by lateral gene transfer (Moreira & López-García, 1998) unnecessary.

The origin of the eukaryotic cyclin-dependent serine/threonine kinase system, which controls both the entry into S-phase and the eukaryotic cell cycle, especially the onset of anaphase and the return to G1, was a key step in eukaryogenesis. Ubiquitin-tagged proteolytic degradation is vital for the centrosome cycle and these replication and mitotic controls. 20S proteasomes are of actinobacterial origin (Cavalier-Smith, 2002a) and were inherited by the neomuran ancestor and retained by archaeobacteria. 26S proteasomes and ubiquitination are eukaryote-specific and must have evolved from 20S proteasomes in the prekaryote or stem eukaryote lineage prior to the eukaryotic cenancestor. I suggest that ubiquitin and ubiquitination of selected proteins, and also the addition of extra proteins to make the more complex 26S proteasome, evolved to label key cell-cycle proteins and target them for temporally specific degradation. Of key importance were the origins of cyclin B, which activates the serine/threonine kinase system, and cohesins, which initially hold together sister chromatids but then are digested at anaphase by the proteasomes. Their evolution and that of their temporally controlled polyubiquitination were key elements in the novel eukaryotic cell-cycle controls that the origin of centrosomes and spindle microtubules necessitated, and which did not occur in archaeobacteria, which ancestrally retained the bacterial FtsZ division mechanism and 20S proteasomes (Cavalier-Smith, 2002a). These innovations created the eukaryotic cell cycle as a bistable cell oscillator able to switch reversibly between the growth phase (G1) and the reproductive phase (S, G2 and M) by means of proteolytic controls of protein kinases that regulate numerous elements of the reproductive machinery. Subsequently, ubiquitination was also used in the ancestral role of proteasomes to scavenge damaged proteins and, after the origin of multicells, selected by kin selection to mediate apoptosis.

These temporal controls were just as important as the origin of the novel nuclear envelope structure and the mechanochemical machinery of the mitotic apparatus (centrosomes, spindle microtubules and the molecular motors dynein and kinesin) in creating the eukaryotic cell. The novel division machinery and cell-cycle controls allowed the multiplication of replicon origins and other fundamentally different features of eukaryote chromosome structure (Cavalier-Smith, 1981a, 1987c, 1993a) compared with the universal bacterial pattern of a single replicon per chromosome that prevails in eubacteria and archaeobacteria (Cavalier-Smith, 2002a). Indirectly, it allowed massive increases in eukaryotic genome size (Cavalier-Smith, 1978b), but is insufficient to explain the patterns of variation of nuclear genome size. We must also postulate a skeletal role for nuclear DNA and the evolutionary coadaptation of nuclear and cell volume to ensure balanced growth of eukaryotic cells (Cavalier-Smith,

1985; Cavalier-Smith & Beaton, 1999). The increased genome size caused by selection for larger nuclei and the origin of sex (see below) together made the rapid intragenomic spread of selfish genetic parasites inevitable in early eukaryotes, notably the spliceosomal introns (Cavalier-Smith, 1991d). However, the spread of such selfish DNA was not the fundamental cause of the genomic expansion. Both such duplicative transposition and normal duplication of genic and non-genic DNA contributed mechanistically to the expansion. Irrespective of the mutational mechanism, increases in genome size were favoured by cellular selection for larger nuclei to ensure balanced growth of the eukaryote cells which, for the first time, separated RNA and protein production into separate compartments. An incidental side effect of this was a vastly expanded cosy habitat for genetic parasites for which sex, also probably incidentally, provided a novel means of spread (Cavalier-Smith, 1978a, 1993a). Although their sequences are selfish, their contribution to the bulk of the nuclei is as positively beneficial to the cellular economy as that of other types of non-coding DNA and genic DNA itself, all three of which contribute to the adaptively significant volume of the chromatin gel upon which nuclear envelopes are assembled (Cavalier-Smith, 1985, 1991e). Thus, calling eukaryotic transposons selfish is a half truth that has incorrectly tempted many into believing partially 'selfish' DNA to be a sufficient explanation for the tremendous variability of eukaryote genome size.

An evolutionarily chimaeric origin for centrosomes is possible, since Hsp90 and Hsp70 and γ -tubulin are key structural constituents (Lange *et al.*, 2000; Scheufler *et al.*, 2000). Although γ -tubulin probably came vertically from the actinobacterium, cytosolic and ER Hsp70 and Hsp90 are probably of proteobacterial affinity, as discussed above. Even if two centrosomal components (cytosolic Hsp70 and 90) had a foreign origin, there is no reason to think that centrioles or cilia arose symbiotically or that either is ancestral to the mitotic spindle (Cavalier-Smith, 1992b), as proposed by the vague and phylogenetically unjustified spirochaete theory of Margulis *et al.* (2000).

Coated vesicles, ercytosis and the origin of the endomembrane system

Classically, the endomembrane system is held to have evolved by invagination and inward separation of vesicles from the plasma membrane (De Duve & Wattiaux, 1964; Stanier, 1970; Margulis, 1970; Cavalier-Smith, 1975, 1980, 1981a, 1987c). Some of the original plasma-membrane proteins remained at the surface and others were internalized as the two membranes became differentiated. The homology of both eukaryotic plasma-membrane proteins and ER proteins (e.g. vacuolar ATPase; Gogarten & Kibak, 1992) to those of the bacterial cytoplasmic membrane supports this historic continuity of membranes, as an example of membrane heredity, despite the development of a topological discontinuity. As De Duve &

Wattiaux (1964) pointed out, evolution of primitive phagotrophy provides both a selective advantage and a cellular mechanism for such internalization. I emphasized (Cavalier-Smith, 1975, 1987c) three key logical features of this mode of origin of the endomembrane system: (i) cell wall loss (here modified to direct conversion to a flexible coat) was a prerequisite, (ii) exocytosis must have evolved concomitantly with membrane internalization by phagocytosis and (iii) differentiation between the protein composition of endomembranes and plasma membranes entails that the membrane budding from the ER ('ercytosis'; Cavalier-Smith, 1987c) and/or the exocytosis process act as a selective 'valve' or 'gate' that allows some proteins to return to the cell surface but traps others permanently in the endomembrane (Cavalier-Smith, 1988b). I stressed the key role of internalization by phagocytosis of the former bacterial plasma membrane's ribosome receptors [both ribophorins and receptors for the signal recognition particle (SRP); Walter *et al.*, 2000] in the differentiation between plasma membrane and RER. As the proteobacterial membrane differs from the eukaryotic ER in that most proteins are inserted or translocated by the SecA or other post-translational chaperone-based mechanisms, I suggested that the switch from predominantly post-translational to predominantly co-translational insertion was selectively advantageous because it discriminated between endomembranes and plasma membrane, thus preventing wasteful secretion of protolysosomal enzymes across the plasma membrane (Cavalier-Smith, 1987c). However, if obligate co-translational protein insertion and translocation had already evolved in the neomuran ancestor shared with archaeobacteria, as now appears to be true (Cavalier-Smith, 2002a), the stem neomuran secretory system was pre-adapted in this key respect for the origin of the endomembrane system by plasma-membrane internalization.

We can now simply interpret the key steps in the evolution of the SRP during the history of life. The simplest signal recognition particle is that of negibacteria, the ancestral cells, with only a short 4.5S SRP RNA and a single bound protein (Ffh), which probably evolved during pre-cellular evolution in the obcell (Cavalier-Smith, 2001). It later increased in complexity three times, each increase plausibly associated with quantum steps in membrane evolution (Cavalier-Smith 2002a). (i) When the negibacterial outer membrane was lost to form the ancestral posibacterium, helices 1–5 were added together with HBSu protein that binds them (homologue of SRP9/14 that mediate the signal arrest domain in eukaryotes); I suggest that this domain and its bound proteins functions thus also in Unibacteria, i.e. in all cells bounded by a single membrane. (ii) The neomuran ancestor added helix 6 and SRP19 that binds it; as this enhances the binding of Ffh, it may, like other neomuran characters, originally have been an adaptation to thermophily (Cavalier-Smith, 2002a). (iii) Loss of helix 1 and

addition of SRP68/72 proteins in the ancestral eukaryote, involved in docking onto the ER membrane. This progressive increase in complexity linked to membrane novelties is much more reasonable than the assumption of eubacterial simplification (Eichler & Moll, 2001) based on the misrooting of the tree by the now refuted myth of archaeobacterial antiquity. Whether archaeobacteria lost the HBSu/SRP9/14 homologue or it is too divergent to be recognised is unclear.

Given such prior evolution of obligate co-translational insertion, phagocytosis and erycytosis alone would have been sufficient to internalize ribosomes permanently in an early prekaryote and to create a primitive RER, provided that direct refusion of phagosomal membranes with the plasma membranes was prevented. Subsequent differentiation of primitive endomembranes into different submembranes (e.g. ER, Golgi, lysosomes and peroxisomes; Cavalier-Smith, 1975) also depended logically on the evolution of additional selective 'valves' between different compartments. We have known for some years that selective coated-vesicle budding is the mechanism for such valves; for example, erycytosis is mediated by COP II coated vesicles, which prevent ribosome receptors reaching the Golgi, and budding from the Golgi is mediated by COP I coated vesicles. The essential steps in evolution of the endomembrane system must have been the origin of these and other types of coated vesicle. As some of the COP protein subunits are related to certain proteins in clathrin-coated vesicles, which mediate vesicle budding from the plasma membrane, endosomes and trans-Golgi network, all three major kinds of coated vesicle probably have a common origin. I suggest that an ancestral COP II evolved first, creating erycytosis at the ER. Fusion of some COP II-generated vesicles with each other probably generated a smooth endomembrane compartment, a proto-Golgi/lysosome intermediate between ER and plasma membrane. Clathrin then evolved, allowing the separation of lysosomes from the Golgi; I suggest its uses in endocytosis and formation of contractile vacuoles were secondary. Finally, COP I evolved, creating the trans-Golgi network as an intermediate compartment between Golgi cisternae and lysosomes. The new rooting of the eukaryotes discussed in a later section shows that Golgi stacking must also have evolved in the cenancestor and have been lost polyphyletically in several derived lineages: contrary to what was previously thought (Cavalier-Smith, 1991a), there are no extant protozoa with primitively unstacked Golgi cisternae. The suggestion that the Golgi constitutes a permanently genetic membrane system distinct from both the ER and plasma membrane (Cavalier-Smith, 1991a, b) has recently been verified (Pelletier *et al.*, 2000; Seemann *et al.*, 2000).

The origin of the ancestral coated vesicle is what enabled 'cytosis' (budding and fusion of endomembrane vesicles; Cavalier-Smith, 1975) to evolve in the first place and thus create a permanent endomembrane system from phagosomal (food vacuole) membranes

temporarily internalized by the more primitive actomyosin-based phagocytosis. Because of the simultaneous establishment of mitochondria providing plentiful ATP, the original plasma membrane V-type ATP synthase (Gogarten & Kiback, 1992) became modified for its new ATP-driven proton-pumping role to acidify the lysosomes and make prey digestion more efficient. Thus, there was a synergy between the lysosomes providing a rich supply of sugars, amino acids and fatty acids for cell growth and the mitochondria using their breakdown products to generate most of the cell's ATP. During this early co-evolution of the cell's digestive system and new power plant, there would have been a burgeoning of new transport proteins in both endomembranes and the mitochondrial envelope to exchange numerous intermediary metabolites.

The evolution of coated-vesicle budding created endomembranes that are topologically discontinuous from the plasma membrane – not continuous with it, as Robertson (1964) mistakenly thought and generations of textbook writers ignorant of the meaning of 'topology' copied: the correct way of saying it is 'developmental precursors of', not 'topologically continuous with'. Cytosis (like cell division) is a topology 'breaker'. The ER lumen is also topologically discontinuous from both cytosol and the external medium. Small GTP-binding proteins of the Rab/Ras/Rho superfamily play key roles in cytosis (e.g. Sar1 in erycytosis); it is possible that they evolved from a clearly related myxobacterial small GTPase (Hartzell, 1997), as Moreira & López-García (1998) suggest, in which case lateral gene transfer played a central and early role in the origin of the endomembrane system; before this can be accepted, we need to establish whether any actinobacteria also have such small GTPases – none has been identified in any archaeobacteria but, as homologues are known from the ϵ -proteobacterium *Aquifex* and the Hadobacteria, they are not diagnostic for δ -proteobacteria and need not have come from them.

Free-living bacteria never harbour living symbionts topologically inside their cytoplasm. Their inability to harbour endosymbionts probably arises partly because almost all bacteria have cell walls and most are too small to harbour other cells, but mainly, I think, because they have no phagocytic machinery to engulf other cells. The numerous hypotheses that assume that one bacterium was the host for an endosymbiont before the beginnings of phagotrophy (e.g. Margulis, 1970; Martin & Müller, 1998) are all mechanistically implausible; to suggest that a bacterium with a normal wall engulfed another cell 'by an unspecified mechanism' (Vellai *et al.*, 1998) is magic, not science. Though conversion of a wall to a flexible coat must have preceded the origin of phagocytosis and the symbiogenetic origin of mitochondria, even a very primitive form of phagotrophy could have led quickly to the uptake of a foreign cells. Indeed, if lysosomal fusion was still inefficient, it would be even easier for engulfed prey to grow enough to burst

the food vacuole before fusion and multiply in the cytoplasm. Often, this would be lethal to the host but, if it became able to tap the symbiont's energy supply, it could enslave it. Gene transfer to the host could begin even before the nucleus was formed; functionally trivial gene replacement of soluble cytosolic enzymes (or beneficial replacement of unduly heat-dependent ones on the present version of the neomuran theory) would be mechanistically much easier than retargeting proteins encoded by transferred genes back into the symbiont.

Martin (1999) suggested recently that the endomembrane system might have arisen not by budding from the plasma membrane but by *de novo* spontaneous assembly of phospholipids in the cytosol made by enzymes donated to an archaeobacterial host by a proteobacterium. An origin by self-assembly is biophysically implausible, as it ignores the problems of the insertion of a functional spectrum of membrane proteins into a novel liposome-like membrane and does not address any of the aspects of membrane evolution and differentiation mentioned above. Not only does Martin's hypothesis not explain the origin of the endomembrane system satisfactorily, but the phylogenetic evidence discussed above tells us that lipid replacement did not actually occur. It is highly probable that the endomembrane system is genetically continuous with the plasma membrane of the bacterial ancestor and that membranes have never arisen *de novo* since the origin of the first pre-cellular ones, as Blobel (1980) and I (Cavalier-Smith, 1987a, c, 1991a, b, 1992a, 2001) have argued. The enzyme that makes acyl ester phospholipids (glycerol-3-phosphate acyltransferase) is itself an integral membrane protein, conserved between eubacteria and eukaryotes and absent from archaeobacteria. Because of the conserved targeting properties of such a membrane protein, it and the phospholipids it made would be inserted into existing membranes and therefore would not generate internal membranes *de novo*. On my interpretation, the originally actinobacterial acyl transferase was selectively excluded from the ancestral eucyctic coated vesicle, making growth of the plasma membrane thereafter dependent on exocytosis, later supplemented by phospholipid exchange proteins, which might first have evolved to allow the outer mitochondrial membrane to grow using lipid made in the ER.

Autogenous origin of peroxisomes

I now revert to the idea that peroxisomes evolved by differentiation from the ER (Cavalier-Smith, 1975) and abandon later suggestions of a symbiogenetic origin (De Duve, 1982; Cavalier-Smith, 1987e, 1990a). This change is prompted primarily by recent evidence that they can arise *de novo* (South & Gould, 1999), as well as other new evidence cited by Martin (1999) for an involvement of the ER in their biogenesis. The primary function of peroxisomes was, I suggest, β -oxidation of fatty acids produced by phospholipid digestion in the lysosomes. Because of their energy-

richness, efficient use of fatty acids would have been a powerful selective force for the separate compartmentation of the enzymes that break them down into acetate for fuelling the mitochondrion. Oxygen consumption was just a necessary input for this, not the primary 'detoxifying' reason for the evolution of peroxisomes, as sometimes suggested. The immediate product, hydrogen peroxide, is more toxic than oxygen itself; its generation as a harmful by-product of β -oxidation was alleviated by the co-localization of catalase to destroy it.

Because of the complexity of peroxisome biogenesis and protein-targeting, a detailed autogenous theory of their origin must be deferred to a separate paper. If peroxisomes themselves are not of symbiotic origin (even now some possibility remains that they were in part), they probably acquired the key enzyme catalase from their neomuran ancestor; although archaeobacteria have been stated to lack catalase (Gupta, 1998a), some actually have it, so it is likely to have been present in the neomuran ancestor. If we now accept the partially overlapping origins of mitochondria and the endomembrane system, this eliminates the idea that peroxisomes preceded mitochondria as a primitive respiratory organelle (De Duve, 1969); instead, their evolution was simultaneous and synergistic (Cavalier-Smith, 2000b). Since the ER is also a respiratory organelle (in the sense of having its own respiratory chain and consuming oxygen, even though not an ATP generator), it is highly probable that the ancestral eukaryote host was a facultative aerobe and that the metabolism of all three respiratory organelles co-evolved. Peroxisome respiration provided acetate to mitochondria, whose respiration yielded ATP; ER respiration was primarily for lipid biosynthesis – oxidation to make sterols and dehydrogenation to make unsaturated fatty acids, probably important for the membrane fluidity on which cytosins depends.

However, this metabolic symbiosis did not initially involve chloroplasts, as I once postulated (Cavalier-Smith, 1987e), for I am now reasonably confident that chloroplasts evolved significantly later (see below), as held by classical serial endosymbiosis theory (Taylor, 1974); thus, photorespiration also evolved later in the ancestral plant. However, having three oxygen-consuming organelles could have enabled these early phagotrophs to harbour cyanobacteria and to tap their photosynthate without converting them into plastids. I suggested previously that many of the so-called phytoplankton in the late Precambrian fossil record might actually be pseudophytoplankton, exploiting the fixed carbon from intracellular cyanobacteria without actually converting them to organelles (Cavalier-Smith, 1990b). Even today, such pseudoalgae can be important primary producers. They probably were much more so prior to the origin of plastids. As I stressed previously (Cavalier-Smith, 1983a, 1987c), the co-cultivation of a proteobacterium and a cyanobacterium by an early eukaryote would have been synergistic, with one providing the CO₂ and the other

the O₂ that its partner needed. Such an intracellular synergy could have contributed greatly to the success of early eukaryotes and is much more plausible as a precursor to real integration than the analogous extracellular synergy (syntrophy) suggested in recent non-phagotrophic hypotheses (Martin & Müller, 1998; Moreira & López-García, 1998); the later origin of plastids is therefore in no way inimical to my thesis that such intracellular metabolic synergy might have played a key role in the symbiogenetic origin of mitochondria and contributed generally to the ecological success of early phagotrophs. As Finlay *et al.* (1996) have shown in modern ecosystems, such consortia can broaden the ecological tolerance of the host.

The symbiogenetic origin of mitochondria

Phagotrophy, helotism, membrane heredity and the origin of mitochondria

As argued previously (Cavalier-Smith, 1983a), the key steps in the origin of mitochondria were fivefold:

(i) Uptake by phagocytosis of a facultatively aerobic α -proteobacterium into the food vacuole of a facultatively aerobic heterotrophic host.

(ii) The accidental breakage of the food-vacuole membrane to liberate the bacterium to multiply freely in the cytosol. The mitochondrial outer membrane is thus derived from the outer membrane of the proteobacterium (Cavalier-Smith, 1983a) (and is traceable back even into pre-cellular evolution, a prime example of membrane heredity over billions; Cavalier-Smith, 2001), not from the food-vacuole membrane, as Schnepf (1964) proposed before the nature of the outer membrane was appreciated. In view of the daily escape of modern symbionts from the food vacuole and the lack of an evident mutational mechanism for losing the outer membrane, it is remarkable that some authors persist (e.g. Rizzotti, 2000) with the mechanistically untenable (and totally unnecessary) assumption that the outer membrane was lost and replaced by the food-vacuole membrane. I assert, contrariwise, that the negibacterial outer membrane was lost only once in the entire history of life – during the origin of the Posibacteria (Cavalier-Smith, 1980, 1987a, 2002a).

(iii) Insertion into the proteobacterial inner membrane of a host-encoded ADP/ATP-exchange protein that allowed the host to extract the proteobacterium's ATP (John & Whatley, 1975). This made it an energy slave of the host. This symbiosis was not mutualism but helotism (enslavement) from the moment this happened.

(iv) Evolution of a generalized protein-import mechanism that allowed any host proteins and any symbiont proteins encoded by gene copies transferred to the nucleus to be imported into the former proteobacterium merely by acquiring a topogenic N-terminal pre-sequence. The moment the Tim/Tom machinery for import evolved, the symbiont was an organelle (Cavalier-Smith & Lee, 1985). This most complex step

was discussed in logical outline as a modification of the proteobacterium's protein translocation system by Cavalier-Smith (1987e) and Pfanner *et al.* (1988), but could now be treated in much more detail; in particular, it appears that the mitochondrion retained the YidC but not the Sec protein-insertion pathway (Stuart & Neupert, 2000; Hell *et al.*, 2001), whereas the chloroplast retained both but lost the SRP RNA.

(v) The generalized protein-import mechanism allowed the rapid transfer to the nucleus of most genes encoding proteins that could easily be retargeted to the mitochondrion by adding N-terminal pre-sequences, which would have been relatively simple (Baker & Schatz, 1987) compared with the complex origin of the Tim/Tom translocation machinery. In response to selection to increase the fraction of the protomitochondrion packed with tricarboxylic acid cycle and other useful enzymes rather than redundant DNA, and to maximize energy and nutrient economy (Cavalier-Smith, 1987e), the proteobacterial genome of around 1600 proteins (or more) was thereby rapidly contracted to about 100 protein genes, if we suppose the *Reclinomonas* mitochondrial genome (Lang *et al.*, 1997) to be a good indicator of that of the cenancestor.

The selective advantage to a host that lacked oxidative phosphorylation of enslaving a facultatively aerobic proteobacterium would have been tremendous. Oxidative phosphorylation by mitochondria yields 34–36 moles of ATP per mole of glucose, whereas glycolysis yields only 2 moles of ATP; this 17- to 18-fold increase in the efficiency of using food translates into several-fold faster cell reproduction (Fenchel & Finlay, 1995). If the host was capable only of glycolysis, as initially assumed (John & Whatley, 1975), but was a facultative aerobe/anaerobe (as first stressed by Hall, 1973), the acquisition of a facultatively aerobic α -proteobacterium plus the ability to tap its ATP supply generated by oxidative phosphorylation would have given this early eukaryote an overwhelming selective advantage over its facultative aerobic congeners that lacked such an ability. However, on the now revised neomuran theory, the host probably already possessed oxidative phosphorylation, so the benefit would be less striking, but very significant.

Compartmentation, evolutionary constraints and mitochondrial origins

The benefits of cell compartmentation, explained previously in the context of the since-disproved autogenous origin of mitochondria and chloroplasts (Cavalier-Smith, 1975, 1977, 1980), are threefold. Two arise from confining a water-soluble enzyme or metabolic pathway to just a small fraction of the cell's volume. This means that a given concentration of enzyme and substrate can be achieved with many fewer copies of each per cell. Since concentration usually determines the rate of production of a product, this is a key advantage to the cellular economy. Since mitochondria occupy about a tenth of the cell, this would give a

tenfold saving in the material necessary to achieve a given concentration. Within each compartment, the maximum concentration possible for a protein is limited and probably cannot exceed about 25% in the general cytoplasm, but can rise to about 45% within the almost solid matrix of an organelle such as a mitochondrion or peroxisome, where cytoplasmic motility is not needed. Given such a limit, the maximum concentration achievable by an individual enzyme depends on the complexity of the protein mixture present. By limiting this to a small fraction of the proteome, the concentration of each can be absolutely higher than would be possible in an uncompartmented cell. Acquiring a protomitochondrion symbiogenetically enabled the early eukaryote to increase the concentration of its pre-existing tricarboxylic acid cycle enzymes severalfold more than would have been possible for its neomuran ancestor. By placing its lipid-oxidizing enzymes in separate, autogenously evolved peroxisomes, it got two respiratory organelles that enabled it to extract food from prey far more efficiently than would have been possible if both sets of enzymes had remained in the cytosol. Segregating the glycolysis enzymes within a separate compartment would also have been advantageous for the ancestral eukaryote, but it did not manage to achieve this. Only protozoa of the euglenozoan class Kinetoplastea ever succeeded in doing this, and evolved glycosomes by modifying peroxisomes (Cavalier-Smith, 1990a). This example nicely refutes the naïve selectionist view that properties will necessarily evolve if they have a sufficient selective advantage. This is only true if the necessary intermediary steps are achievable by a small number of relatively common mutations that are mostly individually advantageous. In practice, evolutionary constraints can be serious. Segregating a whole pathway of many components within a novel compartment is not easy, since it would usually be disadvantageous to segregate one or two of them alone; thus, symbiogenesis can be advantageous even if it does not provide any qualitatively novel properties, merely by providing in a single step a distinct genetic compartment already carrying the whole pathway – individual symbiont components can then easily be replaced by host enzymes (Cavalier-Smith, 1990a).

The third advantage of compartmentation involves the lipid-soluble enzymes of membranes. Putting membrane enzymes that catalyse different functions that do not interact directly on segregated domains within a single membrane or else on topologically distinct membranes concentrates each of them, speeding up the overall reaction. I suggest that the α -proteobacterium was initially a facultative phototroph with inner-membrane invaginations (chromatophores) like those of non-sulphur purple photosynthetic proteobacteria that are made only in the absence of oxygen; a key step in making an efficient mitochondrion would have been the formation of mitochondrial cristae packed with respiratory assemblies under aerobic conditions when photosynthesis was suppressed and eventually lost.

This dense packing of respiratory assemblies into these specialized cristae and of tricarboxylic acid cycle enzymes in the matrix would have allowed far more efficient use of the energy locked in the now super-abundant food provided by phagotrophy than would have been the case if the host had retained its original oxidative phosphorylation machinery in the relatively low-surface-area plasma membrane, which also had to accommodate proteins for many other functions, e.g. those involved in anchoring the cytoskeleton and in phagocytosis. Even the new endomembrane system, with potentially much greater surface area, could not have been as efficient an oxidative phosphorylation system, with its conflicting functions of glycosylation and digestion, as could the mitochondrial cristae. Most purple bacterial chromatophores are tubular; only a minority are flat. Whether this difference is relevant to the initial form of mitochondrial cristae is unclear, as aerobic conditions suppress chromatophores. Possibly, cristae arose *de novo* in two alternative forms as the ancestral non-cristate eukaryote diverged to form the ancestors of opisthokonts (originally with flat cristae) and antero-konts (originally with tubular cristae) – see below.

Thus, on the neomuran hypothesis, it was not the all-or-none addition of oxidative phosphorylation to a cell previously devoid of it (Margulis, 1970) that was the driving force for the establishment of the mitochondrion, but the greater efficiency through division of labour (Ferguson, 1767) that compartmentation and functional specialization allowed – precisely the selective forces assumed by the classical autogenous theory (Cavalier-Smith, 1975, 1977). Thus, the origin of mitochondria did not involve any radical shift in metabolism but a smooth transition driven by selection for more efficient utilization of the phagocytosed prey. Since the more complex spore-forming actinobacteria likely to be most related to eukaryotes are all strongly aerobic, though tolerant of temporary anaerobic conditions, and sterol biosynthesis requires molecular oxygen, it is highly probable that the host for the mitochondrial symbiogenesis was perfectly capable of oxidative phosphorylation and had a complete tricarboxylic acid cycle. Contrary to what is widely assumed, the important thing about the symbiotic origin of mitochondria was not the acquisition of novel genomes, genes, enzymes or biochemical or physiological mechanisms, but getting a novel structure that enabled a cell to do what it was doing already but much more efficiently, by bypassing the evolutionary constraints that make it hard to evolve compartmentalization autogenously. In evolution, structure matters far more than most biochemists realize. It was otherwise for the symbiotic origin of chloroplasts, where the previously heterotrophic eukaryotes were able to acquire a novel mode of nutrition by acquiring numerous novel catalysts: the traditional view of Mereschkovsky (1905) as to its selective advantage was entirely correct, but, even here, the acquisition of novel genetic membranes by a pre-existing structure (Sonneborn, 1963) in

a single step was just as important as getting novel genes and enzymes; without the novel structure, the genes and enzymes would have been totally useless (Cavalier-Smith, 2002a). Although the present phagotrophy theory involves an element of symbiogenesis, it would be misleading to call it a symbiotic theory of the origin of eukaryotes. It is fundamentally autogenous, but with symbiogenesis playing the crucial role in the evolution of compartmentation for one of the many new organelles: the mitochondrion only. Symbiogenesis was also incidentally quantitatively useful to the cell as a whole, but played no part in the origin of the other major qualitative changes in cell organization that make the distinction between bacteria and eukaryotes the most important in the living world. It is, therefore, radically different from the views of Margulis *et al.* (2000), who fail entirely to understand or to explain these features.

It is interesting to note that the negibacterial double envelope pre-adapts bacteria for the development of membrane invaginations to allow more intensive energy metabolism. This happened independently in proteobacterial and cyanobacterial photosynthesis (where the invaginations soon became distinct thylakoids) and in methylotrophic negibacteria. In contrast, unbacteria (posibacteria and archaeobacteria) seem limited in their capacity for such differentiation by their single bounding membrane. However, only a unbacterium with a single bounding membrane could have evolved phagotrophy and the endomembrane system. The partially symbiogenetic origin of eukaryotes from a unimembranous bacterial host (a stem neomuran of actinobacterial ancestry) and an α -proteobacterial symbiont allowed the first eukaryote to combine these otherwise incompatible advantages of the negibacterial and unimembranous condition within a single chimaeric cell: the eukaryote.

Criticisms of the phagotrophy theory of mitochondrial symbiogenesis are invalid

In presenting an alternative 'hydrogen hypothesis' for mitochondrial symbiogenesis, Martin & Müller (1998) caricatured the classical phagotrophy theory of mitochondrial symbiogenesis, claiming it 'carries several tenuous assumptions'. The three listed by them either are not tenuous or else are not assumptions of the theory. Firstly, 'the host was unable to synthesize sufficient ATP by itself'. If the host were purely glycolytic, as traditionally assumed (Stanier, 1970; Cavalier-Smith, 1983a), there would have been a tremendous gain in energy and growth efficiency that is not in the least tenuous. Even if it had some sort of oxidative phosphorylation, as is highly probable, it would also have gained substantially in efficiency if, like many bacteria, it had a lower energy yield than α -proteobacteria. Even if the energy yield per ATP synthetase was identical, there would have been a substantial gain in efficiency of use of food energy and speed of energy generation through compartmentation and specialization, as explained above. As population

geneticists well know, even a 0.01 % gain in efficiency expressed as a proportionate increase in growth rate would be more than sufficient selective advantage to drive a novel mutation to fixation. The actual benefit must have been orders of magnitude greater. The second and third 'assumptions', that 'the symbiont synthesized ATP in excess of its needs' and that the 'symbiont could export ATP to its environment, so that the host could realize this benefit', are absurd ideas that have never been assumptions of any sensible statement of the classical theory. Like many authors, they seem to assume that mitochondriogenesis was an altruistic mutualism, but there is no reason to suppose that the symbiosis was mutualistic. Helotism and parasitism are far commoner in nature than syntrophy.

The classical heterotrophic host theory not only has a cast-iron selective advantage for the enslavement of bacteria, but is more plausible ecologically and much more so cytologically than the hydrogen hypothesis, which suggests instead an obligately anaerobic, hydrogen-producing autotroph like a methanogenic archaeobacterium. A chemoheterotrophic facultative aerobe, whether purely a glycolytic posibacterial chemotroph, as in the classical phagotrophy theories, or one with oxidative phosphorylation (possibly less efficient than mitochondria), as in the revised neomuran theory, is by far the most plausible host for the mitochondrial symbiosis. If the host secreted abundant digestive exoenzymes, as do many posibacteria, it was also pre-adapted for the origin of phagotrophy and the endomembrane system, which methanogens would not have been. When both host and symbiont are facultative aerobes (of which there are many examples), symbiogenesis is ecologically more plausible than the assumption by the hydrogen hypothesis of one between a facultative aerobe and an obligate anaerobe (no examples are known: the symbioses cited by Martin & Müller (1998) are between methanogens and other anaerobes and so are not strictly relevant to their hypothesis); in the hydrogen hypothesis, the postulated driving force of the anaerobic utilization of symbiont H_2 is only temporary and later replaced by a different force (the utility of oxidative phosphorylation, with the attendant difficulty, recognized by the authors, of how aerobic enzymes would have been retained during the earlier anaerobic phase that they unnecessarily postulate). This tortuous replacement of a purely hypothetical initial selective advantage by a later well-established one is a much less straightforward explanation than the classical hypothesis, which assumes only the incontrovertible selective advantage of efficient, compartmented oxidative phosphorylation to a facultatively aerobic and phagotrophic host.

The statement of Doolittle (1998b), that the hydrogen hypothesis accounts more naturally for the presence of eubacterial genes in the nucleus, is unwarranted. The explanation on the classical phagotrophy theory is exactly the same; the neomuran hypothesis in its present form even gives a mild selective advantage for

the substitution of host enzymes. It is also unfortunate that Martin & Müller (1998) called the classical phagotrophy theory the 'Archezoa hypothesis', for two reasons. Firstly, there is no necessary connection between the phagotrophy theory in general and any particular host; the mechanisms of uptake and the basic selective advantages and mechanisms apply to any potential host capable of phagotrophy. Secondly, what Doolittle (1998b) dubbed the 'Archezoa hypothesis' embodied two logically distinct types of hypothesis: (i) a taxonomic hypothesis (implemented by a classificatory act; Cavalier-Smith, 1983b) about relationships among amitochondrial phyla and (ii) an evolutionary hypothesis that some of them, at least, were primitively amitochondrial and related to the host that acquired mitochondria (Cavalier-Smith, 1983a). The latter has been proved false, and has been abandoned by me for several years (Cavalier-Smith, 1998a, b) [but, sadly, not by Margulis *et al.* (2000)]; the former lives on in modified form as the superphylum Archezoa, comprising Metamonada and Parabasalia (Cavalier-Smith, 1998a), a taxon that may yet prove to be holophyletic.

Martin & Müller (1998) implied that the phagotrophy theory is less powerful than it is by asserting that 'the Archezoa hypothesis ... cannot directly account for' the absence of eukaryotic cell structures in archaeobacteria. In fact, the 'neomuran' phagotrophy theory (Cavalier-Smith, 1987c) explained explicitly why this is so. It proposed that archaeobacteria and eukaryotes are sisters and that both evolved from the same wall-less mutant posibacterium; it argued that loss of the wall was the crucial prerequisite for the origin of phagotrophy, which, in turn, caused the origins of the endomembrane system, cytoskeleton and nucleus, and was a prerequisite for the symbiogenetic origin of organelles. I argued explicitly that the key reason why archaeobacteria, the sisters of eukaryotes, did not evolve such structures is that the archaeobacterial ancestor re-evolved a cell wall (with novel chemistry) and this directly prevented the origin of phagotrophy and thereby eukaryotic structures. On the present theory, the fact that all archaeobacteria except *Thermoplasma* retained a cell wall, and so none have evolved phagotrophy, is sufficient to explain the absence of eukaryotic structures, if you accept that phagotrophy was the prime cause of their evolution.

Weaknesses of the hydrogen and syntrophy hypotheses

The hydrogen hypothesis (Martin & Müller, 1998) assumes, on the contrary, that an archaeobacterial host gradually engulfed the α -proteobacterial symbiont, even though no bacteria have ever been observed to engulf other cells and have no known mechanism for doing so. The classical theory assumes that the α -proteobacterium was taken up by phagocytosis, the same mechanism almost certainly used in all other cases of symbiogenetic origin of organelles (Cavalier-Smith, 2000b) and which occurs daily by the billion in

extant protozoa. This assumption of an unknown mechanism for engulfment by the hydrogen hypothesis is a serious defect, making it much less plausible than the phagotrophic-host theory. The syntrophy hypothesis (Moreira & López-García, 1998) also does not provide a physical mechanism for cell engulfment, which is essential for symbiogenesis.

Another serious weakness of the hydrogen hypothesis is the assumption that the host was an obligately anaerobic autotrophic bacterium, not facultatively aerobic and heterotrophic as in the phagotrophy theory. Since the ancestral state for eukaryotes is undoubtedly heterotrophy, the hydrogen hypothesis is driven to postulate a very complex and comprehensive replacement of host metabolism, first by acquiring genes from the α -proteobacterial symbiont for plasma-membrane transporters of organic molecules and for heterotrophic and aerobic intermediary metabolism. Such a wholesale transformation of the metabolism of the host from an autotroph to a heterotroph is much less plausible than the simple substitution of host heat-adapted soluble proteins by symbiont ones catalysing the same reactions. The hydrogen and syntrophy hypotheses also both require the replacement of the archaeobacterial lipids by those from the symbiont and the acquisition from it of all the dozens of proteins present in eubacteria but not archaeobacteria, e.g. Hsp90, plus the unsplitting of the RNA polymerase A and glutamate synthetase genes. Since methanogens are invoked as the archaeobacterial host by both hypotheses and since all studied methanogens have an extra split in the RNA polymerase RpoB gene absent from more primitive archaeobacteria and eubacteria, this split would also have had to be reversed, unless an earlier-diverging methanogen (e.g. *Methanopyrus*) is found to lack it.

Overall, the hydrogen hypothesis is far more complex than the phagotrophic-host theory; even so, it does not explain how a methanogen's wall could become a surface coat, nor explain the origin of the endomembrane system, the real crux of the problem of eukaryogenesis. The 'prediction' of the hydrogen hypothesis (Martin & Müller, 1998) that more archaeobacterial-like genes should be replaced by eubacterial ones in photosynthetic eukaryotes is not specific to the theory, but a logical consequence of the general principle that gene transfer to the nucleus from a symbiogenetic organelle need not be accompanied by the evolution of targeting of the protein to the original organelle, first discussed in detail in connection with the theory, herein set aside, of the symbiogenetic origin of peroxisomes (Cavalier-Smith, 1990a).

Another major problem with the hydrogen hypothesis and other syntrophy hypotheses (Moreira & López-García, 1998; Margulis *et al.*, 2000) is that all ignore the evidence for the presence in actinobacteria of many characters related to those of eukaryotes that are not present in archaeobacteria (Table 1). These are accounted for simply by the neomuran theory of eukar-

yote origins; any theory assuming a direct archaeobacterial ancestry would have to dismiss these similarities as convergences or postulate that chitin, sterols, calmodulin, histone H1 and serine/threonine kinases were all transferred to the prekaryote by lateral gene transfer. It is so much simpler to assume that the ancestor was a derivative of an actinobacterium that had not yet lost these characters, in contrast to the stem archaeobacterium that did. The strength of the neomuran theory is not only that it provides a much simpler explanation for the origin of eukaryotes than any of its competitors, but that it also explains the origins of archaeobacteria in great detail (Cavalier-Smith, 2002a) and, unlike all competing theories, takes full account of and is compatible with the fossil record (Cavalier-Smith, 2002a).

Origin of the nucleus: co-evolution with mitosis

Mereschkovsky (1910) suggested that fungal nuclei evolved autogenously but others evolved from symbiotic bacteria eaten by a mythical anucleate moneran amoeba. In a veritable spate of superficial speculation, Schubert (1988), Sogin (1991), Lake & Rivera (1994), Gupta & Golding (1996), Gupta (1998a), Moreira & López-García (1998) and Margulis *et al.* (2000) have echoed, usually unwittingly, Mereschkovsky's defunct hypothesis that nuclei had a symbiotic origin. It is odd that Rizzotti (2000) takes such ideas seriously, given their cell-biological naïvety. As Martin (1999) correctly argues, none are proper theories; they all lack understanding of nuclear structure and function and fail to suggest any comprehensible steps via functional intermediates by which a symbiont could become a nucleus. All seem unaware of the classical autogenous explanation for the origin of the nucleus by the aggregation of primitive ER cisternae around the DNA (Cavalier-Smith, 1975; Taylor, 1976) or more recent detailed treatments in which the key problem is seen to be the evolution of nuclear pores and the nuclear import and export process (Cavalier-Smith, 1987c, 1988a).

Martin (1999) criticizes past autogenous hypotheses for demanding 'the presence of a cytoskeleton in the prokaryotic ancestor of eukaryotes'. Like his misrepresentation of the classical interpretation of the symbiogenetic origin of mitochondrion, this criticism is simply wrong. Neither Taylor nor I assumed a cytoskeleton in the bacterial ancestor. I have repeatedly argued explicitly that the origin of a cytoskeleton in a bacterium that previously had none was the key set of molecular innovations that led to phagotrophy, the endomembrane system, the nucleus and the cilium (e.g. Cavalier-Smith, 1975, 1980, 1981a, 1987c, 1991a, 1992b). Why are some biochemists so reluctant to accept the radical implications of real innovation or quantum evolution that they think that others have not done so? Martin (1999) also wrongly asserts that his hypothesis for the origin of the endomembrane

system by *de novo* membrane assembly, argued against above, differs from classical ones in invoking processes of nuclear envelope assembly by vesicle fusion; but this is far from new, as vesicle fusion was the very agent of nuclear envelope assembly on the classical theory (Cavalier-Smith, 1975, 1980, 1987c, 1991a). Martin writes as if invagination and vesicle fusion are opposed. They are not; on the classical theory, invagination preceded budding off the plasma membrane to form vesicles, which then fused. Admittedly, there are also naïve diagrams in textbooks and elsewhere that show invaginations becoming nuclear envelopes without intervening budding of vesicles. But these are not accompanied by well-argued explanations of how they might be perpetuated through the cell cycle and are best ignored.

The primary selective advantage of the assembly of a nuclear envelope on the chromatin surface was probably to protect the DNA from shearing damage caused by the novel molecular motors, myosin, dynein and kinesin, involved in phagotrophy, cytokinesis and vesicle transport (Cavalier-Smith, 1987c). The folding of the DNA as an interphase skeleton (Cavalier-Smith, 1982a) and the even more compact folding in mitotic chromosomes would have helped avoid breakage and were furthered by the addition of histones H2a and H2b to the H3 and H4 histones that evolved in the neomuran ancestor and the H1 that was already present in the earlier actinobacterial ancestor. Reversible histone acetylation to mediate mitotic compaction must have arisen in the stem eukaryote or prekaryote, possibly before the nuclear envelope.

A key feature of the origin of the eukaryotic cell from a bacterial ancestor, sadly often ignored, is how bacterial mechanisms for DNA replication and segregation, cell division and cell-cycle controls have been converted into eukaryotic ones. It was not a static bacterium that had to evolve into a static eukaryote but a bacterial cell cycle that had to be converted into a eukaryotic cell cycle while maintaining viability throughout, as I have attempted twice to explain, first assuming the ancestral eukaryote to have been a non-flagellate fungus (Cavalier-Smith, 1980) and later assuming it to be a zooflagellate protozoan (Cavalier-Smith, 1987c), as I still consider most likely (Cavalier-Smith, 2000a). The origin of the nucleus is inseparable from the origin of mitosis. If phagocytosis internalized the membrane regions to which the DNA of the bacterial ancestor was attached, there would be problems for accurate DNA segregation unless a new segregation mechanism (mitosis) based on microtubules evolved. The central logic of my arguments that centrosomes and spindle microtubules must have evolved near the very beginning of eukaryote evolution and that mitosis must have evolved by modifying the bacterial segregation system remains. What has changed is that we understand bacterial DNA segregation and division and mitosis a little more, so we can say more about the nature of the precursors and the product.

Firstly, as I hypothesized (Cavalier-Smith, 1987b), bacterial DNA segregation is not a purely passive process but involves active motors (Sharpe & Errington, 1999), like mitosis. Posibacteria have a family of proteins like the eukaryotic SMC (structural maintenance of chromosomes) proteins that are required for DNA segregation like their eukaryotic relatives; not surprisingly (given my thesis that posibacteria are more related to neomura than are negibacteria), they are more similar to the eukaryotic ones than are the negibacterial equivalents, MukB; all are needed for DNA segregation. The myosin-like large bacterial MukB/SMC protein may be a bacterial motor protein for segregation, and thus an ancestor of myosin, dynein and kinesin, or may be involved instead in chromatin condensation. However, the actual nature of the bacterial motor is unknown; the possibility that it was a DNA helicase instead (Cavalier-Smith, 1987b) is open. Secondly, the bacterial division GTPase FtsZ, which forms a contractile ring and mediates bacterial cell division and the division of chloroplasts and the more primitive mitochondria (Beech & Gilson, 2000), is clearly the ancestor of tubulin, which, as outlined above, could have evolved as soon as the actomyosin contractile ring took sole responsibility for cytokinesis. The great complexification in subunit composition of the eukaryote chaperonin CCT (Cpn 60) compared with its archaeobacterial sister (Archibald *et al.*, 2000) can be attributed to its novel functions as a tubulin chaperone. Though, as outlined above, early centrosomes recruited two proteobacterial chaperones, Hsp90 and Hsp70, the evolutionary origin of other centrosomal proteins, e.g. Ranbpm (almost certainly present in the ancestral centrosome, as it is retained even in cryptomonad nucleomorphs; Zauner *et al.*, 2000), is unclear.

I have argued that the evolution of mitosis and the origin of the nucleus set in train a whole raft of changes to the genome that quite rapidly converted bacterial chromosomes into the substantially different eukaryotic ones (for details, see Cavalier-Smith, 1980, 1987c, 1993a). Thus, genomic change followed rather than led cytological change, contrary to widespread preconceptions. The origin of the centrosome played a key role in the origin of eukaryotes because of its centrality to their mitotic and cell-division mechanisms. Thus, conversion of the neomuran cell wall to a surface coat and the origin of phagotrophy were the key stimulants of the origin of the endomembrane system and the cytoskeleton/cilium that set off a cascade of cellular transformations on a scale not seen before or since in the history of life.

Origin of nuclear pore complexes, import and export and the nucleolus

The primary function of early pore complexes was to prevent total vesicle fusion from creating a continuous double membrane that would seal off the nucleus from the cytoplasm (Cavalier-Smith, 1987c) and to allow passive movement of soluble molecules between cyto-

plasm and nucleoplasm. None of the symbiotic or chimaeric fusion theorists of the origin of the nucleus remotely begins to explain the origin of the pore complex (when I criticized one of them for this, he replied blandly 'That's your job!'; but it is not my job to rescue a stupid theory); none even understands the elementary cell-biological fact that the nuclear envelope is not a double array of two topologically distinct and nested membranes, as in mitochondria and most plastids. Instead, it is topologically a single membrane with three locally differentiated domains; the so-called inner and outer membranes are really inner and outer domains of a continuum. After the envelope evolved from the primitive ER, as explained before (Cavalier-Smith, 1987c), the originally very large pores became plugged with a complex proteinaceous machinery to segregate the RNA- and protein-synthesis machinery in separate compartments and to obtain the advantages of compartmenting water-soluble enzymes; logically, the export/import machinery had to have evolved with substantial efficacy before the passive route was totally closed (Cavalier-Smith, 1988a). It seems unavoidable that there was an intermediate stage with both passive and active topogenically directed exchange co-existing, whether through the same or separate pores. The initial selective advantage of the active process may have been simply to accelerate exchange so that it did not limit the rate of growth.

For protein import, the quantitatively dominant and/or most important proteins that had to be imported were the histones, RNA and DNA polymerases and other DNA-binding proteins. This, I think, explains why nuclear localization sequences (NLS) are runs of basic amino acids; these were the most simple pre-existing shared distinguishing features of the nucleic acid-binding proteins, most essential for the functions of DNA, that distinguish them from the majority of cellular proteins. The topogenic machinery evolved to transport them, rather than their becoming adapted to it: NLS were not added to pre-existing proteins, as in the origin of mitochondrial or plastid import. This explains why NLS can be anywhere in the molecule, not just at an end, and why they are not removed after import. Later, other non-basic proteins could be relocated to the nucleus by adding a basic tail. Can we perhaps distinguish, at least partially, primordial proteins from secondarily nuclear ones, e.g. nucleoplasmin, in this way?

But why are ribosomes assembled in the nucleus not the cytoplasm? Perhaps partly because, as ribosomal proteins are rather basic, it would have been hard for machinery to evolve that could import histones without also importing them. Even more important is the long time taken to make each rRNA; by assembling ribosomes in the nucleolus, and evolving a nucleolus rather than a 'cytolus', ribosomal assembly could begin even during transcription and thereby speed growth. Ribosomal RNA transcript cleavage and modification therefore also had to be nucleolar func-

tions. Together with the fact that 18S rRNA is at the beginning of the transcript and much shorter and requires fewer proteins for assembly, simultaneous assembly and cleavage explain why it is exported more rapidly.

Radical eukaryotic innovations in ribosome biogenesis

Origin of 80S ribosomes: quantum evolution in ribosomes

As all ribosomal RNA has a common ancestry, the fact that eukaryotes are probably over four times younger than eubacteria, when contrasted with the uniformity of eubacterial 70S ribosomes and their RNAs and the marked systematic difference from 80S eukaryote ribosomes and their RNAs, proved long ago that rRNA is not a molecular clock (Cavalier-Smith, 1980). The simultaneous origin of mitochondria and eukaryotes also highlights this; over the same time period, mitochondrial ribosomes and rRNAs have diverged less in several respects from their proteobacterial relatives than cytoplasmic ribosomes have from their archaeobacterial ones. Plastids make the point too; they are probably only about 20% younger than mitochondria, but their ribosomes have diverged very much less from their eubacterial ancestors than have mitochondrial ones. However, although plastids are probably over five times as young as cyanobacteria, their rRNAs have diverged from each other nearly as much as the mutual divergences within cyanobacteria. Of course, both mitochondrial and cytosolic rRNA have diverged even more radically in length and sequence in a non-clock-like way within eukaryotes. However, the shared differences of eukaryotic ribosomes from archaeobacterial ribosomes must have arisen in a bout of quantum evolution (Cavalier-Smith, 1981a); why did this occur? In the past, I suggested three reasons, all based on the fact that ribosomes interact with and must co-evolve with other cell structures that also underwent quantum evolution during the origin of eukaryotes.

First is the obvious point that, in eukaryotes, unlike bacteria, ribosomal subunits are assembled in the nucleus and have to be exported through the nuclear pores and therefore need topogenic sequences/binding sites to allow this (Cavalier-Smith, 1975, 1980, 1981a); the ribosomal proteins had to have NLS for import into nuclei. However, I now think that this could have been and probably was achieved with only relatively small changes. Since ribosomal subunits are the dominant cargos for export, they could themselves have largely dictated the nature of the export machinery, rather than the reverse. For import, as mentioned above, their pre-existing basic nature may have pre-adapted them for import, with only minor substitutions needed to increase efficiency.

Second is another obvious point, that, after spliceosomal introns arose, ribosomes had to be prevented

from translating them before splicing (Cavalier-Smith, 1981a). Shine–Dalgarno sequences may have been lost during eukaryogenesis specifically to prevent nuclear messengers from binding directly to nascent ribosomes (Cavalier-Smith, 1981a); this may explain why capping evolved instead (Cavalier-Smith, 1981a). Together with the rapid export of the small subunit, this may be sufficient, so the resulting overall change to the ribosome was probably relatively small. Keeling & Doolittle (1995) pointed out that, in the ancestral eukaryote, there was also a marked change in protein synthesis initiation factors (IF); apparently, a guanine nucleotide-recycling IF related to EF-2 and -G was replaced by a non-recycling type related to the mitochondrial and eubacterial EF-Tu (Keeling *et al.*, 1998). If the tentative suggestion of a possible origin of eIF-2 from mitochondria (Keeling & Doolittle, 1995) was to be confirmed by more extensive analysis, this would be another interesting contribution of the protomitochondrion to host functions. But it is not clear whether such a shift would have had substantial repercussions on rRNA structure.

However, obvious points are not always the most important. Perhaps the main cause of the shift to 80S ribosomes was the more subtle co-evolutionary impact of the origin of the mitochondrion. The contrast between the conserved bacterial-like plastid ribosomes, the significantly changed mitochondrial ribosomes and the most substantially changed cytosolic ribosomes suggests that mitochondrial and cytosolic ribosomes were both strongly selected to diverge from each other and the ancestral bacterial type, whereas such pressures were largely absent from the chloroplast ribosomes. Why should this be? Although the eukaryote ancestor must have retained a fair number of ribosomal protein genes in its mitochondrial genome, a few dozen were probably transferred to the nucleus. This means that they would have been made in the same cytosol as the cytosolic ribosomal proteins, which would potentially cause problems if they became confused with them (Cavalier-Smith, 1993b). Both nuclear import of ribosomal proteins and their exchange with those on the surface of cytosolic ribosomes would have had to be prevented if both types of ribosomes were to function properly. Both were probably selected for divergence, the cytosolic ones changing their surface properties by inserting extra loops in the RNA and adding extra peripheral proteins, which would make it more difficult for the mitochondrial ones to exchange with them, and the mitochondrial ones by losing any sequences confusable with NLS. Both types of change probably caused some co-adaptive change in other parts of the ribosome.

This interpretation of the shift from 70S to 80S ribosomes in the first eukaryote is supported by the fact that all three phyla that have lost mitochondrial ribosomes totally (Microsporidia, Metamonada, Parabasalia) have secondarily truncated their rRNA to about the same length as in bacteria. Microsporidia have re-evolved 70S ribosomes and the Parabasalia

have reduced the number of proteins almost to the same level as in bacteria (Cavalier-Smith, 1993b). At the same time, all three groups have increased the rate of rRNA evolution drastically, making them among the longest branches on the tree (the longest in the case of microsporidia, which shortened them even more than the other two groups, losing some bits conserved even in bacteria). Both the truncation and the extra-rapid evolution imply a marked reduction in selective constraints, which I consider to have followed the loss of mitochondria. It seems to be universally true that loss of mitochondrial ribosomes is associated with marked increases in the rate of cytosolic rRNA evolution, as all amitochondrial taxa have very long branches on rRNA trees. Though they all clearly have greatly reduced stabilizing selection on rRNA, the consequence is not invariably truncation of the whole rRNA gene: archamoebae show a divergent response. In *Entamoeba*, the rRNA is truncated, whereas, in *Mastigamoeba* (*Phreatamoeba*) *balamuthi*, the 18S rRNA gene vastly expanded to become almost the longest known (Hinkle *et al.*, 1994). Because we know nothing of RNA processing in *Mastigamoeba balamuthi*, we do not know if the rRNA itself is longer; it might even be truncated, since the case of the Euglenozoa shows clearly that the gene and the mature molecules are under distinct evolutionary pressures (see next section). The key point is that the evolutionary pressures on ribosomes are very different in the presence and absence of mitochondria.

The origin of mitochondria may therefore be sufficient explanation for the shift from 70S to 80S ribosomes. A corollary is that this briefly accelerated episode of divergent evolution is the cause of the long stem at the base of the eukaryotic part of the 18S rRNA tree; this stem is probably at least a 100 times longer than it would be if it accurately reflected divergence times. This is the most striking of many cases of transiently wildly accelerated quantum evolution in rRNA (Cavalier-Smith *et al.*, 1996a). If eukaryotes are the same age as archaeobacteria (Cavalier-Smith, 2002a), the fact that overall branch lengths within the eukaryote clade are somewhat greater than in archaeobacteria suggests that their average rate of rRNA evolution has been only a little faster, apart from the exceptional unusually accelerated cases like the amitochondrial phyla and foraminifera.

The reason why plastid ribosomes (except in dinoflagellates) have changed so much less than mitochondrial ones, despite their only slightly lesser antiquity, may be that, prior to the origin of plastids, 80S ribosomes had already changed so much that their proteins could not exchange with those of plastids or bacteria. Those that became nuclear-encoded, however, must have lost or never had NLS. Could an inability to lose all NLS be an impediment to successful transfer? Have chloroplasts retained a lot more ribosomal protein genes than most mitochondria because the larger number of encoded genes (except in dinoflagellates) has ensured stronger stabilizing selection

on ribosome structure (including rRNA) than in mitochondria? With a smaller gene complement, stabilizing selection would be weaker on mitochondria and so the ribosomal proteins could lose NLS more readily and perhaps also be modified more readily for easier re-import to the mitochondria than in plastids. Mitochondria with smaller genomes have much more rapidly evolving rRNA; among chloroplasts, the same is seen even more dramatically in dinoflagellates (Zhang *et al.*, 2000).

Co-evolution with mitochondria is not the only reason why eukaryote rRNA trees sometimes reliably give a radically wrong topology. Indeed, it may not even be the only or even the fundamental reason why the ribosome responded to such co-evolutionary pressures by expansion, as I explain next.

Selfish RNA, expansion segments and the origin of cleavage snoRNPs

Disassociation between the evolutionary pressures on the rRNA gene and on the mature RNAs within the ribosome may be as important. Here, the exemplars are the Euglenozoa, which all have unusually long small-subunit rRNA genes (though much shorter than *Mastigamoeba balamuthi*), which evolved by the expansion of many of the less-conserved internal regions in the stem euglenozoan. Underlying this may be a radical change in rRNA processing. In *Euglena* ribosomes, the rRNA is in many short pieces (Smallman *et al.*, 1996). Their summed length is no greater than that of typical 80S ribosomes. This means that the pre-rRNA expansion segments are removed post-transcriptionally and are not functional in the ribosome. The analogy with introns is intriguing and possibly deeper. It is as if a selfish gene encoding its own excision at the RNA level was inserted randomly into the pre-rRNA genes. It cut itself out, but, unlike introns (also of selfish origin), could not splice the bits together. If inserted where the cut did not affect ribosome assembly and function, the cell could survive despite being lumbered with a longer gene and a fragmented rRNA. If the insertion was in a harmful place, it would die.

I suggest that that is the fundamental evolutionary explanation for the expanded length of euglenozoan genes and that it may also be true for the mycetozoan *Physarum* and the Foraminifera. Like the Euglenozoa, their genes have very poorly conserved expansion segments and, like them, their branches are exceptionally long on trees, implying that the evolutionary rates of the ribosomally functional parts of the genes are greatly elevated by the presence of expansion segments. Protein trees show that *Physarum* is falsely grouped with other long branches on most rRNA trees (Fig. 2 is a notable exception that probably puts the Mycetozoa in approximately the correct position); as the rRNA sequences of the Foraminifera are even more bizarre, their grouping with Percolozoa and Archezoa is almost certainly a long-branch artefact.

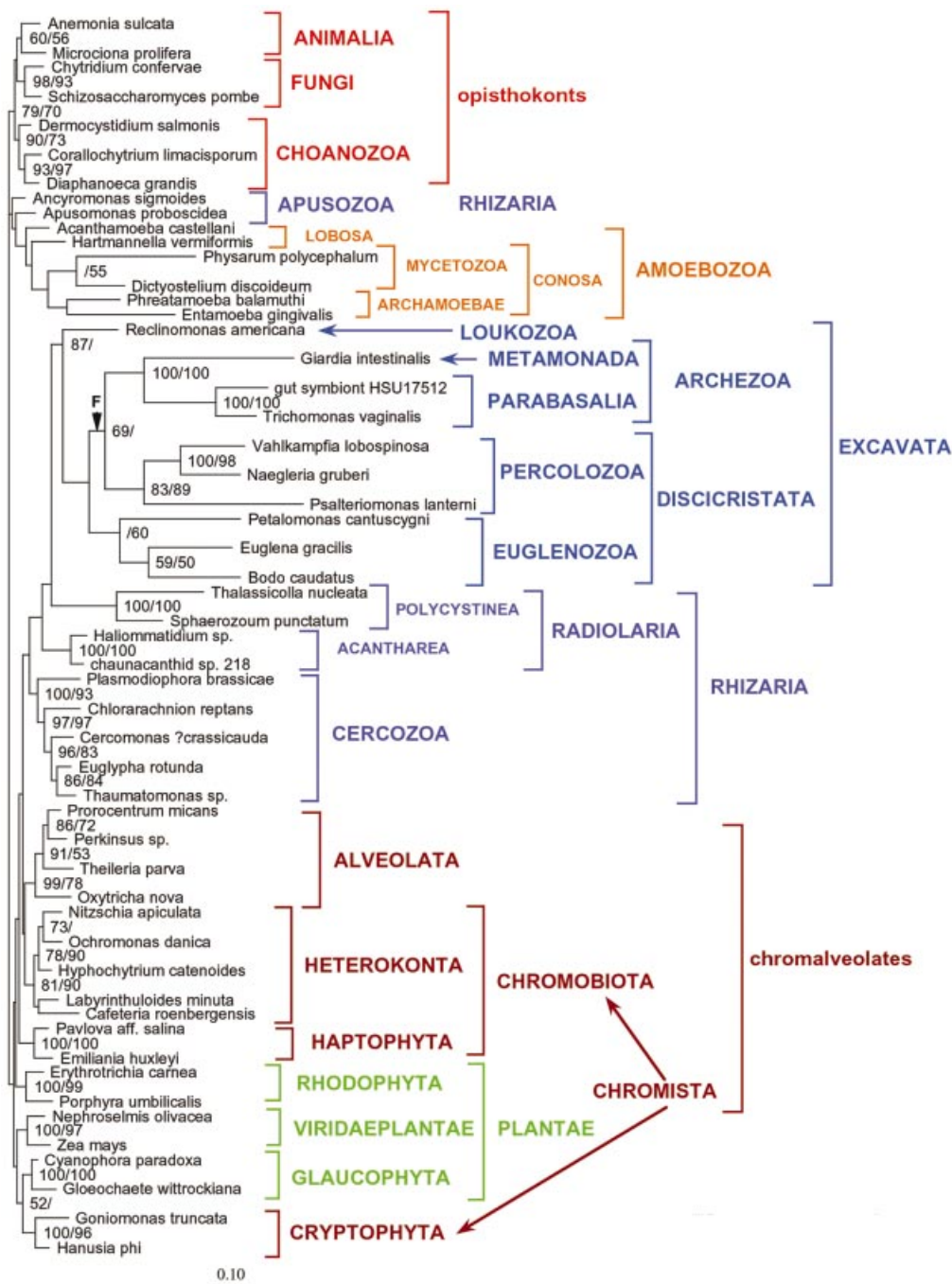


Fig. 2. Maximum-likelihood analysis of 53 nuclear small-subunit rRNA sequences (a representative subset of those shown in the distance tree of Cavalier-Smith, 2000a). From five fastDNAmI analyses (adding taxa in different random orders), this tree had the highest likelihood (ln likelihood = -42739.851). The less likely trees (ln likelihood = -42648.6 to -42755.9) differed primarily in placing Mycetozoa and *Entamoeba* (and once *Phreatamoeba*) within Retaria; *Reclinomonas* was always above Acantharea, but in one (ln likelihood = -42748.6) it was sister to Euglenozoa and, in another (ln likelihood = -42750), it was sister to Euglenozoa and *Giardia*. Bootstrap percentages for 300 pseudo-replicates are shown for separate neighbour-joining (left) and maximum-parsimony (right) analyses. Unbootstrapped parsimony yielded one most-parsimonious tree (10383 steps). Bar, 10% sequence difference. The tree is rooted between opisthokonts and Amoebozoa/bikonta on the assumption that *Phalansterium* and other Amoebozoa are ancestrally uniciliate and unicentriolar and that the root lies between the uniciliate but bicentriolar opisthokonts and the unikont Amoebozoa. Because their sequences are so aberrant and form excessively long branches, Foraminifera were omitted; when included, they branch within Discicristata above Euglenozoa (in the position shown by the arrowhead labelled F), which makes no cytological sense and is probably a long-branch artefact.

All well-studied eukaryotic groups have sporadic examples of one or a few extra cuts in rRNA, especially the longer 28S gene subunit, which has many more poorly conserved regions where such selfish genes could insert with minimal effects. Not only these but the separation of 5.8S RNA from 28S RNA can be explained as the incidental result of the spread of selfish genetic elements able to cleave RNA. If they spread by reverse transcription, they could only do so into transcribed parts of the genome. Organismic selection would ensure that all survivors are found only in parts of transcripts where they do not harm mature function; they would survive most easily in the least essential parts of pre-rRNA genes where breaks can be tolerated, but they could also survive in introns within protein genes if splicing by spliceosomes took precedence over cutting by these hypothetical elements. The small subclass of small nucleolar RNAs (snoRNAs) involved in cleaving pre-rRNA in eukaryotes, but not as far as we know in bacteria, might have originated thus, as I will explain in detail elsewhere (T. Cavalier-Smith, in preparation).

It is possible that such selfish parasites were even the fundamental cause of 70S ribosome expansion to 80S. A sudden burst of insertions could have expanded the rRNA genes; insertions of elements with defective cutting would expand but not fragment them, but others would. In principle, mutations and selection could have eliminated breaks where harmful but not lethal. The microsporidial merger of 5.8S and 28S genes shows that this is possible, but *Euglena* cautions that it may not be easy. Replicational slippage, however, is such an easy way of introducing expansion segments that such insertions might seem unnecessarily complex. However, when genomic evolution is driven by mutation or transposition pressure, unexpectedly bizarre complications like introns or RNA editing do arise (Cavalier-Smith, 1993a).

Origins of marker snoRNPs

Most snoRNAs function not in cleavage of excised regions but to mark positions in the retained rRNA segments by base-pairing so as to create recognition sites for uridylation and methylation by yet unidentified, probably protein enzymes. Methylation and pseudouridylation markers represent two structurally very different families (with C/D- and H/AC-box elements, respectively), each of which contains a minority of the cleavage snoRNAs (Smith & Steitz, 1997). C/D-box snoRNAs bind to the nucleolar protein fibrillarin, which is present in archaeobacteria and associated with rRNA, so they have the methylating part of the nucleolar system (Omer *et al.*, 2000), but it is not known whether any can cleave RNA as a few do in eukaryotes. H/AC-box snoRNAs responsible for marking pseudouridylation sites in 80S ribosomes and additional cleavages are currently unknown in bacteria, and might have arisen as mobile elements in the ancestral eukaryote. If they prove to be present in archaeobacteria, they must have arisen instead in the

neomuran ancestor. In archaeobacteria, four snoRNAs are present in a tRNA intron. In mammals, both types of snoRNA are usually in introns in ribosomal protein genes, but seldom so in yeast, which has secondarily lost most introns (Smith & Steitz, 1997). The fact that they may be in different genes in different vertebrate species is explained most simply (T. Cavalier-Smith, in preparation) by their self-insertion into and hitchhiking on mobile introns. These H/AC snoRNAs could have acquired a useful function for the host even if their origin was selfish. As methylation and pseudouridylation both stabilize RNA (Charette & Gray, 2000), I suggested that secondary structure stabilization by rigidification was their primary function and that the number of modified sites increased in the ancestral neomuran as an adaptation to thermophily (Cavalier-Smith, 2002a) at the same time as the C/D snoRNAs, at least, were evolving.

A conceivable subsidiary function of more-extensive pseudouridylation might be to alter the most conserved regions of cytoplasmic rRNA sufficiently to prevent the binding of mitochondrial ribosomal proteins. Such a general 'labelling' function, dependent on the overall degree of difference it causes rather than site-specific functions, would be consistent with the experimental evidence in yeast that modification marker snoRNAs, but not cleavage snoRNAs, can be deleted individually without affecting growth (Smith & Steitz, 1997). Deleting them all should slow growth significantly. But such 'labelling' can not be the sole function of pseudouridylation and methylation, as it would not explain the persistence of both in the cryptomonad nucleomorph (Douglas *et al.*, 2001), since mitochondria were eliminated from the surrounding periplastid space long ago. Might they be retained to stop nuclear- and nucleomorph-encoded chloroplast ribosomal proteins exchanging with periplastid ribosomal proteins during their transit through the periplastid space? This is doubtful, since the chloroplast ribosomal proteins should be sufficiently different not to cause confusion. Nucleomorph methylation and uridylation are probably retained simply for their putatively ancestral function of structural stabilization; this would predict their retention in all secondarily amitochondrial eukaryotes like the Parabasalia and Metamonada. It would be interesting to know how many sites are modified in nucleomorph rRNA.

Mosaic and quantum evolution: keys to archaeobacterial and eukaryote origins

The term mosaic evolution was invented by the zoologist De Beer (1954) to express the fact, thoroughly established by palaeontology for morphological evolution, that different parts of an organism often evolve at radically different rates; this makes organisms appear as a mosaic of primitive and derived characters. Quantum evolution is the generalization, made by Simpson (1944, 1953), that sometimes, for short historical periods, especially when a new

body plan arises, evolutionary rates can be temporarily highly accelerated. Elsewhere, I have explained in detail how these two phenomena can produce apparently incompatible trees (Cavalier-Smith, 2002a). My 1987 neomuran theory of a common origin of eukaryotes and archaeobacteria from a eubacterial ancestor applied the principle of quantum evolution, by arguing that the common features of archaeobacteria and eukaryotes (e.g. in transcription and translation that seemed to differ so greatly from those of eubacteria) had evolved very suddenly in a short period in their common ancestor before the two groups diverged from each other, and thereafter had changed relatively much less.

The drastic changes that took place in cell walls and membranes and in the informational molecules, and the innovations in protein secretion, during the origin of neomura from an actinobacterium and the immediately following origins of eukaryotes and archaeobacteria are prime examples of rapid quantum evolution. Neither lateral gene transfer nor symbiogenesis can explain real innovation; they can only move existing things from one place to another. On very rare occasions, symbiogenesis radically increased complexity, most strikingly in the origins of eukaryote algae (Cavalier-Smith, 1995a, 2000b). But it is the exception, not the rule. Most increases in complexity and origins of major groups involved quantum and mosaic evolution, but not symbiogenesis. The origin of eukaryotes was unusual in involving both, but, even here, autogenous quantum changes caused the most radical and most numerous biologically significant innovations. As explained elsewhere (Cavalier-Smith, 2002a), the cellular changes during this neomuran revolution have turned out to be even more extensive than I imagined.

Ribosome quantum evolution and inter-lineage biases severely distort the tree of life

I have detailed how short-term quantum evolution during the origin of 80S ribosomes distorted the rRNA tree by vastly stretching the stem at the base of eukaryotes and how the loss of mitochondrial ribosomes and introduction of expansion segments have been associated with great long-term increases in the rates of evolution of 18S and 28S rRNA in the Microsporidia, Archezoa, Archamoebae, Euglenozoa, Foraminifera and *Physarum*, which cause these long-branch taxa all to group together on unrooted rRNA trees. However, I have no explanation for the long branches of the Percolozoa. While it might be that the Percolozoa are the 'early diverging eukaryotes' that we have been seeking, it is far more likely (given the complexity of their kinetids and the evidence of protein trees) that they are no more basal than any other excavates. Concatenated mitochondrial protein trees (Gray *et al.*, 1999) show a very short stem at the base of eukaryotes (though it might be compressed because of saturation) and a bush-like radiation consistent with my rooting of Figs 2–4 using kinetid properties.

But the neatness of that mitochondrial tree owes something to the fact that the really problematic, ultra-long-branch taxa (Euglenozoa and Sporozoa) were excluded. Mitochondrial proteins are not clock-like either; mitochondrial rRNA is even worse.

Nor are nuclear proteins clock-like. EF-1 α accelerates dramatically in ciliates because of a second function in the cytoskeleton, while tubulins accelerate greatly in all taxa that have lost cilia and centrioles because of the removal of the constraints imposed by their interactions with numerous other ciliary proteins. The logical impeccability of rooting duplicated genes using sister paralogues is always compromised to a greater or lesser degree by quantum evolution immediately following their divergence to yield a stretched common stem, and thus a long outgroup branch that will tend to attract the longest branches in each subtree (for a discussion of this serious problem, see Cavalier-Smith, 2002a). But, as the problems are usually gene-specific and lineage-specific, they can be sorted out.

The fact that rRNA is highly interactive (Woese, 1998) does not mean that its trees are reliable; interactivity reduces the risks only of lateral gene transfer. It prevents neither quantum evolution nor grossly systematically biased rates between lineages. High interactivity with a specific subset of the cell's molecules actually increases the risk of gross systematic biases through co-evolutionary effects when the other molecules change substantially. This is certainly true for tubulins, EF-1 α and rRNAs. I have shown elsewhere that this was true not only for eukaryotic ribosomes, but also for bacterial ribosomes during the shift from the simple eubacterial SRP without a translation-arrest domain or SRP19 protein to a more-complex 7S SRP with novel helices 5 and 6 and SRP19p in the stem neomuran (Cavalier-Smith, 2002a), which explains why the segment linking the base of archaeobacteria to the eubacteria is several times longer than the total depth from the shorter-branch members of either group. It has never been sensible to suppose that these dimensions are a realistic representation of evolutionary time. This long stem is not an 'artefact' but simply the product of the quantum evolutionary changes induced by the origin of neomuran 7S SRP RNA and obligate co-translational secretion. But it does cause the dimensions of the tree to be misinterpreted by believers in a molecular clock and, as a phylogenetic artefact, attracts long-branch negibacteria like *Aquifex* and so misroots the eubacterial tree and causes it to fail to place the archaeobacteria among the eubacteria, where other evidence suggests they belong (Cavalier-Smith, 2002a). As ribosomal proteins are strongly interactive with rRNA, they should produce similarly misleading trees. A concatenated 29 protein-family tree for bacteria dominated by ribosomal proteins (Teichmann & Mitchison, 1999) clearly separates archaeobacteria and eubacteria by a very deep stem but, like rRNA, does not resolve the relative branching order of the eubacterial phyla.

Table 2. Revised classification of kingdom Protozoa and its 13 phyla

Modified from Cavalier-Smith (1999, 2000a) by (i) adopting subkingdoms Gymnomyxa and Corticata as the primary protozoan subdivision rather than Eozoa and Neozoa; (ii) establishing infrakingdoms Excavata and Rhizaria; (iii) decreasing the ranks of infrakingdoms Loukozoa, Discicristata and Archezoa to superphyla, Retaria [diagnosis in Cavalier-Smith (1999) (p. 349)] to phylum, Foraminifera and Radiolaria to subphyla and subphylum Ascetosporea to class Ascetosporea; (iv) increasing the ranks of Choanozoa (originally a phylum; Cavalier-Smith, 1981b) and Apusozoa from subphylum to phylum and abandoning Neomonada; (v) narrowing Sarcomastigota; and (vi) transferring Ascetosporea from Sporozoa; Miozoa and Alveolata to Cercozoa; and Oxymonadida and *Stephanopogon* from Percolozoa to Loukozoa and Cercozoa, respectively.

Taxon	Diagnosis/constituent groups	Mitochondria*	Plastids
Subkingdom 1. GYMNOMYXA† Lankester 1878 stat. nov. emend.	Cell cortex soft, often with pseudopodia or axopodia; lacking localized cytostome or cytopharynx; ancestrally and typically with radiating conose or spherically symmetrical centrosomal microtubules		
Infrakingdom 1. Sarcomastigota† emend.	Reticulopodia absent; often with lobose or filose pseudopods; typically uniciliate, often and probably ancestrally unicentriolar; ciliary transformation typically absent, restricted to biciliate myxogastrid Mycetozoa, where posterior cilium is younger		
Phylum 1. Choanozoa	Flat cristae (usually). Choanoflagellata, Corallochytrata, Ichthyosporea, Cristidiscoidea	+	—
Phylum 2. Amoebozoa	Tubular cristae. Lobosa, Conosa (Mycetozoa and Archamoebae), Phalansterea	—/+	—
Infrakingdom 2. Rhizaria infraregnum nov.	Etymology: Gr. <i>rhizo</i> - root, because of their commonly root-like reticulose or filose pseudopodia and the inclusion of many of the original rhizopods (Dujardin, 1841), notably the euglyphid Cercozoa and the Foraminifera, plus a euphonious, meaningless suffix as in Retaria and Radiolaria, two included groups. Lobopodial locomotion absent; often with reticulopodia and/or filopodia or axopodia; ancestrally and typically bikont; each centriole ancestrally with a single root of a microtubular band or fan; mitochondrial cristae ancestrally tubular, sometimes secondarily flattened; extrusomes are often kinetocysts		
Phylum 3. Apusozoa stat. nov.	Ancestrally with dense 'thecal' layer (plastron-like) inside dorsal plasma membrane. Thecomonadea: Ancyromonadida, Apusomonadida, Hemimastigida	+	—
Phylum 4. Cercozoa	Subphylum 1. Monadofilosa (classes Sarcomonadea, Testaceafilosea, Ramicristea); Subphylum 2. Reticulofilosa, e.g. <i>Spongomonas</i> , <i>Chlorarachnion</i> , <i>Gymnophrys</i> ; Subphylum 3. Endomyxa subphylum nov. Etymology Gr. <i>endo</i> within; Gr. <i>myx</i> - slime, because they are typically plasmodial endoparasites of other eukaryotes. Classes Phytomyxea (e.g. <i>Plasmodiophora</i> , <i>Spongospora</i>) and Ascetosporea classis nov.; diagnosis as phylum Ascetosporea Sprague 1979 (p. 42) (orders Haplosporida; Paramyxida)	+	—/+
Phylum 5. Retaria phylum nov.	Foraminifera; Radiolaria (euradiolarians; acantharians)	+	—
Phylum 6. Heliozoa‡	Centrohelea (flat cristae); Nucleohelea (possibly really belong elsewhere)	+	—
Subkingdom 2. CORTICATA† Lankester 1878 stat. nov. emend.	Ancestrally bikont with asymmetrical ciliary roots of microtubular bands; often with additional cortical microtubules; ciliary transformation general, with younger anterior cilium, ancestrally associated with a single, broad, dorsal microtubular fan; older posterior cilium with two dissimilar microtubular bands, one each side of its centriole; phagotrophy localized ancestrally to ventral groove or cytostome; rims of groove/cytostome supported by the two posterior microtubular roots		

Table 2 (cont.)

Taxon	Diagnosis/constituent groups	Mitochondria*	Plastids
Infrakingdom 1. Excavata infraregnum nov.	Ancestrally with a single anterior/dorsal ciliary microtubular root and two ventral roots; additional cortical microtubules; without cortical alveoli		
Superphylum 1. Loukozoa stat. nov.	Etymology. Gr. <i>loukos</i> groove; Gr. <i>zoa</i> animals; Loukozoa the groovy animals. Biciliate or tetraciliate; mitochondria or hydrogenosomes; diagnosis as phylum and infrakingdom Loukozoa Cavalier-Smith 1999		
Phylum 7. Loukozoa	Anaeromonadea Cavalier-Smith 1997 emend. (Oxymonadida, <i>Trimastix</i> §); <i>Malawimonas</i> , <i>Carpediemonas</i> ; Jakobea, e.g. <i>Reclinomonas</i> ; Diphyllleida	+ /H	—
Superphylum 2. Discicristata stat. nov.	Cristae ancestrally discoid; ancestrally biciliate with subparallel centrioles connected by striated fibres; cytostome/cytopharynx supported by a microtubular band		
Phylum 8. Percolozoa	Typically tetrakont; Golgi unstacked. Lyromonads, <i>Percolomonas</i> , Heterolobosea	+ /— /H	—
Phylum 9. Euglenozoa	Typically bikont; Golgi stacked. Euglenoids, diplomonids, kinetoplastids, <i>Postgaardi</i>	+	— /+
Superphylum 3. Archezoa stat. nov.	Tetrakont; three subparallel anterior and one orthogonal posterior cilium; no mitochondria		
Phylum 10. Metamonada	Golgi unstacked; intranuclear spindle. Retortamonads and diplomonads	—	—
Phylum 11. Parabasalia	Golgi stacked; extranuclear spindle. Trichomonads and hypermastigotes	H	—
Infrakingdom 2. Alveolata	Ancestrally bikonts with cortical alveoli		
Phylum 12. Miozoa	Typically haploid Dinozoa (ellobiopsids and dinoflagellates), Protalveolata and Sporozoa	+ /—	+ /—
Phylum 13. Ciliophora	Diploid micronuclei; multiploid macronuclei. Ciliates and suctorians	+ /— /H	—

* H, Hydrogenosomes, which evolved from mitochondria (several times, independently).

† Probably paraphyletic.

‡ Possibly polyphyletic; as nucleohelid microtubules nucleate on the nuclear envelope not the centrosome, at least some (e.g. actinophryids with tubular mitochondrial cristae) may be pedinellid chromists (Karpov, 2000), not Protozoa.

§ For evidence of the relationship between *Trimastix* and Oxymonadida, see Dacks *et al.* (2001).

Co-evolution of cilia and the cytoskeleton

The autogenous origin of cilia

Was the first eukaryote an amoeba or a heliozoan without any cilia or a flagellate with one or more cilia? Since centrosomes are the nucleating sites for centrioles, which in turn nucleate ciliary growth, and since DNA segregation is much more basically essential for cell viability than ciliary motility, they probably, at least slightly, preceded the origin of the vastly more complex cilia (probably needing about 1000 genes). The long drawn-out love affair of Margulis (1970) with the notion that cilia evolved from motile bacterial ectosymbionts (Kozo-Polyansky, 1924) implausibly assumes the reverse, but is unperturbed by this or by the total absence of any chemical, functional or phylogenetic evidence for its basic assumption of a connection between spirochaete and ciliary motility (Cavalier-Smith, 1978a, 1982b, 1992b); its latest reincarnation (Margulis *et al.*, 2000) is as devoid as earlier ones of any recognition of the scientific necessity to be explicit about the structural and functional changes postulated in evolutionary transformations or the utility of Occam's razor. I agree with Margulis only on the ancientness of the connection between nuclei and cilia, seen so well in her favourite complex hairy flagellates. I have long argued that indirect attachment of a single cilium to the nucleus via the centrosome was the ancestral state for all eukaryotes with cilia (Cavalier-Smith, 1982b, 1987c, 1991c, d, 1992c). I argued that a single cilium arose in association with the origin of the nucleus prior to the eukaryotic cenancestor and, thus, postulated that there are no extant primitively non-ciliate eukaryotes. I shall not add to those earlier detailed treatments of the autogenous (non-symbiogenic) origin of cilia, but will concentrate on the phylogenetic implications of ciliary root structure in the light of increased recognition of the fundamental importance of the remarkable phenomenon of ciliary transformation for understanding eukaryote cell evolution (Cavalier-Smith, 2000a; Moestrup, 2000).

Unikonty and the roots of cilia and the eukaryote tree

Microtubular ciliary roots, better called centriolar roots, as they are actually attached to the centrioles (also known as ciliary or flagellar basal bodies or kinetosomes), are of central importance for eukaryote phylogeny. They form the most structurally distinctive part of the cytoskeleton and are sufficiently well conserved to help define major groups and phylogenetic relationships. In essence, they define the body plan of protists, analogously to the importance of the vertebrate endoskeleton or the arthropod exoskeleton in classical zoology. Moestrup (2000) reviewed the major variations and attempted to provide a uniform terminology. He assumed that the ancestral state was biciliate with cruciate roots having two microtubular bands per centriole. Cruciate roots predominate and

may well be the ancestral state for the kingdoms Plantae and Chromista, to which, as a phycologist, he has devoted most attention. However, this interpretation cannot be correct for the Protozoa, the basal eukaryotic kingdom, from which the four higher kingdoms evolved. Neither of the other two derived kingdoms (Animalia, Fungi) has cruciate roots and they are very rare among the Protozoa. In fact, they are found only in a minority of taxa in two protozoan phyla (Miozoa and Cercozoa) and are almost certainly a derived condition in both. Thus, none of the 13 protozoan phyla recognized here (Table 2) had cruciate roots ancestrally, so they were not the ancestral condition for the eukaryote cell. It is also unlikely that cruciate roots are strictly homologous in plants and chromists. Whether the first flagellates were biciliate, as Moestrup assumes, or uniciliate, as I consider much more probable, is crucial for locating the root of the eukaryotic tree, but remains to be established by molecular evidence (Cavalier-Smith, 2000a). At present, we cannot rule out the possibility that ancestral eukaryote flagellates were uniciliate, but that the eukaryote cenancestor was actually biciliate and the biciliate condition evolved once only in a stem eukaryote. We need to focus phylogenetic attention on all unikont protozoa and determine which ones, if any, are primitively uniciliate.

In order to orient the reader in the complex discussion that follows, Fig. 3 summarizes my present interpretation, according to which the root of the tree lies near those uniciliate protozoa that are good candidates for being primitively uniciliate: the zooflagellate *Phalansterium* and some of the amoeboid flagellate amoebozoia. I shall argue that the biciliate condition evolved twice independently at the base of the clades shown by the yellow boxes. I here designate one of these clades the bikonts (Greek for 'two oars'), as it includes the vast majority of the ancestrally biciliate eukaryotes: the kingdoms Plantae and Chromista and the new protozoan subkingdom Corticata and infraphylum Rhizaria established here, which together include the nine most speciose protozoan phyla (Table 2 shows the defining characters of these novel higher taxa and which phyla each includes). The biciliate condition also evolved in the Mycetozoa, which are nested within the more basal Amoebozoa with only a single centriole per kinetid. I define unikonty as the state of having just a single centriole as well as a single cilium per kinetid. A kinetid with two centrioles per kinetid is not regarded as unikont whether it bears two cilia or just one (as in opisthokonts and in several secondarily uniciliate bikonts, e.g. pedinellid chromists or some uniciliate prasinophyte algae among plants). Thus, I distinguish between two kinds of uniciliate cells: those with two centrioles, often attesting to their biciliate ancestry, and those that have only one, the unikonts, which are candidates for being primitively uniciliate. A multiciliate cell may have unikont kinetids with single cilia (e.g. the pelobiont amoeba *Pelomyxa*, the lobosan amoeba *Multicilia* or the pseudociliate *Stephanopogon*)

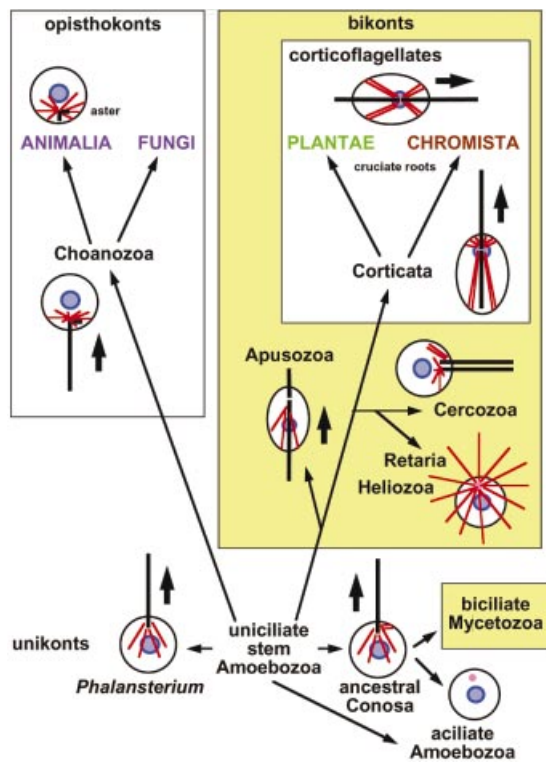


Fig. 3. A simple interpretation of eukaryote phylogeny, emphasizing the diversification of the microtubular cytoskeleton. The tree is rooted among putatively unikont flagellates or amoeboid flagellates for the reasons explained in the text. The biciliate condition may have evolved twice independently in the cenozoans of the clades enclosed by yellow boxes (bikonts, biciliate Mycetozoa). Protozoan taxa are in black and the four higher kingdoms in upper-case in colour. The nucleus is blue, centriolar microtubular roots are red and the cilia and barren centrioles are shown as thick black lines. Typical swimming directions are shown by the thicker black arrows. In some lineages within most taxa, secondary multiplication of centrioles and cilia or losses of one cilium or all cilia have led to deviations from the root structures depicted. Among bikonts, reorientation of cilia between the anisokont and isokont states also modified the patterns; those shown are the predominant and/or putatively ancestral pattern for each group. Corticata comprise the infrakingdoms Alveolata (biciliate Miozoa and multiciliate Ciliophora) and Excavata (biciliate Loukozoa and Euglenozoa; tetrakont Percolozoa and Archezoa). The ancestral corticate pattern of two posterior roots and a broad anterior fan probably originated in the ancestral excavate in association with the origin of their feeding groove, the cruciate patterns of plants and chromists being derived independently from it when their ancestors became photosynthetic (see text). Earlier patterns are all simply derivable from the conical microtubular array of *Phalansterium*. For Cercozoa, the diagram is for the isokont *Spongomonas*; the anisokont sarcomonads have more complex roots (see text). Retaria include Radiolaria and Foraminifera. For Heliozoa, only the pattern in centrohelids (the great majority) is shown; the affinities of nucleohelids are unclear; as their microtubules are nucleated by the nuclear envelope, not the centrosome, some or even all nucleohelids (e.g. actinophryids) might not belong in phylum Heliozoa but with the pedinellid chromists, where this is also the case (Karpov, 2000).

or may have bicentriolar kinetids (the ancestral and majority state for ciliates, where the second centriole is always barren).

Though Margulis *et al.* (2000) implicitly adopt my earlier thesis (Cavalier-Smith, 1991c) that the most likely first eukaryote was a unikont *Mastigamoeba* (class Pelobiontea), they incorrectly classify pelobionts with Metamonada and Parabasalia in a polyphyletic phylum Archaeoprotista, ignoring molecular-phylogenetic evidence (Cavalier-Smith, 2000a; Roger, 1999) that mastigamoebids are not directly related to Metamonada and Parabasalia (i.e. superphylum Archezoa of my present protozoan system; Table 2); contrary to the incorporation into their syntrophic cocktail of my earlier phylogeny (Cavalier-Smith, 1992b) that postulated a mastigamoebid ancestor for Archezoa, mastigamoebids almost certainly evolved from an aerobic amoeboid flagellate with mitochondria. Mastigamoebids (including *Phreatamoeba*, now regarded as a synonym of *Mastigamoeba*; Simpson *et al.*, 1997) are related to the secondarily non-ciliate entamoebas and do not branch with Archezoa on rRNA (Fig. 2; Cavalier-Smith & Chao, 1996) or RNA polymerase trees (Hirt *et al.*, 1999; Stiller *et al.*, 1998). There is good evidence that *Entamoeba* lost aerobic respiration secondarily by converting mitochondria to a minute relict organelle, the mitosome (Tovar *et al.*, 1999; Roger, 1999). Such a loss probably occurred in the common ancestor of *Entamoeba* and Pelobiontea, classified together as the amoebozoan infraphylum Archamoebae (Cavalier-Smith, 1998a), which is often monophyletic on rRNA trees (Cavalier-Smith & Chao, 1996, 1997); in addition to this and the absence of mitochondria, entamoebas and pelobionts both have unique neoinositol polyphosphates instead of the *myo*-inositol polyphosphates of other eukaryotes (Martin *et al.*, 2000).

I selected mastigamoebids as likely early eukaryotes because they alone of amitochondrial eukaryotes appeared to be primitively unikont (Cavalier-Smith, 1991c), which I considered the likely ancestral state for eukaryotes (Cavalier-Smith, 1982b, 1987c, 1992b); others since have thought their characters are primitive (Simpson *et al.*, 1997). Most eukaryote groups are ancestrally bicentriolar; though only some are ancestrally biciliate (e.g. Plantae, Chromista, Alveolata, Cercozoa, Excavata), others (notably the important opisthokont clade) are ancestrally uniciliate but bicentriolar. Fig. 4 is a more detailed eukaryotic tree than Fig. 3, emphasizing the chief congruences between the latest rRNA tree (Fig. 2; for a more taxon-rich tree see Cavalier-Smith, 2000a) and most protein trees (e.g. Baldauf *et al.*, 2000); in places (e.g. the chromist clade), its resolution is exaggerated compared with sequence trees to indicate affinities strongly supported by ultrastructural and biochemical evidence (the length of the segment above and below the base of the excavates is expanded merely because there are too many corticoflagellate taxa to show side by side). In every biciliate group that has been well studied, one of the two cilia is one or more cell generations older than the other and cilia undergo ciliary transformation, i.e. the first-formed, younger cilium is structurally and

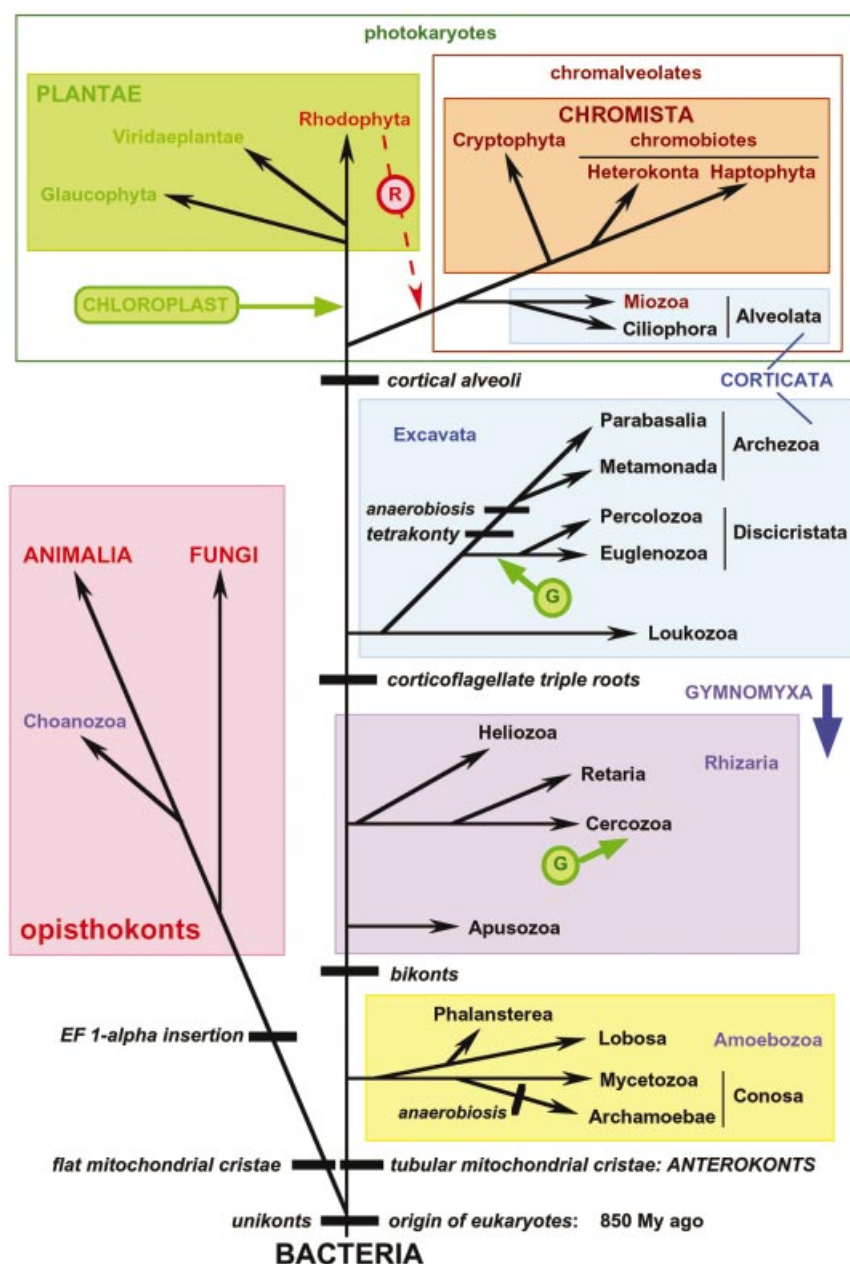


Fig. 4. Proposed phylogenetic relationships between the major eukaryotic groups. Some rRNA trees also show Retaria as paraphyletic or polyphyletic, but this may be an artefact of the exceptionally long branches of Foraminifera and the much longer branches of euryzoaria compared with Acantharia. In addition to the primary symbiogenetic origin of chloroplasts from cyanobacteria to create the ancestral plant, the three secondary symbiogeneses are shown: a red alga (R) was enslaved to form the ancestor of chromalveolates and two different green algae were enslaved to form photosynthetic euglenoids and chlorarachnean Cercozoa (G). Two major secondarily anaerobic groups are indicated (Archezoa and Archamoebae); additional losses of mitochondrial oxidative phosphorylation are not shown (e.g. in Fungi, the microsporidia and rumen fungi; multiply in ciliates). Multiple losses of plastids in chromalveolates and Euglenozoa and the ancestral percolozoan are not shown. It is unclear whether Archezoa are sisters to Percolozoa, as Fig. 2 weakly suggests, or to discicristates, as shown here.

functionally different from the older one into which it is transformed one cell cycle after it was first assembled. Normally, the centriolar roots attached to the young centriole (basal body) are structurally different from those of the older one and are also radically changed during transformation. As centriolar roots are often

the most important part of protist cytoskeletons, their transformation means that cytoskeletal assembly, like ciliary assembly in bikonts, is spread over two cell generations. The great complexity of this process is a strong reason for thinking that a bikont like the jakobid *Reclinomonas* (phylum Loukozoa) cannot

Table 3. Kishino–Hasegawa tests of alternative maximum-likelihood trees

User trees were constructed that differed in topology from Fig. 2 only in the respects specified, and their ln likelihood was calculated by fastDNAML. The significance of their differences in likelihood from the most likely tree (Fig. 2; ln likelihood = -42739.851) was assessed by the Kishino–Hasegawa test using the empirical transition/transversion ratio; the Felsenstein and Hasegawa–Kishino–Yano models gave the same results. The same test was also done under maximum parsimony: the third tree (euglenoids holophyletic) was the shortest; trees marked with an asterisk were significantly worse at the $P < 0.05$ level and none was worse at the $P < 0.01$ level.

ln Likelihood	Likelihood difference from Fig. 2 (rounded)	Branching differences from Fig. 2	Significantly less likely than Fig. 2 tree?	
			$P < 0.05$	$P < 0.01$
-42740-02275003	0.17174	Radiolaria holophyletic	No	No
-42740-21654497	0.36553	Radiolaria, euglenoids holophyletic	No	No
-42741-84527286	1.99426	Euglenoids holophyletic	No	No
-42744-53531401	4.68430	Thecomonadea holophyletic	No	No
-42744-95779726	5.10679	Radiolaria, euglenoids, Thecomonadea holophyletic	No	No
-42749-16915213	9.31814	Radiolaria, euglenoids, Discicristata holophyletic	No*	No
-42749-99986307	10.14885	Euglenoids, Discicristata holophyletic	No*	No
-42765-90120272	26.05019	<i>Reclinomonas</i> sister to Neomonada	No	No
-42770-99223012	31.14122	Holophyletic Discicristata below <i>Reclinomonas</i> ; Radiolaria, euglenoids holophyletic	No*	No
-42797-8956284	58.04462	Chromista holophyletic	Yes	No
-42805-20288052	65.35187	Chromista, Thecomonadea, Radiolaria holophyletic	Yes*	No
-42816-06914706	76.21813	Chromista, Plantae holophyletic (sisters)	Yes*	Yes
-42820-75716959	80.90616	Chromista, Plantae, Thecomonadea, Radiolaria holophyletic	Yes*	Yes

really be a primitive eukaryote, however primitive its mitochondrial genome appears (Lang *et al.*, 1997).

Since tetrakont kinetids develop over three cell cycles with two successive ciliary transformations and the Archezoa are probably ancestrally tetrakont, they are even less plausible as ancestral eukaryotes than biciliates and must be derived (Cavalier-Smith, 1992b), as indicated on Figs 2 and 4. The evidence that the Archamoebae may group either with Mycetozoa (with which they were later classified as the amoebozoan subphylum Conosa; Cavalier-Smith, 1998a) or with amoebozoan Lobosa (Cavalier-Smith, 1993b, 1995a; Cavalier-Smith & Chao, 1995), nowhere near the base of the rRNA tree (Cavalier-Smith & Chao, 1996), was initially confusing. Protein trees then came to the rescue, showing beyond serious question that Mycetozoa were misplaced on earlier rRNA trees (Baldauf & Doolittle, 1997) and that microsporidia were even more drastically misplaced, belonging not near the rRNA root but among the fungi (Edlind *et al.*, 1996; Keeling & Doolittle, 1996; Hirt *et al.*, 1999; Müller, 1997; Roger, 1999; Keeling *et al.*, 2000), where they have at last found their proper taxonomic home (Cavalier-Smith, 1998a, 2000c). Concomitant critical appraisal of long-branch artefacts in rRNA trees (Philippe & Adoutte, 1996, 1998; Stiller & Hall, 1999) reinforced earlier suspicions of their unreliability. Although proteins also suffer from long-branch problems (Philippe & Adoutte, 1998), now that rRNA trees are dethroned from their position of primacy we can

seek for congruence between trees with fewer preconceptions. Note that Conosa and Amoebozoa are both holophyletic on Fig. 2, as they also would be on Fig. 2 of Cavalier-Smith & Chao (1997) (and Conosa alone on their Fig. 1) if they were similarly rooted.

If we allow for the common misplacement of Mycetozoa and Microsporidia on rRNA trees (not universal: when more sophisticated methods are used, small-subunit trees can place Mycetozoa correctly, as in Fig. 2, and large-subunit trees can place microsporidia correctly; Van de Peer *et al.*, 2000) and ignore the position of the root, there is actually substantial overall congruence between the broad patterns shown by rRNA and protein trees, implying that, apart from the serious problem of rooting, they are not grossly in error. Thus, Fig. 2 is congruent with the concatenated four-protein tree of Baldauf *et al.* (2000) except for the lower position of the very-long-branch Archezoa on their tree. As Table 3 indicates, however, Fig. 2 should not be used to argue against the holophyly of Discicristata, Thecomonadea, Radiolaria or Chromista. All trees, including Fig. 2, can be partitioned cleanly into two halves: opisthokonts on the one hand and a huge group comprising Amoebozoa (in which I now include *Phalansterium* in addition to Lobosa and Conosa) and the bikonts on the other. Bikonts comprise Plantae, Chromista and Alveolata (collectively designated photokaryotes, as each is probably ancestrally photosynthetic; Cavalier-Smith, 1999) plus Discicristata (Euglenozoa and Percolozoa), Archezoa (Meta-

monada and Parabasalia) and Rhizaria. There is little doubt that photokaryotes and Discicristata are all ancestrally biciliate; as they share the same pattern of ciliary transformation, with the anterior cilium younger (Moestrup, 2000), their common ancestor almost certainly was also. Three other important ancestrally biciliate bikont groups (Cercozoa, Retaria and Loukozoa) apparently branch among them in a poorly resolved bush, though only sparse molecular sequence data and no studies of ciliary transformation are yet available for them. Although there is always high bootstrap support for the bipartition between opisthokonts and bikonts/Amoebozoa, the branching order within the basal radiations of these two groups is poorly resolved on both rRNA and protein trees. If the concatenated protein tree were rooted as advocated here, Amoebozoa would be sisters to bikonts and Archezoa would be sisters to all other bikonts, not to Percolozoa as on the rRNA tree; the latter is cytologically more reasonable, since Percolozoa and Archezoa might have had a tetrakont common ancestor.

The pseudociliate *Stephanopogon* was once thought to be related to the Percolozoa because of its flat, somewhat discoid cristae and absence of Golgi stacks (Cavalier-Smith, 1993d); although some authors have included it in the other discicristate phylum, Euglenozoa, instead (Hausmann & Hülsmann, 1996), this is not generally accepted (Simpson, 1997). However, as mentioned above, Golgi unstacking has arisen polyphyletically; moreover, discoid cristae are themselves polyphyletic, being found not only in the Discicristata, but also in the Cristidiscoidea (Choanozoa; Cavalier-Smith, 2000a). The marked resemblance between the centriolar cups of *Stephanopogon* (Lipscomb & Corliss, 1982) and those of the biciliate cercozoan flagellate *Spongomonas* (Hibberd, 1976), not previously noted, indicates a definite affinity between them. I also suggest that, if the *Spongomonas* cilium bearing the more band-like root was lost and that bearing the fan-shaped ciliary root retained when the probably similarly biciliate ancestor of *Stephanopogon* multiplied its kinetids, the fan would probably become a symmetric cone like that in *Stephanopogon*. The complexity of the roots and the uniqueness of the centriolar cup among protists make them likely to be reliable phylogenetic characters. Therefore, I now remove the Pseudociliatida from Discicristata altogether and place it as a second order within the cercozoan class Spongomonadea (Cavalier-Smith, 2000a). If this position is correct, *Stephanopogon* must be secondarily unikont. As similar secondary unikonty evolved in the cenancesor of the multiciliate opalinid chromists and of the apusozoan Hemimastigida mentioned above, and within the ciliates, it is a common evolutionary consequence of the multiciliate condition.

The only currently established groups that are reasonable candidates for being primitively unikont are the Amoebozoa and the zooflagellate *Phalansterium*, so I have suggested that the eukaryotic root may lie among

them (Cavalier-Smith, 2000a). Although the amoebozoan Archamoebae and *Multicilia* are all unikont, the Mycetozoa are more problematic, some being unikont and some bicentriolar (and others secondarily akont); Moestrup (2000) assumes that the Mycetozoa are ancestrally bikont, whereas I argued the reverse, partly because their bikont members differ from photokaryotes/discicristates in that the anterior cilium is the older one, suggesting that they may have evolved bikonty independently (Cavalier-Smith, 2000a). The derived myxogastrid Mycetozoa are bicentriolar and usually biciliate, but the non-fruiting *Hyperamoeba* derived from them (Cavalier-Smith & Chao, 1999; Cavalier-Smith, 2000a) is uniciliate. Kinetids in the ancestral mycetozoa (Protostelea) may be unikont or bikont and uniciliate, biciliate or multiciliate. Myxogastrids and some protostelids have a cone of microtubules attaching the kinetid to the nucleus. Because of this and their similar pseudopodial motility and the unikont character of all ciliated archamoebae, the Mycetozoa are classified with the Archamoebae as the amoebozoan subphylum Conosa (named after the shared microtubular cone; Cavalier-Smith, 1998a) and are probably sisters. Conosa are related to the subphylum Lobosa, mainly comprising aciliate amoebae; though this is shown only rarely on rRNA trees (e.g. Fig. 2), it is obvious on actin and on concatenated protein trees (Baldauf *et al.*, 2000). Since the multiciliated lobosan amoeba *Multicilia* also has unikont kinetids with microtubular cones (Mikrjukov & Mylnikov, 1998), unikonty is probably the ancestral state for amoebozoa. I suggest that the absence of the cone in a few protostelids is secondary, possibly resulting in some cases from their multiciliarity. Unikont Mycetozoa have three extra microtubular roots and biciliate ones yet another, associated with the posterior cilium, additional to the putatively ancestral cone; cladistic arguments would suggest that these extra complexities of a subgroup nested within the phylum where the outgroups have a simpler arrangement are derived. To attempt to homologize them with the kinetids of bikonts (Karpov, 1997; Moestrup, 2000) is probably a mistake. Even the simplest amoebozoan ciliary roots are more complex than that of *Phalansterium*, which has a simple cone of singlet microtubules identical to that postulated to be ancestral for all eukaryotes (Cavalier-Smith, 1982b, 1987c, 1992c). In addition to such a cone, archamoebae and *Multicilia* (Mikrjukov & Mylnikov, 1998) have a transverse bipartite microtubular band associated with the centriole. As *Phalansterium* has the simpler kinetid, it is a better model for the ancestral eukaryote (Cavalier-Smith, 2000a). However, on my own unpublished 18S rRNA trees that allow for intersite variation by a gamma model, *Phalansterium* clearly branches within the Amoebozoa, so I now transfer it to that phylum and restrict the Apusozoa to the Thecomonadea.

The uniciliate character of opisthokonts was previously thought to be secondarily derived, as their apparent immediate outgroup, the apusozoan (the-

comonad) flagellates (the biciliate Ancyromonadida and Apusomonadida and the multiciliate Hemimastigida; Cavalier-Smith, 2000a), is ancestrally biciliate. However, the initially high bootstrap support for a sister relationship between *Apusomonas* and opisthokonts (Cavalier-Smith & Chao, 1995) is not found in recent trees with additional members of the Apusozoa (Cavalier-Smith, 2000a; Fig. 2). In my own unpublished trees, with even better taxon sampling among protozoa and using better phylogenetic methods, allowing for intersite variation by a gamma model, the Apusozoa are simply part of a very poorly resolved bikont/amoebozoan radiation and are not specifically related to opisthokonts. Coupled with the current uncertainty about the position of the eukaryotic root, this makes me question seriously the traditional assumption that the fundamentally uniciliate Choanozoa (true both for choanoflagellates and Ichthyosporea: the other two classes are entirely non-ciliate) once had biciliate ancestors. The presence of a second dormant, non-ciliated centriole is poor evidence for such ancestry. In contrast to the barren centrosome of uniciliate heterokonts, which is present because it bore a cilium in the previous cell cycle, there is no evidence for the transformation of cilia, centrioles or their roots from one cell cycle to the next in any opisthokonts. The second centriole of choanoflagellates may simply be formed in the preceding cell cycle so as to be ready to grow a cilium immediately the cell divides, and not a relic from a biciliate ancestor, persisting uselessly for over 500 My. For chytrid fungi, which lack centrioles during vegetative growth, such early assembly of centrioles would help to facilitate the rapid production of zoospores during sporogenesis by rapid multiple fission. In the mature zoospores, which do not need to divide, such centrioles would be dispensable. However, only some chytrid fungi have lost them (Barr, 2000); most retain short centrioles. These may reasonably be regarded as relics of the useful presence of a second centriole during zoosporogenesis, but they do not provide evidence for a biciliate ancestor. Since the ciliary roots of choanozoa and most chytrid fungi are cones of microtubules similar to those of the unikont Amoebozoa, I suggest that opisthokonts also are primitively uniciliate and that the root of the eukaryote tree lies between opisthokonts and the Amoebozoa.

On this view, the fundamentally uniciliate Sarcomastigota are the ancestral eukaryotes and all bikonts with ciliary transformation over two or more cell cycles must be derived. Whether ciliary transformation occurs in the Apusozoa and Cercozoa or not is currently unknown and needs to be established. It may be a fundamental feature of all bikonts. As the cilium is anterior in all uniciliate Amoebozoa, the first bikont could have evolved from an amoebozoan-like ancestor by adding a second, posterior cilium. This would have yielded an anisokont amoeboflagellate that crawled on surfaces like the apusomonad *Amastigomonas* or the cercomonads; the common ancestor of Apusozoa and Cercozoa probably had such habits and form. Because

amoebzoa and bikonts both have anterior flagella, I designate them collectively anterokonts, so as to contrast them with the opisthokonts. On the present view, therefore, the primary split among eukaryotes is between opisthokonts, with a posterior cilium, and anterokonts, with an anterior one. The first eukaryotes were uniciliate and split at an early stage into the only two primitively uniciliate phyla: Choanozoa and Amoebozoa. I suggest that this split represents the two most basic ways of being a predator for a ciliated eukaryote: a sessile filter feeder (Choanozoa) and a mobile raptorial feeder (Amoebozoa). Opisthokonts and anterokonts generate water currents that bear prey in opposite directions. Uniciliate amoebzoa such as *Mastigamoeba* creep along on surfaces and use their anterior cilium to help to pull prey towards them and engulf them in pseudopods. The sessile choanoflagellates, with a cilium undulating from base to tip, push water away from the base of the cilium, thereby sucking in water from the side where suspended bacteria are caught by the collar of microvilli and then moved down to the cell surface for phagocytosis; by retaining the same ciliary beat during dispersal, their cilia are necessarily posterior when swimming – the opisthokont condition. Thus, the posterior cilia of opisthokonts probably arose co-adaptively with the filopodia/microvilli that were used to make the collar for filter feeding. By contrast, the anterior cilium of amoebzoa is co-adaptive with the classical amoeboid locomotion of amoebae with lobose pseudopods. Exploiting these two broad adaptive zones was achieved by the primary bifurcation of the first (uniciliate) eukaryotes. Since sarcomonad cercozoans have ciliary roots with a perinuclear microtubular cone like those of amoebzoa including *Phalansterium* (Karpov, 1997), this cone was probably present in the ancestral ciliated eukaryote and was lost by the ancestor of photokaryotes and discicristates, which only have microtubular bands and non-microtubular striated ciliary roots. This cladistic reconstruction adds credibility to my interpretation of the origin of cilia from microtubules radiating from a centrosome (Cavalier-Smith, 1980, 1982b, 1992b). It does not, however, prove that this happened in the first eukaryote; I argued that it did because both the origin of cilia and the origin of mitotic division require that centrosomes are attached to the cell surface, and it seemed economical to suppose that a single attachment triggered both. However, the first eukaryote might instead have been like a non-ciliate amoeba that evolved cilia subsequently from cell projections (Cavalier-Smith, 1978a). However, this would only be possible if the Amoebozoa are paraphyletic and the root of the tree lies within this phylum, contrary to the arguments just presented.

Ever since I stressed the need for cell-cycle continuity between prokaryotic division and segregation dependent on a rigid cell surface and the mitotic system dependent on rigid microtubules, I have not favoured the view that the ancestral eukaryote was a simple soft-

surfaced amoeba (Cavalier-Smith, 1980) and have regarded amoeboid locomotion, as in the lobosan and the heterolobosean amoebae, as an advanced character and treated these taxa as secondarily non-ciliate. Although one lobosan, *Multicilia* (Mikrjukov & Mylnikov, 1998), has plenty of unikont cilia, we cannot yet be sure that the ancestral amoebozoan was ciliated. We can be confident that heterolobosean amoebae and amoebflagellates are not primitive, but are derived from bikont or tetrakont ancestors, but cannot rule out on phylogenetic grounds the possibility that some lobosans are ancestrally akont. However, lobosans do not have cytoplasmic microtubules and only assemble them for mitosis; coupled with their ever-changing shape, this makes them implausible ancestors for the origin of cilia, so I predict that all will eventually be shown to be secondarily akont.

Akont heliozoa would be much more plausible as primitively aciliate ancestors of the first zooflagellate, since their radiating axopodia have a microtubular skeleton that could have been a precursor for ciliary axonemes. There are two types, often thought to be unrelated: centrohelids, with axopodia radiating from a centrosome, and nucleohelids, where the axopodia are nucleated by the nuclear envelope. If cilia evolved in a heliozoan, the first eukaryote would probably be uniciliate if the ancestor was centrohelid, but multiciliate with each kinetid unikont, like *Multicilia*, if the ancestor was a nucleohelid. I consider that heliozoa probably evolved from flagellate ancestors by the loss of cilia and that their axopodia were derived from the centrosomal radiating microtubules that characterize Sarcomastigota and Cercozoa; one 'heliozoan' (*Clathulina*) is biciliate, but it is unclear whether it is really related to the nucleohelids. The presence of kinetocysts in heliozoans and some cercozoans might suggest an affinity between them, which would favour a biciliate ancestry for Heliozoa. The thecomonad apusozoan *Ancyromonas* also has kinetocysts; unless they are convergent structures, this suggests that the Apusozoa, which are fundamentally bikont, may be distantly related to the Cercozoa and implies that kinetocysts were present in the common ancestor of thecomonads, Heliozoa and Cercozoa, i.e. very early in bikont evolution. Kinetocysts are widespread in the new infrakingdom Rhizaria, but are also found in histionid Loukozoa; this need not mean that they are convergent, since it is possible that they arose in the ancestral bikont and are not a synapomorphy for Rhizaria. Our own studies of centrohelid heliozoan molecular phylogeny (T. Cavalier-Smith and E. E. Chao, unpublished) indicate that they branch among the bikonts, indicating that they are not primitively non-ciliate.

The re-rooting of the eukaryote tree advocated here (Fig. 4) may be compatible with the idea that ancestral mitosis was closed with an intranuclear spindle and that open mitosis evolved polyphyletically to allow membrane rearrangement in larger cells (Cavalier-Smith, 1982c). However, mitosis needs studying in

apusozoans, mastigamoebids and cercozoans; an ancestral semi-open mitosis now seems mechanistically more likely since, in all organisms near the putative root, the centrosome is cytoplasmic and paranuclear.

Cytoskeletal evolution and eukaryote diversification

A major innovation appears to have occurred in eukaryote cell evolution at the bar labelled 'cortico-flagellate triple roots' in Fig. 4. All taxa below this point have centrosomes with radiating microtubules (somewhat like the astral microtubules of animals) and generally rather soft cell surfaces, not supported by cortical microtubules, and a great propensity to form filose, lobose or reticulose pseudopods and/or axopodia. Most taxa above this point have a relatively rigid cell cortex, often supported by microtubules, some of which originate as ciliary roots made of distinctive bands of aggregated microtubules, but lack evenly radiating single microtubules resembling asters. These 'corticate' taxa typically ingest prey in a localized region near the base of the ventral/posterior cilium, which may be a simple groove or pocket or a more complex cytostome and gullet or cytopharynx. By contrast, the taxa below the bar typically ingest prey diffusely anywhere on their cell surface and never have a discrete cytostome. Since such diffuse ingestion is less specialized than that with a localized cytostome, which requires a more complex cortical structure with a basic asymmetry associated with complex ciliary transformation, I argue that the latter is derived and that the diffuse feeding pattern is the primitive one. Since the tree is rooted using the more fundamental criterion for primitiveness of unikonty, the fact that diffuse feeding and radiating centrosomal microtubules also appear basal strongly suggests that all three structural characters were historically associated and constitute the ancestral state for eukaryotes. Interestingly, this distinction in mode of feeding was made long ago by Saville Kent (1880), who referred to such diffusely feeding protozoa as Panstomata, a group that corresponds roughly to the subkingdom Gymnomyxa as defined here (Table 2). To emphasize the major importance of the evolution of the rigid cortex and associated pattern of ciliary transformation within the Protozoa, I adopt the Corticata and Gymnomyxa ('naked slime') of Lankester (1878) as the names for the protozoan subkingdoms (both necessarily paraphyletic) in my revised system (Table 2).

If the corticate character of the taxa above the bar is indeed derived, as this argument indicates, we can treat it as a synapomorphy to define a major eukaryotic clade, which I designate the corticoflagellates. The name refers to the fact that their cell cortex is generally semi-rigid and strengthened by microtubules, typically in the form of three or four distinct ciliary roots consisting of bands of parallel, particularly stable microtubules. The name Corticoflagellata was used originally to designate a putative major group (Cavalier-Smith, 1978a) roughly equivalent to alveolates plus opisthokonts. Since that grouping was not

soundly based and the name now defunct, it is available usefully to designate the clade comprising the corticate protozoa plus the two kingdoms (Plantae and Chromista) derived from them, in which cortical bands of laterally adhering microtubules play a fundamental role except in those subgroups that are secondarily non-ciliate. Ancestrally, corticoflagellates have a fundamental cell asymmetry such that, in all biciliate members of the group, the anterior cilium is the younger one and, in most that have become secondarily uniciliate (e.g. pedinellids, centric diatoms, hyphochytrids), it is the anterior cilium that is retained. The opisthokonts are one of the few really major eukaryotic clades that are well corroborated by sequence trees, sequence signatures (Baldauf, 1999) and ultrastructural cladistics, so its validity is almost indubitable (Patterson, 1999). I predict that the corticoflagellate, bikont and anterokont clades will all eventually become as well supported.

One corticoflagellate phylum, the Parabasalia, is an apparent exception that proves the rule. Parabasalia have highly complex internal bands of aggregated microtubules and a soft, semi-amoeboid cell surface that can ingest anywhere. As I suggested previously (Cavalier-Smith, 1992b), this is almost certainly a secondary internalization of the formerly cortical microtubule bands, as an adaptation to dwelling in animal guts where superabundant food particles surround them on all sides; the ancestral cytostome that evolved in the ancestral anterokont to enable it to predate unidirectionally was no longer advantageous and was thus abandoned. Their remarkably complex internal skeleton is thus a relic of a former cortical skeleton inherited from their free-living ancestors. One parabasalid, *Dientamoeba*, carried the reduction to its logical conclusion by abandoning both cilia and their microtubular bands, becoming an amoeba. Secondary evolution of amoebae also occurred in another corticate phylum, Percolozoa: their ancestors were probably purely flagellates with cortical skeleton and localized ingestion, as in *Percolomonas*, but, early on, they evolved a temporary amoeboid phase with eruptive pseudopods very different from the typically non-eruptive ones of the Gymnomyxa. Many heterolobosean percolozoans dispensed with the temporary ciliate phase to become obligate amoebae.

As Fig. 4 and Table 2 make clear, I have now grouped five corticate phyla (Metamonada, Parabasalia, Percolozoa, Euglenozoa and Loukozoa) together as a new infrakingdom, Excavata. They are characterized ancestrally by having two cilia, a single broad anterior centriolar microtubular fan and two lateral posterior centriolar bands, typically predominantly cortical. In the Parabasalia, this pattern is obscured, I suggest, by secondary tetrakonty and cytoskeletal internalization; the term excavate was applied originally (Simpson & Patterson, 1999) only to the three groups that show clear evidence of an 'excavated' feeding groove (Metamonada, Percolozoa, Loukozoa), though, even then, it embraced heterolobosean amoebae that had second-

arily lost it. My inclusion of Parabasalia and Euglenozoa is made because, on several molecular trees, they appear related to Metamonada and Percolozoa, respectively, and thus have secondarily lost both the feeding grooves and the ciliary flanges used initially, together with details of the ciliary root patterns, to characterize excavates. The characteristic excavate three ciliary roots are also obvious in the Euglenozoa. In contrast to the Excavata, most photokaryotes have cruciate ciliary roots, where the anterior cilium has two flanking microtubular bands like the posterior one (Moestrup, 2000). The rRNA tree rooted as in Fig. 2 is consistent with the monophyly of Excavata as defined here and with their closer relationship to photokaryotes than to opisthokonts.

A concatenated α - and β -tubulin tree places the two jakobid loukozoans for which we have rRNA sequences (*Reclinomonas americana* and *Jakoba libera*; Cavalier-Smith, 2000a) as a sister clade to a monophyletic Discicristata (Euglenozoa and Percolozoa) (Edgcomb *et al.*, 2001), albeit with low bootstrap support, consistent with the grouping of all three phyla within Excavata. However, on the tubulin trees, unlike the rRNA tree, the amitochondrial Archezoa are long branches and do not group with these three mitochondrial excavate phyla. The β -tubulin tree (if rooted as in Fig. 2) shows Archezoa as a long-branch clade, sister to all other bikonts, while those two jakobids nest within an apparently paraphyletic Discicristata. Valyl-tRNA synthetase (Hashimoto *et al.*, 1998) and Cpn60 (Roger *et al.*, 1998) both show an archezoan clade, like rRNA and β -tubulin and the concatenated tubulin/actin/EF-1 α tree (Baldauf *et al.*, 2000). The α -tubulin tree, however, puts Archezoa as a paraphyletic group at the base of the bikonts/Amoebozoa, but *Jakoba libera* and *Reclinomonas* separate and no longer grouped with Discicristata (but are not far removed). Moreover, two other loukozoans (*Malawimonas* and *Jakoba incarcerata*) do not group with the first two or with the discicristates on the tubulin trees, but occupy separate, low positions among the long-branch bikonts/Amoebozoa (Edgcomb *et al.*, 2001). I consider that this non-holophyly of both Loukozoa and Excavata on the tubulin trees and, with α -tubulin, the paraphyly of Archezoa, where Metamonada and Parabasalia are separated by *Jakoba incarcerata*, are more likely to be artefacts of the long branches of Archezoa and, to a lesser extent, of *Malawimonas* and *Jakoba incarcerata* than genuine indications of paraphyly or polyphyly of Excavata, Loukozoa and Archezoa. Despite these inconsistencies, when rRNA and tubulin trees are both rooted as in Figs 2 and 3, they are more congruent with each other and with the other protein trees mentioned above than when rooted arbitrarily on the metamonads (Edgcomb *et al.*, 2001). This greater congruence of unrelated molecular trees supports the arguments based on the microtubular skeleton that the Archezoa must be highly derived compared with the unikonts (Fig. 3) and that cytoskeletal evolution is a sounder basis for rooting the

eukaryotic tree than rRNA, which suffers immensely from the exceptionally long branches of most excavates seen in Fig. 2 (Cavalier-Smith, 2002a). Future protein trees will probably be easier to compare with each other and rRNA trees if both are rooted as advocated here; I propose that a root between opisthokonts and anterokonts be used as the null hypothesis until a better one is found.

The single anterior fan of excavates may be homologous with the centrosomal cone of the *Gymnomyxa*, which became restricted to a single dorsal segment, instead of a 360-degree cone, as a result of the evolution of the second centriole in an early biciliate ancestor of excavates. For the other cilium, ancestrally associated with the ventral feeding groove, the microtubules were rearranged to form two longitudinal bands, one on each side of the ventral centriole, to support the rims of the groove. On this functional interpretation, the single broad microtubular band attached to the younger centriole and cilium (C1) is the ancestral state and the two narrow longitudinal microtubular bands associated with the older centriole (C2) is the derived one. Thus, in this instance of ciliary transformation in excavates, ontogeny does appear to recapitulate phylogeny, which would please Haeckel.

Moestrup (2000) suggests that cruciate roots are the ancestral state for alveolates, but cladistic reasoning contradicts this. Ciliophora with bikinetids all have two roots associated with the mature cilium and only one associated with the younger cilium, as in excavates. In the Miozoa, the roots of protalveolates, the basal and ancestral group, are poorly known, except for *Parvilucifera*, where they are not cruciate: there are only three roots, two anterior ones each of only a single microtubule and one posterior one (Moestrup, 2000). The ciliary roots of the Sporozoa are poorly characterized. Those of dinoflagellates are cruciate, but three of them comprise but a single microtubule. Much more work is needed on protalveolate roots, especially on *Colponema*, which has a putatively ancestral lateral anisokont arrangement. The apical and backward-pointing arrangement of the apicomonad and perkinsid cilia is probably a derived adaptation to the evolution of predatory myzocytotic habits, not the ancestral condition for alveolates; it might be expected to entail much modification of the roots. Given the likelihood that excavates are the outgroup to alveolates and the presence of an excavate-like pattern of three roots in ciliate bikinetids, the simplest interpretation would be that this was the ancestral state for alveolates and Corticata as a whole, as indicated in Fig. 3.

Figs 3 and 4 show a novel group, Rhizaria, comprising Apusozoa, Heliozoa, Cercozoa and Retaria, as the outgroup to excavates and cruciates. Apusozoa, Cercozoa and Retaria are ancestrally biciliate. Little is known about ciliary roots in Retaria (Radiolaria, including Acantharea; Foraminifera), but their gametes/zoospores are anisokont like sarcomonad cercozoans. All the Retaria have reticulose pseudopods, as do several cercozoans (*Chlorarachnion*, *Gymnophrys*,

Penardia), which are absent from any other well-characterized group. Some radiolarian zoospores seem to have a perinuclear microtubular cone emanating from their bikinetid, similarly to cercomonads (Hollande, 1974), and no apparent affinity with corticoflagellates. For these reasons and the relative closeness of Cercozoa and Radiolaria on rRNA trees (e.g. Fig. 2), I think they are probably related. Thorough studies of ciliary roots in zoospores of the Retaria would be a valuable test of this grouping. I argue that the common ancestor of corticoflagellates and the Rhizaria was a biciliate that evolved from a unikont ancestor similar to an aerobic amoebozoan. I therefore designate the putative clade comprising corticoflagellates and the Rhizaria the bikonts. I omitted Foraminifera from the rRNA tree of Fig. 2 and Cavalier-Smith (2000a) since their 18S rRNAs are so bizarre (Pawlowski *et al.*, 1997) that their long branches would have distorted them. However, I find that, on gamma-corrected distance trees omitting the long-branch Archezoa, Foraminifera branch within Radiolaria and this retarian clade is sister to Cercozoa. On such trees, when *Ascetosporea* are also included, they are monophyletic and sisters to the classical Cercozoa. As they share a unique rRNA signature sequence with them, I now place them within Cercozoa as a new class, *Ascetosporea*. Keeling (2001) now has evidence from actin trees that Foraminifera are indeed related to Cercozoa, which partially supports the holophyly of Rhizaria: protein data are needed for Radiolaria, Apusozoa and Heliozoa for a stronger test. I suggest that the axopodia of the Radiolaria (now including Acantharea as a distinct class; Cavalier-Smith, 1999) and the centrohelid Heliozoa, which both typically radiate from centrosome-like structures, are independent derivatives of the ancestral gymnomyxan centrosomally radiating microtubules that originated in a *Phalansterium*-like flagellate. The reticulopodia of foraminiferans are supported by microtubules and may have had a similar origin. Planktonic foraminifera are derived and benthic ones much more diverse and ancestral. Conceivably, however, the ancestral stem foraminiferan was planktonic, like the Radiolaria, with stiff, radiating axopodia, and modified them by developing anastomoses to form a feeding net as an adaptive shift to a benthic habitat prior to the foraminiferal cenancestor. This might explain why their reticulopodia are intermediate in some respects between axopodia and the simplest microtubule-free reticulopodia of other groups.

Ciliary diversification among the Cercozoa

Centriolar root structure is more diverse among the Cercozoa than in other phyla. They are quite different in the three subphyla, suggesting early mutual divergence following the origin of the bikont condition just prior to their common ancestor; their early divergence on the rRNA tree is consistent with this. The cercozoan class Spongomonadea (subphylum Reticulofilosa; Cavalier-Smith, 2000a) has two simple roots, easily derivable from the ancestral condition.

One of the centrioles of the spongomonadid *Rhipidodendron* has a diverging cortical band of microtubules, whereas the other has a horizontal fan of microtubules (Hibberd, 1976) that resembles one segment of the radially symmetrical microtubular skirt of choanoflagellates and chytridiomycetes of the order Monoblepharidales. The latter root somewhat resembles the cone of *Phalansterium*, as Karpov (1990) pointed out. While the similarity is not sufficient to justify Karpov's placement of *Phalansterium* in the Spongomonadida (Karpov, 1990), I argue that the *Rhipidodendron* fan-like microtubular root could have been derived from such a symmetrical skirt-like root or the more cone-like one of *Phalansterium*, as proposed above for the anterior cilium of excavates, to which I suggest it is homologous. Roots in the other three reticulofilosan classes are poorly characterized. They are simplified in chlorarachnean algae, where the *Chlorarachnion* zoospore has only one microtubular band (Hibberd, 1990), which is understandable as their ancestor became secondarily uniciliate and the flagellate phase probably became secondarily non-phagotrophic: both changes typically greatly alter the architecture of ciliated cells. It seems that the *Chlorarachnion* zoospore lost the anterior cilium with its microtubular fan, retaining only the posterior one, with its simple microtubular band.

In contrast to the largely sessile spongomonadids, which are fundamentally isokont with two parallel cilia, the subphylum Monadofilosa includes the sarcomonads, which are fundamentally anisokont, the akont Filosea, derived from them by the loss of cilia, and the Ramicrstea, which are modified anisokonts. The markedly anisokont sarcomonads have more complex roots than spongomonads, giving them a cellular asymmetry specialized for phagotrophy while swimming or gliding on surfaces. *Heteromita* and *Cercomonas* have a microtubular cone subtending the nucleus, like *Phalansterium* and other amoebzoa, and also a dorsal cortical microtubular cape and microtubular band associated with the anterior centriole. In some species, the posterior centriole has two microtubular bands. There are therefore distinct similarities to the biciliate Mycetozoa. Karpov (1997, 2000) and Moestrup (2000) interpret them as homologies. I cannot concur, for they ignore the extensive phylogenetic evidence that sarcomonads and biciliate Mycetozoa are not directly related; each has sister groups that do not share these structures and is nested relatively shallowly within two very different phyla: Cercozoa and Amoebozoa. I therefore consider these similarities to be partly plesiomorphic and partly convergent. If the eukaryote tree of Fig. 4 is essentially correct, and the root lies between opisthokonts and anterokonts, then the Cercozoa and Amoebozoa had a common ancestor that was not shared with the opisthokonts. As discussed above, the ancestral state for the Amoebozoa was probably a microtubular cone subtending the nucleus plus a transverse microtubular band. If both were present in the common ancestor of

Cercozoa and Amoebozoa, two of their shared microtubular arrays would be plesiomorphic. The dorsal cortical cape of microtubules in the Mycetozoa is not found in their amoebzoan outgroups, nor in the cercozoans outgroups of the sarcomonads, so probably evolved independently. The two posterior microtubule bands found in some sarcomonads and some members of the Mycetozoa (only one is present in others) are probably also convergent, being absent in their respective outgroups.

The ramicristean helioflagellates *Dimorpha* and *Tetradimorpha* have axopodial bundles of microtubules in quincunx or random array radiating from the centrosomal region (Brugerolle & Mignot, 1984). These feeding devices bearing kinetocysts are convergent with those of centrohelid Heliozoa and Radiolaria. Like them, they probably evolved from the microtubular cone of the ancestral gymnomyxan. In contrast to heliozoa and radiolaria, dimorphids often spread their axopodia out on surfaces and retain their cilia during feeding to waft food towards them. This meant that the ancestral centriolar cone could only be converted into a radiating set of axopodia by the grotesque elongation of their centrioles so as to place the microtubule-nucleating centrosome surrounding its basal region much more deeply within the cell. This explains why the dimorphids have by far the longest centrioles of any pauciciliate eukaryote; the only longer ones are those of certain hypermastigote parabasalids. Each *Dimorpha* cilium also has a single microtubule band as its root. I conjecture that a single such band per cilium and centriolar cone around the nucleus is the ancestral state for Monadofilosa, Cercozoa and the bikonts as a whole. If their unikont ancestor had a cone plus one band, as in pelobionts, simply doubling the cilium and its root would yield this very pattern.

Thus, the microtubular patterns of the Monadofilosa and Reticulofilosa can be understood as divergent specialization of the ancestral gymnomyxan pattern caused by the origin of the biciliate condition and its specialization for three modes of feeding (sessile isokonty, mobile anisokonty and temporarily sessile axopodial) plus the secondary loss of the anterior cilium and of phagotrophy on chlorarachnean algal zoospores following the symbiogenetic acquisition of a green-algal plastid (see below). As this centriolar root diversity can be understood as divergent adaptations of an ancestral body plan, it casts no doubt on the relationship between Monadofilosa and Reticulofilosa supported consistently by rRNA (Cavalier-Smith, 2000a), tubulin and actin trees (Keeling, 2001) and by unique rRNA signatures.

The origin of the simple cruciate root pattern of *Plasmodiophora* (class Phytomyxea) is less obvious. In photokaryotes and dinoflagellates, the only other taxa with cruciate roots, the anterior and posterior roots are visibly different. Only in plasmodiophorids, the most phylogenetically divergent lineage within Cer-

cozoa, are they identical (Moestrup, 2000). This difference is consistent with their having evolved entirely independently of the cruciate roots of photokaryotes. Their simplicity may be related to the fact that they are the skeleton of a zoospore for a parasite whose ancestors abandoned phagotrophy. Understanding their origin is hampered by the absence of molecular or detailed ciliary root structure for phagotrophic phytomyxans. Since the plasmodial non-ciliate haplosporidian and paramyxid parasites of invertebrates share a deletion of a G in a highly conserved region of 18S rRNA uniquely with the Phytomyxea (T. Cavalier-Smith and E. E. Chao, unpublished results), predominantly plasmodial parasites of plants and protists, and virtually all other cercozoans, it is highly probable that they are secondarily non-ciliate cercozoans that have lost ciliary roots entirely. I therefore group Phytomyxea and Ascetosporea together as a new cercozoan subphylum, Endomyxa. The claim that paramyxids and haplosporidia are unrelated (Berthe *et al.*, 2000) has been firmly refuted by a more thorough phylogenetic analysis of 18S rRNA (T. Cavalier-Smith and E. E. Chao, unpublished results).

Evolution of photosynthetic eukaryotes

Symbiogenesis and the single origin of chloroplasts and the Plantae

Evidence for a single symbiogenetic origin only of chloroplasts (Cavalier-Smith, 1982d) is now compelling, as detailed elsewhere (Cavalier-Smith, 2000b). Monophyly of the kingdom Plantae *sensu* Cavalier-Smith 1981b, comprising Viridaeplantae (land plants and Chlorophyta), Rhodophyta and Glaucophyta, is almost equally solidly established (Moreira *et al.*, 2000), but denser taxon sampling is needed for concatenated protein trees before we can be sure that the Plantae are holophyletic (as I think) rather than paraphyletic. It is probable that Viridaeplantae and Rhodophyta are sisters and Glaucophyta is the outgroup (Cavalier-Smith, 2000b; Moreira *et al.*, 2000). The host was undoubtedly an aerobic, phagotrophic, biciliate anterokont. If glaucophyte cortical alveoli are homologues of those of alveolates and the ciliary root multilayered structures (MLS) of plants and dinoflagellates are also homologues, then the host for the symbiosis must also have had cortical alveoli and an MLS, as postulated originally (Cavalier-Smith, 1982d), since alveolates are clearly an outgroup to Plantae. Incidentally, if this is true, cortical alveoli are not a synapomorphy for alveolates, as is often incorrectly asserted. If cortical alveoli of raphidophyte heterokonts are homologous and the haptophyte haptonemal reticulum is also related, cortical alveoli might have originated in the ancestral photokaryote (Fig. 4) and were lost in cryptophytes and the common ancestors of red algae and green plants. It appears that plastids have never been lost in Plantae, even when photosynthesis is lost, presumably because their ancestor lost the host's fatty acid synthetase and became

dependent on that of the former cyanobacterial endosymbiont instead (Cavalier-Smith, 1993c).

Secondary symbiogenesis and the origin of chromalveolates

Plastids of chromists (Cavalier-Smith, 1986) and alveolates (Cavalier-Smith, 1991f) originated by the symbiogenetic uptake of a eukaryote alga by a biciliate anterokont host, probably a heterotroph (Cavalier-Smith, 1982d, 1986). I have argued that chromists and alveolates are sister groups, that the clade comprising them (chromalveolates) originated in a single secondary symbiogenetic event by the uptake of a unicellular red alga and that the resulting eukaryote chimaera evolved chlorophyll c_2 prior to diverging into chromists and alveolates (Cavalier-Smith, 1999, 2000a, b); thus, all chromophytes (algae with one or more of chlorophyll c_1 , c_2 and c_3) are descended vertically from a common photosynthetic ancestor. Overconfident assertions of chromist polyphyly based on chloroplast 16S rRNA trees (Oliveira & Bhattacharya, 2000) are entirely unconvincing, as this molecule suffers from weak resolution and systematic biases (Zhang *et al.*, 2000); in fact, the maximum-likelihood chloroplast 16S rRNA tree of Tengs *et al.* (2000) does show chromists as monophyletic! The recent discovery by Fast *et al.* (2001) that a duplicate of the gene for the cytosolic version of glyceraldehyde-phosphate dehydrogenase has been retargeted to the plastid and replaced the red-algal gene in dinoflagellates, sporozoa, cryptomonads and heterokonts virtually proves that chromalveolates are monophyletic. If chromalveolates are sisters of Plantae, as indicated in Fig. 4, then photokaryotes also are monophyletic.

If this is true, and glaucophyte and alveolate cortical alveoli are homologous, then the ancestral photokaryote also had cortical alveoli. The use of the term photokaryotes (Cavalier-Smith, 1999) for chromalveolates and plants together was not meant to imply that their common ancestor was photosynthetic. Although Häuber *et al.* (1994) proposed that the host for the red-algal symbiosis was a photosynthetic alga like *Cyanophora* (kingdom Plantae), I have always thought it more likely that the chromalveolate's host was a heterotroph, both because the selective advantage of its enslaving a red alga would have been far greater and because phagotrophy is almost entirely absent in the plant kingdom. However, this is not yet firmly established and might possibly be incorrect, in which case photokaryotes would be ancestrally photosynthetic.

Heterochrony, cell symmetry and origins of cruciate roots

If alveolates ancestrally had three ciliary roots, as argued above, and are also sisters of chromists, as is now solidly established (Fast *et al.*, 2001), then it follows that cruciate roots originated independently in plants and chromists. The ancestor of both groups

would have had the three-root condition, as in excavates: a single fan-like root on the younger centriole and two roots on the older one. This condition could evolve into the cruciate pattern merely by advancing the development of the two roots so that they formed directly in association with the younger root instead of replacing a single anterior fan-shaped root. Thus, it can be seen as an example of heterochrony: changing the timing of developmental events rather than evolving a radically new structure. The selective advantage of this change would be the economy resulting from the simplification of development, reducing the need to disassemble the anterior root and replace it by a different pattern. In the ancestral phagotrophic lousozoan that first evolved ciliary and ciliary root transformation, the replacement was essential for giving an asymmetric cell structure with feeding groove and heterodynamic cilia to direct food into it. After the symbiogenetic origin of chloroplasts, when early plants gave up phagotrophy, this advantage was removed, but its developmental cost remained. By neotenually advancing the maturity of the younger cilium, they evolved a more symmetrical cruciate arrangement of roots. Given the probable ease of such heterochronic evolution of cruciate roots and a selective advantage for it when photosynthesis replaced or augmented phagotrophy, it is not surprising that the roots are predominantly cruciate in both plants and chromists and reasonable to argue that this condition arose convergently in the two photokaryote kingdoms. In both cases, their roots probably became cruciate when their excavate-like ancestors discontinued feeding via a posterior groove and moved their subapical centrioles closer to the cell equator, making their anisokont cilia laterally inserted, as in the glaucophyte *Cyanophora*, the prasinophyte *Nephroselmis*, the cryptomonad *Hemiselmis* and the heterokont *Pseudobodo*.

It is likely that the plant cruciate condition evolved once only in the ancestral plant; the plant MLS roots are probably derived from similar structures that first evolved in the bands supporting the lips of the excavate groove, as in retortamonads. Chromists were ancestrally photophagotrophs with two dissimilar cilia, but may have depended more on photosynthesis than phagotrophy. Many have retained phagotrophy, but the three groups do it in ways different from each other and from the putatively ancestral posterior excavate groove. Cryptists evolved an anterior gullet and ejectisomes and seem to have dispensed with the posterior right root, r2. Ancestral heterokonts necessarily retained this r2 root, as they modified it into an active entrapment device to catch prey propelled to the base of the anterior cilium by the reversed flow caused by its rigid tripartite hairs. Haptophytes evolved a haptonema for catching prey and so became very different from their probably sister heterokonts, losing the tubular hairs that characterize most other chromistan cilia (Cavalier-Smith, 1994). Hair loss was essential for the evolution of functionally correlated forward-pointing haptonema and homodynamic iso-

kont cilia (Cavalier-Smith, 1994). Predatory prymnesiophyte haptophytes retain this condition and cruciate roots, but the purely photosynthetic Pavlovophyceae became secondarily anisokont, moving the kinetid to the cell apex and therefore losing the ciliary roots associated with the anterior cilium.

Novel trophic methods were important not only for the fundamental variants of the chromist body plan, but also in further variation. Within heterokonts, Dictyochea evolved axopodial feeding by removing microtubular roots altogether from the centrosome and nucleating microtubule triads instead on the nuclear envelope. Independently, the diatoms lost all microtubular roots as a consequence of the evolution of silica frustules and total abandonment of phagotrophy. Within the ancestrally phagotrophic and anisokont motile Chrysophyceae, synurids lost phagotrophy and became pure phototrophs, rearranging their cilia to become often sessile or colonial algae with parallel centrioles, necessarily greatly modifying their roots in the process. In the Opalinata, the movement of the ciliary hairs onto the body of *Proteromonas* and their loss in *Karotomorpha* were adaptively associated with the switch from phagotrophy to osmotrophy as gut symbionts (Cavalier-Smith, 1998a). The associated movement of the kinetid to the cell apex and the reorientation of the centrioles to direct both cilia posteriorly entailed a fundamental change in cell symmetry, probably necessitating the loss of the two anterior roots of the ancestral cruciate pattern.

The simpler cruciate roots of dinoflagellates almost certainly evolved independently, but probably also through the movement of the kinetid to a more equatorial position, in association with the evolution of the dinokont ciliary pattern with transverse and longitudinal grooves—in contrast to the apical insertion of apicomonads and Sporozoa, a more markedly derived condition within Miozoa compared with the putatively ancestral lateral arrangement in *Colponema*. The difference between the ancestral excavate pattern, with a single anterior centriolar fan, and the multiply derived cruciate patterns is primarily the result of the shift from a highly asymmetrical cell with subapical kinetid, necessitated by the posterior feeding groove, to a more symmetrical cell with a lateral kinetid, for which the cruciate pattern provides a more suitable skeletal support. It is this change in cell symmetry, rather than the various trophic innovations that caused it, that is the primary reason for the origins of cruciate roots. Symmetrical cruciate roots, once evolved, remain compatible with secondary isokonty and symmetrical apical insertion of forward-directed cilia, as in chlorophyte and ulvophyte green algae and many prymnesiophytes.

Centriolar roots are not just arbitrarily differing three-dimensional structures, to be described and labelled (Moestrup, 2000). Protist roots are subject to occasional radical evolutionary transformations that yield a functionally significant diversity of microtubular patterns that we can understand, if we appreciate

their fundamentally adaptive changes in relation to adaptive trophic shifts (Sleigh, 2000) and changes in cell symmetry. Functional insights also enable us better to assess their respective weights in phylogenetic reconstruction and to establish a convincing overall picture of eukaryote cell diversification, especially if we root the tree correctly.

Multiple plastid losses and replacements

The chromalveolate theory assumes multiple plastid losses within the group, in marked contrast to Plantae. We have solid molecular-phylogenetic evidence for three already in chromists (Cavalier-Smith *et al.*, 1995, 1996b) but, if I am right, at least three more must have occurred in the kingdom and many more in alveolates. I estimate that at least 18 independent plastid losses must have occurred in chromalveolates. We have demonstrated at least eight independent losses of peridinin-containing chloroplasts in dinoflagellates (Saldarriaga *et al.*, 2001). Three involve replacements by foreign plastids: from a green alga in *Lepidodinium*, a haptophyte in the *Gymnodinium breve* group (Delwiche, 1999; Tengs *et al.*, 2000) and a diatom in *Peridinium balticum* and *Peridinium foliaceum*. The first two cases are genuine examples of tertiary symbiogenesis, in which the host must have evolved novel protein-targeting mechanisms into the new chloroplast, as in primary and secondary symbiogenesis. However, the diatoms might instead be simply obligate symbionts. Whether the chloroplasts of cryptomonad affinity in *Dinophysis*, which branches on rRNA trees among the Peridinea (Saunders *et al.*, 1997), are genuine tertiary plastid replacements or temporarily stolen plastids (kleptochloroplasts) is unclear.

Green secondary symbiogenesises

Though I tentatively suggested that the chloroplasts of green-algal origin in chlorarachnean and euglenoid algae may have had a common secondary symbiogenetic origin (Cavalier-Smith, 1999), the re-evaluation of eukaryote cytoskeletal evolution outlined above now leads me firmly to reject this 'cabozoan' hypothesis. If the Euglenozoa are nested well within the excavates, as Figs 3 and 4 imply, but Cercozoa are outgroups not merely to excavates but to corticoflagellates as a whole (Fig. 4), then these two phyla cannot form a clade on their own or including Percolozoa (which seem to be related to Euglenozoa; Baldauf *et al.*, 2000), as the cabozoan thesis (Cavalier-Smith, 1999) postulated. Therefore, I now accept the traditional view that these were two separate secondary symbiogenesises. There were therefore probably no chloroplast losses within the Cercozoa; the green-algal plastid (Ishida *et al.*, 1999) was probably acquired by a secondarily uniciliate cercozoan flagellate with a simplified cytoskeleton and a propensity to form filose

pseudopods (like some purely phagotrophic cercozoans).

Though the euglenozoan symbiogenesis could have taken place within the euglenoids, as traditionally assumed, it might have occurred earlier, as we know that plastid loss has occurred in euglenoids (Linton *et al.*, 1999); as additional losses might have occurred within Euglenozoa or even in stem Percolozoa, the possibility that euglenozoans or discicristates as a whole were ancestrally photosynthetic (Cavalier-Smith, 1999) cannot be excluded. Indeed, the simplest interpretation of a 6-phosphogluconate dehydrogenase gene of cyanobacterial affinity in the percolozoan *Naegleria* (Andersson & Roger, 2002) is that it is a relic of an early implantation of the green-algal plastid prior to the divergence of Euglenozoa and Percolozoa. Even the much more divergent homologue found in trypanosomes (Andersson & Roger, 2002) might also be. The presence of apparently laterally transferred cyanobacterial-like glucokinase genes in diplomonads and parabasalids (Wu *et al.*, 2001), which may also be (more distantly) related to euglenoids (Figs 3 and 4; unless the relationship shown by some trees is a long-branch artefact), raises the possibility that green-algal chloroplasts might have been implanted in a common ancestor of Euglenozoa and Archezoa and that archezoan ancestors lost plastids as well as peroxisomes and mitochondria (or converted them into hydrogenosomes) when becoming anaerobic. The absence of a homologous glucokinase from chloroplasts and all other eukaryotes does not favour this interpretation, however, unless it was lost subsequent to the euglenoid symbiogenesis. More likely, this was a case of lateral transfer from bacterial food of an aerobic common ancestor of Parabasalids and Metamonada. If a homologous gene cannot be found in any other excavate (or any other eukaryote), this lateral transfer event will be compelling cladistic evidence for the holophyly of Archezoa (*sensu* Cavalier-Smith 1998a and subsequently). A green-plant-like vacuolar H⁺-pyrophosphatase present also in kinetoplastids and alveolates has been cited as possibly derived from secondary symbiogenesis (Drozdowicz & Rea, 2001). However, although this enzyme appears to be absent from animals and fungi, homologues are present in both posibacteria and archaeobacteria, so might have entered eukaryotes vertically and been lost by opisthokonts. Both type-I and -II enzymes are present in eobacteria, so might even date back to the first cell – pyrophosphate metabolism probably preceded that based on ATP (Cavalier-Smith, 2001). The possibility that kinetoplastids got it from the green-algal symbiont now persisting in euglenoids cannot currently be excluded, but the fact that it has not yet been found in cyanobacteria, despite being found sporadically in all but one other bacterial phyla, casts doubt on the idea in its simplest form. The sporadic distribution of this enzyme may owe more to frequent differential losses than to lateral transfers as assumed by Drozdowicz & Rea (2001).

Eukaryote phylogeny: the big picture

Are there any early-diverging eukaryotes?

To summarize the foregoing discussion, we now have a remarkably simple picture of the eukaryote tree (Fig. 3). It is basically bipartite: on one side is a purely heterotrophic and probably ancestrally uniciliate clade, the opisthokonts (i.e. Choanozoa, Animalia, Fungi). On the other, we have a large, nutritionally heterogeneous clade, the anterokonts, comprising Amoebozoa (probably ancestrally unikont), and the bikonts (probably ancestrally biciliate). Bikonts comprise two major protozoan groups (Rhizaria and Corticata) and the ancestrally photosynthetic anterokont kingdoms Plantae and Chromista. Thus, the basal bifurcation led to animals and fungi on the one hand and plants, chromalveolates, excavates and the Rhizaria on the other. Not only does this give a simple interpretation of ciliary and cytoskeletal evolution, but it is consistent with a more qualified version of the long-standing idea of an early split between eukaryotes with flat mitochondrial cristae and those with tubular mitochondrial cristae, immediately following the origin of mitochondria (Taylor, 1978; Cavalier-Smith, 1982d). Opisthokonts ancestrally had flat cristae and anterokonts had tubular cristae. However, cristall morphology has not been invariant in either group: opisthokonts have secondarily evolved tubular cristae twice within animals, vesicular cristae in the Ichthyophonida within the Ichthyosporidia and discoid cristae in the Cristidiscoidea. Similarly, the ancestrally tubular cristae of anterokonts became secondarily flattened within the Apusozoa (in *Ancyromonas*), Cercozoa (e.g. in biomyxids) and in early Cryptista, Discicristata and partially to make the irregularly flattened cristae of the Plantae.

How confident can we be in this new interpretation? Although it is not yet certain that the root lies between opisthokonts and anterokonts or instead on the bikont side of (or within) the Amoebozoa, it cannot lie within the opisthokonts [because all have a unique derived insertion in protein synthesis elongation factor EF 1- α (Baldauf, 1999) as well as uniquely derived rRNA signature sequences] or within the corticoflagellates (as they have a derived fusion of thymidylate synthase and dihydrofolate reductase genes; Philippe *et al.*, 2000) and it is almost certainly not within the Cercozoa (all of which have a derived rRNA signature sequence). What is now evident is that the common description of the Archezoa and Euglenozoa as 'early-diverging eukaryotes' (e.g. Bouzat *et al.*, 2000; Watanabe & Gray, 2000) is definitely wrong. The root cannot be among the Corticata and must lie within the Gymnomyxa, very probably within the Sarcomastigota. It is highly improbable that the root can be among the bikonts, because of the complexity of their ciliary/cytoskeletal differentiation. Thus, the major outstanding question about the eukaryote tree, apart from the precise position(s) of the Heliozoa, the position of minor taxa such as the Kathablepharida and Disco-

celida (both of which might be either corticate protozoa or degenerate chromists) and whether rhizaria are paraphyletic or holophyletic, is whether the amoebozoa are sisters of bikonts, as I think, or sisters to opisthokonts or to opisthokonts plus bikonts. If amoebozoa prove to be holophyletic, then either amoebozoa as a whole are the earliest-diverging eukaryotes or else there are none. Only if amoebozoa prove to be paraphyletic and basal to all eukaryotes would there be any extant eukaryotes (certain amoebozoa) that diverged prior to the fundamental bifurcation between the two major derived clades: opisthokonts and bikonts.

To reduce these residual uncertainties, we are currently focusing our attention on the Apusozoa, Amoebozoa and Heliozoa in order to pinpoint the position of the root and of Amoebozoa more precisely and establish more confidently that there are no extant early-diverging eukaryotes.

A general implication of the present view is that, in order to reconstruct the nature of the cenacestral eukaryote, we should look for all those features that are shared between animals and/or fungi on the one hand and plants/chromalveolates/excavates/amoebozoa on the other. Simply comparing animals and fungi like yeasts alone can be misleading, as both are opisthokonts, representing only one side of the eukaryotic tree. Higher fungi are radically derived by having lost cilia and simplifying the cytoskeleton (like flowering plants) and by secondarily unstacking their Golgi complex. Yeasts are also simplified in other less obvious ways, e.g. in cell-cycle controls: the bikont green alga *Chlamydomonas*, like animals but unlike yeasts, has a retinoblastoma protein involved in cell-size-related cycle control (Umen & Goodenough, 2001) that has presumably been lost by ascomycete yeasts; this gene must have been in the cenacestral eukaryote cell. Yeasts are not 'lower eukaryotes', but highly advanced, specialized ones. If my rooting is correct, the amoebozoan *Dictyostelium* is on the plant side of the tree, making a comparison between it and plants useful in reconstructing the cenacestral anterokont; but, because it also has lost cilia, it will tell us much less about the nature of the ancestral cytoskeleton than would a flagellate like *Phalansterium* or *Mastigamoeba*. If the eukaryotic root is between opisthokonts and anterokonts, we can say confidently that the cenacestral eukaryote was an aerobic, uniciliate, sexual protozoan with all the cell organelles now found in *Phalansterium* and that all anaerobic or non-ciliate eukaryotes are secondarily derived by losses or drastic modification of basic organelles.

Late polyphyletic origins of hydrogenosomes

Against this backdrop, let us look at the origin of hydrogenosomes. The polyphyly of anaerobic eukaryotes is now well established (Roger, 1999), as is the polyphyletic origin of hydrogenosomes (Müller & Martin, 1999). Despite their polyphyly, it seems that

all evolved from mitochondria and not merely most of them, as I originally suggested (Cavalier-Smith, 1987e). Clearly, mitochondria have a marked evolutionary propensity to switch permanently from aerobic to anaerobic ATP generation, using hydrogenase to generate molecular hydrogen, when oxygen is unavailable to accept waste electrons and protons. Hydrogenosomes have evolved in this way in the common ancestor of the Parabasalia, several times in ciliates, several times in excavates [in the Percolozoa (*Psalteriomonas*) and very likely (but biochemical confirmation is wanting) in *Postgaardi* in Euglenozoa and *Trimastix* and *Carpediemonas* in Loukozoa] and in chytridiomycete fungi.

All these origins are so much after the primary diversification of aerobic eukaryotes that it is unlikely that all their hydrogenases were inherited vertically from the proteobacterial ancestor of mitochondria, since, although it is present in a few photosynthetic eukaryotes, it is unknown in heterotrophs. The origin of pyruvate-ferredoxin oxidase is more problematic; all eukaryotic enzymes are related and, though of eubacterial origin (Horner *et al.*, 1999), the possibility that it is of protomitochondrial origin is neither excluded nor demonstrated. As it is found in some aerobes, it might have lurked cryptically with a minor function for millions of years and then resurfaced later during the secondary evolution of anaerobiosis. Whether lateral transfer of genes from other anaerobes played a major or a minor role in the multiple origins of hydrogenosomes remains unclear. However, we are now confident that, in the biochemically well-studied fungi and trichomonads, the organelle as a whole did not evolve through separate symbiogenesis, as suggested some time ago (Whatley *et al.*, 1979).

Thus, there are only seven well-demonstrated cases of successful symbiogenesis in the entire history of life: mitochondria from α -proteobacteria, chloroplasts from cyanobacteria, chromalveolate plastids from red algae, euglenoid and chlorarachnean plastids separately from green algae and the tertiary replacements of dinoflagellate peridinin-containing plastids by green (*Lepidodinium*) and haptophyte (e.g. *Gymnodinium breve*) plastids. Whether mitochondria originated in a phagotrophic host, as traditionally argued, or in an autotrophic methanobacterium, as suggested by the hydrogen hypothesis (Martin & Müller, 1998), is entirely irrelevant to understanding either the conversion of mitochondria into hydrogenosomes or the equally polyphyletic loss of mitochondria in type-I eukaryote anaerobes like the Archamoebae, Metamonada and Microsporidia. The assertion of Doolittle (1998b), that the hydrogen hypothesis more readily explains the energy metabolism of amitochondrial eukaryotes, is totally unjustified. If we accept that all anaerobic eukaryotes evolved from aerobic mitochondrial ancestors, as I now do (Cavalier-Smith, 1998a, b, 2000a), then both the phagotrophy theory and the hydrogen hypothesis accept that hydrogenosomes evolved from mitochondria and that other eukaryotic

anaerobes have totally lost them. The dilemma of whether to explain the origins of anaerobe-specific enzymes by vertical descent or lateral transfer is exactly the same for both.

Evolution of eukaryotic life-cycles: starvation, cysts and sex

Origin of protozoan cysts

Every protozoan phylum is able to make resting cysts or spores, so it is highly probable that this capacity was present in the ancestral eukaryote. I suggested (Cavalier-Smith, 1987c) that the ability to make a cyst wall and to undergo cytoplasmic differentiation into a resting stage with low metabolism was inherited directly from their actinobacterial ancestors, which typically make exospores (Cavalier-Smith, 1987c). However, too little is known about cyst-wall chemistry and biosynthetic enzymes and the molecular biology of encystment and excystment in protozoa to test this hypothesis in detail. Chitin is present in many animals, most fungi and some chromists, as well as in the actinobacterium *Streptomyces*, and may therefore have been present in the cysts of the eukaryote ancestor. The enzymes that make it are homologous in animals and fungi and were obviously present in the ancestral opisthokont; others are not certainly identified, though several bacteria including actinobacteria have distantly related glycosyl transferases. Cellulose is found not only in plants and chromalveolates, but also in some animals (tunicates), rarely in fungi and in some amoebozoa such as the mycetozoan *Dictyostelium* and the lobosan *Acanthamoeba*. If the eukaryote root is between opisthokonts and anterokonts, cellulose might also have been present in the ancestor. The *Dictyostelium* cellulose synthase is related to those of bacteria and plants (Blanton *et al.*, 2000), but conservation is poor and evolutionary analysis complicated because not all related glycosyl transferases need make the same product. Unfortunately, we know nothing about cyst-wall chemistry in key groups such as the Apusozoa, Choanozoa or Cercozoa, which makes it hard to reconstruct the ancestral state. However, it is highly probable that the ancestor was a naked flagellate or amoeboid flagellate in its trophic stage that differentiated into a resting cyst when starvation conditions set in. This is especially likely if it lived in freshwater or soil, where most species form resting cysts, but, even in marine protists, resistant walled resting stages are phylogenetically widespread and may be an ancestral condition.

Origin of sex: cell fusion and syngamy

Whether sex was necessary, important (as often assumed, e.g. Schopf, 1970) or relatively trivial for the origin of eukaryotes is unclear to me. But, without phagotrophy, the cytoskeleton, molecular motors and the endomembrane system, evolution would have achieved nothing of interest to us, as we would not be

here. However, the present rerooting of the eukaryotic tree makes it likely that syngamy and meiosis had already evolved by the time of the eukaryotic cenancestor. If the root of the tree is between the opisthokonts and anterokonts, this is certainly true, since the cenancestor of animals and plants would be one and the same as the cenancestor of all eukaryotes and there would be no earlier-diverging extant eukaryotes. If the root is within the Amoebozoa, then there might be an earlier-diverging eukaryote group, such as *Phalansterium* or pelobionts or even one group of naked lobose amoebae (gymnamoebae). Since sex is unknown for *Phalansterium* or pelobionts or most gymnamoebae, the possibility remains that one amoebozoan group may be primitively asexual. Even if one such group is truly early-diverging, and also truly asexual (equally uncertain), I consider it more likely that sex evolved in the cenancestor almost as soon as the origin of phagotrophy created a flexible cell surface. Both the absence of a cell wall and the presence of an internal cytoskeleton would have facilitated the membrane fusion and cell merger that is the basis of syngamy.

Developing a proper understanding of the origin of sex depends on having a realistic picture of the nature of the life-cycles of the first eukaryotes. We should not think of an early, pre-sexual eukaryote as being like a vegetative fission yeast, with nothing to its life but a binary-fission cell cycle, growth, replication and division, all tightly coupled. Two factors, often neglected, may be important to give a more realistic picture: syncytia and cysts.

Cell fusion is relatively widespread among protozoa. Some protozoan life-cycles involving cell fusion also have nuclear fusion and meiosis, whereas others apparently do not and are referred to as agamic (Seravin & Goodkov, 1999). One example of these is the cellular networks formed by the fusion of filopodia in the cercozoan alga *Chlorarachnion*, the haptophyte *Reticulosphaera* or the protist *Leucodictyon*; this unusual morphology evolved independently at least in *Reticulosphaera* and *Chlorarachnion* (Cavalier-Smith *et al.*, 1996c). Such cellular nets (reticulo-meroplasmodia; Grell & Schüller, 1991) are probably adaptations for phagotrophic feeding on surfaces. They might be a way of sharing sparse prey among kin, analogous to the widespread searching for and sharing of food by ants. Several cercozoan amoeboid flagellates, e.g. some species of *Thaumatomonas* or *Cercomonas*, can undergo cell fusion to form temporary multinucleate cells that can feed as a microplasmodium and later divide to form uninucleate ones. Several cercozoans and most retaria have reticulopodia, the formation of which requires membrane fusion between different filopodia of the same cell. For such a cell, it should be very easy to evolve the capacity to fuse with another cell of the same species. Several amoebozoa can undergo cell fusion to form multinucleate cells (e.g. *Leptomyxa reticulata*) or extensive plasmodia, as in slime moulds, which can be metres across. The diverse examples of

cell fusion, many apparently not part of sexual cycles, are important for understanding the origin of sex, since they imply that there are several selective advantages for evolving cell fusion in protozoa that need have nothing to do with recombination. They also indicate that, for a gymnomyxan protozoan with a soft, naked surface, cell fusion is mechanistically easy to evolve and has probably done so on numerous occasions.

For an early eukaryote with such a high propensity for membrane fusion, it is the temporal control of cell fusion, rather than its basic mechanism, that is evolutionarily most interesting. It is traditional to view organisms as either unicellular or multicellular. But this is an oversimplification. Protozoa can be uninucleate or multinucleate plasmodia or syncytia (the distinction is rather arbitrary). Thus, there are three kinds of eukaryote organisms, unicells, multicells and plasmodia, which are adapted to different broad adaptive zones. We do not know whether sex evolved in a unicellular or in a plasmodial protozoan. But the trophic advantages of plasmodia are such that, even in the prekaryote phase, there would probably have been niches for strictly uninucleate unicells and for others with more complex life-cycles including syncytial phases. Syncytia can form either by cell fusion or by nuclear division without cytokinesis. We may regard primitive eukaryote life-cycles as being made up of four processes: cell growth, nuclear division, cytokinesis and cell fusion. Early life-cycles evolved by the temporal coupling or partial uncoupling of these four processes. An organism with a syncytial phase must have a more flexible coupling between these processes, which it can control facultatively, than a strictly uninucleate unicell. But even the latter may sometimes form multinucleate cells by accidental failures of cytokinesis. Unless the cell had a correction mechanism, such errors would accumulate and syncytia evolve. The complex life-cycle of a syncytial organism like a slime mould is the result of two opposing selective forces: selection for feeding and growth and selection for reproduction. In the myxogastrid slime mould adaptive-zone, selection favours giant cells with billions of nuclei during the growth phase, when they spread over the forest floor, engulfing bacteria by the million. Reproduction favours the formation of millions of separate uninucleate spores to generate as many propagules as possible, most of which will die, in order to colonize new food supplies. But, in the more aquatic *Cercomonas* habitat, more likely the ancestral state, only much smaller, temporary syncytia are favoured. Given that, under certain conditions, syncytia may be advantageous, whether these are produced by cell fusion or by delayed cytokinesis, the relative advantages of unicells and syncytia may depend on the relative density of predators and prey. High prey densities and low predator densities might favour growth and delayed cytokinesis and high predator and lower prey densities might favour predator-cell fusion. It is likely that well-regulated, par-

tially syncytial life-cycles and strictly binary-fission life-cycles are both evolved characteristics and that the original eukaryotic cell cycle was both simpler and sloppier. The yeast and mammalian cell cycles involve numerous checkpoints to ensure that one does not end up with anucleate or multinucleate daughters. A syncytial life-cycle must also have checkpoints.

Whether the cenancestor was unicellular or syncytial, it probably had a walled resting cyst (see above), so karyogamy and meiosis were probably inserted into a relatively complex life-cycle adapted to a feast-and-famine existence, not into a simple binary-fission cell cycle always supplied with food.

A failure in mitosis or an accidental nuclear fusion would make polyploid nuclei. Since the fundamental and universal function of meiosis is ploidy reduction, I have suggested that it may have evolved primarily to repair such mistakes (Cavalier-Smith, 1995b). Mathematical modelling shows that ploidy reduction could have been an effective selective force (Kondrashov, 1994). Its advantage would be stronger the more often polyploidy occurred. However, a key question is whether it evolved before or after the evolution of the nuclear envelope. If it evolved beforehand, there would be no distinction between multinuclearity and polyploidy. Thus, one could envisage an intermediate stage with meiosis and cell fusion but no nuclear fusion. Nuclear fusion need not have evolved till later.

Origin of meiosis

If the root of the tree is between opisthokonts and anterokonts, the ancestral meiosis was almost certainly a two-step meiosis. The idea that a single-step meiosis was the ancestral state (Cleveland, 1947) is almost certainly incorrect. Although the Miozoa were once thought to have single-step meiosis, they actually have a normal two-step one. Whether microsporidia or the amitochondrial flagellates studied by Cleveland (1956) have a single-step meiosis or a normal two-step one remains unclear (Cavalier-Smith, 1995b). It is, however, irrelevant to the origin of meiosis, as microsporidia are highly derived fungi and the amitochondrial flagellates are all excavates, which we now know are derived, not early-diverging, eukaryotes. Whether a single-step meiosis exists in any organism is seriously open to question. If we view meiosis as a modification of a mitotic cell cycle, the nature of these cell-cycle controls would make it much easier to evolve a two-step meiosis than a single-step one (Cavalier-Smith, 1981a).

To evolve ploidy reduction by meiosis requires four things: (i) a homology-search mechanism to enable chromosome pairing; (ii) a delay in the splitting of sister centromeres until the second meiotic division; (iii) blocking DNA replication between meiosis I and II; and (iv) reorienting sister centromeres to face the same pole. Homology search is probably mediated primarily by base-pairing between homologues and is

fundamentally a renaturation process. I have suggested that this occupies a major part of meiotic prophase and is the fundamental determinant of the proportionality between meiotic duration and genome size (Cavalier-Smith, 1995b). I suggested previously that the key step in the evolution of meiosis was the blocking of centromere separation in meiosis I. I postulated that, because of the nature of cell-cycle controls, this would ensure that the second meiotic division proceeds without an intervening DNA replication, if centromere splitting is a necessary signal for the switch from the mitotic state, where replication is inhibited, to the growth state, where it is allowed (Cavalier-Smith, 1981a). The molecular basis of this is still only partially elucidated. Sister chromatids are held together by cohesin proteins, which cross-link them by binding to mcm proteins. In mitosis, chromosome splitting and centromere splitting are both caused by digestion of the cohesins by a protease. The switch from the mitotic state to the growth state (G1 phase) of the cell cycle, where replication is allowed, is also mediated by proteolytic digestion – of the cyclin B attached to the cyclin-dependent kinase, which is the switch (Iwabuchi *et al.*, 2000). However, a direct causal connection between centromere splitting and cyclin digestion has not yet been demonstrated. If there is such a connection, then the essentials of meiosis could have originated by a single key change, as I proposed (Cavalier-Smith, 1981a), simply by the blockage of centromeric cohesin digestion in meiosis I.

It is known that all eukaryotes have homologous cohesins, which must have arisen in their cenancestor during the origin of mitosis. Opisthokonts at least have distinct meiotic cohesins. These necessarily assemble during premeiotic S phase (Smith *et al.*, 2001), which makes meiosis necessarily two-step (Watanabe *et al.*, 2001). It would have been easier for meiosis-specific cohesins to be inserted by the normal mechanisms used for mitosis onto sister centromeres than for a novel cell-cycle control to evolve that allowed a reductional division without a preceding S phase. In meiosis I, these cohesins are digested in the chromosome arms, but not at the centromeres, which therefore remain unsplit (Buonomo *et al.*, 2000). Likewise, cyclin B remains partially undigested in the period between meiosis I and II (Iwabuchi *et al.*, 2000). Thus, this intervening period is not a true interphase and would not be competent to initiate replication; as I assumed (Cavalier-Smith, 1981a), the cell essentially remains in M phase until anaphase II, when the centromeres split, cyclin B is digested and the cell reverts to G1. Thus, present understanding fits the view that evolution of a block to centromere splitting would be a necessary and sufficient mechanism for the ancestral mechanism of meiosis (Cavalier-Smith, 1981a) and explains why it is two-step. In baker's yeast, centromere reorientation to ensure that sister centromeres become attached to spindle fibres attached to the same pole is mediated by a protein known as monopolin (Toth *et al.*, 2000): as this lacks clear homologues in other organisms, it is

unclear whether it is a general mechanism or one of many special simplified characteristics of baker's yeast.

In most eukaryotes, chiasmata and/or a synaptonemal complex are also essential for ensuring accurate meiotic chromosome disjunction during ploidy reduction. The odd cases where one or both of these is dispensed with seem to be phylogenetically derived, so both would have been present in the eucenozoic ancestor. Chiasmata are initiated by double-strand breaks and require exonucleases to generate single-stranded DNA, a RecA protein homologue to allow this to invade a homologous duplex to form hybrid DNA and a Holliday junction resolvase and repair enzymes to complete the crossing over. The double-strand breaks are made by Spo11, part of a heterodimeric enzyme homologous with topoisomerase VI of archaeobacteria, which probably evolved from DNA gyrase in the eucenozoic ancestor (Cavalier-Smith, 2002a). The RecA homologue is Rad51, which also evolved from a eubacterial enzyme in the eucenozoic ancestor, as did the characteristic eucenozoic Holliday junction resolvase (Cavalier-Smith, 2002a). Thus, the basic molecular machinery for crossing over was already present in the eucenozoic ancestor, where it is likely to have been used for recombinational repair by sister-chromatid exchange as in bacteria (Cavalier-Smith, 2002b). Although many of these genes are now meiosis-specific and must have special meiosis promoters, this need not have been so initially; they could have been constitutive. The precise role for the synaptonemal complex is unclear. Since ploidy reduction can occur in its absence, it is probably there to increase the efficiency and reliability of meiosis and so could have been added after it began; it was clearly present in the eukaryote eucenozoic ancestor. Its proteins evolve too rapidly to be very useful for phylogeny, but two of them have coiled-coil motifs similar to myosin tails, suggesting that the complex might have originated by recruiting cytoskeletal and/or motor proteins.

It is likely that the recombination caused by sex is an incidental consequence of crossing over that evolved primarily to ensure disjunction during ploidy reduction, rather than the major selective advantage for the origin of sex. Although recombination does have some advantages, they are probably small compared with the advantages of ploidy reduction, whether through correcting the errors in division or accidental fusion. The high frequency of loss of sex in protists is consistent with the theoretical conclusion that recombination has a clear advantage over clonal reproduction only under special conditions (Lenormand & Otto, 2000). It can speed up evolution by combining rare mutations together. I suggest that meiosis originated as a cell-cycle repair mechanism to correct accidental polyploidy when the eukaryotic cell cycle was first evolving. It became temporally coupled to the induction of Spo11 to make double-strand breaks and allow recombinational repair prior to the germination of resting cysts. The incidental recombinational effect of this could itself also have been beneficial if this

occurred during the other processes of eukaryogenesis discussed above, since mutations in hundreds, perhaps thousands of genes might simultaneously have been subject to directional selection. Thus, sex could have begun during the largest transformation in cell organization in the history of life and been swept to fixation by the success of the phagotrophs that its recombinational side-effect helped establish, by combining valuable novelties in a single cell and enabling the culling of hopeless monsters more efficiently and at lower cost. Perhaps it did play a key and positive role. Its wide persistence in protists may have as much to do with epigenetic constraints caused by its causal linking to cyst formation and excystment as with its recombinational advantages, which are not so overwhelming as to prevent frequent losses.

Palaeontology and megaevolution

Fossils, bushes and the two biological big bangs

Earlier (Cavalier-Smith, 1987a, 1991a), I argued that the unresolved bush at the base of the eubacterial tree represents the primary diversification of eubacteria into the different niches in the first microbial mats, immediately after the origin of the first photosynthetic cell, about 3.5 Gy ago. I still think that this is essentially true, though I have argued that Eobacteria (Chlorobacteria and Hadobacteria) probably diverged just prior to the major radiation (Cavalier-Smith, 2002a). Eukaryotic protein trees also show the major lineages as emerging from a poorly resolved bush. Thus, instead of three domains of life, we actually have just two bushes: an ancient bacterial bush and a modern eucenozoic bush, each representing the primary adaptive radiation of the two fundamental types of living organism: bacteria and eukaryotes. All talk of 'early-diverging bacteria' or 'early-diverging eukaryotes' is probably misleading. 'Bush' is a better metaphor than 'tree'. Trees with long, unbranched stems are usually signs of quantum evolution, not accurate pictures of the past (Cavalier-Smith, 2002a). The bush-like character of eukaryotic radiation is corroborated well palaeontologically by the Cambrian explosion. For the bacterial explosion, we have to rely on the molecular trees, making allowances for obvious systematic biases. Archaeobacteria and all the other bacterial phyla show early radiations too.

An early Archaean eubacterial big bang, nearly three billion years of stasis and a late Proterozoic eucenozoic big bang should only be counter-intuitive to 18th-century uniformitarians. Those who have thought most deeply about evolution, among whom Simpson (1953) really stands out, argue that this is exactly what one should expect. As soon as a major new body plan with a distinctive ecological role evolves, it radiates rapidly in the absence of competition from any previous organisms in the same broad adaptive zone. In so doing, its descendants create new niches for themselves and other organisms. As Simpson (1944) showed, this

rapid early radiation is not only theoretically predictable but is empirically the general rule for every group of organisms with a good fossil record. The only significant exception is that the major radiation of mammals was delayed, long after their origin, until dinosaurs became extinct; however, if dinosaurs really were warm-blooded, this actually proves the rule, for mammals did not occupy a unique adaptive zone (warm-blooded, terrestrial vertebrates) until the late Cretaceous extinction eliminated the dinosaurs. This emphasizes that it is not so much novelty in body plan as in adaptive zone that triggers massive radiation. In the case of eukaryotes, the novelty was phagotrophy. After billions without it, this triggered the late Proterozoic/Cambrian explosion of protists (Cavalier-Smith, 2002a). However, the new body plan was important. The key innovations were the cytoskeleton and its molecular motors and the endomembrane system; without these, neither complex protozoa nor animals, plants, fungi or chromists were possible. Elsewhere, I outlined the evidence that this occurred around 800–850 My ago (Cavalier-Smith, 2002a), essentially the same time as estimated originally (800 My ago) for the mitochondrial symbiogenesis (Cavalier-Smith, 1983a). My present interpretation differs in two key respects: (i) the mitochondrial origin was essentially simultaneous with the origin of the pre-eukaryote host and (ii) the chloroplast symbiogenesis was distinctly later than that for mitochondria, as long assumed by the serial symbiogenesis theory (Taylor, 1974), not simultaneous with it (Cavalier-Smith, 1983a, 1987e).

The origin of chloroplasts as the trigger for the Cambrian explosions?

Animals could not have evolved prior to the evolution of eukaryotes. As they appear to have evolved scarcely any later than the major protozoan and algal lineages, the animal Cambrian explosion is best seen as a simple extension of the basic protist big bang, not a separate phenomenon requiring its own explanation. Contrary to what many have recently gratuitously assumed, there was no slow-burning fuse involving unfossilized animals radically older than those revealed in the rocks. If one were to seek an external factor beyond the creation of a zooflagellate ancestor, it would be the origin of chloroplasts and eukaryotic algae (Cavalier-Smith, 1983a, 1987e). It is hard to believe that the animal explosion could have been either as extensive or as rapid if the only photosynthesizers were bacteria, whether free-living or symbionts in protozoa. From the protein trees (and Fig. 2), it appears that the bases of the plant and animal kingdoms were roughly contemporaneous. As the oldest indubitable animal fossils are around 570 My old, 580 My ago is a reasonable estimate for the origin of plastids. Geological evidence indicates that, during this Cryogenian period in the late Proterozoic, the Earth underwent more dramatic upheavals in climate and ecology than in any period in the preceding 2000 My. There was a succession of ice ages and vast fluctuations in carbon

isotope ratios, suggesting that photosynthesis and/or the rates of burial of its products were fluctuating hugely over several hundred million years (Brasier, 2000). Might there be some connection between these unprecedented geological upheavals and the biologically unprecedented origin of eukaryotes? Recent interpretations suggest that two of the ice ages (Sturtian, ~ 760–700 My ago; Varangerian, ~ 620–580 My ago) were so dramatic in extent that most, if not all, of the Earth, including the oceans, was buried several miles deep in ice (Runnegar, 2000). Proponents of this snowball-Earth theory (Hoffman *et al.*, 1998) wonder how eukaryotic algae could have survived this giant freeze-up for millions of years. But, if my estimate is correct, there is no problem, as plastids evolved just after the Varangerian snowball Earth melted. If this is true, only bacterial photosynthesizers need have survived the near global glaciations and eukaryotic algae could have originated and radiated immediately after the climate rewarm, with animals following hard on their heels in the Vendian and Cambrian. I consider that Vendian fossils may all be radiate animals and that the Cambrian explosion of fossils itself accurately represents the ultra-rapid diversification into 17 distinct phyla (Cavalier-Smith, 1998a) of the first bilaterian that evolved from an anthozoan about 545 My ago. The explosive, very-late Proterozoic algal and protozoan radiations, for which proteins provide evidence, were predicted on general evolutionary grounds as the inevitable result of the origin of plastids (Cavalier-Smith, 1978a, 1982d) before rRNA trees confused the picture.

Whether the Earth was entirely ice-bound in the two great Cryogenian glaciations except for a few patches around high volcanoes, or instead had an equatorial oceanic band of surface water (Hyde *et al.*, 2000), is a minor quibble in the face of the evidence for the virtual elimination of primary production for millions of years. The Sturtian ice age probably greatly reduced bacterial and protozoan biodiversity; when it ended (700 My ago), the thaw opened new environmental possibilities that facilitated the diversification of eukaryotes. It was probably just coincidence that the neomuran revolution took place and eukaryotes originated about 100 My before the first well-established Cryogenian glaciations (Cavalier-Smith, 2002a); the huge delay after the origin of cells simply reflects the inherent improbability of the steps need to make a eukaryote. However, there are some indications that there may have been one or two other major glaciations just before the possible origin of eukaryotes (Brasier, 2000). This was also the time of the break-up of the great supercontinent Rodinia. By disrupting the long-term stability of global ecosystems, these major changes possibly stimulated the origin of the ancestral thermophilic neomura (Cavalier-Smith, 2002a). Whether these geological changes had such an influence, or the timing was purely coincidental, it is likely that the glaciations in the period 750–680 My ago inhibited the origin of chloroplasts (or extirpated earlier successful

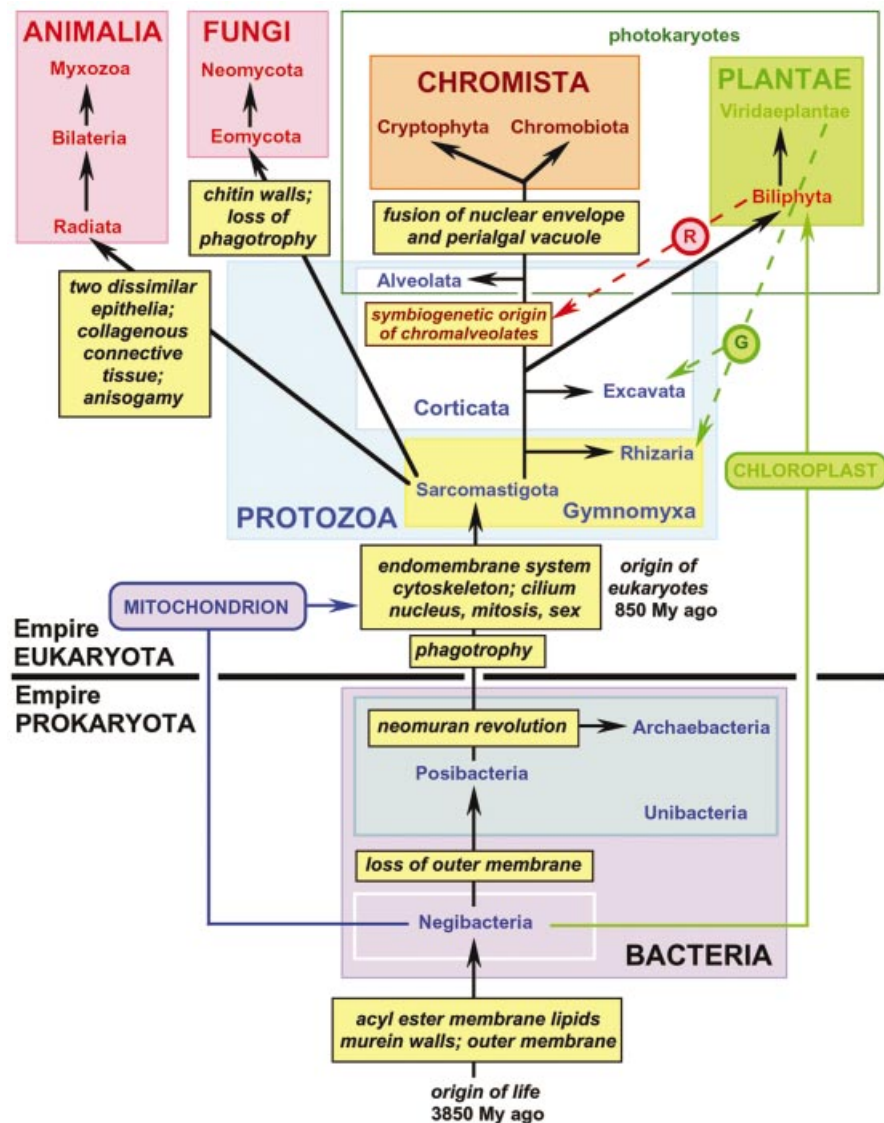


Fig. 5. Revised six-kingdom phylogeny of life showing all 13 subkingdoms. The five major symbiogenetic events in the history of life are shown: primary symbiogenesis with solid coloured lines (origins of mitochondria from an α -proteobacterium and chloroplasts from a cyanobacterium); secondary symbiogenetic enslavement of plant cells to form eukaryote/eukaryote chimaeras shown by dashed lines [the origin of chromalveolates when a biciliate corticate protozoan incorporated a red alga (R) and the independent acquisition of plastids (G) from different green algae by euglenoids and chlorarachneans]. Tertiary chloroplast replacements within dinoflagellates are not shown. The most important non-symbiogenetic events in the tree of life are shown in the black and yellow boxes. For Unibacteria, both phyla are shown. For the basal eukaryotic kingdom, Protozoa, the four infrakingdoms as well as the two subkingdoms are shown.

attempts) and also delayed the origin of higher eukaryotes such as animals. If these Cryogenian glaciations had not occurred, it is likely that the Cambrian explosion would have occurred distinctly earlier, closer to the time of the origin of eukaryotes.

On this view, there was a delay, possibly caused jointly by the Cryogenian environmental upheavals and the absence of true eukaryotic algae, of about 200–270 My between the origin of eukaryotes 850–800 My ago and the essentially simultaneous radiation of the opisthokonts and bikonts about 580 My ago. The fact that it

is very easy to resolve the bipartition between opisthokonts and anterokonts with high bootstrap support on both rRNA trees (Cavalier-Smith, 2000a; also Fig. 2) and protein trees (Baldauf *et al.*, 2000) is consistent with and is simply explained by such a long delay. Likewise, the greater difficulty of resolving the basal branching order within the opisthokonts and within the bikonts is consistent with a very rapid radiation of both groups immediately after the melting of the Varangerian ice. By using the term 'big bang', I do not imply that the radiation were instantaneous, unresol-

vable or incomprehensible, but merely that both radiations occurred in a geologically relatively short time, constituting only a small fraction (probably between 1 and 10%) of the total history of the two groups. The fossil record attests to the reality of this qualitatively dramatic quantum radiation. It was predicted by Darwin (1859), who wrote that, once a new adaptation is perfected, 'a comparatively short time would be necessary to produce many divergent forms'. There is no evidence whatever that stands up to critical examination of either animals or bikonts before 570 My ago (Cavalier-Smith, 2002a). Neontologists should accept the compelling fossil evidence for the simultaneous radiation of animals and protists about 570 My ago, which decisively refutes Darwin's assumption that Precambrian fossiliferous rocks are missing from the record, and reject the naïve 18th-century uniformitarian assumptions of steady rates of morphological or molecular change, which are amply refuted by the direct fossil evidence for both microbes and macrobes (Cavalier-Smith, 2002a).

Whether all extant amoebozoan lineages also initially radiated at the same time as bikonts and opisthokonts is less clear. I suspect that they may have done and that all the eukaryotic fossils between 850 and 580 My ago may have belonged to stem Choanozoa and Amoebozoa or now-extinct sarcomastigote lineages, probably including various 'pseudophytoplankton', protozoa that harboured photosynthetic symbionts that had not made the transition from symbiont to organelle (Cavalier-Smith, 1990b). It is probable that only two lineages from the primary eukaryotic radiation roughly 850 My ago still survive, the opisthokonts and the anterokonts. The primary bifurcation between them may date back to the initial radiation of the first unikont eukaryote.

Envoi: the two Empires of life

In summary, the most far-reaching and difficult steps in the history of life were the origin of bacteria and the origin of eukaryotes (Fig. 5). From the point of view of the fossil record, the history of life is fundamentally bipartite (Schopf, 1994), just as living organisms are fundamentally of only two kinds, bacteria and eukaryotes. Archaeobacteria are of great intrinsic interest as a basically hyperthermophilic bacterial phylum, but they are fundamentally just a special kind of bacterium, the last bacterial phylum to have evolved (Cavalier-Smith, 2002a). The neomuran revolution was a springboard for the subsequent evolution of both eukaryotes and archaeobacteria. But the changes occurring during the origin of eukaryotes were far more radical and far more important for the subsequent evolution of the biosphere than the origin of archaeobacteria. There are very few archaeobacteria-specific characters apart from their membrane lipids and special flagellar-shaft proteins; almost all the main differences between archaeobacteria and eubacteria arose in the common ancestor of archaeobacteria and eukaryotes during the neomuran revolution and are neomuran novelties, not archae-

bacterial ones (Cavalier-Smith, 2002a). However, the existence and characterization of archaeobacteria has been very important for our understanding of eukaryote evolution, since it enables us to make the problem more manageable and comprehensible by breaking it down into two successive phases: the neomuran revolution (discussed elsewhere; Cavalier-Smith, 2002a) and the origin of eukaryotes-proper from an early neomuran bacterium, as outlined here. But, to reconstruct the nature of our bacterial ancestors more thoroughly, we also need to focus more intensively on the actinobacteria and their remarkable structural and chemical diversity.

If the neomuran revolution had not occurred and eukaryotes had never evolved, the world would be very different indeed. But, if archaeobacteria had never evolved, the large-scale structure of the biosphere would be very similar to what it is now. As others also argue (Mayr, 1998; Gupta, 1998b), we must now conclude that the idea of the early divergence of the three 'domains' of life and the view that the distinction between archaeobacteria and eubacteria is more important than that between bacteria and eukaryotes were serious conceptual mistakes, fostered by molecular myopia, ignorance of palaeontology and unjustified faith in a mythical molecular clock (Ayala, 1999) unaffected by quantum evolution (Cavalier-Smith, 2002a).

NOTE ADDED IN PROOF

Two notes added in proof are available as supplementary material in IJSEM Online (<http://ijs.sgmjournals.org/>).

ACKNOWLEDGEMENTS

I thank NERC for a Professorial Fellowship and research grant, Ema Chao for technical assistance, P. J. Keeling for performing the Kishino-Hasegawa tests on PAUP* (by courtesy of D. W. Swofford) and M. Müller for information prior to publication and general encouragement. I thank A. J. Roger for stimulating discussions and many valuable and perceptive comments and suggestions and the Evolutionary Biology Programme of the Canadian Institute for Advanced Research for fellowship support.

REFERENCES

- Andersson, J. O. & Roger, A. J. (2002). A cyanobacterial gene in non-photosynthetic protists – an early chloroplast acquisition in eukaryotes? *Curr Biol* **12**, 115–119.
- Aravind, L. & Koonin, E. V. (2001). Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res* **11**, 1365–1374.
- Archibald, J. M., Logsdon, J. M. & Doolittle, W. F. (1999). Recurrent paralogy in the evolution of archaeal chaperonins. *Curr Biol* **9**, 1053–1056.
- Archibald, J. M., Logsdon, J. M., Jr & Doolittle, W. F. (2000). Origin and evolution of eukaryotic chaperonins: phylogenetic evidence for ancient duplications in CCT genes. *Mol Biol Evol* **17**, 1456–1466.

- Archibald, J. M., Cavalier-Smith, T., Maier, U. & Douglas, S. (2001).** Molecular chaperones encoded by a reduced nucleus: the cryptomonad nucleomorph. *J Mol Evol* **52**, 490–501.
- Av-Gay, Y. & Everett, M. (2000).** The eukaryotic-like Ser/Thr protein kinases of *Mycobacterium tuberculosis*. *Trends Microbiol* **8**, 238–244.
- Ayala, F. J. (1999).** Molecular clock mirages. *Bioessays* **21**, 71–75.
- Baker, A. & Schatz, G. (1987).** Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. *Proc Natl Acad Sci U S A* **84**, 3117–3121.
- Baldauf, S. L. (1999).** A search for the origins of animals and fungi: comparing and combining molecular data. *Am Nat* **154**, S178–S188.
- Baldauf, S. L. & Doolittle, W. F. (1997).** Origin and evolution of the slime molds (Mycetozoa). *Proc Natl Acad Sci U S A* **94**, 12007–12012.
- Baldauf, S. L., Palmer, J. D. & Doolittle, W. F. (1996).** The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. *Proc Natl Acad Sci U S A* **93**, 7749–7754.
- Baldauf, S. L., Roger, A. J., Wenk-Siefert, I. & Doolittle, W. F. (2000).** A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* **290**, 972–977.
- Barr, D. J. S. (2000).** 6. Chytridiomycota. In *The Mycota*, vol. VII, part A, pp. 93–112. Edited by D. J. McLaughlin, E. G. McLaughlin & P. A. Lemke. Berlin: Springer.
- Beech, P. L. & Gilson, P. R. (2000).** FtsZ and organelle division in protists. *Protist* **151**, 11–16.
- Berthe, F. C. J., Le Roux, F., Peyretailade, E., Peyret, P., Rodriguez, D., Gouy, M. & Vivares, C. P. (2000).** Phylogenetic analysis of the small subunit ribosomal RNA of *Marteilia refringens* validates the existence of phylum Paramyxea (Desportes and Perkins, 1990). *J Eukaryot Microbiol* **47**, 288–293.
- Blanton, R. L., Fuller, D., Iranfar, N., Grimson, M. J. & Loomis, W. F. (2000).** The cellulose synthase gene of *Dictyostelium*. *Proc Natl Acad Sci U S A* **97**, 2391–2396.
- Blobel, G. (1980).** Intracellular protein topogenesis. *Proc Natl Acad Sci U S A* **77**, 1496–1500.
- Bouzat, J. L., McNeil, L. K., Robertson, H. M. & 8 other authors. (2000).** Phylogenomic analysis of the alpha proteasome gene family from early-diverging eukaryotes. *J Mol Evol* **51**, 532–543.
- Brasier, M. D. (2000).** The Cambrian explosion and the slow burning fuse. *Sci Prog* **83**, 77–92.
- Brocks, J. J., Logan, G. A., Buick, R. & Summons, R. E. (1999).** Archean molecular fossils and the early rise of eukaryotes. *Science* **285**, 1033–1036.
- Brown, J. R. & Doolittle, W. F. (1997).** Archaea and the prokaryote-to-eukaryote transition. *Microbiol Mol Biol Rev* **61**, 456–502.
- Brugerolle, G. & Mignot, J. (1984).** Les caractéristiques ultra-structurales de l'hélioflagellé *Dimorpha mutans* Gruber (Sarcodina–Actinopoda) et leur intérêt phylétique. *Protistologica* **20**, 97–112.
- Buonomo, S. B., Clyne, R. K., Fuchs, J., Loidl, J., Uhlmann, F. & Nasmyth, K. (2000).** Disjunction of homologous chromosomes in meiosis I depends on proteolytic cleavage of the meiotic cohesin Rec8 by separin. *Cell* **103**, 387–398.
- Cavalier-Smith, T. (1975).** The origin of nuclei and of eukaryote cells. *Nature* **256**, 463–468.
- Cavalier-Smith, T. (1977).** Mitochondri e cloroplasti: un problema evolutivo. In *Scienza e Technica 1977*, pp. 303–318. Milan: Mondadori.
- Cavalier-Smith, T. (1978a).** The evolutionary origin and phylogeny of microtubules, mitotic spindles and eukaryote flagella. *Biosystems* **10**, 93–114.
- Cavalier-Smith, T. (1978b).** Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA C-value paradox. *J Cell Sci* **34**, 247–278.
- Cavalier-Smith, T. (1980).** Cell compartmentation and the origin of eukaryote membranous organelles. In *Endocytobiology: Endosymbiosis and Cell Biology, a Synthesis of Recent Research*, pp. 893–916. Edited by W. Schwemmler & H. E. A. Schenk. Berlin: de Gruyter.
- Cavalier-Smith, T. (1981a).** The origin and early evolution of the eukaryotic cell. In *Molecular and Cellular Aspects of Microbial Evolution* (Society for General Microbiology Symposium no. 32), pp. 33–84. Edited by M. J. Carlile, J. F. Collins & B. E. B. Moseley. Cambridge: Cambridge University Press.
- Cavalier-Smith, T. (1981b).** Eukaryote kingdoms: seven or nine? *Biosystems* **14**, 461–481.
- Cavalier-Smith, T. (1982a).** Skeletal DNA and the evolution of genome size. *Annu Rev Biophys Bioeng* **11**, 273–302.
- Cavalier-Smith, T. (1982b).** The evolutionary origin and phylogeny of eukaryote flagella. In *Prokaryotic and Eukaryotic Flagella* (35th Symposium of the Society of Experimental Biology), pp. 465–493. Edited by W. B. Amos & J. G. Duckett. Cambridge: Cambridge University Press.
- Cavalier-Smith, T. (1982c).** The evolution of nuclear matrix and envelope. In *The Nuclear Envelope and Matrix*, pp. 307–318. Edited by G. G. Maul. New York: Alan R. Liss.
- Cavalier-Smith, T. (1982d).** The origins of plastids. *Biol J Linn Soc* **17**, 289–306.
- Cavalier-Smith, T. (1983a).** Endosymbiotic origin of the mitochondrial envelope. In *Endocytobiology II*, pp. 265–279. Edited by W. Schwemmler & H. E. A. Schenk. Berlin: de Gruyter.
- Cavalier-Smith, T. (1983b).** A 6-kingdom classification and a unified phylogeny. In *Endocytobiology II*, pp. 1027–1034. Edited by W. Schwemmler & H. E. A. Schenk. Berlin: de Gruyter.
- Cavalier-Smith, T. (1985).** Cell volume and the evolution of genome size. In *The Evolution of Genome Size*, pp. 211–251. Edited by T. Cavalier-Smith. Chichester: Wiley.
- Cavalier-Smith, T. (1986).** The kingdom Chromista: origin and systematics. In *Progress in Phycological Research*, vol. 4, pp. 309–347. Edited by F. E. Round & D. J. Chapman. Bristol: Biopress.
- Cavalier-Smith, T. (1987a).** The origin of cells: a symbiosis between genes, catalysts, and membranes. *Cold Spring Harb Symp Quant Biol* **52**, 805–824.
- Cavalier-Smith, T. (1987b).** Bacterial DNA segregation: its motors and positional control. *J Theor Biol* **127**, 361–372.
- Cavalier-Smith, T. (1987c).** The origin of eukaryote and archaeobacterial cells. *Ann N Y Acad Sci* **503**, 17–54.
- Cavalier-Smith, T. (1987d).** Eukaryotes with no mitochondria. *Nature* **326**, 332–333.
- Cavalier-Smith, T. (1987e).** The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. *Ann NY Acad Sci* **503**, 55–71.
- Cavalier-Smith, T. (1987f).** Glaucophyceae and the origin of plants. *Evol Trends Plants* **2**, 75–78.
- Cavalier-Smith, T. (1987g).** The origin of Fungi and pseudofungi. In *Evolutionary Biology of the Fungi* (Symposium of the British Mycological Society no. 13), pp. 339–353. Edited by A. D. M. Rayner, C. M. Brasier & D. Moore. Cambridge: Cambridge University Press.
- Cavalier-Smith, T. (1988a).** Origin of the cell nucleus. *Bioessays* **9**, 72–78.
- Cavalier-Smith, T. (1988b).** Eukaryote cell evolution. In *Proceedings of the 14th International Botanical Congress*, pp. 203–223. Edited by W. Greuter & B. Zimmer. Königstein: Koeltz.
- Cavalier-Smith, T. (1990a).** Symbiotic origin of peroxisomes. In *Endocytobiology IV*, pp. 515–521. Edited by P. Nardon, V. Gianinazzi-Pearson, A. M. Grenier, L. Margulis & D. C. Smith. Paris: Institut National de la Recherche Agronomique.
- Cavalier-Smith, T. (1990b).** Microorganism megaevolution: integrating the living and fossil evidence. *Rev Micropaleontol* **33**, 145–154.
- Cavalier-Smith, T. (1991a).** The evolution of cells. In *Evolution of Life*, pp. 271–304. Edited by S. Osawa & T. Honjo. Tokyo: Springer-Verlag.
- Cavalier-Smith, T. (1991b).** The evolution of prokaryotic and eukaryotic cells. In *Fundamentals of Medical Cell Biology*, vol. I, pp. 217–272. Edited by G. E. Bittar. Greenwich, CT: J.A.I. Press.

- Cavalier-Smith, T. (1991c).** Archamoebae: the ancestral eukaryotes? *Biosystems* **25**, 25–38.
- Cavalier-Smith, T. (1991d).** Intron phylogeny: a new hypothesis. *Trends Genet* **7**, 145–148.
- Cavalier-Smith, T. (1991e).** Co-evolution of vertebrate genome, cell and nuclear size. In *Symposium on the Evolution of Terrestrial Vertebrates* (Selected Symposia in Monographs U. Z. I.), pp. 51–86. Edited by G. Ghiara and others. Modena: Mucchi.
- Cavalier-Smith, T. (1991f).** Cell diversification in heterotrophic flagellates. In *The Biology of Free-living Heterotrophic Flagellates*, pp. 113–131. Edited by D. J. Patterson & J. Larsen. Oxford: Oxford University Press.
- Cavalier-Smith, T. (1992a).** Bacteria and eukaryotes. *Nature* **356**, 570.
- Cavalier-Smith, T. (1992b).** Origin of the cytoskeleton. In *The Origin and Evolution of the Cell*, pp. 79–106. Edited by H. Hartman & K. Matsuno. Singapore: World Scientific Publishers.
- Cavalier-Smith, T. (1992c).** Origins of secondary metabolism. In *Secondary Metabolites: Their Function and Evolution* (CIBA Foundation Symposium no. 171), pp. 64–87. Edited by D. J. Chadwick & J. Whelan. Chichester: Wiley.
- Cavalier-Smith, T. (1993a).** Evolution of the eukaryotic genome. In *The Eukaryotic Genome*, pp. 333–385. Edited by P. Broda, S. G. Oliver & P. Sims. Cambridge: Cambridge University Press.
- Cavalier-Smith, T. (1993b).** Kingdom Protozoa and its 18 phyla. *Microbiol Rev* **57**, 953–994.
- Cavalier-Smith, T. (1993c).** The origin, losses and gains of chloroplasts. In *Origin of Plastids: Symbiogenesis, Prochlorophytes and the Origins of Chloroplasts*, pp. 291–348. Edited by R. A. Lewin. New York: Chapman & Hall.
- Cavalier-Smith, T. (1993d).** Percolozoa and the symbiotic origin of the metakaryotic cell. In *Endocytobiology V*, pp. 399–406. Edited by H. Ishikawa, M. Ishida & S. Sato. Tübingen: Tübingen University Press.
- Cavalier-Smith, T. (1994).** Origin and relationships of Haptophyta. In *The Haptophyte Algae*, pp. 413–435. Edited by J. C. Green & B. S. C. Leadbeater. Oxford: Clarendon Press.
- Cavalier-Smith, T. (1995a).** Membrane heredity, symbiogenesis, and the multiple origins of algae. In *Biodiversity and Evolution*, pp. 75–114. Edited by R. Arai, M. Kato & Y. Doi. Tokyo: National Science Museum Foundation.
- Cavalier-Smith, T. (1995b).** Cell cycles, diplokaryosis and the archezoan origin of sex. *Arch Protistenkd* **145**, 198–207.
- Cavalier-Smith, T. (1997).** Amoeboflagellates and mitochondrial cristae in eukaryote evolution: megasystematics of the new protozoan subkingdoms Eozoa and Neozoa. *Arch Protistenkd* **147**, 237–258.
- Cavalier-Smith, T. (1998a).** A revised six-kingdom system of life. *Biol Rev Camb Philos Soc* **73**, 203–266.
- Cavalier-Smith, T. (1998b).** Neomonada and the origin of animals and fungi. In *Evolutionary Relationships Among Protozoa*, pp. 375–407. Edited by G. H. Coombs, K. Vickerman, M. A. Sleight & A. Warren. London: Kluwer.
- Cavalier-Smith, T. (1999).** Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J Eukaryot Microbiol* **46**, 347–366.
- Cavalier-Smith, T. (2000a).** Flagellate megaevolution: the basis for eukaryote diversification. In *The Flagellates*, pp. 361–390. Edited by J. R. Green & B. S. C. Leadbeater. London: Taylor & Francis.
- Cavalier-Smith, T. (2000b).** Membrane heredity and early chloroplast evolution. *Trends Plant Sci* **4**, 174–182.
- Cavalier-Smith, T. (2000c).** What are Fungi? In *The Mycota*, vol. VII, pp. 1–37. Edited by D. J. McLaughlin, E. C. McLaughlin & P. A. Lemke. Berlin: Springer.
- Cavalier-Smith, T. (2001).** Obcells as proto-organisms: membrane heredity, lithophosphorylation, and the origins of the genetic code, the first cells, and photosynthesis. *J Mol Evol* **53**, 555–595.
- Cavalier-Smith, T. (2002a).** The neomuran origin of archaeobacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int J Syst Evol Microbiol* **52**, 7–76; erratum in IJSEM Online (<http://ijs.sgmjournals.org/>).
- Cavalier-Smith, T. (2002b).** Origins of the machinery of recombination and sex. *Heredity* **89** (in press).
- Cavalier-Smith, T. & Beaton, M. J. (1999).** The skeletal function of non-genic nuclear DNA: new evidence from ancient cell chimaeras. In *Structural Biology and Functional Genomics*, pp. 1–18. Edited by E. M. Bradbury & S. Pongor. Dordrecht: Kluwer.
- Cavalier-Smith, T. & Chao, E. E. (1995).** The opalozoan *Apusomonas* is related to the common ancestor of animals, fungi and choanoflagellates. *Proc R Soc Lond B Biol Sci* **261**, 1–6.
- Cavalier-Smith, T. & Chao, E. E. (1996).** Molecular phylogeny of the free-living archezoan *Trepomonas agilis* and the nature of the first eukaryote. *J Mol Evol* **43**, 551–562.
- Cavalier-Smith, T. & Chao, E. E. (1997).** Sarcomonad ribosomal RNA sequences, rhizopod phylogeny, and the origin of euglyphid amoebae. *Arch Protistenkd* **147**, 227–236.
- Cavalier-Smith, T. & Chao, E. E. (1999).** *Hyperamoeba* rRNA phylogeny and the classification of the phylum Amoebozoa. *J Eukaryot Microbiol* **46**, 5A.
- Cavalier-Smith, T. & Lee, J. J. (1985).** Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. *J Protozool* **32**, 376–379.
- Cavalier-Smith, T., Chao, E. E. & Allsopp, M. T. E. P. (1995).** Ribosomal RNA evidence for chloroplast loss within Heterokonta: pedinellid relationships and a revised classification of ochristan algae. *Arch Protistenkd* **145**, 209–220.
- Cavalier-Smith, T., Allsopp, M. T. E. P., Chao, E. E., Boury-Esnault, N. & Vacelet, J. (1996a).** Sponge phylogeny, animal monophyly and the origin of the nervous system: 18S rRNA evidence. *Can J Zool* **74**, 2031–2045.
- Cavalier-Smith, T., Chao, E. E., Thompson, C. & Hourihane, S. (1996b).** *Oikomonas*, a distinctive zooflagellate related to chrysoomonads. *Arch Protistenkd* **146**, 273–279.
- Cavalier-Smith, T., Allsopp, M. T. E. P., Häuber, M. M., Rensing, S. A., Gothe, G., Chao, E. E., Couch, J. A. & Maier, U.-G. (1996c).** Chromobionte phylogeny: the enigmatic alga *Reticulosphaera japonensis* Grell is an aberrant haptophyte, not a heterokont. *Eur J Phycol* **31**, 315–328.
- Charette, M. & Gray, M. W. (2000).** Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**, 341–351.
- Cleveland, L. R. (1947).** The origin and evolution of meiosis. *Science* **105**, 287–289.
- Cleveland, L. R. (1956).** Brief accounts of the sexual cycles of the flagellates of *Cryptocercus*. *J Protozool* **3**, 161–180.
- Dacks, J. B., Silberman, J. D., Simpson, A. G. B., Moriya, S., Kudo, T., Ohkuma, M. & Redfield, R. J. (2001).** Oxymonads are closely related to the excavate taxon *Trimastix*. *Mol Biol Evol* **18**, 1034–1044.
- Darwin, C. (1859).** *The Origin of Species*. London: Murray.
- De Beer, G. (1954).** Archaeopteryx and evolution. *Adv Sci* **42**, 160–170.
- De Duve, C. (1969).** Evolution of the peroxisome. *Ann N Y Acad Sci* **168**, 369–381.
- De Duve, C. (1982).** Peroxisomes and related particles in historical perspective. *Ann N Y Acad Sci* **386**, 1–4.
- De Duve, C. & Wattiaux, R. (1964).** Functions of lysosomes. *Annu Rev Physiol* **28**, 435–492.
- Delwiche, C. F. (1999).** Tracing the thread of plastid diversity through the tapestry of life. *Am Nat* **154**, S164–S177.
- von Dohlen, C. D., Kohler, S., Alsop, S. T. & McManus, W. R. (2001).** Mealybug β -proteobacterial endosymbionts contain γ -proteobacterial symbionts. *Nature* **412**, 433–436.
- Doolittle, W. F. (1998a).** You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet* **14**, 307–311.

- Doolittle, W. F. (1998b). A paradigm gets shifty. *Nature* **392**, 15–16.
- Douglas, S., Zauner, S., Fraunholz, M. & 7 other authors. (2001). The highly reduced genome of an enslaved algal nucleus. *Nature* **410**, 1091–1096.
- Drozdowicz, Y. M. & Rea, P. A. (2001). Vacuolar H⁺ pyrophosphatases: from the evolutionary backwaters into the mainstream. *Trends Plant Sci* **6**, 206–211.
- Dujardin, F. (1841). In *Histoire Naturelle des Zoophytes*, pp. 240–259. Paris: Librairie Encyclopédique de Roret.
- Edgcomb, V. P., Roger, A. J., Simpson, A. G., Kysela, D. T. & Sogin, M. L. (2001). Evolutionary relationships among 'jakobid' flagellates as indicated by alpha- and beta-tubulin phylogenies. *Mol Biol Evol* **18**, 514–522.
- Edlind, T. D., Li, J., Visvesvara, G. S., Vodkin, M. H., McLaughlin, G. L. & Katiyar, S. K. (1996). Phylogenetic analysis of beta-tubulin sequences from amitochondrial protozoa. *Mol Phylogenet Evol* **5**, 359–367.
- Eichler, J. & Moll, R. (2001). The signal recognition particle of Archaea. *Trends Microbiol* **9**, 130–136.
- Embley, T. M. & Hirt, R. P. (1998). Early branching eukaryotes? *Curr Opin Genet Dev* **8**, 624–629.
- van den Ent, F., Amos, L. A. & Lowe, J. (2001). Prokaryotic origin of the actin cytoskeleton. *Nature* **413**, 39–44.
- Errington, J., Bath, J. & Wu, L. J. (2001). DNA transport in bacteria. *Nat Rev Mol Cell Biol* **2**, 538–545.
- Fast, N. M., Kissinger, J. C., Roos, D. S. & 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.
- Fenchel, T. & Finlay, B. J. (1995). *Ecology and Evolution in Anoxic Worlds*. Oxford: Oxford University Press.
- Feng, D. F., Cho, G. & Doolittle, R. F. (1997). Determining divergence times with a protein clock: update and reevaluation. *Proc Natl Acad Sci USA* **94**, 13028–13033.
- Ferguson, A. (1767). *An Essay on the History of Civil Society*. Reprinted, 1996. Edinburgh: Edinburgh University Press.
- Finlay, B. J., Maberly, S. C. & Esteban, G. (1996). Spectacular abundance of ciliates in anoxic pond water: contribution of symbiont photosynthesis to host respiratory oxygen requirements. *FEMS Microbiol Ecol* **71**, 221–237.
- Fitch, W. M. & Upper, K. (1987). The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. *Cold Spring Harb Symp Quant Biol* **52**, 759–767.
- Galtier, N., Tourasse, N. & Gouy, M. (1999). A nonhyperthermophilic common ancestor to extant life forms. *Science* **283**, 220–221.
- Gogarten, J. P. & Kibak, H. (1992). The bioenergetics of the last common ancestor and the origin of the eukaryotic endomembrane system. In *The Origin and Evolution of the Cell*, pp. 131–162. Edited by H. Hartman & K. Masuno. Singapore: World Scientific Publishers.
- Golding, G. B. & Gupta, R. S. (1995). Protein-based phylogenies support a chimeric origin for the eukaryotic genome. *Mol Biol Evol* **12**, 1–6.
- Gray, M. W., Burger, G. & Lang, B. F. (1999). Mitochondrial evolution. *Science* **283**, 1476–1481.
- Grell, K. G. & Schüller, S. (1991). The ultrastructure of the plasmodial protist *Leucodictyon marimum* Grell. *Eur J Protistol* **27**, 168–177.
- Gupta, R. S. (1998a). Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeobacteria, eubacteria, and eukaryotes. *Microbiol Mol Biol Rev* **62**, 1435–1491.
- Gupta, R. S. (1998b). Life's third domain (Archaea): an established fact or an endangered paradigm? *Theor Popul Biol* **54**, 91–104.
- Gupta, R. S. & Golding, G. B. (1996). The origin of the eukaryotic cell. *Trends Biochem Sci* **21**, 166–171.
- Hall, J. B. (1973). The nature of the host in the origin of the eukaryotic cell. *J Theor Biol* **31**, 501–509.
- Hartzell, P. L. (1997). Complementation of sporulation and motility defects in a prokaryote by a eukaryotic GTPase. *Proc Natl Acad Sci USA* **94**, 9881–9886.
- Hashimoto, T., Sánchez, L. B., Shirakura, T., Müller, M. & Hasegawa, M. (1998). Secondary absence of mitochondria in *Giardia lamblia* and *Trichomonas vaginalis* revealed by valyl-tRNA synthetase phylogeny. *Proc Natl Acad Sci USA* **95**, 6860–6865.
- Häuber, M. M., Müller, S. B., Speth, V. & Maier, U.-G. (1994). How to evolve a complex plastid? – a hypothesis. *Bot Acta* **107**, 383–386.
- Hausmann, K. & Hülsmann, N. (1996). *Protozoology*, 2nd edn. Stuttgart: Thieme.
- Hell, K., Neupert, W. & Stuart, R. A. (2001). Oxa1p acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA. *EMBO J* **20**, 1281–1288.
- Henze, K., Badr, A., Wettern, M., Cerff, R. & Martin, W. (1995). A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. *Proc Natl Acad Sci USA* **92**, 9122–9126.
- Hibberd, D. (1976). The fine structure of the colonial flagellates *Rhipidodendron splendidum* Stein and *Spongomonas uvella* Stein with special reference to the flagellar apparatus. *J Protozool* **23**, 374–385.
- Hibberd, D. J. (1990). Phylum Chlorarachnida. In *Handbook of Protoctista*, pp. 288–292. Edited by L. Margulis, J. O. Corliss, M. Melkonian & D. J. Chapman. Boston: Jones & Bartlett.
- Hinkle, G., Leipe, D. D., Nerad, T. A. & Sogin, M. L. (1994). The unusually long small subunit ribosomal RNA of *Phreatamoeba balamuthi*. *Nucleic Acids Res* **22**, 465–469.
- Hirt, R. P., Logsdon, J. M., Jr, Healy, B., Dorey, M. W., Doolittle, W. F. & Embley, T. M. (1999). Microsporidia are related to Fungi: evidence from the largest subunit of RNA polymerase II and other proteins. *Proc Natl Acad Sci USA* **96**, 580–585.
- Hoffman, P. F., Kaufman, A. J., Halverson, G. P. & Schrag, D. P. (1998). A neoproterozoic snowball Earth. *Science* **281**, 1342–1346.
- Hollande, A. (1974). Données ultrastructurales sur les isospores des radiolaires. *Protistologica* **X**, 567–572.
- Horner, D. S., Hirt, R. P. & Embley, T. M. (1999). A single eubacterial origin of eukaryotic pyruvate:ferredoxin oxidoreductase genes: implications for the evolution of anaerobic eukaryotes. *Mol Biol Evol* **16**, 1280–1291.
- Hyde, W. T., Crowley, T. J., Baum, S. K. & Peltier, W. R. (2000). Neoproterozoic 'snowball Earth' simulations with a coupled climate/ice-sheet model. *Nature* **405**, 425–429.
- Ishida, K., Green, B. R. & Cavalier-Smith, T. (1999). Diversification of a chimaeric algal group, the chlorarachniophytes: phylogeny of nuclear and nucleomorph small-subunit rRNA genes. *Mol Biol Evol* **16**, 321–331.
- Iwabuchi, M., Ohsumi, K., Yamamoto, T. M., Sawada, W. & Kishimoto, T. (2000). Residual Cdc2 activity remaining at meiosis I exit is essential for meiotic M–M transition in *Xenopus* oocyte extracts. *EMBO J* **19**, 4513–4523.
- Jefferies, R. S. (1979). The origin of chordates: a methodological essay. In *The Origin of Major Invertebrate Groups*, pp. 443–477. Edited by M. R. House. London: Academic Press.
- John, P. & Whatley, F. R. (1975). *Paracoccus denitrificans* and the evolutionary origin of the mitochondrion. *Nature* **254**, 495–498.
- Karpov, S. (1990). Analysis of the orders Phalansteriida, Spongomonadida and Thaumatomonadida. *Zool Zh* **69**, 5–12 (in Russian).
- Karpov, S. A. (1997). Cercomonads and their relationship to the myxomycetes. *Arch Protistenkd* **148**, 297–307.
- Karpov, S. A. (2000). Flagellate phylogeny: an ultrastructural approach. In *The Flagellates*, pp. 336–360. Edited by J. R. Green & B. S. C. Leadbeater. London: Taylor & Francis.
- Kasinsky, H. E., Lewis, J. D., Dacks, J. B. & Ausio, J. (2001). Origin of H1 linker histones. *FASEB J* **15**, 34–42.
- Keeling, P. J. (1998). A kingdom's progress: Archezoa and the origin of eukaryotes. *Bioessays* **20**, 87–95.

- Keeling, P. J. (2001).** Foraminifera and Cercozoa are related in actin phylogeny: two orphans find a home? *Mol Biol Evol* **18**, 1551–1557.
- Keeling, P. J. & Doolittle, W. F. (1995).** Archaea: narrowing the gap between prokaryotes and eukaryotes. *Proc Natl Acad Sci U S A* **92**, 5761–5764.
- Keeling, P. J. & Doolittle, W. F. (1996).** Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. *Mol Biol Evol* **13**, 1297–1305.
- Keeling, P. J. & Doolittle, W. F. (1997).** Evidence that eukaryotic triosephosphate isomerase is of alpha-proteobacterial origin. *Proc Natl Acad Sci U S A* **94**, 1270–1275.
- Keeling, P. J. & McFadden, G. I. (1998).** Origins of microsporidia. *Trends Microbiol* **6**, 19–23.
- Keeling, P. J., Fast, N. M. & McFadden, G. I. (1998).** Evolutionary relationship between translation initiation factor eIF-2 γ and selenocysteine-specific elongation factor SELB: change of function in translation factors. *J Mol Evol* **47**, 649–655.
- Keeling, P. J., Luker, M. A. & Palmer, J. D. (2000).** Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. *Mol Biol Evol* **17**, 23–31.
- Klenk, H. P., Meier, T. D., Durovic, P., Schwass, V., Lottspeich, F., Dennis, P. P. & Zillig, W. (1999).** RNA polymerase of *Aquifex pyrophilus*: implications for the evolution of the bacterial rpoBC operon and extremely thermophilic bacteria. *J Mol Evol* **48**, 528–541.
- Knoll, A. H. (1992).** The early evolution of eukaryotes: a geological perspective. *Science* **256**, 622–627.
- Kohl, W., Gloe, A. & Reichenbach, H. (1983).** Steroids from the myxobacterium *Nannocystis exedens*. *J Gen Microbiol* **129**, 1629–1635.
- Kondrashov, A. S. (1994).** The asexual ploidy cycle and the origin of sex. *Nature* **370**, 213–216.
- Koonin, E. V., Mushegian, A. R. & Bork, P. (1996).** Non-orthologous gene displacement. *Trends Genet* **12**, 334–336.
- Kozo-Polyansky, B. M. (1924).** *A New Principle of Biology. Essay on the Theory of Symbiogenesis*. Moscow (in Russian).
- Kyrpides, N. C. & Olsen, G. J. (1999).** Archaeal and bacterial hyperthermophiles: horizontal gene exchange or common ancestry? *Trends Genet* **15**, 298–299.
- Lake, J. A. & Rivera, M. C. (1994).** Was the nucleus the first endosymbiont? *Proc Natl Acad Sci U S A* **91**, 2880–2881.
- Lamb, D. C., Kelly, D. E., Manning, N. J. & Kelly, S. L. (1998).** A sterol biosynthetic pathway in *Mycobacterium*. *FEBS Lett* **437**, 142–144.
- Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M. & Gray, M. W. (1997).** An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* **387**, 493–497.
- Lange, B. H. M., Bachi, A., Wilm, M. & Gonzalez, C. (2000).** Hsp90 is a core centrosomal component and is required at different stages of the centrosome cycle in *Drosophila* and vertebrates. *EMBO J* **19**, 1252–1262.
- Lankester, E. R. (1878).** In *Gegenbaur's Elements of Anatomy*, Preface to the English translation of the 2nd edn. London: Macmillan.
- Lenormand, T. & Otto, S. P. (2000).** The evolution of recombination in a heterogeneous environment. *Genetics* **156**, 423–438.
- Linton, E. W., Hittner, D., Lewandowski, C., Auld, T. & Triemer, R. E. (1999).** A molecular study of euglenoid phylogeny using small subunit rDNA. *J Eukaryot Microbiol* **46**, 217–223.
- Lipscomb, D. L. & Corliss, J. O. (1982).** *Stephanopogon*, a phylogenetically important 'ciliate', shown by ultrastructural studies to be a flagellate. *Science* **215**, 303–304.
- Logsdon, J. M., Jr. (1998).** The recent origins of spliceosomal introns revisited. *Curr Opin Genet Dev* **8**, 637–648.
- Margulis, L. (1970).** *Origin of Eukaryotic Cells*. New Haven, CT: Yale University Press.
- Margulis, L. (1981).** *Symbiosis in Cell Evolution*. San Francisco: W. F. Freeman.
- Margulis, L., Dolan, M. F. & Guerrero, R. (2000).** The chimeric eukaryote: origin of the nucleus from the karyomastigont in amitochondriate protists. *Proc Natl Acad Sci U S A* **97**, 6954–6959.
- Martin, W. (1996).** Is something wrong with the tree of life? *Bioessays* **18**, 523–527.
- Martin, W. (1998).** Endosymbiosis and the origins of chloroplast-cytosolic isoenzymes: revising the product-specificity corollary. In *Horizontal Gene Transfer*, pp. 363–379. Edited by M. Syvanen & C. Kado. London: Chapman & Hall.
- Martin, W. (1999).** A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not. *Proc R Soc Lond B Biol Sci* **266**, 1387–1395.
- Martin, W. & Müller, M. (1998).** The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41.
- Martin, W. & Schnarrenberger, C. (1997).** The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. *Curr Genet* **32**, 1–18.
- Martin, J. B., Laussmann, T., Bakker-Grunwald, T., Vogel, G. & Klein, G. (2000).** Neo-Inositol polyphosphates in the amoeba *Entamoeba histolytica*. *J Biol Chem* **275**, 10134–10140.
- Mayr, E. (1998).** Two empires or three? *Proc Natl Acad Sci U S A* **95**, 9720–9723.
- Mereschkovsky, C. (1905).** Über Natur und Ursprung der Chromatophoren in Pflanzenreiche. *Biol Zentbl* **25**, 593–604.
- Mereschkovsky, C. (1910).** Theorie der Zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol Zentbl* **30**, 278–303, 321–347, 353–367.
- Mikrjukov, K. A. & Mylnikov, A. P. (1998).** The fine structure of a carnivorous multiflagellar protist, *Multicilia marina* Cienkowski 1881 (Flagellata incertae sedis). *Eur J Protistol* **34**, 391–401.
- Moestrup, Ø. (2000).** The flagellate cytoskeleton. In *The Flagellates: Unity, Diversity and Evolution* (Systematics Association Special Volume Series no. 59), pp. 69–94. Edited by B. S. C. Leadbeater & J. C. Green. London: Taylor & Francis.
- Møller-Jensen, J., Jensen, R. B. & Gerdes, K. (2000).** Plasmid and chromosome segregation in prokaryotes. *Trends Microbiol* **8**, 313–320.
- Moreira, D. & López-García, P. (1998).** Symbiosis between methanogenic archaea and δ -proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J Mol Evol* **47**, 517–530.
- Moreira, D., Le Guyader, H. & Philippe, H. (2000).** The origin of red algae and the evolution of chloroplasts. *Nature* **405**, 69–72.
- Müller, M. (1997).** What are microsporidia? *Parasitol Today* **13**, 455–456.
- Müller, M. & Martin, W. (1999).** The genome of *Rickettsia prowazekii* and some thoughts on the origin of mitochondria and hydrogenosomes. *Bioessays* **21**, 377–381.
- Nesbø, C. L., L'Haridon, S., Stetter, K. O. & Doolittle, W. F. (2001).** Phylogenetic analyses of two 'archaeal' genes in *Thermotoga maritima* reveal multiple transfers between archaea and bacteria. *Mol Biol Evol* **18**, 362–375.
- Niki, H., Jaffe, A., Imamura, R., Ogura, T. & Hiraga, S. (1991).** The new gene *mukB* codes for a 177 kDa protein with coiled-coil domains involved in chromosome partitioning in *E. coli*. *EMBO J* **10**, 183–193.
- Oliveira, M. C. & Bhattacharya, D. (2000).** Phylogeny of the Bangiophycidae (Rhodophyta) and the secondary endosymbiotic origin of algal plastids. *Am J Bot* **87**, 482–492.
- Omer, A. D., Lowe, T. M., Russell, A. G., Ebhardt, H., Eddy, S. R. & Dennis, P. P. (2000).** Homologs of small nucleolar RNAs in Archaea. *Science* **288**, 517–522.
- Pace, N. R., Olsen, G. J. & Woese, C. R. (1986).** Ribosomal RNA phylogeny and the primary lines of evolutionary descent. *Cell* **45**, 325–326.
- Patterson, D. J. (1999).** The diversity of eukaryotes. *Am Nat* **154**, S96–S124.

- Pawlowski, J., Bolivar, I., Fahrni, J. F., De Vargas, C., Gouy, M. & Zaninetti, L. (1997). Extreme differences in rates of molecular evolution of foraminifera revealed by comparison of ribosomal DNA sequences and the fossil record. *Mol Biol Evol* **14**, 498–505.
- Pelletier, L., Jokitalo, E. & Warren, G. (2000). The effect of Golgi depletion on exocytic transport. *Nat Cell Biol* **2**, 840–846.
- Pfanner, N., Hartl, F.-U. & Neupert, W. (1988). Import of proteins into mitochondria: a multi-step process. *Eur J Biochem* **175**, 205–212.
- Philippe, H. & Adoutte, A. (1996). How reliable is our current view of eukaryotic phylogeny? In *Protistological Actualities*, pp. 17–33. Edited by G. Brugerolle & J.-P. Mignot. Clermont-Ferrand: Université Blaise Pascal de Clermont-Ferrand.
- Philippe, H. & Adoutte, A. (1998). The molecular phylogeny of Eukaryota: solid facts and uncertainties. In *Evolutionary Relationships Among Protozoa*, pp. 25–56. Edited by G. H. Coombs, K. Vickerman, M. A. Sleight & A. Warren. London: Kluwer.
- Philippe, H., Lopez, P., Brinkmann, H., Budin, K., Germot, A., Laurent, J., Moreira, D., Müller, M. & 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.
- Reeve, J. N., Sandman, K. & Daniels, C. J. (1997). Archaeal histones, nucleosomes, and transcription initiation. *Cell* **89**, 999–1002.
- Ribeiro, S. & Golding, G. B. (1998). The mosaic nature of the eukaryotic nucleus. *Mol Biol Evol* **15**, 779–788.
- Rivera, M. C. & Lake, J. A. (1992). Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* **257**, 74–76.
- Rivera, M. C., Jain, R., Moore, J. E. & Lake, J. A. (1998). Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci U S A* **95**, 6239–6244.
- Rizzotti, M. (2000). *Early Evolution*. Basel: Birkhäuser.
- Robertson, J. D. (1964). Unit membranes: a review with recent new studies of experimental alterations and a new subunit structure in synaptic membranes. In *Cellular Membranes in Development*, pp. 1–81. Edited by M. Locke. New York: Academic Press.
- Roger, A. J. (1999). Reconstructing early events in eukaryotic evolution. *Am Nat* **154**, S146–S163.
- Roger, A. J., Keeling, P. J. & Doolittle, W. F. (1994). Introns, the broken transposons. *Soc Gen Physiol Ser* **49**, 27–37.
- Roger, A. J., Svard, S. G., Tovar, J., Clark, C. G., Smith, M. W., Gillin, F. D. & Sogin, M. L. (1998). A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. *Proc Natl Acad Sci U S A* **95**, 229–234.
- Runnegar, B. (2000). Loophole for snowball Earth. *Nature* **405**, 403–404.
- Saldarriaga, J. F., Taylor, F. J. R., Keeling, P. J. & Cavalier-Smith, T. (2001). Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements. *J Mol Evol* **53**, 204–213.
- Sánchez, M., Valencia, A., Ferrándiz, M.-J., Sander, C. & Vicente, M. (1994). Correlation between the structure and biochemical activities of FtsA, an essential cell division protein of the actin family. *EMBO J* **13**, 4919–4925.
- Sandman, K. & Reeve, J. N. (1998). Origin of the eukaryotic nucleus. *Science* **280**, 502–503.
- Saunders, G. W., Hill, D. R. A., Sexton, J. P. & Andersen, R. A. (1997). Small-subunit ribosomal RNA sequences from selected dinoflagellates: testing classical evolutionary hypotheses with molecular systematic methods. *Plant Syst Evol Suppl* **11**, 237–259.
- Saville Kent, W. (1880). *A Manual of the Infusoria*. London: David Bogue (published in parts, 1880–1881).
- Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F. U. & Moarefi, I. (2000). Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70–Hsp90 multichaperone machine. *Cell* **101**, 199–210.
- Schnepf, E. (1964). Zur Feinstruktur von *Geosiphon pyriforme*. Ein Versuch zur Deutung cytoplasmatischer Membranen und Kompartimente. *Arch Mikrobiol* **49**, 112–131.
- Schopf, J. W. (1970). Precambrian micro-organisms and evolutionary events prior to the origin of vascular plants. *Biol Rev Camb Philos Soc* **45**, 319–352.
- Schopf, J. W. (1994). Disparate rates, differing fates: tempo and mode of evolution changed from the Precambrian to the Phanerozoic. *Proc Natl Acad Sci U S A* **91**, 6735–6742.
- Schubert, I. (1988). Eukaryotic nuclei of endosymbiotic origin? *Naturwissenschaften* **75**, 89–91.
- Searcy, D. G., Stein, D. B. & Searcy, K. B. (1981). A mycoplasma-like archaeobacterium possibly related to the nucleus and cytoplasm of eukaryotic cells. *Ann N Y Acad Sci* **361**, 312–324.
- Seemann, J., Jokitalo, E., Pypaert, M. & Warren, G. (2000). Matrix proteins can generate the higher order architecture of the Golgi apparatus. *Nature* **407**, 1022–1026.
- Seravin, L. N. & Goodkov, A. V. (1999). *Agamic Fusion of Protists and the Origin of the Sexual Process*. St Petersburg: Omsk (in Russian).
- Sharpe, M. E. & Errington, J. (1999). Upheaval in the bacterial nucleoid. An active chromosome segregation mechanism. *Trends Genet* **15**, 70–74.
- Simpson, G. G. (1944). *Tempo and Mode in Evolution*. New York: Columbia University Press.
- Simpson, G. G. (1953). *The Major Features of Evolution*. New York: Columbia University Press.
- Simpson, A. G. B. (1997). The identity and composition of the Euglenozoa. *Arch Protistenkd* **148**, 318–328.
- Simpson, A. G. B. & Patterson, D. J. (1999). The ultrastructure of *Carpodemonas membranifera* (Eukaryota) with reference to the ‘excavate hypothesis’. *Eur J Protistol* **35**, 353–370.
- Simpson, A., Bernard, C., Fenchel, T. & Patterson, D. J. (1997). The organisation of *Mastigamoeba schizophrentia* n. sp.: more evidence of ultrastructural idiosyncrasy and simplicity in pelobiont protists. *Eur J Protistol* **33**, 87–98.
- Sleight, M. (2000). Trophic strategies. In *The Flagellates*, pp. 147–165. Edited by J. R. Green & B. S. C. Leadbeater. London: Taylor–Francis.
- Smallman, D. S., Schnare, M. N. & Gray, M. W. (1996). RNA:RNA interactions in the large subunit ribosomal RNA of *Euglena gracilis*. *Biochim Biophys Acta* **1305**, 1–6.
- Smith, C. M. & Steitz, J. A. (1997). Sno storm in the nucleolus: new roles for myriad small RNPs. *Cell* **89**, 669–672.
- Smith, K. N., Penkner, A., Ohta, K., Klein, F. & Nicolas, A. (2001). B-type cyclins CLB5 and CLB6 control the initiation of recombination and synaptonemal complex formation in yeast meiosis. *Curr Biol* **11**, 88–97.
- Sogin, M. L. (1991). Early evolution and the origin of eukaryotes. *Curr Opin Gen Dev* **1**, 457–463.
- Sonneborn, T. M. (1963). Does preformed cell structure play an essential role in cell heredity? In *The Nature of Biological Diversity*, pp. 165–221. Edited by J. M. Allen. New York: McGraw–Hill.
- South, S. T. & Gould, S. J. (1999). Peroxisome synthesis in the absence of preexisting peroxisomes. *J Cell Biol* **144**, 255–266.
- Sprague, V. (1979). Classification of the haplosporidia. *Mar Fish Res* **41**, 41–44.
- Stanier, R. Y. (1970). Some aspects of the biology of cells and their possible evolutionary significance. In *Organization and Control in Prokaryotic and Eukaryotic Cells* (Society for General Microbiology Symposium no. 20), pp. 1–38. Edited by H. P. Charles & B. C. J. G. Knight. Cambridge: Cambridge University Press.
- Stanier, R. Y. & Van Niel, C. B. (1962). The concept of a bacterium. *Arch Mikrobiol* **42**, 17–35.
- Stiller, J. W. & Hall, B. D. (1999). Long-branch attraction and the rDNA model of early eukaryotic evolution. *Mol Biol Evol* **16**, 1270–1279.
- Stiller, J. W., Duffield, E. C. S. & Hall, B. D. (1998). Amitochondriate amoebae and the evolution of DNA-dependent RNA polymerase II. *Proc Natl Acad Sci U S A* **95**, 11769–11774.

- Stuart, R. A. & Neupert, W. (2000).** Making membranes in bacteria. *Nature* **406**, 575–577.
- Taylor, F. J. R. (1974).** Implications and extensions of the serial endosymbiosis theory of the origin of eukaryotes. *Taxon* **23**, 229–258.
- Taylor, F. J. R. (1976).** Autogenous theories for the origin of eukaryotes. *Taxon* **25**, 377–390.
- Taylor, F. J. R. (1978).** Problems in the development of an explicit hypothetical phylogeny of lower eukaryotes. *Biosystems* **10**, 67–89.
- Taylor, F. J. R. (1999).** Ultrastructure as a control for protistan molecular phylogeny. *Am Nat* **154**, S125–S136.
- Teichmann, S. A. & Mitchison, G. (1999).** Is there a phylogenetic signal in prokaryote proteins? *J Mol Evol* **49**, 98–107.
- Tengs, T., Dahlberg, O. J., Shalchian-Tabrizi, K., Klaveness, D., Rudi, K., Delwiche, C. F. & Jakobsen, K. S. (2000).** Phylogenetic analyses indicate that the 19'-hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. *Mol Biol Evol* **17**, 718–729.
- Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S. & van Dijk, J. M. (2000).** Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* **64**, 515–547.
- Toth, A., Rabitsch, K. P., Galova, M., Schleiffer, A., Buonomo, S. B. & Nasmyth, K. (2000).** Functional genomics identifies monopolin: a kinetochore protein required for segregation of homologs during meiosis I. *Cell* **103**, 1155–1168.
- Tovar, J., Fischer, A. & Clark, C. G. (1999).** The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol Microbiol* **32**, 1013–1021.
- Turner, S., Pryer, K. M., Miao, V. P. & Palmer, J. D. (1999).** Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol* **46**, 327–338.
- Umen, J. G. & Goodenough, U. W. (2001).** Control of cell division by a retinoblastoma protein homolog in *Chlamydomonas*. *Genes Dev* **15**, 1652–1661.
- Van de Peer, Y., Ben Ali, A. & Meyer, A. (2000).** Microsporidia: accumulating molecular evidence that a group of amitochondriate and suspectedly primitive eukaryotes are just curious fungi. *Gene* **246**, 1–8.
- Van Valen, L. M. & Maiorana, V. C. (1980).** The archaeobacteria and eukaryotic origins. *Nature* **287**, 248–250.
- Vellai, T., Takács, K. & Vida, G. (1998).** A new aspect to the origin and evolution of eukaryotes. *J Mol Evol* **46**, 499–507.
- Vossbrinck, C. R. & Woese, C. R. (1986).** Eukaryotic ribosomes that lack a 5-8S RNA. *Nature* **320**, 287–288.
- Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. & Woese, C. R. (1987).** Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature* **326**, 411–414.
- Walter, P., Keenan, R. & Schmitz, U. (2000).** SRP – where the RNA and membrane worlds meet. *Science* **287**, 1212–1213.
- Watanabe, Y. & Gray, M. W. (2000).** Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: cbf5p in *Euglena gracilis*, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaeobacteria. *Nucleic Acids Res* **28**, 2342–2352.
- Watanabe, Y., Yokobayashi, S., Yamamoto, M. & Nurse, P. (2001).** Pre-meiotic S phase is linked to reductional chromosome segregation and recombination. *Nature* **409**, 359–363.
- Whatley, J. M., John, P. & Whatley, F. R. (1979).** From extracellular to intracellular: the establishment of mitochondria and chloroplasts. *Proc R Soc Lond B Biol Sci* **204**, 165–187.
- Woese, C. R. (1998).** The universal ancestor. *Proc Natl Acad Sci U S A* **95**, 6854–6859.
- Woese, C. R., Kandler, O. & Wheelis, M. L. (1990).** Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**, 4576–4579.
- Wu, G., Henze, K. & Müller, M. (2001).** Evolutionary relationships of the glucokinase from the amitochondriate protist, *Trichomonas vaginalis*. *Gene* **264**, 265–271.
- Zauner, S., Fraunholz, M., Wastl, J., Penny, S., Beaton, M., Cavalier-Smith, T., Maier, U.-G. & Douglas, S. (2000).** Chloroplast protein and centrosomal genes, a tRNA intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph. *Proc Natl Acad Sci U S A* **97**, 200–205.
- Zhang, Z., Green, B. R. & Cavalier-Smith, T. (2000).** Phylogeny of ultra-rapidly evolving dinoflagellate chloroplast genes: a possible common origin for sporozoan and dinoflagellate plastids. *J Mol Evol* **51**, 26–40.
- Zillig, W., Schnabel, R. & Stetter, K. O. (1985).** Archaeobacteria and the origin of the eukaryotic cytoplasm. *Curr Top Microbiol Immunol* **114**, 1–18.
- Zillig, W., Klenk, H.-P., Palm, P., Leffers, H., Pühler, G., Gropp, F. & Garrett, R. A. (1989).** Did eukaryotes originate by a fusion event? *Endocytobiosis Cell Res* **6**, 1–25.