The Origin of Eukaryotes: The Difference between Prokaryotic and Eukaryotic Cells

Tibor Vellai; Gabor Vida


Stable URL: http://links.jstor.org/sici?sici=0962-8452%2819990807%29266%3A1428%3C1571%3ATOOETD%3E2.0.CO%3B2-K

*Proceedings: Biological Sciences* is currently published by The Royal Society.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/rsl.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.
The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells

Tibor Vellai* and Gábor Vida

1Institute for Advanced Study, Collège Budapest, Budapest, H-1014. Szentléány s. u. 2., Hungary
2Department 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 archaebacteria (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 megaevolutionary theory concerns the prokaryote—eukaryote transition (Brown & Doolittle 1997; Gupta 1998; Gray et al. 1999). Merechowsky (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 archaebacterial-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 mitochondriate protists.

Since the beginning of the 1990s, the so-called chimera hypotheses (Sogin 1991; Zillig et al. 1991; Moreira & López-Garcia 1998) suggested that the eukaryotic cell did not directly evolve from an archaebacterial-like ancestor. Rather, the eukaryotic nucleus has been formed from the cellular fusion between an archaebacterium 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, Meta-monada 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).
in both Gram-negative α-proteobacteria and mitochondria of eukaryotes, but not in archaeobacterial or Gram-positive eu- bacterial species, were found in all amitochondriate protozoan taxa (Reiner et al. 1992; Gupta et al. 1994; Clark & Roger 1995; Horner et al. 1996; Gernot 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-bran- ching eukaryotes (Embley & Hirt 1998). There is also doubt whether amitochondriate protozoa really diverged earlier than other eukaryotes.

These data imply that mitochondria may have origi- nated 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 trichomono- 

In 1998, two hypotheses, based on the comparative biochemistry of energy metabolism (Martin & Muller 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 iden- tical to the origin of eukaryotes. The proposed scenarios assume that some ancient aerobic 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 respira- tion, as previously thought), the symbiont 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 cyto- skeleton. According to this proposal, the cellular merger leading to the origin of eukaryotes might have taken place simultaneously among a eubacterium, an archaeabacterium 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 cyto- skeleton, whereas prokaryotes do not. In many cases,
Table 1. Major differences in genome organization between prokaryotes and eukaryotes

<table>
<thead>
<tr>
<th>characteristics of genome organization</th>
<th>prokaryotes</th>
<th>eukaryotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>genome size</td>
<td>600 kb–9.5 Mb</td>
<td>3 Mb–140 000 Mb</td>
</tr>
<tr>
<td>average gene size/number</td>
<td>950 bp/4300a</td>
<td>2500 bp/19 000b</td>
</tr>
<tr>
<td>operon-like regulatory unit</td>
<td>general</td>
<td>no</td>
</tr>
<tr>
<td>horizontal gene transfer rate of non-coding sequences</td>
<td>significant</td>
<td>negligible</td>
</tr>
<tr>
<td>intron</td>
<td>rarec</td>
<td>general</td>
</tr>
<tr>
<td>redundant gene number</td>
<td>rare²</td>
<td>general</td>
</tr>
<tr>
<td>ploidy level</td>
<td>haploidy</td>
<td>haploidy–polyplody</td>
</tr>
<tr>
<td>chromosome number</td>
<td>onef</td>
<td>more than one</td>
</tr>
<tr>
<td>heterozygosity</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

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 tRNA operons.
e Group II self-splicing introns in cyanobacteria and proteobacteria.
f For example, tRNA genes.

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 organization 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 metabolic 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. 1990). 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 intra-cellular parasitic lifestyle (Andersson & Kurland 1998), are able to achieve a relatively faster reproduction rate, thus strengthening the so-called r-strategy (Carlini 1982). Selection for genome streaming 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'Donnell 1994); the DNA 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 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 polyploidy 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 in the free-living diplozoan Trepomonas agilis, which is believed to have diverged earlier in the eukaryotic lineage than Giardia (Cavaler-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
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 archaeabacteria 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 archaebacteria (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 archaeabacteria 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 archaeabacterial host able to reorganize its energy metabolism via compartmentalization. The transfer of the originally plasma-membrane-bound 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 Éórs Szathmáry, Michael Klein, József Pálfy and Andráš 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


The origin of eukaryotes  T. Vellai and G. Vida  1577


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