

The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells

Tibor Vellai^{1*} and Gábor Vida²

¹*Institute for Advanced Study, Collegium Budapest, Budapest, H-1014. Szentháromság u. 2., Hungary*

²*Department of Genetics, Eötvös Loránd University, Budapest, H-1088. Múzeum krt. 4/A., Hungary*

Eukaryotes have long been thought to have arisen by evolving a nucleus, endomembrane, and cytoskeleton. In contrast, it was recently proposed that the first complex cells, which were actually proto-eukaryotes, arose simultaneously with the acquisition of mitochondria. This so-called symbiotic association hypothesis states that eukaryotes emerged when some ancient anaerobic archaeobacteria (hosts) engulfed respiring α -proteobacteria (symbionts), which evolved into the first energy-producing organelles. Therefore, the intracellular compartmentalization of the energy-converting metabolism that was bound originally to the plasma membrane appears to be the key innovation towards eukaryotic genome and cellular organization. The novel energy metabolism made it possible for the nucleotide synthetic apparatus of cells to be no longer limited by subsaturation with substrates and catalytic components. As a consequence, a considerable increase has occurred in the size and complexity of eukaryotic genomes, providing the genetic basis for most of the further evolutionary changes in cellular complexity. On the other hand, the active uptake of exogenous DNA, which is general in bacteria, was no longer essential in the genome organization of eukaryotes. The mitochondrion-driven scenario for the first eukaryotes explains the chimera-like composition of eukaryotic genomes as well as the metabolic and cellular organization of eukaryotes.

Keywords: genome organization; origin of eukaryotes; symbiotic association; intracellular compartmentalization; phagocytosis

1. AN OUTLINE OF TRADITIONAL VIEWS ON EUKARYOTIC ORIGIN

As recent reviews on the origin of eukaryotes confirm, one of the most outstanding problems in recent mega-evolutionary theory concerns the prokaryote–eukaryote transition (Brown & Doolittle 1997; Gupta 1998; Gray *et al.* 1999). Mereschowsky (1910) was among the first to suggest the symbiotic origin of certain organelles found in eukaryotic cells. Sixty years later Margulis (1970) postulated that mitochondria and chloroplasts, the energy-converting organelles of eukaryotes, descended from originally free-living bacteria. According to her, the first eukaryotic cells may have originated as communities of interacting microbial entities (Margulis 1970). She supposed that the precursors of mitochondria might have been established in nucleus-containing host cells.

In the mid-1980s, it was commonly accepted, based on the ultrastructural characteristics of eukaryotic cells (Cavalier-Smith 1987) and small subunit (SSU) rDNA-based phylogeny (Woese 1987), that the first eukaryotic cell arose directly from an archaeobacterial-like ancestor by evolving a nucleus, endomembrane and cytoskeleton. The development of phagocytosis was assumed as a key innovation leading to the complex cell type. This hypothesis argued that the earliest eukaryotes would have lacked mitochondria and other organelles as they were acquired by endosymbiosis after the divergence of amitochondriate protists.

Since the beginning of the 1990s, the so-called chimera hypotheses (Sogin 1991; Zillig *et al.* 1991; Moreira & López-García 1998) suggested that the eukaryotic cell did not directly evolve from an archaeobacterial-like ancestor. Rather, the eukaryotic nucleus has been formed from the cellular fusion between an archaeobacterium and a Gram-negative eubacterium, explaining the chimeric nature of eukaryotic genomes, as revealed by various universal protein gene phylogenies.

Summing up these different scenarios, the evolution of complex cells could have been divided into two major phases. First, the emergence of the eukaryotic cell possessing a nucleus, endomembrane, and cytoskeleton. Second, the symbiotic origin of mitochondria resulting in the appearance of the first energetically efficient aerobic protozoan. The extant descendants of the first phase could be the present-day anaerobic protozoa, the Microsporidia, Metamonada and Parabasalia. These protozoa lacking mitochondria as well as other organelles, such as Golgi dictyosomes, hydrogenosomes, peroxisomes and plastids, are thought to represent the earliest divergent eukaryotic lineages.

2. THE LACK OF EVIDENCE IS NOT THE EVIDENCE OF LACK

Contrary to the traditional views outlined above, several studies have recently provided evidence for the secondary loss of mitochondria in early-branching eukaryotes believed to be primarily amitochondriate. Genes encoding highly conserved proteins that have homologues

*Author for correspondence (vellai@falco.elte.hu).

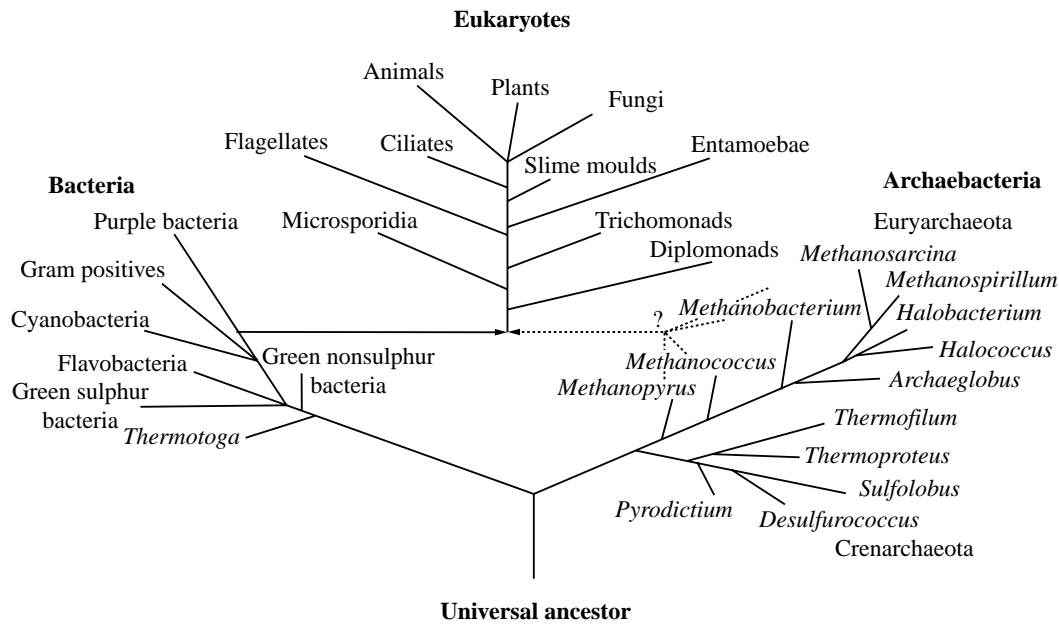


Figure 1. The newly proposed phylogenetic relationships among known life forms. The symbiotic association leading to the emergence of eukaryotes (arrows) has occurred between α -proteobacteria and certain anaerobic archaeobacteria that were probably methane-generating (dotted arrow). The origin of plastids from cyanobacteria is not shown.

in both Gram-negative α -proteobacteria and mitochondria of eukaryotes, but not in archaeobacterial or Gram-positive eubacterial species, were found in all amitochondriate protozoan taxa (Reiner *et al.* 1992; Gupta *et al.* 1994; Clark & Roger 1995; Horner *et al.* 1996; Germot *et al.* 1996, 1997; Roger *et al.* 1996, 1998; Keeling & Doolittle 1997; Hashimoto *et al.* 1998). On the other hand, it was suggested most recently that Microsporidia are probably related to fungi rather than being deep-branching eukaryotes (Embley & Hirt 1998). There is also doubt whether amitochondriate protozoa really diverged earlier than other eukaryotes.

These data imply that mitochondria may have originated at the same time or even earlier than the nucleus. As a consequence of this assumption, the former ancestors of present-day amitochondriate protozoa must once have harboured mitochondria that were either lost secondarily or converted to hydrogenosomes (or hydrogen-producing mitochondria; Embley & Martin 1998), as in trichomonads known to have diverged in a very early phase of eukaryotic evolution, due to specialization to anaerobic environments (Bui *et al.* 1996; Akhmanova *et al.* 1998).

Another possible explanation for the presence of genes of α -proteobacterial origin in early diverging amitochondriate eukaryotes may indicate an ancient, massive horizontal gene transfer from Gram-negative bacteria, presumably through nutrition (Doolittle 1998c). This so-called food hypothesis argues that some resident eukaryotic genes were replaced with genes derived from bacterial food. However, this proposal does not explain that there are only two, strikingly different inheritance patterns of eukaryotic genomes which are of proteo-, cyano- and archaeobacterial origin (Rivera *et al.* 1998). It seems unlikely that ancestral protozoa fed exclusively on α -proteobacteria, cyanobacteria (precursors of chloroplasts) and methane-generating archaeobacteria.

In 1998, two hypotheses, based on the comparative biochemistry of energy metabolism (Martin & Müller 1998) and the comparison of fundamental differences in genome organization between prokaryotes and eukaryotes (Vellai *et al.* 1998), led independently to the conclusion that the origin of the intracellular symbiont being the precursor of the first energy-producing organelle is identical to the origin of eukaryotes. The proposed scenarios assume that some ancient anaerobic archaeobacteria could engulf respiring α -proteobacteria, although not by a phagocytosis-like process (see §5). The cellular merger between two different bacteria resulted in the appearance of the first complex cells (figure 1) that, contrary to the traditional predictions, had probably no nucleus; thus, they were actually proto-eukaryotes. After a special initial advantage of the symbiotic association (that was presumably not ATP export to cytosol by virtue of oxygen respiration, as previously thought), the symbionts developed into mitochondria (aerobic or anaerobic), whereas the hosts evolved into real nucleus-bearing eukaryotic cells.

Interestingly, improving the symbiotic association theory, a complex chimeric merger was proposed (Doolittle 1998a) to account for the origin of the eukaryotic cytoskeleton. According to this proposal, the cellular merger leading to the origin of eukaryotes might have taken place simultaneously among a eubacterium, an archaeobacterium and a third, cytoskeleton-bearing prokaryote.

3. THE FUNDAMENTAL DIFFERENCES BETWEEN PROKARYOTES AND EUKARYOTES

The difference between a prokaryotic and a eukaryotic cell structure is obvious (Doolittle 1998b). Fundamentally, all eukaryotic cells have a membrane-bounded true nucleus, a complex endomembrane system and cytoskeleton, whereas prokaryotes do not. In many cases,

Table 1. Major differences in genome organization between prokaryotes and eukaryotes

characteristics of genome organization	prokaryotes	eukaryotes
genome size	600 kb–9.5 Mb	3 Mb–140 000 Mb
average gene size/ number	950 bp/4300 ^a	2500 bp/19 000 ^b
operon-like regulatory unit ^c	general	no
horizontal gene transfer	significant	negligible
rate of non-coding sequences	low ^d	high
intron	rare ^e	general
redundant gene number	rare ^f	general
ploidy level	haploidy	haploidy– polyploidy
chromosome number	one ^g	more than one
heterozygosity	no	yes

^a Average gene size of *Escherichia coli* and the number of its genes.

^b Average gene size of *Caenorhabditis elegans* and the number of its genes.

^c It results in polycistronic messengers.

^d For instance, spacer regions between rRNA operons.

^e Group II self-splicing introns in cyanobacteria and proteobacteria.

^f For example, rRNA genes.

^g Sometimes with additional replicons (plasmids) approaching the size of the genome.

however, to determine those features exclusive to eukaryotes is not so unambiguous as thought previously, because numerous characteristics formerly believed to be exclusive to eukaryotic cells have recently been discovered in prokaryotes as well. Some bacteria have been reported to possess a nucleus (Fuerst & Webb 1991), a cytoskeleton-like structure (Margolin 1998) and perhaps mitotic-like apparatus (Wheeler & Shapiro 1997), a complex internal membrane system (Prescott *et al.* 1996), a linear chromosome (Baril *et al.* 1989), introns (Ferat & Michel 1993), a eukaryotic-like gene regulation mechanism (Errington 1993), or multiple genomes (Battista 1997). Furthermore, our knowledge about the microbial world is constantly increasing, thus further eukaryotic characters are expected to be discovered in prokaryotes (the most recent example is the giant bacterial cell size; Schulz *et al.* 1999). Although it seems that prokaryotes invented most of what characterize the complex cells, there are at least two cytological features that have not yet been found in bacteria: (i) they lack any energy-producing organelle, i.e. their metabolism is not compartmentalized, and (ii) they are not capable of carrying out endocytosis, a complex and highly energy-requiring process.

Actually, the differences in genome organization between prokaryotes and eukaryotes are as fundamental (table 1) as those in their ultrastructural characters. The genome of prokaryotes can be characterized by limited size ranging from 600 kb to 9.5 Mb (Fonstein & Haselkorn 1995), whereas that of eukaryotes by a significant expansion in size (up to 140 000 Mb; Maynard Smith & Szathmáry 1995) and complexity. The general mechanism of the unexpectedly dynamic genome organ-

ization of prokaryotes, as revealed by the analysis of completely sequenced bacterial genomes (Andersson *et al.* 1998; Lawrence & Ochman 1998), involves two contrasting processes: genome loading through acquisition of exogenous DNA and genome decay by DNA loss (deletion). Presumably, the latter process, in part, is a direct consequence of the prokaryotic-like, i.e. non-compartmentalized, energy metabolism. When bacteria exist under exhausted conditions, which is predominant in natural habitats, their nucleotide synthetic apparatus is subsaturated with ATP as a substrate and catalytic component (Jensen & Pedersen 1990). Due to the metabolite limitation in DNA synthesis and selection for a faster reproduction rate, one of the major driving forces in bacterial genome organization is size reduction through deletion (Andersson & Kurland 1998; Lawrence & Ochman 1998; Vellai *et al.* 1998). Prokaryotic cells losing a part of their genetic information (which is not expressed under conditions of prolonged starvation (Moyer & Morita 1989; Seto & Miyata 1998) or intracellular parasitic lifestyle (Andersson & Kurland 1998)), are able to achieve a relatively faster reproduction rate, thus strengthening the so-called *r*-strategy (Carlile 1982). Selection for genome streamlining in bacteria has resulted in the most regressive character of their genome, the limited genetic potential.

In prokaryotes, DNA replication is limited to a single replication origin. Furthermore, the initiation of their DNA replication is controlled directly by ATP availability (Kelman & O'Donell 1994); the DnaA origin-binding protein binds ATP to form the open complex by unwinding DNA, and subsequently the helicase couples ATP hydrolysis to melt the DNA duplex into single-stranded DNAs. This regulatory mechanism makes DNA replication possible even under extremely adverse conditions. This can explain why, for example, respiring prokaryotes could not have developed eukaryotic-size genomes, although they can be considered as ATP-producing organelles like mitochondria. Notably, the availability of ATP is one factor, but not necessarily the only factor, because almost all of the reductions in genome size and complexity can also be explained by adaptation to parasitic lifestyle (Andersson & Kurland 1998). Limitation in DNA content might have been the barrier against the further increase in cellular complexity. This may explain why present-day prokaryotes are presumably no more complex than their ancestors living two billion years ago (Maynard Smith & Szathmáry 1995).

The genome loading mechanism is also essential in prokaryotes: they are known as the only organisms able to take up exogenous DNA actively and recombine it into their genome. The capability of cells for recovering the original genetic potential is probably a result of adaptation to constantly changing environments. The DNA uptake often takes place between distantly related taxa resulting in the general genetic mixing.

In contrast, an expansion of genetic material has occurred during the evolution of eukaryotes. As a possible explanation, the establishment of mitochondria may have resulted in the novel genome organization that was no longer limited by the subsaturation of ATP; an efficient ATP–ADP exchange between the mitochondria and the cytosol maintains the cell's ATP pool as highly charged

(Klingenberg 1979). Furthermore, the initiation of DNA replication in eukaryotes is regulated by a more complex cell cycle control than that in prokaryotes, which involves many signal transduction pathways. Eukaryotic cells can enter into S-phase only if they have already passed a growing, energy accumulating phase (G1). Therefore, when eukaryotic cells duplicate their genetic information, their nucleotide synthetic apparatus must be supplied by relatively rich energy sources. Contrary to bacterial cells possessing a relatively immortal reproduction potential, the majority of complex cells can be lost by programmed cell death. Moreover, most of the physiological cell death in eukaryotes takes place, in part, through a mitochondrion-regulated active process. Presumably, the initiation of cell death due to cellular stress takes place before the energy level of cells decreases under the limits of DNA synthesis.

The significant increase in the amount of genetic information through the redundant duplication of genes and entire genomes (di- and polyploidy), as well as the spreading of non-coding DNA sequences, might have been the basis for most of the further evolutionary changes in cellular complexity. Switching from haploidy to diploidy provided immediate selective advantages by protecting against the expression of recessive deleterious mutations (Kondrashov & Crow 1991). Chromosome multiplicity serves to ensure functionality when one or more copies of a gene lose their function. Thus, more reliable functioning cells and organisms could appear by the parallel buffering mechanisms of genetic subsystems. Eventually, the novel organization of genetic information led to real heterozygosity and redundant copy-number of many genes and regulatory units. Thus, the expansion of genetic information might have been the basis for the further increase in cellular complexity during the evolution of eukaryotes.

Since eukaryotes were not obliged to minimize the size of their genomes, the lateral gene transfer, a genome loading mechanism (though it occurred when many cyanobacterial genes were incorporated into nuclear DNA in the origin of plastids), was no longer essential in the genome organization of eukaryotes.

In spite of the significant DNA expansion, extremely reduced (prokaryotic-like) genomes can be found in certain eukaryotes, as a result of secondary reduction due to parasitism. For example, Microsporidia, all of which are obligate intracellular parasites of higher eukaryotes (Canning & Lom 1986), are known that have a 2.9 Mb genome (Biderre *et al.* 1996), smaller than that of *Escherichia coli*. Although these primitive eukaryotes possess a nucleus, endomembrane and cytoskeleton, but no other organelles, they entirely rely on the metabolism of hosts, suggesting the secondary loss of their own genes involved in metabolism.

4. GENOMES OF EUKARYOTES WITHOUT MITOCHONDRIA

The genome size of the binucleate protozoan *Giardia lamblia*, which has no energy-converting organelles, was estimated at 12 Mb distributed onto five chromosomes (TIGR Microbial Database 1998). How could these protozoa secondarily lose their mitochondria while

maintaining their complex and relatively large genomes? An explanation can be that most protozoan parasites entirely lack purine synthesis. Moreover, *G. lamblia* and *Trichomonas vaginalis* also lack *de novo* synthesis of pyrimidines; they depend only on the salvage of exogenously synthesized nucleic acids and nucleotides (Adam 1991). The secondary loss of mitochondria can be explained even in the free-living diplozoan *Trepomonas agilis*, which is believed to have diverged earlier in the eukaryotic lineage than *Giardia* (Cavalier-Smith & Chao 1996), though there is no direct experimental evidence yet. *Trepomonas* are phagotrophic, implying that the highly energy consuming *de novo* pathway of nucleotide synthesis from amino acids, tetrahydrofolate derivatives, CO₂, and ribose 5-phosphate could be replaced by a much less costly salvage reaction (Stryer 1988), in which the free purine bases are formed directly by the hydrolytic degradation of nucleic acids and nucleotides derived from other organisms. Thus, amitochondriate protozoa could have maintained their relatively large genomes in the absence of mitochondria because they have relied only on the salvage of exogenous nucleic acids and nucleotides.

5. THE ORIGIN OF EUKARYOTES IN CELL EVOLUTION

The first living organisms are believed to have arisen more than 3.85 billion years ago (Holland 1997) (figure 2). Presumably, the metabolic capability and genetic system of the earliest cellular entities were more simple and rudimentary than that of any modern cell (Woese 1998). The transfer of genetic information may have been pervasive and non-specific among the primitive genomes, making each entity capable of acquiring any innovation occurring within the entire biotic community. Indeed, the absence of barriers to DNA exchange (general genetic mixing) may have characterized the first main stage of cellular evolution.

Later, parallel with the increase in genetic and cellular complexity, some genetic subsystems became non-compatible with exogenous parts. When lateral gene transfer was restricted significantly, the first prokaryotes emerged, thereby initiating the second major stage of cellular evolution. This was the time of the origin of individuality and speciation, presumably through partial reproductive isolation. Despite the limitation in formerly general genetic mixing, the horizontal gene transfer has continued to be essential in the speciation dynamics of prokaryotes (Lawrence & Ochman 1998). Ancestral prokaryotes, the first modern cells, developed novel biochemical pathways for ATP production, such as the present-day forms of fermentation, and then oxidative phosphorylation. A considerable step was the development of photosynthesis, probably more than three billion years ago (Schopf 1993). The oxygen-producing photosynthesis resulted in, at least in part, the appearance of molecular O₂ in the atmosphere. However, the increase in atmospheric oxygen level was very slow at first, because the free O₂ was used in converting Fe(II) to Fe(III) (Schopf *et al.* 1983). The availability of free O₂, however, made the development of aerobic bacteria possible. About two billion years ago, slightly before the origin of eukaryotes (Feng *et al.* 1997), a rapid atmospheric

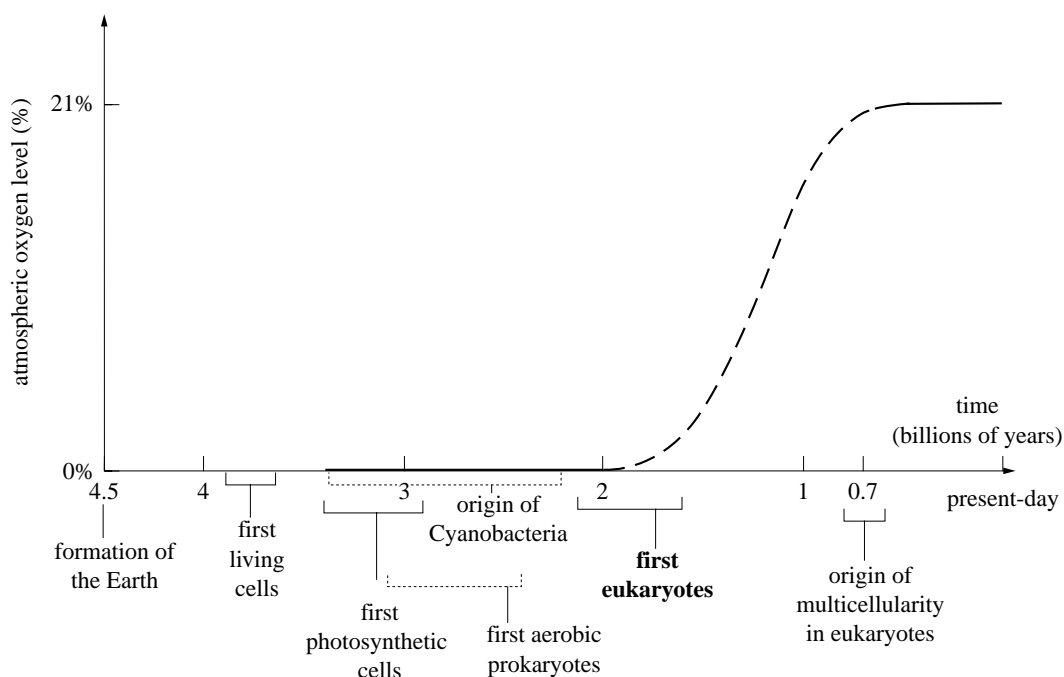


Figure 2. A putative scenario of the main steps in cellular evolution. Before the emergence of the first photosynthetic cells the amount of free O_2 produced abiotically by photodissociation of H_2O was negligible. The divergence time of cyanobacteria (indicated by a dotted line) is highly contentious (Schopf 1993; Feng *et al.* 1997).

oxygen accumulation began (Kasting 1993) resulting from the photosynthetic activity of cyanobacteria and the exhaustion of $Fe(II)$ in the oceans.

Under these selective conditions, the anaerobic cells either died due to this photosynthetic by-product highly toxic to them or retreated to anaerobic niches. Based on the observations that the O_2 first reached a significant level in the atmosphere slightly more than two billion years ago (Kasting 1993), and the divergence of eukaryotes almost coincided with this (Feng *et al.* 1997), another possibility of the survival of anaerobic cells was proposed (Martin & Müller 1998; Vellai *et al.* 1998). According to this idea, certain anaerobic archaeobacteria might have survived in environments that were becoming oxidizing by the incorporation of respiring proteobacteria. The engulfed symbionts using O_2 for respiration may have decreased the intracellular oxygen concentration of host cells. In parallel, the symbiont could get into a new niche. Thus, this hypothesis argues that the initial advantage of this symbiotic association was not ATP export from the symbiont into the cytosol of the host through respiration (as proposed previously by the oxygen hypothesis for the acquisition of mitochondrial precursors by ancestral anaerobic protozoa; Margulis 1970, 1996), but rather clearing the host from O_2 and its free radicals. The gradual, slow change in the composition of the atmosphere made the maintenance and stabilization of the mutually beneficial association between the two different prokaryotes possible, leading to their co-evolution. Eventually, the symbionts evolved into mitochondria reorganizing the original metabolism of hosts, whereas the host cells developed a nuclear membrane around their expanded genetic material. So, according to this proposal, (proto)eukaryotes may have acquired a mitochondrial symbiont before they evolved any of their other typical cytological features like the nucleus, cytoskeleton and

endomembrane system. Going back again into anaerobic environments, however, mitochondria could have been lost secondarily or developed into anaerobic energy-producing organelles, such as hydrogenosomes (Akhmanova *et al.* 1998) or anaerobic mitochondria (Embley & Martin 1998), by loss of organellar genes responsible for respiration (Bui *et al.* 1996). The development of complex cells appears to be the last great innovation in cellular evolution, initiating its third major stage. The lateral gene transfer occurring especially, but not exclusively, through endosymbioses was no longer essential in the organization of eukaryotic genomes. Thus, their subsystems could have been changed primarily through mutations.

Alternatively, the so-called 'hydrogen hypothesis' (Martin & Müller 1998) suggests that anaerobic archaeobacteria (hosts) were originally strictly autotrophic and hydrogen dependent. Later, they were transferred under conditions characterized by the absence of hydrogen sources. This selective force made starving archaeobacteria establish tight physical association with heterotrophic proteobacteria able to produce molecular H_2 through anaerobic fermentation. The symbionts feeding the hosts were surrounded by them completely, and were then lost or evolved into mitochondria, or into hydrogenosomes, explaining the anaerobic energy metabolism in eukaryotes. According to this hypothesis, the initial advantage of cellular merger was the excretion of molecular H_2 produced by the symbionts.

Even though the so-called symbiotic association hypotheses are not supported by evidence regarding the mechanism of cellular merger, the traditional view on the role of the key importance of phagocytosis in the prokaryote-eukaryote transition (Cavalier-Smith 1987) seems to be implausible. Phagocytosis is an extremely costly and complex process. What kind of inventions would be essential for a prokaryote to evolve into a cell capable of endocytosis? First, it

would need to take off the cell wall and, consequently, the periplasm. To exist without a cell wall, however, is a special venture for most bacteria, generally accompanied by a dramatic restriction in both genome and cell size. Second, this prokaryote would need to possess a complex cytoskeleton and an internal membrane system. The latter must be able to isolate digestive enzymes intracellularly (lysosomes), contrary to bacteria that secrete digestive enzymes to the cell exterior, implying the presence of a primitive Golgi apparatus as well. However, the cell possessing all of these instruments can already be regarded as a complex cell. In the presence of an endomembrane, a nuclear membrane could be developed endogenously resulting in a nucleus-bearing real eukaryote.

It is more likely that a still not well-defined acquisition of the mitochondrial precursor made the archaeobacterial host able to reorganize its energy metabolism via compartmentalization. The transfer of the originally plasma-membrane-bounded metabolism into organelles allowed the wall-covered cells to lose the cell wall and the periplasm. Parallel to the opening of the cell surface, the increased genetic potential even in a single cell could now encode the cellular tools required for endocytosis, which was thus a eukaryotic invention. Some eukaryotes were eventually able to acquire photosynthetic prokaryotes by phagocytosis, which have evolved into chloroplasts.

We thank Eörs Szathmáry, Michael Klein, József Pálffy and András Takács-Sánta for their helpful comments on the manuscript, and the Collegium Budapest for financial support. We are very grateful to two anonymous reviewers for critical comments on the first version of the manuscript.

REFERENCES

- Adam, R. D. 1991 The biology of *Giardia* spp. *Microbiol. Rev.* **55**, 706–732.
- Akhmanova, A., Voncken, F., Van Alen, T., Van Hoek, A., Boxma, B., Vogels, G., Veenhuis, M. & Hackstein, J. H. P. 1998 A hydrogenosome with a genome. *Nature* **396**, 527–528.
- Andersson, S. G. E. & Kurland, C. H. 1998 Reductive evolution of resident genomes. *Trends Microbiol.* **6**, 263–268.
- Baril, C., Richaud, C., Baranton, G. & Saint-Girons, I. S. 1989 Linear chromosome of *Borrelia burgdorferi*. *Res. Microbiol.* **140**, 570–576.
- Battista, J. R. 1997 Against all odds: the survival strategies of *Deinococcus radiodurans*. *A. Rev. Microbiol.* **51**, 203–224.
- Bidre, C., Pagés, M., Méténier, G., Canning, E. U. & Vivares, C. P. 1996 Evidence for the smallest nuclear genome (2.9 Mb) in the microsporidian *Encephalitozoon cuniculi*. *Mol. Biochem. Parasitol.* **74**, 229–231.
- Brown, J. R. & Doolittle, W. F. 1997 *Archaea* and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev.* **61**, 456–502.
- Bui, E. T. N., Bradley, P. J. & Johnson, P. J. 1996 A common evolutionary origin for mitochondria and hydrogenosomes. *Proc. Natl Acad. Sci. USA* **93**, 9651–9656.
- Canning, E. U. & Lom, J. 1986 *The microsporidia of vertebrates*. London: Academic Press.
- Carlile, M. 1982 Prokaryotes and eukaryotes: strategies and successes. *Trends Biochem. Sci.* **7**, 128–130.
- Cavalier-Smith, T. 1987 Eukaryotes with no mitochondria. *Nature* **326**, 332–333.
- 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.
- Clark, C. G. & Roger, A. J. 1995 Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. *Proc. Natl Acad. Sci. USA* **92**, 6518–6521.
- Doolittle, R. F. 1998a Microbial genomes opened up. *Nature* **392**, 339–342.
- Doolittle, W. F. 1998b A paradigm gets shifty. *Nature* **392**, 15–16.
- Doolittle, W. F. 1998c You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* **14**, 307–311.
- Embley, T. M. & Hirt, R. P. 1998 Early branching eukaryotes? *Curr. Opin. Genet. Dev.* **8**, 624–629.
- Embley, T. M. & Martin, W. 1998 A hydrogen-producing mitochondrion. *Nature* **396**, 517–519.
- Errington, J. 1993 *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. *Microbiol. Rev.* **57**, 1–33.
- 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**, 13 028–13 033.
- Ferat, J.-L. & Michel, F. 1993 Group II self-splicing introns in bacteria. *Nature* **364**, 358–361.
- Fonstein, M. & Haselkorn, R. 1995 Physical mapping of bacterial genomes. *J. Bacteriol.* **177**, 3361–3369.
- Fuerst, J. A. & Webb, R. I. 1991 Membrane-bounded nucleoid in the eubacterium *Gemmata obscuriglobus*. *Proc. Natl Acad. Sci. USA* **88**, 8184–8188.
- Germot, A., Philippe, H. & Le Guyader, H. 1996 Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. *Proc. Natl Acad. Sci. USA* **93**, 14 614–14 617.
- Germot, A., Philippe, H. & Le Guyader, H. 1997 Evidence for loss of mitochondria in Microsporidia from a mitochondrial-type HSP70 in *Nosema locustae*. *Mol. Biochem. Parasitol.* **87**, 159–168.
- Gray, M. W., Burger, G. & Lang, B. F. 1999 Mitochondrial evolution. *Science* **283**, 1476–1481.
- Gupta, R. S. 1998 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., Aitken, K., Falah, M. & Singh, B. 1994 Cloning of *Giardia lamblia* heat shock protein HSP70 homologs: implications regarding origin of eukaryotic cells and of endoplasmic reticulum. *Proc. Natl Acad. Sci. USA* **91**, 2895–2899.
- Hashimoto, T., Sanchez, 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.
- Holland, H. D. 1997 Evidence for life on Earth more than 3850 million years ago. *Science* **275**, 38–39.
- Horner, D. S., Hirt, R. P., Kilvington, S., Lloyd, D. & Embley, T. M. 1996 Molecular data suggest an early acquisition of the mitochondrion endosymbiont. *Proc. R. Soc. Lond.* **B263**, 1053–1059.
- Jensen, K. F. & Pedersen, S. 1990 Metabolic growth rate control in *Escherichia coli* may be a consequence of subsaturation of the macromolecular biosynthetic apparatus with substrates and catalytic components. *Microbiol. Rev.* **54**, 89–100.
- Kasting, J. F. 1993 Earth's early atmosphere. *Science* **259**, 920–926.
- Keeling, P. J. & Doolittle, W. F. 1997 Evidence that eukaryotic triosephosphate isomerase is of alpha-proteobacterial origin. *Proc. Natl Acad. Sci. USA* **94**, 1270–1275.
- Kelman, Z. & O'Donnell, M. 1994 DNA replication: enzymology and mechanisms. *Curr. Opin. Genet. Dev.* **4**, 185–195.
- Klingenberg, M. 1979 The ADP, ATP shuttle of the mitochondrion. *Trends Biochem. Sci.* **4**, 249–252.

- Kondrashov, A. & Crow, J. F. 1991 Haploidy or diploidy: which is better? *Nature* **351**, 314–315.
- Lawrence, J. G. & Ochman, H. 1998 Molecular archeology of the *Escherichia coli* genome. *Proc. Natl Acad. Sci. USA* **95**, 9413–9417.
- Margolin, W. 1998 A green light for the bacterial cytoskeleton. *Trends Microbiol.* **6**, 233–238.
- Margulis, L. 1970 *Origin of eukaryotic cells*. New Haven, CT: Yale University Press.
- Margulis, L. 1996 Archaeal–eubacterial mergers in the origin of Eukarya: phylogenetic classification of life. *Proc. Natl Acad. Sci. USA* **93**, 1071–1076.
- Martin, W. & Müller, M. 1998 The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41.
- Maynard Smith, J. & Szathmáry, E. 1995 *The major transitions in evolution*. Oxford and New York: Freeman.
- Mereschowsky, C. 1910 Theorie der Zwei Pflanzenarten als Grundlage der Symbiogenese, einer neuen Lehre der Entstehung der Organismen. *Biol. Zentralbl.* **30**, 278–367.
- 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.
- Moyer, C. L. & Morita, R. Y. 1989 Effect of growth rate and starvation–survival on cellular DNA, RNA, and protein of a psychrophilic marine bacterium. *Appl. Environ. Microbiol.* **55**, 2710–2716.
- Prescott, L. M., Harley, J. P. & Klein, D. A. 1996 *Microbiology*. 3rd edn. Dubuque, IA: W. C. Brown Publishers.
- Reiner, D. S., Shinnick, T. M., Ardeshir, F. & Gillin, F. D. 1992 Encystation of *Giardia lamblia* leads to expression of antigens recognized by antibodies against conserved heat shock proteins. *Infect. Immun.* **60**, 5312–5315.
- Rivera, M. C., Jain, R., Moore, J. E. & Lake, J. A. 1998 Genomic evidence for two functionally distinct gene classes. *Proc. Natl Acad. Sci. USA* **95**, 6239–6244.
- Roger, A. J., Clark, C. G. & Doolittle, W. F. 1996 A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. *Proc. Natl Acad. Sci. USA* **93**, 14 618–14 622.
- 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. USA* **95**, 229–234.
- Schopf, J. W. 1993 Microfossils of the early archean apex chert: new evidence of the antiquity of life. *Science* **260**, 640–646.
- Schopf, J. W., Hayes, J. M. & Walter, M. R. 1983 Evolution of Earth's earliest ecosystems: recent progress and unsolved problems. In *Earth's earliest biosphere: its origin and evolution* (ed. J. W. Schopf), pp. 361–384. Princeton University Press.
- Schulz, H. N., Brinkhoff, T., Ferdelman, T. G., Hernández Mariné, M., Teske, A. & Jorgensen, B. B. 1999 Dense populations of a giant sulfur bacterium in namibian shelf sediments. *Science* **284**, 493–495.
- Seto, S. & Miyata, M. 1998 Cell reproduction and morphological changes in *Mycoplasma capricolum*. *J. Bacteriol.* **180**, 256–264.
- Sogin, M. L. 1991 Early evolution and the origin of eukaryotes. *Curr. Opin. Genet. Dev.* **1**, 457–463.
- Stryer, L. 1988 *Biochemistry*. New York: Freeman.
- TIGR Microbial Database (<http://www.tigr.org/tdb/mdb/mdb/html>) The Institute for Genomic Research.
- Vellai, T., Takács, K. & Vida, G. 1998 A new aspect to the origin and evolution of eukaryotes. *J. Mol. Evol.* **46**, 499–507.
- Wheeler, R. T. & Shapiro, L. 1997 Bacterial chromosome segregation: is there a mitotic apparatus? *Cell* **88**, 577–579.
- Woese, C. R. 1987 Bacterial evolution. *Microbiol. Rev.* **51**, 221–271.
- Woese, C. R. 1998 The universal ancestor. *Proc. Natl Acad. Sci. USA* **95**, 6854–6859.
- Zillig, W., Palm, P., Klenk, H.-P., Pühler, G., Gropp, F. & Schleper, C. 1991 Phylogeny of DNA-dependent RNA-polymerases: testimony for the origin of eukaryotes. In *General and applied aspects of halophilic microorganisms* (ed. F. Rodriguez-Valera), pp. 321–332. New York: Plenum Press.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.

