THE ROYAL SOCIETY

Published online 18 December 2002

On the origin of mitochondria: a genomics perspective

Siv G. E. Andersson*, Olof Karlberg, Björn Canbäck and Charles G. Kurland

Department of Molecular Evolution, University of Uppsala, Uppsala S-75124, Sweden

The availability of complete genome sequence data from both bacteria and eukaryotes provides information about the contribution of bacterial genes to the origin and evolution of mitochondria. Phylogenetic analyses based on genes located in the mitochondrial genome indicate that these genes originated from within the α-proteobacteria. A number of ancestral bacterial genes have also been transferred from the mitochondrial to the nuclear genome, as evidenced by the presence of orthologous genes in the mitochondrial genome in some species and in the nuclear genome of other species. However, a multitude of mitochondrial proteins encoded in the nucleus display no homology to bacterial proteins, indicating that these originated within the eukaryotic cell subsequent to the acquisition of the endosymbiont. An analysis of the expression patterns of yeast nuclear genes coding for mitochondrial proteins has shown that genes predicted to be of eukaryotic origin are mainly translated on polysomes that are free in the cytosol whereas those of putative bacterial origin are translated on polysomes attached to the mitochondrion. The strong relationship with α-proteobacterial genes observed for some mitochondrial genes, combined with the lack of such a relationship for others, indicates that the modern mitochondrial proteome is the product of both reductive and expansive processes.

Keywords: glycolysis; hydrogenosomes; mitochondria; respiration; yeast

1. INTRODUCTION

The mitochondrion is a cellular organelle that serves as the 'power-house' of the eukaryotic cell. ATP production from glucose and oxygen is in most mitocondriate eukaryotes the result of two catabolic processes. Glucose is first converted to pyruvate via the glycolytic pathway in the cytoplasm; pyruvate is then imported into the mitochondrion and oxidized to CO_2 and H_2O via two interconnected and highly specialized functions: the tricarboxylic-acid cycle and the respiratory chain complex. Aerobic respiration is strictly dependent on oxygen, and during conditions of oxygen deprivation ATP synthesis is halted, eventually resulting in cell death. However, yeast and some other eukaryotes can survive by fermentation in the absence of oxygen, during which glucose is converted to lactate or ethanol via the glycolytic pathway.

It is known that there is a trade-off between the yield (moles of ATP per mol of substrate) and the rate (moles of ATP per unit time) of ATP production (Stucki 1980). The balance between these two parameters is different for fermentation and respiration. Fermentation supports a high rate of ATP production, but at the cost of a lower total yield (only 2 mol of ATP per mol of glucose). By contrast, respiration supports a low rate of ATP production but results in a high yield of ATP (more than 30 mol of ATP per mole of glucose; Stryer 1995). Thus, from a thermodynamic perspective, aerobic respiration is the preferred pathway for ATP production in oxygenic environments, especially if the substrate is a limiting resource.

Not surprisingly, aerobically respiring mitochondria are present in all multicellular organisms that live in oxygen-rich environments, including humans, arthropods, vertebrates and plants (Gray 1992). However, some organisms, mainly unicellular protists and parasitic nematodes, have adapted to environments that are poor in or devoid of oxygen. These organisms rely on terminal acceptors other than O_2, such as NO_3^- and NO_2^- (Finlay *et al.* 1983; Kobayashi *et al.* 1996; Zumft 1997; Takaya *et al.* 1999). Nitrate respiration by NiR results in the production of NO_2^- and NO rather than water as in aerobic respiration. A different type of anaerobic mitochondria is found in parasitic helmiths where malate is imported from the cytoplasm and the electron transport chain is linked to the reduction of fumarate to succinate by fumarate reductase (Tielens 1994; Tielens & Van Hellemond 1998; Kita *et al.* 2002).

Another type of anaerobic ATP-producing organelle, the hydrogenosome, has been described in a wide spectrum of anaerobic protists (Mueller 1993; Embley *et al.* 1995; Akhmanova *et al.* 1998; Hackstein *et al.* 1999; Voncken *et al.* 2002; Van der Giezen *et al.* 2002). In these organelles, which lack membrane-associated electron-transport chain complexes, ATP production is linked to H_2 production with the aid of enzymes such as PFO and hydrogenase. ATP production by anaerobic mitochondria and hydrogenosomes results in a lower total yield than aerobic respiration. For example, the reduction of fumarate to succinate yields only 5 mol of ATP per mol of degraded glucose.

*Author for correspondence (siv.andersson@ebc.uu.se).

One contribution of 21 to a Discussion Meeting Issue 'Chloroplasts and mitochondria: functional genomics and evolution'.

The origin of the eukaryotic cell as well as its different organelles has been a topic of discussion ever since the differences between these and the prokaryotes were revealed. The discovery of mitochondrial and plastid genomes revived the 19th-century view of plastids as domesticated bacterial endosymbionts (Mereschowsky 1910). However, it was not until the more recent developments in molecular biology that the theory really gained ground (Margulis 1970, 1981). The similarities between bacterial respiration and mitochondrial functions as well as the similarities between cyanobacterial photosynthesis and chloroplast functions in plants taken together with molecular phylogenies have provided convincing evidence of the bacterial origins of these modern cellular organelles (Gray 1992). Nevertheless, the big questions remain: from where, why and how?

2. MITOCHONDRIAL GENOME DIVERSITY

First, we need to summarize briefly the content and sizes of the genetic material present in mitochondria. Despite the extreme difference in overall genome sizes, ranging from only 16 kb in humans to more than 500 kb in some plant species (Hanson & Folkerts 1992; Unseld *et al.* 1997; Gray *et al.* 1998), most mitochondrial genomes contain between 12 and 20 protein-coding genes. The mitochondrial genome of *Plasmodium falciparum*, which contains only two protein-coding genes, represents one extreme, and the other is *Reclinomonas americana* with as many as 67 genes (Lang *et al.* 1997). Most of these code for key components of aerobic respiration and translation, such as cytochrome oxidases and ribosomal protein genes.

The mitochondrial genome of *Saccharomyces cerevisiae* is 85 779 bp and encodes two rRNAs, 24 tRNAs and 30 proteins. One-third of the latter genes encode vital mitochondrial functions such as subunits of the ATP synthase complex, cytochrome *b* and the cytochrome *c* oxidase complex (Foury *et al.* 1998). The yeast mitochondrial genome is somewhat atypical in that as many as 10 genes code for endonucleases, reverse transcriptases and mRNA maturases. Another 10 genes are of unknown function. In striking contrast to the small number of proteins encoded in the yeast mitochondrial genome, as many as 400 or more proteins with assigned functions in the yeast mitochondrion (Hodges *et al.* 1999) are encoded in the nuclear genome and subsequently imported into the mitochondrion (www.proteome.com).

Due to the rapid rate of sequence evolution for the small and large rRNA gene sequences in animal and human mitochondria, these have evolved in diverse ways that complicated the interpretations of the first phylogenetic analyses (Gray *et al.* 1989; Gray 1992). However, improvements in phylogenetic methods as well as a larger sampling of mitochondrial genes for phylogenetic reconstructions have by now provided convincing evidence that the various mitochondrial genomes have a single, monophyletic origin (Gray *et al.* 1999).

3. FROM WHERE?

Phylogenetic reconstructions based on mitochondrially encoded proteins have placed the mitochondrial ancestor in the α-proteobacterial subdivision, and several reconstructions point specifically towards the Rickettsiaceaea family (Yang *et al.* 1985; Olsen *et al.* 1994; Viale & Arakaki 1994; Gray & Spencer 1996; Sicheritz-Ponten *et al.* 1998; Karlberg *et al.* 2000; Kurland & Andersson 2000). For example, phylogenetic studies based on mitochondrial proteins such as cytochrome oxidase subunits and cytochrome *b* indicate that the respiratory chain complex were derived from an ancestral endosymbiont belonging to the α-proteobacteria (figure 1; Sicheritz-Ponten *et al.* 1998; Gray *et al.* 1999). Likewise, an interpretation of phylogenetic reconstructions based on ribosomal proteins is that a significant part of the translation system is also derived from the α-proteobacteria (Karlberg *et al.* 2000). A series of α-proteobacterial genomes, with sizes ranging from 1 Mbp in *Rickettsia* to *ca.* 8 Mbp in *Mesorhizobium loti*, are now completely sequenced, providing a rich source of material for addressing questions concerning the specifics of mitochondrial origins.

The smallest genomes of the α-proteobacteria are found within Rickettsiaceaea, which belongs to the order Rickettsiales and contains three tribes Ehrlichieae, Rickettsieae and Wolbachieae (National Center for Biotechnology Information taxonomy database, htttp://www.ncbi.nlm.nih.gov/ (Wheeler *et al.* 2000)), all with obligate intracellular lifestyles. Two genomes have been sequenced in this group: that of *Rickettsia prowazekii*, the causative agent of epidemic typhus (Andersson *et al.* 1998) and that of *Rickettsia conorii*, the causative agent of Mediterranean Spotted Fever (Ogata *et al.* 2001). In addition to different disease symptoms, the two *Rickettsia* species use different arthropod vectors for transmission. *R. prowazekii* is transmitted by the human body louse *Pediculus humanus*, whereas *R. conorii* is transmitted by the dog brown tick *Ixodes*.

The genome of *R. prowazekii* comprises 1.1 Mbp encoding 834 proteins (Andersson *et al.* 1998), whereas the *R. conorii* genome is 1.3 Mbp with *ca.* 1300 annotated genes (Ogata *et al.* 2001). Both genomes are in a process of evolutionary degradation as reflected in a high content of non-coding DNA (19–24% relative to a typical non-coding content of 10% in bacteria) (Andersson *et al.* 1998; Andersson & Andersson 1999*a*,*b*, 2001; Ogata *et al.* 2001). Despite their compactness, both genomes contain a complete set of genes for aerobic respiration as well as for ATP-transport functions. Thus, both are excellent producers of ATP as well as extractors of ATP from their host cells. Due to their obligate intracellular lifestyles, compact genomes, respiration capabilities and specific relationship with mitochondria as inferred from genes such as cytochrome oxidase (figure 1), they represent a particularly good model system for studies of the evolutionary relationship between bacteria and mitochondria.

The genomes of two other intracellular parasites within the α-proteobacteria, *Bartonella quintana* and *Bartonella henselae*, have recently been sequenced (C. M. Alsmark and 11 others, unpublished data). These bacteria, which cause trench fever and cat-scratch disease, respectively, can be cultivated on artificial media although growth rates are higher in the intracellular environment. The genomes have sizes of 1.6 and 1.9 Mbp, which is reflected in a larger metabolic capacity than for their obligate intracellular relatives of the genus *Rickettsia* (C. M. Alsmark and 11 others, unpublished data). For example, they contain a set of glycolytic genes that are not present in *Rickettsia*,

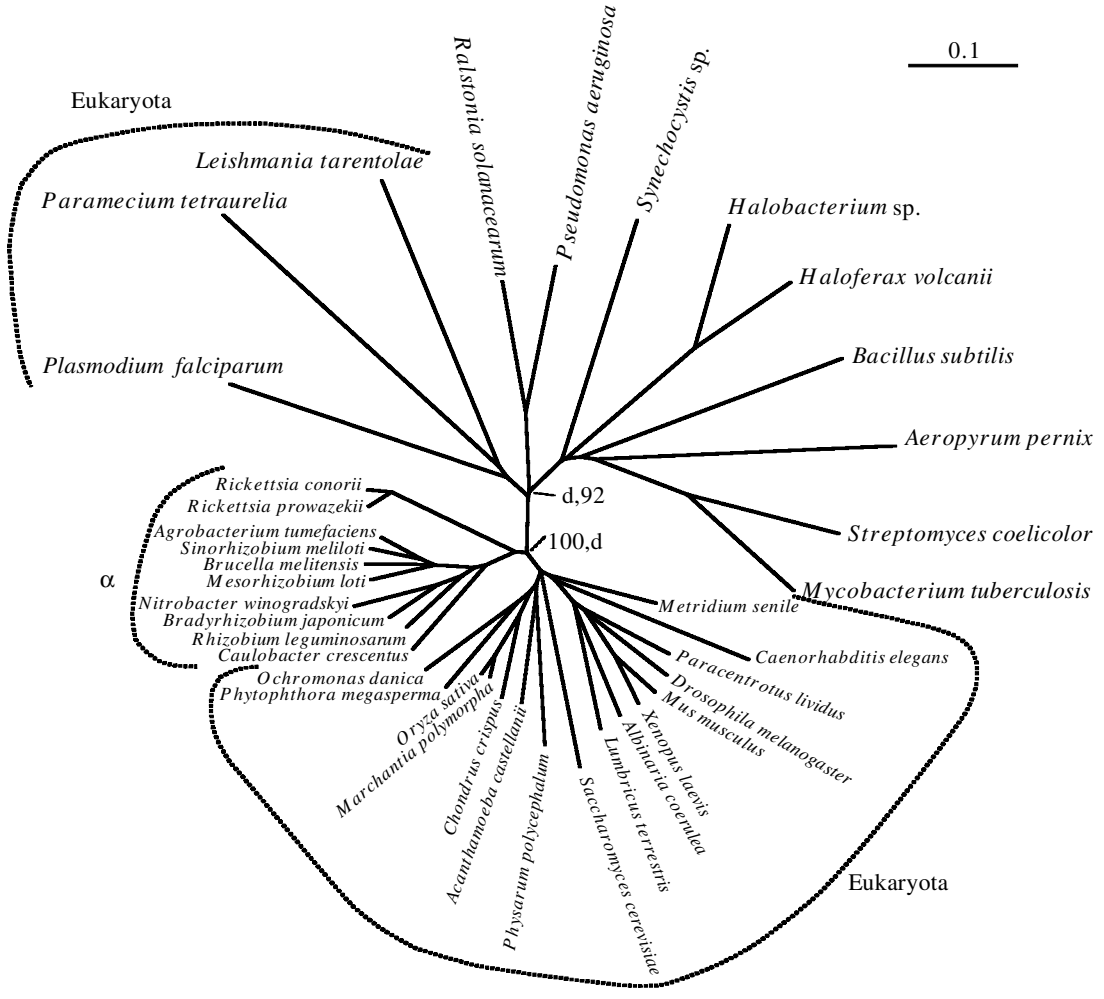

Figure 1. The ME tree for cytochrome *c* oxidase subunit I based on amino-acid sequences where ambiguous regions were removed from the alignment. The topology and branch lengths correspond to a ME tree calculated with random addition order and 500 repetitions, using distances based on the raw fraction of amino acids, equal rates and no invariable sites. The first bootstrap value corresponds to a ME value, the second to a maximum-parsimony value. In both cases 500 bootstrap replicates were produced with a number of 10 repetitions per replicate. The symbol 'α' denotes α-proteobacteria; a 'd' in the bootstrap values denotes that a different topology was produced. Accession numbers clockwise starting with *P. falciparum*: Q02766, P05489, P14544, NP_518484, AAG03496, Q06473, P33518, AAK14377, 24010, NP_147500, NP_631290, O53290, Q35101, P24893, P12700, P00399, P00397, P00398, P48887, Q34941, P00401, Q07434, NP_042527, P48866, NP_054454, P14578, Q02211, NP_066450, NP_422200, Q08855, P31833, CAA61744, BAB53582, NP_540382, NP_385011, NP_531468, O54069, NP_360190. Scale bar shows substitutions per site.

which have been useful for studies on the origin and evolution of glycolysis (Canbäck *et al.* 2002). In contrast to the aerobic respiration system (figure 1), the glycolytic system of the eukaryotes seems not to trace its origin back to the α-proteobacteria (figure 2), indicating that it was not acquired from the mitochondrial endosymbiont.

Another completely sequenced genome within this group is that of *Brucella melitensis*, a facultative intracellular parasite like *B. henselae and B. quintana* but with an even larger genome. This genome of 3.2 Mbp is divided into two megareplicons (DelVecchio *et al.* 2001). Several genomes of free-living bacterial species have also been sequenced within the α-proteobacteria. These include *Caulobacter crescentus*, an aquatic organism with a complex life cycle that thrives in a nutrient poor environment. It has one chromosome of 4 Mbp, containing more than 3700 genes (Nierman *et al.* 2001). *Agrobacterium tumefaciens* is a plant pathogen that causes crown gall disease in the rose family. This genome has a size of 5.7 Mbp along with one of the most remarkable structures known to date: this consists of two main chromosomes, one of which is linear, and two plasmids (Goodner *et al.* 2001; Wood *et al.* 2001). *Sinorhizobium meliloti* lives in a symbiotic relationship with plants forming nitrogen-fixing nodules and has a genome of 6.7 Mbp, divided into one chromosome and two mega-replicons (Galibert *et al.* 2001). Finally, the largest genome sequenced so far within this group is that of *Mesorhizobium loti*, which consists of one main chromosome of 7 Mbp and two smaller plasmids (Kaneko *et al.* 2000).

The broad distribution of genome sizes in the α-proteobacteria is most probably an expression of both reductive and expansive processes. For example, the small genome sizes of obligate intracellular parasites are mainly explained by reductive evolution (Andersson & Kurland 1998; Andersson & Andersson 1999*a*,*b*), with a biased loss of genes involved in biosynthetic and regulatory functions. For example, most or all of the amino-acid biosyn-

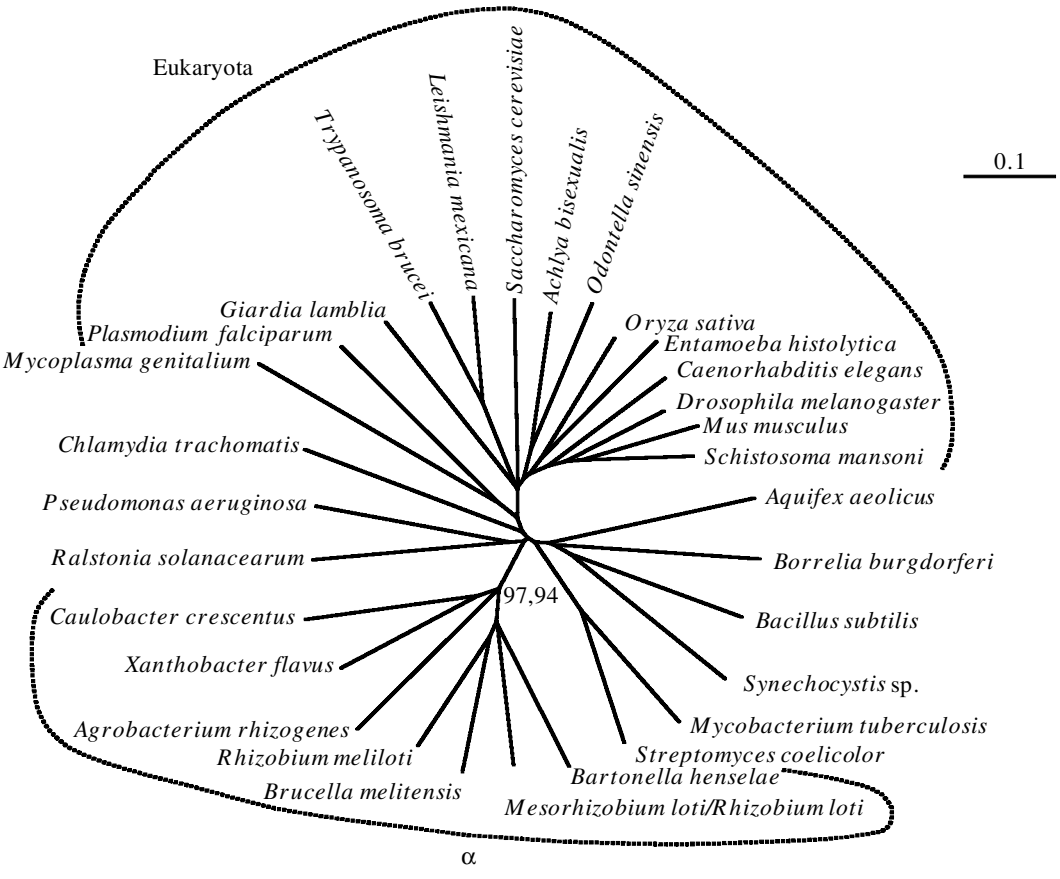

Figure 2. The ME tree for triose-phosphate isomerase based on amino-acid sequences. Abbreviations, methods, bootstrap values and other details are as described in the legend to figure 1. Accession numbers clockwise starting with *Mycoplasma genitalium*: U39725, L01654, L02120, X03921, P48499, J01366, AF063107, Q9M7R5, M87064, Y13387, U23081, P29613, P17751, M83294, AE000685, U28760, P27876, NC_000911, O08408, Q9Z520, AAL74288, AP002995/Q98ME7, Q8YHF5, Q92QA1, AB039932, U77930, AE005863, Q8XXP9, Q9HV51, AE001305. Scale bar shows substitutions per site.

thetic genes have been discarded and replaced by the uptake of amino acids from the host cytoplasm. Likewise, nucleoside monophosphates are imported directly from the cytosol. Thus, we can assume that the α-proteobacterial endosymbiont that contributed genes to the eukaryotic cell had a gene repertoire that was considerably larger than that seen in the modern, small rickettsias. Such coding sequence purges may have been tolerated because they are compensated by the uptake of compounds such as amino acids and nucleoside monophosphates directly from the cytosol. In more general terms, coding sequences are lost when neutralized by the organism's environment (Andersson & Kurland 1998; Berg & Kurland 2000). As the mitochondria resemble *Rickettsia* in that they inhabit the same intracellular environment, it seems probable that many of the same functions were lost in both genomic lineages. Examples of convergent losses are the glycolytic genes. Of course, it is also possible that the purging of some neutralized sequences had begun in the α-proteobacterial ancestor of both mitochondrial and *Rickettsia* lineages.

Likewise, the large genome sizes of free-living α-proteobacteria are also the result of evolutionary processes that have occurred subsequent to the divergence of the mitochondrial clade. In this case, the principal evolutionary thrust was novel sequence acquisition via duplications or transfers of genes required for their environmental adaptations (Kaneko *et al.* 2000; Goodner *et al.* 2001; Wood *et al.* 2001; Galibert *et al.* 2001). Thus, the gene complement of the ancestral endosymbiont was most probably smaller than what we see in the large genomes of modern, free-living α-proteobacteria.

Even the modern mitochondrial proteome is the result of both reductive and expansive processes. There are at least three identifiable tendencies in the evolution of the mitochondrial proteome (Karlberg *et al.* 2000). First, many ancestral mitochondrial genes have been simply purged from the mitochondrial genome and lost. Second, the coding sequences of a smaller number of mitochondrial proteins have been expelled from the mitochondrial genome and transferred to the nuclear genome. These nuclear genes are α-proteobacterial descendents. Third, the largest number of mitochondrial proteins are also imports that are coded by nuclear sequences. However, these have no bacterial or archaeal orthologues. Such novel 'eukaryotic' mitochondrial proteins have evidently evolved from the nuclear genome by duplication and functional divergence of sequences that predate the arrival of the endosymbiont ancestor of mitochondria (Karlberg *et al.* 2000).

The transformation of the ancestral α-proteobacterium into the modern mitochondrion was obviously not instantaneous. The loss of sequences from the ancestral genomes and the expansion of its proteome with novel contributions from the eukaryotic nuclear genome are incremental processes that are presumably spread over

millions of years. This has the consequence that any attempts to trace the phylogenetic backgrounds of the mitochondrial proteins with different organisms are not simple. In particular, the loss of proteins from the α-proteobacterial lineages and the introduction of novel eukaryotic proteins should not be synchronous in the ancestors to all modern lineages (Karlberg *et al.* 2000). Alternatively, we expect this lack of synchrony to be useful in tracing the detailed history of the organelle. Nevertheless, some mitochondrial genes involved in bioenergetic and basic information processes represent a relatively stable set of organellar genes that are nearly universally present in most α-proteobacterial genomes. These are very probably part of the ancestral coding complement. Such sequences are unambiguously descendants of the α-proteobacteria (Karlberg *et al.* 2000).

The origin of ATP-generating organelles without genomes, such as anaerobic mitochondria and hydrogenosomes, has been more elusive. These not only lack their own genetic sequences but their proteomes are on the whole very different from those of aerobic mitochondria. There are two possible scenarios for the evolution of these organelles. First, they may derive from a facultative anaerobic symbiont that initially functioned like the modern hydrogenosomes and were later transformed into mitochondria with the help of genes for aerobic respiration also present in the symbiont (Martin & Mueller 1998; Rotte *et al.* 2000). Alternatively, they may derive from aerobic organelles as an adaptation to anaerobic environments by modification of the primary aerobic organelle (Embley *et al.* 1997; Van der Giezen *et al.* 1997*a*,*b*; Andersson & Kurland 1999). Here, the anaerobic gene functions may have been acquired by the transfer of genes from bacteria other than the ancestral α-proteobacterial endosymbiont.

The data currently available support versions of the second scenario. Most importantly, a number of key enzymes, such as the fumarate reductase in anaerobic mitochondria, have been shown to be more closely related to the mitochondrial SDH than to bacterial forms of FRD (figure 3; Van Hellemond & Tielens 1994; Van Hellemond *et al.* 1995; Kita *et al.* 2002). Likewise, chaperonin 60 and ADP/ATP transporters of the anaerobic, hydrogenosome-bearing fungus *Neocallimastix* sp., cluster with high bootstrap support with their homologous mitochondrial proteins from aerobic yeasts and fungi (Voncken *et al.* 2002; Van der Giezen *et al.* 2002). Finally, the first description of a hydrogenosome with a genome revealed similarities between hydrogenosomal and mitochondrial rRNA sequences (Akhmanova *et al.* 1998; Van Hoek *et al.* 2000).

A particularly interesting case is the hydrogenase that seems to contain segments derived from bacterial genes combined with segments derived from mitochondrial genes (Akhmanova *et al.* 1998). Likewise, the PFO characteristic of hydrogenosomes seems to be derived by horizontal gene transfer from bacteria other than the α-proteobacteria (Horner *et al.* 1999). Taken together, all available data indicate that anaerobic organelles represent secondary adaptations of aerobic mitochondria to anaerobic environments (Bui *et al.* 1996; Biagini *et al.* 1997; Finlay & Fenchel 1989; Horner *et al.* 1999; Hackstein *et al.* 1999; Voncken *et al.* 2002; Van der Giezen *et al.* 2002). To our knowledge, there are no data suggesting that the origin of mitochondria was the result of selection for metabolic pathways other than aerobic respiration.

4. ORIGINS OF MITOCHONDRIA: WHY?

Judged by ubiquity of the organelle and the ability of eukaryotes to form large, complex multicellular organisms, the acquisition of mitochondria must in retrospect be viewed as a wonderful success. However, an examination of modern mitochondrial functions may not necessarily tell us what kind of relationship initiated the symbiotic event. Certainly, it was not based on the prospect of creating large multicellular organisms in a distant future.

The origins of the mitochondrial lineages are conventionally associated with a global disaster about two billion years ago. At that time atmospheric oxygen levels started to rise in the atmosphere, presumably from the accumulated activity of oceanic photosynthetic cyanobacteria (Holland 1994). Before this, the atmosphere of the Earth was not aggressively oxidative and consisted mainly of carbon dioxide along with reduced forms of sulphur, nitrogen as well as carbon with very low concentrations of oxygen (Rubey 1951; Holland 1994). The accumulation of substantial levels of oxygen in the atmosphere is likely to have provoked a major crisis for early life on Earth. Anaerobic life forms unable to protect themselves from the toxicity of oxygen or unable to find a suitable anaerobic microenvironment probably became extinct. Of course, one form of protection involves the reduction of local oxygen concentrations by metabolism. For example, the consumption of oxygen as an electron acceptor in the electron transport chain of a bacterium reduces the ambient oxygen tension in the immediate vicinity of that bacterium as well as producing 'energy' in the form of ATP.

A conventional view of the symbiotic relationship between the host and the endosymbiont that seeded the mitochondrial lineage is that ATP produced by the endosymbiont was exchanged for carbohydrates made available by the host cell. However, ATP import or export by free-living bacteria is unknown. Furthermore, as described in the following paragraphs, the mitochondrial ATP/ADP transport system is not bacterially derived but is instead a member of a large family of transport systems that are found in eukaryotes. Thus, the bacterial ancestor of the mitochondria could not easily enter into a symbiosis based on ATP exchange (Andersson *et al.* 1998; Karlberg *et al.* 2000). However, local depletion of the oxygen tension by the respiration of an aerobic bacterium could provide a basis for a symbiotic relationship with an anaerobic partner (Andersson & Kurland 1999; Kurland & Andersson 2000). If this scenario were relevant, anaerobic organelles with affinities to the mitochondria would be viewed as secondarily derived from the aerobic organelles.

Alternative evolutionary scenarios have also been described. The syntrophic theory indicates that the nucleated eukaryote was formed through symbiosis between methanogenic archaea and δ-proteobacteria (Moreria & Lopez-Garcia 1998). This suggestion rests on two observations. One is that the syntrophic symbioses between archaeal methanogens and bacteria have been observed in nature (Hanson & Hanson 1996). The other

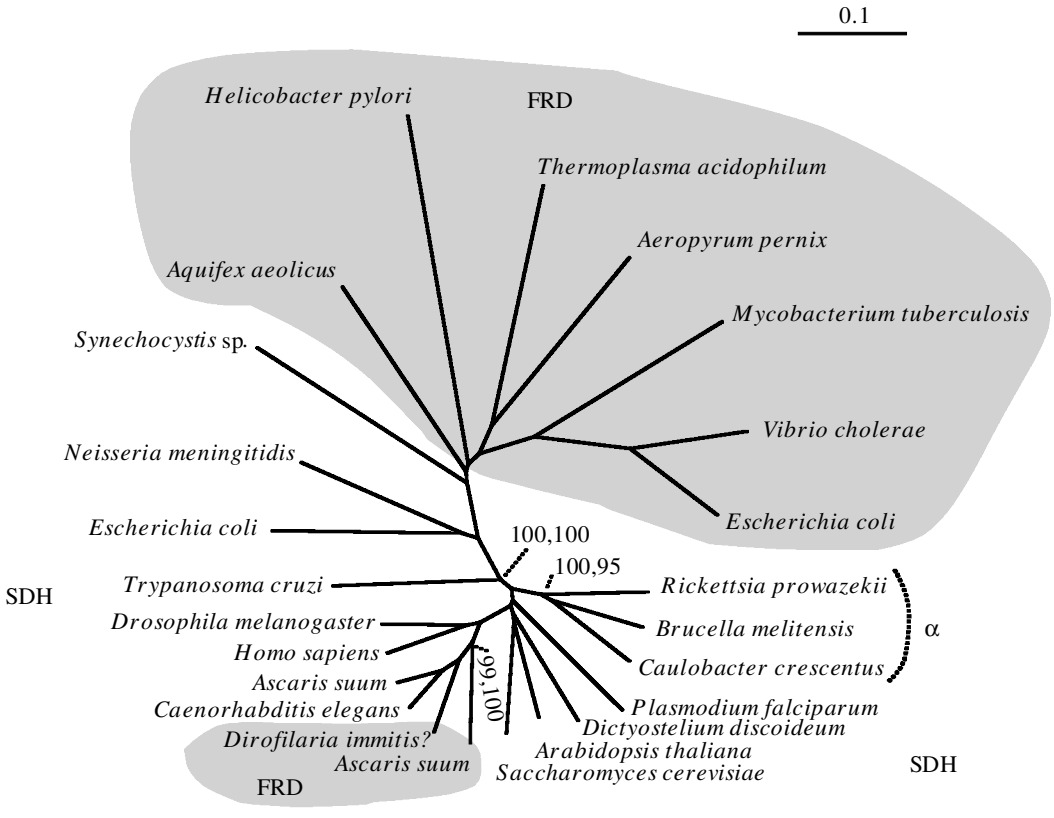

Figure 3. The combined ME tree for FRD and SDH based on amino-acid sequences. Note that the sequence from *Dirofilaria immitis* has not been explicitly annotated as a FRD. Abbreviations, methods, bootstrap values and other details are as described in the legend to figure 1. Accession numbers clockwise starting with *Aquifex aeolicus*: O66855, O06913, Q9HJG7, Q9YDG3, Q10760, Q9KNS5, P00363, AAA16097, NP_539079, NP_422321, BAA13119, AAF21045, O82663, Q00711, Q33862, AAB34901, Q09508, BAB84191, P31040, Q94523, BAA84681, P78282, Q9JUT3, P73479. Scale bar shows substitutions per site.

is that there are α-proteobacteria with the capability to oxidize methane under both aerobic and anaerobic conditions. In that case, mitochondria should have been acquired for their methanotrophic metabolism. Here, the capacity to utilize oxygen is secondary, or may even have evolved after the mitochondrial organelle had been established.

A similar hypothesis based on similar assumptions is the hydrogen hypothesis (Martin & Mueller 1998). The major difference between the two theories is that the hydrogen hypothesis uses α-proteobacteria as the bacterial partner, which circumvents the need for two endosymbiotic events to explain the origin of eukaryotes with mitochondria. In this case, it was the α-proteobacterium that contributed the molecular hydrogen needed by the archaeal methanogen. Martin & Mueller (1998) reject the possibility that the methane produced by the methanogen could be used by the α-proteobacteria because all contemporary methanotrophs are strictly aerobic, while contemporary methanogens are strict anaerobes.

In both hypotheses, it is suggested that the mitochondrial ancestor was selected because of anaerobic functions that are not found in contemporary mitochondria (Moreria & Lopez-Garcia 1998; Martin & Mueller 1998). Instead, it is assumed that the potential for aerobic metabolism was ancillary to the association and integration process of the endosymbiont. In addition, strong selection for gene transfers of other metabolic processes is required. In light of what is left of the original endosymbiont genome today such a process seems improbable. Without strong selective pressure, the respiratory functions of protomitochondria would most probably have been lost.

The hydrogen hypothesis explicitly predicts that modern eukaryotic genes encoding glycolysis are the descendants of the α-proteobacterial endosymbiont that had been transferred to the nuclear genome (Martin & Mueller 1998). Contradicting this clear prediction is the observation that the nuclear genes encoding glycolytic enzymes in eukaryotes do not branch specifically with the α-proteobacteria, indicating that they were not derived from the ancestral endosymbiont (figure 2) or with any other bacterial phylum (Canbäck *et al.* 2002). Alternatively, phylogenetic reconstructions for single gene sequences encoding α-proteobacterial glycolytic functions do cluster with strong bootstrap support (Canbäck *et al.* 2002) as expected from phylogenetic reconstructions based on rRNA gene sequences (Olsen *et al.* 1994). Such clustering shows that the phylogenetic signal for the glycolytic genes has, in most cases, not been distorted by an excessive frequency of gene transfers, nor by atypical rates or modes of sequence evolution. In other words, the data clearly contradict the expectations of the hydrogen hypothesis.

There is no evidence that any of the non-standard mitochondrial gene functions present in anaerobic mitochondria and hydrogenosomes are the primary descendants of the corresponding gene functions in the α-proteobacteria. By contrast, the available data indicate that these are secondary adaptations and have evolved from pre-existing

mitochondrial and nuclear gene functions present in the mitochondriate host cell (Van Hellemond & Tielens 1994; Van Hellemond *et al.* 1995; Akhmanova *et al.* 1998; Van Hoek *et al.* 2000; Voncken *et al.* 2002; Van der Giezen *et al.* 2002). Indeed, even eukaryotic cells without organelles often contain one or a few nuclear genes encoding functions that are of mitochondrial origin. Common examples are the *hsp*60 and *hsp*70 genes encoding 60 and 70 kDa heat-shock proteins (Clark & Roger 1995; Germot *et al.* 1996, 1997; Roger *et al.* 1996, 1998; Hirt *et al.* 1997; Peyretaillade *et al.* 1998*a*,*b*). Finally, recent phylogenetic reconstructions indicate that most eukaryotes without mitochondria, such as Microsporidia, are highly derived forms of other eukaryotes (Embley & Hirt 1998; Hirt *et al.* 1999). While these observations do not preclude the existence of eukaryotes that diverged prior to the acquisition of mitochondria, all the available data indicate that the anaerobic eukaryotes described so far are descendants of mitochondriate ancestors.

One possible, albeit speculative scenario for the acquisition of mitochondria arises from the capacity of an oxygen-respiring organism to reduce local oxygen concentrations (Andersson & Kurland 1999; Kurland & Andersson 2000). In this way, an aerobic α-proteobacteria might help an anaerobic, or less effective aerobic, eukaryote to survive in oxygen-rich environments. By accommodating such an aerobe in the endocellular compartment, the host would obtain continuous, efficient protection against toxic oxygen. Subsequent evolution of a large set of novel nuclear genes could then transform the endosymbiont into a more efficient organelle (Andersson & Kurland 1999; Karlberg *et al.* 2000; Kurland & Andersson 2000). These novelties include, for example, the ATP/ADP translocases that enabled the host cell to exploit the large amounts of ATP produced by the endosymbiont. We discuss this second stage of evolution in more detail in § 5.

5. TRANSFER AND RECRUITMENT OF NUCLEAR GENES

To account for the fact that only a small fraction of the proteins required for mitochondrial functions are encoded by the mitochodrial genome, it has been indicated that there was a massive transfer of genes from the original endosymbiont to the nuclear genome of the host (Gray *et al.* 1999). For example, the mitochondrial genome of the freshwater protozoan *Reclinomonas americana*, which is the mitochondrial genome with the largest gene content yet found, encodes only 67 proteins (Lang *et al.* 1997). This is not much even when compared with the smallest known genome of a free-living organism, that of *Mycoplasma genitalium*, which encodes approximately 470 proteins (Fraser *et al.* 1995). The low coding potential of the mitochondrial genomes can be explained by two mechanisms working together, gene loss and gene transfer to the nuclear genome. Thus, to understand the origin and evolution of the mitochondrial proteome, it is necessary to examine the complete set of mitochondrial proteins encoded by the eukaryotic cell. Here, we take into account those encoded by the mitochondrial genome as well as those encoded by the nuclear genome. The questions are as follows. How many genes have been lost from the ancestral endosymbiont's genome? How many have been transferred to the nuclear genome? How many have been recruited from the nuclear genome?

To answer these questions, we needed a eukaryote for which the mitochondrial proteome and the complete genome sequence were completely defined. The first such organism was the budding yeast *Saccharomyces cerevisiae*, with a 12 Mbp genome. Owing to its importance to the food industry and its small genome size, it is by far the most well characterized eukaryote to date. The Yeast Proteome Database (www.proteome.com) contains 6281 proteins, of which 4127 are characterized genetically or biochemically. Of these proteins, 2225 are known by homology to other proteins and 1902 are of unknown function (Hodges *et al.* 1999). More to the point, approximately 400 have been associated experimentally with mitochondrial functions. At present, this is the only well characterized mitochondrial proteome available, which makes the yeast mitochondrion a unique model system for studies of the origin of mitochondrial proteins.

Briefly, *ca.* 50% of the nucleus-encoded mitochondrial proteins have bacterial homologues, whereas the other 50% appear to be uniquely present in eukaryotes (Karlberg *et al.* 2000). Thus, the mitochondrial proteome has to a first approximation a dual origin, with some proteins of putative bacterial origin and others of putative eukaryotic origin. The two protein sets are not random subsets, but rather highly biased fractions of the different functional categories. Thus, genes with bacterial homologues are predominantly associated with categories such as translation, energy and small-molecule biosynthesis, whereas most of the membrane, transport and regulatory proteins appear to have no bacterial homologues. However, even within the functional groups for translation and bioenergetic processes, a fraction of the genes appear to be of eukaryotic origin.

Out of the *ca.* 200 nuclear yeast genes with bacterial homologues, only *ca.* 20% can with certainty be traced back to the α-proteobacteria, of which as many as 24 are present in the mitochondrial genome of *R. americana* (Karlberg *et al.* 2000). Half of the proteins traced with high support to the α-proteobacteria are involved in energy metabolism and the remaining half are ribosomal proteins. The transfer of genes to the nuclear genome is particularly well illustrated with a set of ten ribosomal protein genes that are nuclearly located in yeast but mitochondrially located in *R. americana*. Here, phylogenetic analyses reveal a very close relationship between the mitochondrial ribosomal proteins from the various eukaryotes and their α-proteobacerial homologues, whether the genes are located in the mitochondrial genome or in the nuclear genome (Karlberg *et al.* 2000). Similar results have been inferred from phylogenetic reconstructions based on components of the SDH complex (figure 3) and subunits of the ATP synthase complex.

The data motivate three conclusions. First, many mitochondrial activities, such as those of the translation and respiratory systems, are derived from the ancestral α-proteobacteria. Second, genes from the α-proteobacterial ancestor have been transferred from the mitochondrial to the nuclear genome. Third, the presence of these α-proteobacterial genes in the *R. americana* mitochondrial genome indicates that transfer to the nucleus took place in

the other lineages subsequent to their divergence from *R. americana*.

Other genes coding for mitochondrial functions also have a strong phylogenetic resemblance with their α-proteobacterial homologues, even though they are nuclearly encoded in all currently examined species. Most of these encode proteins of the bioenergetic and translation complexes, including, for example, the PDH complex. As many of these crucial gene functions evolve under strong functional constraints, they tend to be more slowly evolving than genes of less functional importance. Thus, it is possible that the phylogenetic signal for these proteins has been less distorted by time than the signal for other genes. In total, there are roughly 150 yeast nuclear genes that may be α-proteobacterial descendants, but the phylogenetic signal for these is too weak to make a definitive statement about their origin. Another 40 are unambiguously α-proteobacterial descendants. In no single case have we observed a strong phylogenetic signal between nucleus-encoded mitochondrial proteins and proteins from a phylum other than the α-proteobacteria. However, a mitochondrial-like aconitase has recently been identified in *Bacteriodes fragilis*, indicative of a paralogous gene family for this enzyme or gene transfer between mitochondria and the Cytophaga–Flavobacterium–Bacteriodes group (Baughn & Malamy 2002).

Another set of about 200 genes tends to be exclusively of eukaryotic origin; these include genes involved in the regulation of gene expression, mRNA stability and splicing. Likewise, many proteins of the mitochondrial membrane and transport complexes seem to have been recruited from the nuclear genome for mitochondrial functions at some time after the ancestral endosymbiont established itself in the eukaryotic cell. In addition, many species-specific protein subunits have been added to protein complexes formed around a set of ubiquitous proteins that descended from the endosymbiont. Striking examples of these latter proteins are found in the yeast mitochondrial ribosome. Thus, *ca*. 60% of the ribosomal proteins have recognizable bacterial homologues, some of which form a close relationship with α-proteobacteria. Our interpretation is that all or most of the remaining ribosomal proteins (40%) with no bacterial homologues were recruited subsequent to the appearance of the symbiont/mitochondrion. Indeed, the yeast mitochondrial ribosome contains roughly 10 more proteins than the bacterial ribosome, so additional proteins were presumably added after the integration of the endosymbiont into the host cell. Perhaps this was to compensate for the loss of nucleotides in the small and large rRNAs encoded by the mitochondrial genomes (T. Obrien, personal communication).

A mixed evolutionary origin is observed also for many other protein complexes in the mitochondrion. For example, the cytochrome *bc*1 complex consists of 10 proteins, of which the three core proteins (cytochrome *b*, cytochrome *c* and the Rieske iron–sulphur protein) are clearly α-proteobacterial descendants in phylogenetic reconstructions. However, the remaining seven proteins in this complex have no bacterial homologues. Indeed, some of the 'extra' proteins seem to have been acquired quite recently. For example, of the seven additional proteins in the cytochrome *bc*1 complex, three are virtually unique to *S. cerevisiae*; only one is found among other fungi (Karlberg *et al.* 2000). Likewise, of the 400 mitochondrial proteins as many as 64 were not identified in any other eukaryotes, and among the 125 with external homologues, many are found only among fungi.

Indeed, the eukaryotic cell seems to be highly robust in its ability to utilize existing or modified gene functions for novel purposes. For example, the yeast nuclear genome contains three copies of a gene coding for malate dehydrogenase, one of which is utilized in the mitochondrion, another in the cytosol and the third in the peroxisome (Karlberg *et al.* 2000). These appear to have originated by recent gene duplication events in the yeast genome. Again, three aminoacyl-tRNA synthetases are encoded by similar, multiple copies in the yeast nuclear genome. These include the mitochondrial and cytoplasmic arginyl-tRNA synthetases that cluster together in phylogenetic reconstructions. There are another four instances in which a single gene in yeast is responsible for producing aminoacyl-tRNA synthetases that are used in both the cytosol and the mitochondrion.

The mechanism of the transfer of coding sequences from mitochondrial to nuclear genomes is no mystery. For example, it has been shown that gene migration in plants has occurred via RNA intermediates that have been converted into DNA prior to insertion into the nuclear genome (Nugent & Palmer 1991; Covello & Gray 1992; Adams *et al.* 1999). The process of gene transfer has also been studied in the laboratory using yeast as a model system (Thorsness & Fox 1993). The rates to the nuclei were at least 100 000 times greater than rates in the opposite direction. Thorsness & Fox (1990) indicated that phagolysosomes consumed the mitochondria and released DNA fragments from the organelle that could transform the nuclear genome. It was demonstrated by modelling this process that a systematic transfer of the coding sequences from the genome of the endosymbiont or the mitochondria to nuclei was inevitable for all coding sequences that are not selected to remain in the organelle for functional reasons (Berg & Kurland 2000).

Given that it should, in principle, be easy to displace mitochondrial gene functions with homologous proteins that are present in the cytosol or vice versa, it is surprising that so few replacements seem to have taken place. Curiously, many of the shared gene functions in the mitochondrion and the cytoplasm, such as amino acylation, are even today carried out by genes of different origins even though two or more gene copies have coexisted in the nuclear genome for hundreds of million of years (Karlberg *et al.* 2000).

Thus, if the ancestral endosymbiont contained one or a few thousand genes, the conclusion is that a majority was lost and that only a small fraction, at most a few hundred genes, were transferred and selectively maintained in the nuclear genome of the host. Another few hundred genes seem to have been recruited for mitochondrial functions by modification of existing nuclear functions or evolved *de novo* inside the nuclear genome. Thus, the host cell in effect took complete control over the endosymbiont.

6. PROTEIN IMPORT AND PHYLOGENY

Nucleus-encoded mitochondrial proteins are translated in the cytosol and subsequently imported into the mito-

chondrion. The machinery required for the import of the mitochondrial proteins synthesized in the cytosol is complicated and not yet fully understood (Baker & Schatz 1991; Schatz 1996; Schatz & Dobberstein 1996; Neupert 1997). However, several protein complexes involved in this process have been identified. The TOM protein complex transfers the proteins across the outer membrane, the TIM23 protein complex imports them into the matrix, the TIM22 complex integrates proteins destined for the inner membrane and the two complexes TIM9/10 and TIM8/13 guide the precursor proteins from TOM to TIM22 (Koehler 2000). None of these complexes have recognizable bacterial homologues, indicating that they originated subsequent to the integration of the endosymbiont into the eukaryotic cell (Karlberg *et al.* 2000).

Another subset of nuclearly encoded mitochondrial proteins is translated on polysomes that are directly attached to the mitochondrial membrane and co-translationally imported into the mitochondrion. Despite extensive research, it is still an open question why some mitochondrial proteins are translated on polysomes that are attached to the mitochondrion whereas others are translated on polysomes that are free in the cytosol. To resolve these issues, the identity of genes translated on the two types of polysomes were analysed using yeast as the model system (Marc *et al.* 2002). To this end, cytosolic polysomes were separated from polysomes attached to mitochondria and mRNA was isolated from the two different fractions. Hybridization experiments employing microarrays with genes coding for mitochondrial proteins were used to score each gene according to whether it was solely translated on cytosolic polysomes (score = 1) or whether its product was exclusively found on polysomes attached to mitochondrion (score = 0). These scores were then used to search for a correlation between the different modes of translation/import and the physical parameters associated with each of the experimental yeast mitochondrial proteins.

The only correlation found was that between the microarray scores and the predicted origin of the yeast mitochondrial genes. Thus, a general trend was observed in which genes classified as eukaryotic (Karlberg *et al.* 2000) were primarily translated on polysomes free in the cytosol (Marc *et al.* 2002), while those classified as bacterial (Karlberg *et al.* 2000) were mainly translated on polysomes attached to the mitochondrion (Marc *et al.* 2002). Given this remarkable correlation, it is tempting to speculate that genes originally transferred to the nuclear genome initially utilized a bacterial co-translational secretion system for import from the cytosol into the mitochondrion. Later on alternative mitochondrial import system such as the TIM/TOM system evolved from nuclear genes of the host cell. After this novelty was introduced, only an appropriate addressing signal was required to direct a protein from the cytosol to the mitochondrion.

Once such a system had been invented it became possible to decorate the mitochondrial core complexes with a variety of new proteins. Some of these new proteins were modifications of gene products already present in the cytosol while others may have evolved *de novo*. With the development of such a general import system, the foundation was laid for a continual addition of components and further fine tuning of the functional interaction between the mitochondrion and its host.

7. THE ATP/ADP TRANSPORTERS

A fascinating import/export system is used by the mitochondrion to mediate the exchange of ATP for ADP across the organelle's membrane. There is now convincing evidence that these transporters originated by duplication and divergence from a more general type of eukaryotic phosphate transporters (Kuan & Saier 1993; Walker & Runswick 1993). They have since evolved vertically in the conventional mode. The recent identification of mitochondrial-type ADP/ATP transporters in the hydrogenosomes of the anaerobic fungus *Neocallimastix* sp., confirm that these organelles are secondarily derived from aerobic mitochondria (Van der Giezen *et al.* 2002; Voncken *et al.* 2002). Accordingly, these genes provide an example of how novel mitochondrial functions may have been recruited from pre-existing sequences in the nuclear genome. That mitochondrial ATP/ADP transporters were not acquired from the endosymbiont was to be expected since free-living bacteria are in fact not capable of transporting ATP and ADP across their membranes.

Nevertheless, phylogenetically unrelated ATP/ADP translocases have been identified in obligate intracellular parasites, such as *Rickettsia* (Winkler 1976; Williamson *et al.* 1989; Andersson 1998; Andersson *et al.* 1998) as well as *Chlamydia* (Hatch *et al.* 1982; Stephens *et al.* 1998; Kalman *et al.* 1999). Similar ATP/ADP transport systems have also been identified in the inner membrane of plastids (Heldt 1996; Pozueta-Romero *et al.* 1991; Neuhaus *et al.* 1993; Schunemann *et al.* 1993; Kampfenkel *et al.* 1995; Möhlmann *et al.* 1998) and more recently in the eukaryotic parasite *Encephalitozoon cuniculi* (Katinka *et al.* 2001). Both the bacterial and plastid ATP/ADP translocases seem to import ATP from the host cell cytoplasm in exchange for ADP, i.e. they seem to function normally with a polarity that is opposite that of the mitochondrial translocases, which are ATP exporters.

The reflexive explanation for the presence of related proteins in phylogenetically unrelated organisms is gene transfer. Indeed, it has been suggested that the presence of ATP/ADP transporters in *Rickettsia, Chlamydia* and plastids was associated with two horizontal gene transfer events (Wolf *et al.* 1999). What is unusual here is that the phylogeny of the species involved is not significantly deeper than the phylogeny of the transferred gene. Thus, the divergences of the plastid/parasite type of transporters present in unrelated species appear to be as deep as the divergences of the mitochondrial ATP/ADP transporters that branch in a species-specific manner (Amiri *et al.* 2002). This indicates that both types of transporters originated at an early stage of evolution, possibly shortly after the acquisition of the pre-mitochondrion. However, it cannot be excluded that the rate of sequence evolution for parasite/plastid transporters was extremely high at an early stage of their evolution.

While it is easy to understand that there was a requirement for a transporter protein shuffling ATP into the cytoplasm, the evolutionary pressure that triggered the development of the second ATP/ADP transporter is less clear. It is conceivable that the gene encoding the plastid/parasite type of ATP/ADP transporter also originated in the early mitochondriate cell to support the pumping of ATP from the host into the ancestral endo-

symbiont. Thus, the ancestral endosymbiont may have had a parasitic mode in which it functioned like *Rickettsia* as an energy parasite.

8. MITOCHONDRIA AND MULTICELLULARITY

It can be assumed that the early mitochondriate cell was capable of both fermentation (possibly mediated only by enzymes encoded by the host genome) and aerobic respiration (possibly mediated only by enzymes encoded by the endosymbiont genome). These two processes may initially have been separated by the membrane structures of the interacting cells. As long as the bioenergetic systems of the host and the endosymbiont were structurally separated, they would have competed internally for the uptake of substrate. So, under which conditions would the fermentation system of the host have been favoured over the oxidative phosphorylation system of the endosymbiont?

The concentration of oxygen is an obvious factor that would have favoured the aerobic respiration system of the endosymbiont. Pfeiffer *et al.* (2001) have also pointed out that under conditions of low rates of resource influx, aerobic respiration should have been the preferred pathway for the early mitochondriate cell. Indeed, respiration is the preferred pathway in yeast when glucose is limiting (Pronk *et al.* 1996). Under these conditions, carbon is shunted to the mitochondrial tricarboxylic cycle, thereby increasing electron transport and respiration (Pronk *et al.* 1996) as well as lifespan (Lin *et al.* 2002).

However, under conditions of high substrate concentrations glucose would be consumed at a higher rate by the fermentation system of the early mitochondriate cell, making it competitive with other fermenting cells (Pfeiffer *et al.* 2001). If such conditions were to exist for prolonged periods in anaerobic environments, aerobic respiration would simply disappear due to mutational attrition. Indeed, we know that there are anaerobes in which the aerobic respiration system of the mitochondrion has been lost or converted into other metabolic systems that are more suitable for anaerobic conditions.

Simulations have shown that a cell population using a pathway with low yield but high rate will invade and gradually replace a cell population using a pathway with high yield and low rate (Pfeiffer *et al.* 2001). Thus, if cells are in direct competition for a shared resource, evolution may work to select the less efficient pathway (Pfeiffer *et al.* 2001).

By contrast, at low cell diffusion rates, respirers because of their greater efficiencies would outcompete fermentors. Thus, it would be preferable to avoid direct competition by controlling the influx of substrate to the different partners. However, this requires a spatially structured environment in which the diffusion rates of the resource can be controlled. Accordingly, it has been suggested that multicellularity evolved as a strategy to facilitate an efficient and collaborative use of the available substrate by aerobic respiration (Pfeiffer *et al.* 2001).

A prerequisite for such a cooperative strategy is that ATP can be shared between the endosymbiont and its host. Hence, there must have been a strong selection for the evolution of ATP/ADP transport systems so as to obtain high yields of ATP per mole of glucose consumed. The presumed advantage of high ATP yields would also explain the presence of undifferentiated multicellular organisms. Thus, most other evolutionary theories that explain multicellularity by associating it with cell differentiation and specialization (Pfeiffer *et al.* 2001). The theory of Pfeiffer seems to us to provide a more basic account of the origins of multicellularity, a simple form of which is even observed among some bacteria. Thus, the acquisition of mitochondria may not only have formed the basis for the eukaryotic cell as we know it today, but also stimulated interactions among cells to form multicellular organisms.

9. CONCLUSIONS

Contemporary aerobic mitochondria represent a monophyletic clade for which all the available evidence identifies the α-proteobacteria as the sole ancestors. Anaerobic mitochondria, however, appear to have originated secondarily by modifications of previous mitochondrial gene functions. A growing body of evidence indicates that known amitochondriate eukaryotes may once have had a mitochondrion. If so, the origin of mitochondria represents the most successful of all symbiotic relationships described to date. The question is why?

A key notion is that aerobic respiration in eukaryotes evolved simultaneously with or following the rise of oxygen in the atmosphere. Accordingly, it is tempting to speculate that this symbiotic relationship enabled the anaerobic or 'weakly' aerobic eukaryotes to survive conditions under which they would otherwise have gone extinct. The subsequent evolution of ATP/ADP transporter systems by modification of transporters already present in the host made it possible for the mitochondriate cell to fully exploit the benefits of the high ATP yields obtained by aerobic respiration (Andersson & Kurland 1999). The evolution of ATP-transport systems (Karlberg *et al.* 2000) may even have been critical for the transition from single cells to early, undifferentiated forms of multicellular eukaryotes (Pfeiffer *et al.* 2001).

During the millions of years that have passed since this symbiotic relationship was first established, almost all of the original gene complement has been lost from the endosymbiont/mitochondrial genome. A few hundred genes were transferred into the nuclear genome and another few hundred novel eukaryotic sequences were eventually recruited from the nuclear genome for service in the organelle. Among the latter, we find those coding for novel mitochondrial import systems that enabled proteins freely available in the cytosol to be retargeted to the mitochondrion. It is clear that even in recent times, gene transfers and protein re-targeting from one compartment to another has taken place in the eukaryotic cell. Understanding the fate of the endosymbiont genes that were transferred to the nucleus of the anerobic partner is important if we are to understand the evolutionary dynamics that have shaped the evolution of the eukaryotic cell. However, the outstanding unanswered question is the identity of the cell that served as the host for the endosymbiont that eventually evolved into the mitochondrion.

This work is supported by the Swedish Natural Research Council, the Foundation for Strategic Research, the European Union and the Knut and Alice Wallenberg Foundation.

REFERENCES

Adams, K. L., Song, K., Roessler, P. G., Nugent, J. M., Doyle, J. L., Doyle, J. J. & Palmer, J. D. 1999 Intracellular gene transfer in action: dual transcription and multiple silencings of nuclear and mitochondrial *cox2* genes in legumes. *Proc. Natl Acad. Sci. USA* **96**, 13 863–13 868.

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.

Amiri, H., Karlberg, O. & Andersson, S. G. E. 2002 Deep origin of plastid/parasite ATP/ADP translocases. *J. Mol. Evol*. (In the press.)

Andersson, J. O. & Andersson, S. G. E. 1999*a* Insights into the evolutionary process of genome degradation. *Curr. Opin. Genet. Dev.* **9**, 664–671.

Andersson, J. O. & Andersson, S. G. E. 1999*b* Genome degradation is an ongoing process in *Rickettsia*. *Mol. Biol. Evol.* **16**, 1178–1191.

Andersson, J. O. & Andersson, S. G. 2001 Pseudogenes, junk DNA, and the dynamics of Rickettsia genomes. *Mol. Biol. Evol.* **18**, 829–839.

Andersson, S. G. E. 1998 Bioenergetics of the obligate intracellular parasite *Rickettsia prowazekii*. *Biochim. Biophys. Acta* **1365**, 105–111.

Andersson, S. G. E. & Kurland, C. G. 1998 Reductive evolution of resident genomes. *Trends Microbiol.* **6**, 263–268.

Andersson, S. G. E. & Kurland, C. G. 1999 Origin of mitochondria and hydrogenosomes. *Curr. Opin. Microbiol.* **2**, 535–541.

Andersson, S. G. E. (and 10 others) 1998 The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**, 133–140.

Baker, K. P. & Schatz, G. 1991 Mitochondrial proteins essential for viability mediate protein transport into yeast mitochondria. *Nature* **349**, 205–208.

Baughn, A. D. & Malamy, M. H. 2002 A mitochondrial-like aconitase in the bacterium *Bacteriodes fragilis*: implications for the evolution of the mitochondrial Krebs cycle. *Proc. Natl Acad. Sci. USA* **99**, 4662–4667.

Berg, O. & Kurland, C. G. 2000 Why mitochondrial genes are most often found in nuclei. *Mol. Biol. Evol.* **17**, 951–961.

Biagini, G. A., Hayes, A. J., Suller, M. T. E., Winters, C., Finlay, B. J. & Lloyd, D. 1997 Hydrogenosomes of *Metopus contortus* physiologically resemble mitocondria. *Microbiology* **143**, 1623–1629.

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.

Canbäck, B., Andersson, S. G. E. & Kurland, C. G. 2002 The global phylogeny of glycolytic enzymes. *Proc. Natl Acad. Sci. USA* **99**, 6097–6102.

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.

Covello, P. S. & Gray, M. W. 1992 Silent mitochondrial and active nuclear genes for subunit 2 of cytochrome c oxidase (cox2) in soybean: evidence for RNA-mediated gene transfer. *EMBO J.* **11**, 3815–3820.

DelVecchio, V. G. (and 24 others) 2001 The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. *Proc. Natl Acad. Sci. USA* **8**, 443–448.

Embley, T. M. & Hirt, R. P. 1998 Early branching eukaryotes? *Curr. Opin. Gen. Dev.* **8**, 624–629.

Embley, T. M., Finlay, B. J., Dyal, P. L., Hirt, R. P., Wilkinson, M. & Williams, A. G. 1995 Multiple origin of anaerobic ciliates with hydrogenosomes within the radiation of aerobic ciliates. *Proc. R. Soc. Lond.* B **262**, 87–93.

Embley, T. M., Horner, D. A. & Hirt, R. P. 1997 Anaerobic eukaryote evolution: hydrogenosomes as biochemically modified mitochondria. *Trends Ecol. Evol.* **12**, 437–441.

Finlay, B. J. & Fenchel, T. 1989 Hydrogenosomes in anaerobic protozoa resemble mitochondria. *FEMS Microbiol. Lett.* **65**, 311–314.

Finlay, B. J., Span, A. S. W. & Harman, J. M. P. 1983 Nitrate respiration in primitive eukaryotes. *Nature* **303**, 333–336.

Foury, F., Roganti, T., Lecrenier, T. & Purnelle, B. 1998 The complete sequence of the mitochondrial genome of *Saccharomyces*. *FEBS Lett.* **440**, 325–331.

Fraser, C. M. (and 29 others) 1995 The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397–403.

Galibert, F. (and 55 others) 2001 The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**, 668–672.

Germot, A., Phillipe, H. & Le Guyader, H. 1996 Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* indicates a very early mitochondrial endosymbiosis in eukaryotes. *Proc. Natl Acad. Sci. USA* **93**, 14 614–14 618.

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.

Goodner, B. (and 30 others) 2001 Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**, 2323–2328.

Gray, M. W. 1992 The endosymbiont hypothesis revisited. *Int. Rev. Cytol.* **141**, 233–357.

Gray, M. W., Cedergren, R., Abel, Y. & Sankoff, D. 1989 On the evolutionary origin of plant mitochondrion and its genome. *Proc. Natl Acad. Sci. USA* **86**, 2267–2271.

Gray, M. W. & Spencer, D. F. 1996 In *Evolution of microbial life* (ed. D. M. Roberts, P. M. Sharp, G. Alderson & M. Collins), pp. 109–126. Cambridge University Press.

Gray, M. W. (and 14 others) 1998 Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res.* **26**, 865–878.

Gray, M. W., Burger, G. & Lang, B. F. 1999 Mitochondrial evolution. *Nature* **283**, 1476–1481.

Hackstein, J. H., Akhmanova, A., Boxma, B., Harhangi, H. R. & Voncken, F. G. 1999 Hydrogenosomes: eukaryotic adaptations to anaerobic environments. *Trends Microbiol.* **7**, 441–447.

Hanson, M. R. & Folkerts, O. 1992 Structure and function of the higher plant mitochondrial genome. In *Mitochondrial genomes* (ed. D. R. Wolstenholme & K. W. Jeon), pp. 129–173. San Diego, CA: Academic.

Hanson, R. S. & Hanson, T. E. 1996 Methanotrophic bacteria. *Microbiol. Rev.* **60**, 439–471.

Hatch, T. P., Al-Houssaine, E. & Silvermann, J. A. 1982 Adenine nucleotide and lysine transport in *Chlamydia psittaci*. *J. Bacteriol.* **150**, 662–670.

Heldt, H. W. 1996 Adenine nucleotide translocation in spinach chloroplasts. *FEBS Lett.* **5**, 11–14.

Hirt, R. P., Healy, B., Vossbrink, C. R., Canning, E. U. & Embley, T. M. 1997 A mitochondrial HSP70 orthologue in *Vairimorpha necatrix*: molecular evidence that microsporidia once contained mitochondria. *Curr. Biol.* **7**, 996–998.

Hirt, R. P., Logsdon, J. M., Healy, B., Dorby, 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.

Hodges, P. E., McKee, A. H. Z., Davis, B. P., Pasyne, W. E. B. & Garrels, J. I. 1999 Yeast Protein Database (YPD): a model for the organization and presentation of genome-wide functional data. *Nucleic Acids Res.* **27**, 69–73.

Holland, H. D. 1994 Early proterozoic atmospheric change. In *Early life on earth* (ed. S. Bengtson), pp. 237–244. New York: Columbia University Press.

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.

Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, L., Hyman, R. W., Olinger, L., Grimwood, L., Davis, R. W. & Stephens, R. S. 1999 Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. *Nature Genet.* **21**, 385–389.

Kampfenkel, K., Möhlman, T., Batz, O., Van Montagu, M., Inze, D. & Neuhaus, H. E. 1995 Molecular characterization of an *Arabidopsis thaliana* cDNA encoding a novel putative adenylate translocator of higher plants. *FEBS Lett.* **374**, 351–355.

Kaneko T. (and 12 others) 2000 Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res.* **7**, 331–338.

Karlberg, O., Canbäck, B., Kurland, C. G. & Andersson, S. G. E. 2000 The dual origin of the yeast mitochondrial proteome. *Yeast Comp. Funct. Genom.* **17**, 170–187.

Katinka, M. D. (and 16 others) 2001 Genome sequence and gene compaction of the eukaryotic parasite *Encephalitozoon cuniculi*. *Nature* **414**, 450–453.

Kita, K., Hirawake, H., Miyadera, H., Amino, H. & Takeo, S. 2002 Role of complex II in anaerobic respiration of the parasitic mitochondria from *Ascaris suum* and *Plasmodium falciparum*. *Biochim. Biophys. Acta* **1553**, 123–139.

Kobayashi, M., Matsuo, Y., Takimoto, A., Suzuki, S., Maruo, F. & Shoun, H. 1996 Denitrification, a novel type of respiratory metabolism in fungal mitochondrion. *J. Biol. Chem.* **271**, 16 263–16 267.

Koehler, C. 2000 Protein translocation pathways of the mitochondrion. *FEBS Lett.* **476**, 27–31.

Kuan, J. & Saier, M. H. 1993 The mitochondrial carrier family of transportproteins: structural, functional, and evolutionary relationships. *Crit. Rev. Biochem. Mol. Biol.* **28**, 209–233.

Kurland, C. G. & Andersson, S. G. E. 2000 Origin and evolution of the mitochondrial proteome. *Microbiol. Mol. Biol. Rev.* **64**, 786–820.

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.

Lin, S.-J., Kaeberlein, M., Andalis, A. A., Sturz, L. A., Defossez, P.-A., Culotta, V. C., Fink, G. R. & Guarente, L. 2002 Calorie restriction extends *Saccharomyces cerevisiae* lifespan by increasing respiration. *Nature* **418**, 344–348.

Marc, P., Mageot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M. & Jacq, C. 2002 Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* **3**, 159–164.

Margulis, L. 1970 *Origin of eukaryotic cells*. New Haven, CT: Yale University Press.

Margulis, L. 1981 *Symbiosis in cell evolution*. San Francisco, CA: Freeman.

Martin, W. & Mueller, M. 1998 The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41.

Mereschowsky, C. 1910 *Biol. Zbl.* **30**, 278–367.

Moreira, D. & Lopez-Garcia, P. 1998 Symbiosis between methanogenic archaea and δ-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J. Mol. Evol.* **47**, 517–530.

Mueller, M. 1993 The hydrogenosome. *J. Gen. Microbiol.* **139**, 2879–2889.

Möhlman, T., Tjaden, J., Schwöppe, C., Winkler, H. H., Kampfenkel, K. & Neuhaus, H. E. 1998 Occurrence of two plastidic ADP/ATP transporters in *Arabidopsis thaliana*. *Eur. J. Biochem.* **252**, 353–359.

Neuhaus, H. E., Henrichs, G. & Scheibe, R. 1993 Characterization of glucose-6-phosphate incorporation into starch by isolated intact cauliflower-bud plastids. *Plant Physiol.* **101**, 573–578.

Neupert, W. 1997 Protein import into mitochondria. *A. Rev. Biochem.* **66**, 683–717.

Nierman, W. C (and 37 others) 2001 Complete genome sequence of *Caulobacter crescentus*. *Proc. Natl Acad. Sci. USA* **98**, 4136–4141.

Nugent, J. M. & Palmer, J. D. 1991 RNA-mediated transfer of the gene coxII from the mitochondrion to the nucleus during flowering plant evolution. *Cell* **66**, 473–481.

Ogata, H. (and 10 others) 2001 Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* **293**, 2093–2098.

Olsen, G. J., Woese, C. R. & Overbeek, R. 1994 The winds of evolutionary change: breathing new life into microbiology. *J. Bacteriol.* **176**, 1–6.

Peyretaillade, E., Biderre, C., Peyret, P., Duffieux, F., Metenier, G., Gouy, M., Michot, B. & Vivares, C. 1998*a* Microsporidian *Encephalitozoon cuniculi*, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. *Nucleic Acids Res.* **26**, 3513–3520.

Peyretaillade, E., Broussolle, V., Peyret, P., Metenier, G., Gouy, M. & Vivares, C. P. 1998*b* Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin. *Mol. Biol. Evol.* **15**, 683–689.

Pfeiffer, T., Schuster, S. & Bonhoeffer, S. 2001 Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292**, 504–507.

Pozueta-Romero, J., Frehner, M., Viale, A. M. & Akazawa, T. 1991 Direct transport of ADP-glucose by an adenylate translocator is linked to startch biosynthesis in amyloplasts. *Proc. Natl Acad. Sci. USA* **88**, 5769–5773.

Pronk, J. T., Yde Steensma, J. & Van Dijken, J. P. 1996 Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast* **12**, 1607–1633.

Roger, A. J., Clark, C. G. & Doolittle, W. F. 1996 A possible mitochondrial gene in the early-branching amitochondrial 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.

Rotte, C., Henze, K., Muller, M. & Martin, W. 2000 Origins of hydrogenosomes and mitochondria. *Curr. Opin. Microbiol.* **3**, 481–486.

Rubey, W. W. 1951 Geological history of seawater: an attempt to state the problem. *Geol. Soc. Am. Bull.* **62**, 1111–1148.

Schatz, G. 1996 The protein import system of mitochondria. *J. Biol. Chem.* **271**, 31 763–31 766.

Schatz, G. & Dobberstein, B. 1996 Commom principles of protein translation across membranes. *Science* **271**, 1519–1526.

Schunemann, D., Borchert, S., Flugge, U. I. & Heldt, H. W. 1993 ADP/ATP translocator from pea root plastids. Comparison with translocators from spinach chloroplasts and pea leaf mitochondria. *Plant Physiol.* **103**, 131–137.

Sichertz-Ponten, T., Kurland, C. G. & Andersson, S. G. E. 1998 A phylogenetic analysis of the cytochrome *b* and cytochrome *c* oxidase I genes supports an origin of mitochondria from within the Rickettsiaceae. *Biochim. Biophys. Acta* **1365**, 545–551.

Stephens, R. S. (and 10 others) 1998 Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**, 754–759.
Stryer, L. 1995 *Biochemistry*. New York: Freeman.
Stucki, J. W. 1980 The optimal efficiency and the economic degrees of coupling of oxidative phosphorylation. *Eur. J. Biochem.* **109**, 269–283.
Takaya, N., Suzuki, M. S., Kuwazaki, S., Shoun, H., Maruo, F., Yamaguchi, M. & Takeo, K. 1999 Cytochrome p450nor, a novel class of mitochondrial cytochrome P450 involved in nitrate respiration in the fungus *Fusarium oxysporum*. *Arch. Biochem. Biophys.* **372**, 340–346.
Thorsness, P. E. & Fox, T. D. 1990 Escape of DNA from mitochondria to nucleus in *Saccharomyces cerevisiae*. *Nature* **346**, 376–379.
Thorsness, P. E. & Fox, T. D. 1993 Nuclear mutations in *Saccharomyces cerevisiae* that affect the escape of DNA from mitochondria to the nucleus. *Genetics* **134**, 21–28.
Tielens, A. G. M. 1994 Energy generation in parasitic helminths. *Parasitol. Today* **10**, 346–352.
Tielens, A. G. M. & Van Hellemond, J. J. 1998 The electron transport chain in anaerobically functioning eukaryotes. *Biochim. Biophys Acta* **1365**, 71–78.
Unseld, M., Marienfeld, J. R., Brandt, P. & Brennicke, A. 1997 The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366 924 nucleotides. *Nature Genet.* **15**, 57–61.
Van der Giezen, M., Sjollema, K. A., Artz, R. R., Alkema, W. & Prins, R. A. 1997*a* Hydrogenosomes in the anaerobic fungus *Neocallimastix frontalis* have a double membrane but lack an associated organelle genome. *FEBS Lett.* **408**, 147–150.
Van der Giezen, M., Rechinger, K. B., Svendsen, I., Durand, R., Hirt, R. P., Embley, T. M. & Prins, R. A. 1997*b* A mitochondrial-like targeting signal on the hydrogenosomal malic enzyme from the anaerobic fungus *Neocallimastix frontalis*: support for the hypothesis that hydrogenosomes are modified mitochondria. *Mol. Microbiol.* **23**, 11–21.
Van der Giezen, M., Slotboom, D. J., Horner, D. S., Dyal, P. L., Harding, M., Xue, G. P., Embley, T. M. & Kunji, E. R. 2002 Conserved properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles. *EMBO J* **21**, 572–579.
Van Hellemond, J. J. & Tielens, A. G. M. 1994 Expression and functional properties of fumarate reductase. *Biochem. J.* **304**, 321–331.
Van Hellemond, J. J., Klockierwics, M., Gaasenbeck, C. P., Roos, M. H. & Tielens, A. G. 1995 Rhodoquinone and complex II of the electron transport chain in anaerobically functioning eukaryotes. *J. Biol. Chem.* **270**, 31 065–31 070.
Van Hoek, A. H. A. M., Akhmanonva, A. S., Huynen, M. A. & Hackstein, J. H. P. 2000 A mitochondrial ancestry of the hydrogensomes of *Nyctotherus ovalis*. *Mol. Biol. Evol.* **17**, 202–206.
Wheeler, D., Chappey, C., Lash, A., Leipe, D., Madden, T., Schuler, G., Tausova, T. & Rapp, B. 2000 Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **28**, 10–14.
Viale, A. & Arakaki, A. K. 1994 The chaperone connection to the origins of the eukaryotic organelles. *FEBS Lett.* **341**, 146–151.
Voncken, F. (and 11 others) 2002 Multiple origins of hydrogenosomes: functional and phylogenetic evidence from the ADP/ATP carrier of the anaerobic chytrid *Neocallimastix* sp. *Mol. Microbiol.* **44**, 1441–1454.
Walker, J. E. & Runswick, M. J. 1993 The mitochondrial transport protein superfamily. *J. Bioenerg. Biomem.* **25**, 435–446.
Williamson, L. R., Plano, G. V., Winkler, H. H., Krause, D. C. & Wood, D. O. 1989 Nucleotide sequence of the *Rickettsia prowazekii* ATP/ADP translocase-encoding gene. *Gene* **80**, 269–278.
Winkler, H. H. 1976 Rickettsial permeability: an ADP-ATP transport system. *J. Biol. Chem.* **251**, 389–396.
Wolf, Y. I., Aravind, L. L. & Koonin, E. V. 1999 Rickettsiae and Chlamydiae: evidence of horizontal gene transfer and gene exchange. *Trends Genet.* **15**, 173–175.
Wood, D. W. (and 50 others) 2001 The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**, 2317–2323.
Yang, D., Oyaizu, Y., Oyaizu, H., Olsen, G. J. & Woese, C. R. 1985 Mitochondrial origins. *Proc. Natl Acad. Sci. USA* **82**, 4443–4447.
Zumft, W. G. 1997 Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* **61**, 533–616.

Discussion

F. R. Whatley (*Department of Plant Sciences, University of Oxford, Oxford, UK*). I was almost expecting you to try to charm us into believing that *Rickettsia* was an organelle, almost.

S. G. E. Andersson. Well I would not call them an organelle. It is clear that they cannot live by themselves, but they seem to have all the genes that would be required for free-living growth, so it may be that it is just a regulatory phenomenon. Perhaps, the bacterium needs some type of regulatory stimulus from its eukaryotic host to start growing. In humans, *Rickettsia prowazekii* is a parasite that confers no benefits to its host, also distinguishing it from an organelle. *Buchnera* is a symbiont and thereby more like an organelle, contributing amino acids to its host. This organism has also lost a lot of genes, essential genes, and even if it were to get the appropriate stimulus for growth it probably would not be able to grow in any case. However, in contrast to *Rickettsia prowazekii* and organelles, *Buchnera* grows in specialized cells and not freely in the cytoplasm.

M. van der Giezen (*School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey, UK*). The AACs you showed, the parasitic AACs from Rickettsia and the normal mitochondrial AACs that excrete ATP—I thought that they were structurally unrelated. I mean, they do the same thing but they do it in the other direction.

S. G. E. Andersson. Yes, in the mitochondria ADP–ATP translocases are used for exporting ATP from the mitochondria into the cytoplasm, whereas the parasitic ATP–ADP translocases are used for taking ATP from the cytoplasm into the bacterium.

M. van der Giezen. Because you put the trees on top of each other to correlate them? But I thought they were both completely independent proteins.

S. G. E. Andersson. Yes, they are completely independent proteins and there is no similarity among them. The trees were not derived from the same alignment, but were independent alignments and independent trees. We only put them on top of each other to compare the phylogenetic depth by calibrating the rate of sequence evolution with the help of plant species present in both trees. Here, we found that the depths were approximately similar, which contradicts the idea of horizontal transfers among the parasitic and plastid lineages, even though they are only present in these particular lineages.

F. R. Whatley. To ask the question again, essentially the difference between *Rickettsia* and the organelles is which way they exchange ATP?

S. G. E. Andersson. Yes, you could say that.

F. R. Whatley. But they are by different detailed mechanisms, different proteins rather?

S. G. E. Andersson. Yes, completely different proteins.

The above is an edited transcript of the discussion at the meeting. The following are invited, written questions and answers addressed specifically to Professor Andersson's manuscript.

W. Martin (*Institute of Botany III, Heinrich-Heine Universität, Düsseldorf, Düsseldorf, Germany*). In your listing of three current scenarios for the origin of hydrogenosomes, you credited me as having suggested that 'they may be the descendants of a bacterial ancestor that was acquired prior to, simultaneously with, or subsequent to the acquisition of the mitochondrial ancestor', but to be absolutely sure I never suggested anything of the sort; rather I suggested that they descend from the same endosymbiont as gave rise to mitochondria.

S. G. E. Andersson. Sorry, I remember now that you were more specific about the origin of hydrogenosomes and that you have in fact suggested that both organelles descend from one and the same endosymbiont that was later transformed into mitochondria in some lineages. We agree that mitochondria and hydrogenosomes probably descend from the same ancestral endosymbiont, but we disagree about the putative selective pressure that triggered the integration of this endosymbiont. Whereas you speculate that it was anaerobic, hydrogenosomal-like gene functions, I speculate that it was aerobic, mitochondrial-like gene functions.

A. G. M. Tielens (*Faculty of Biology, Utrecht University, Utrecht, The Netherlands*). I am pleased that our earlier proposals (Tielens & Van Hellemond 1998) that anaerobic mitochondria originate from classical aerobic mitochondria, and that this happened by adaptation to anaerobic functioning, are included in your discussion of the evolutionary origin of mitochondria. I am also pleased that your phylogenetic analysis of fumarate reductases is indeed in full agreement with earlier ones, and that you agree with the suggestion that eukaryotic fumarate reductases are more closely related to succinate dehydrogenases of aerobic eukaryotes than to fumarate reductases from prokaryotes (Kita *et al*. 2002; Tielens & Van Hellemond 1998; Van Hellemond *et al*. 1995; Van Hellemond & Tielens 1994). I can add that we have just written a new review on the origin of non-classical mitochondria, and in this review we of course again paid attention to this relationship between fumarate reductases and succinate dehydrogenases (Tielens *et al*. 2002).

On the other hand, I do not agree with your conclusion that—similar to anaerobic mitochondria—hydrogenosomes are also derived from aerobic mitochondria and thus represent secondary adaptations to anaerobic environments, after a prior development into an aerobic organelle. It is clear that although there are many differences between hydrogenosomes and mitochondria, they are related organelles with a common ancestor. However, apart from the fact that it is now also becoming clear that different types of hydrogenosomes exist, I wondered what evidence or indications you have that, for instance, the hydrogenosomes of trichomonads are secondary adaptations of mitochondria after the ancestral type had adapted into the aerobic direction?

S. G. E. Andersson. We seem to agree that mitochondria and hydrogenosomes are related organelles with a common ancestor. The question is, what did this ancestor look like? We know that most of the gene functions unique to mitochondria are present in the α-proteobacteria, and phylogenetic reconstructions based on these enzymes confirm that they are related. My interpretation of this strong relationship is that these gene functions were already present in the invading endosymbiont. By contrast, phylogenetic reconstructions inferred from enzymes that are unique to hydrogenosomes, such as the hydrogenase, reveal no specific relationship with the α-proteobacteria, indicating that they may not have been present in the ancestral endosymbiont. Therefore, my interpretation is that the hydrogenosome-specific functions were acquired at a later stage of evolution, possibly when the organism adapted to an anaerobic environment.

W. Martin. Martin and Müller observed on the basis of data then available that eukaryotes possess a eubacterial, rather than an archaebacterial, glycolytic pathway, and they predicted subsequent sequence data to uphold that view (Martin & Müller 1998). Being reasonably familiar with glycolytic enzymes, I find it fair to state that said prediction has been quite firmly borne out by genome data. Can you comment on the patterns of sequence similarity observed for archaebacterial glycolytic enzymes relative to their eukaryotic and eubacterial counterparts, and contrast those sequence similarities with the corresponding observations for rRNA? In your phylogenies for glycolytic enzymes do you see any evidence for horizontal gene transfer, or are those phylogenies identical to each other and to the phylogenies for rRNA from the same species?

S. G. E. Andersson. With a few exceptions, Archaea do not contain the type of glycolytic enzymes found in bacteria and eukaryotes. One explanation is that these genes were present in the common ancestor of bacteria, Archaea and eukaryotes but were lost from the lineage leading to Archaea and replaced by other pathways. The alternative explanation is that the last common ancestor relied on a yet undiscovered pathway for production of ATP and that glycolysis was invented in the bacterial lineage subsequent to its divergence from the Archaea and eukaryotes. To explain the presence of glycolysis in eukaryotes one must then postulate that they were transferred from bacteria to eukaryotes via a mitochondrial intermediate, as suggested, for example, in the hydrogen hypothesis. However, phylogenetic reconstructions do not reveal any specific affiliation with the α-proteobacteria, as expected from the hypothesis. The first explanation, i.e. that glycolysis was present already in the last universal common ancestor, seems therefore more likely. Because many glycolytic genes are absent from the Archaea, it is not always possible to make phylogenetic reconstructions based on these genes that are directly comparable to that of the universal rRNA tree. We do observe anomalies in some trees, but these cannot *a priori* be taken as evidence for horizontal gene transfers. It is possible that gene paralogy, differences in functional constraints and atypical rates and patterns of sequence evolution may distort the phylogenetic signal in some parts of the trees. However, surprisingly many genes yield phylogenies that are consistent with each other and with the corresponding parts of the rRNA tree.

J. F. Allen (*Plant Biochemistry, Lund University, Lund, Sweden*) and J. A. Raven (*Department of Biological Sciences, University of Dundee, Dundee, UK*). The complete genome sequence of *Rickettsia prowazekii1* (Andersson *et al.* 1998) tells us that this obligate intracellular parasite has lost many genes during its evolution from a free-living ancestor. Many of the remaining genes code for proteins involved in energy transduction. *R. prowazekii* and mitochondria share both this emphasis on energy transduction and details of the mechanisms it involves. In the case of mitochondria, many of these genes have been transferred to the cell nucleus. The deduced amino acid sequences of 18S rRNA and of NADH dehydrogenase I show that *R. prowazekii* and the putative mitochondrial ancestor are members of the α-proteobacteria (Andersson 1998; Andersson *et al.* 1998; Gray 1998), but these data do not furnish evidence on whether reduction of the two genomes occurred independently or as successive stages in a single evolutionary sequence. The latter possibility has been developed (Andersson 1998), with the conclusion that the reduction in gene content of the genomes reflects a very close phylogenetic relationship. *R. prowazekii* might then be expected to contain a subset of its ancestor's genes, while the mitochondrion contains a subset of that subset.

We wish to draw attention to a plausible alternative explanation, which is that the genomes of *Rickettsia* and mitochondria have been reduced independently but by the same selective forces, giving merely a genomic version of classical convergent evolution. Their resemblance then no more argues their kinship than the cephalopod eye makes the squid a vertebrate. Since it is beyond the bounds of coincidence (Gray 1998) for the same core of genes for oxidative phosphorylation and ancillary processes to be retained by random loss, there must be reasons why some genes have been lost and others retained. We have made a testable suggestion (Allen & Raven 1996), which is that decreased mutation frequency in a specialized and stable information-storage compartment (the cell nucleus) is a clear advantage for most genes, the exceptions being those with an overriding requirement for direct redox regulatory control. Thus mitochondrial, plastidic, and now intracellular parasitic genomes may consist of genes that cannot effectively be regulated unless they stay close to their gene products (Allen 1993). Do you have any comments on these suggestions?

S. G. E. Andersson. While it is true that the mitochondrial genomes contain a subset of the *R. prowazekii* subset of genes, we have never suggested that the mitochondrial genomes have originated from a genome that was identical in size to the modern *R. prowazekii* genome. On the contrary, we have in numerous publications emphasized that gene loss and genome degradation are still ongoing processes in *Rickettsia*. For example, we have estimated that *R. prowazekii* has lost several hundred genes only since its divergence from *R. conorii* some 30–50 million years ago. There is no doubt that *Rickettsia* and mitochondria share a common ancestor, as inferred from phylogenetic studies. If the rate of gene loss in the distant past was comparable with or larger than that we see today in *Rickettsia*, the genome of this common ancestor that lived some 1500 million years ago must have been much larger than the modern *R. prowazekii* genome. But we do not know how much larger.

Yes, I agree that all evidence points to decreased substitution rates for genes located in the cell nucleus compared to genes located in the mitochondrial genome (at least for animal mitochondrial genomes). However, there is no evidence that rickettsial genes have been transferred into the nuclear genome of the host with the corresponding gene products being imported back into *Rickettsia* upon infection. Since *Rickettsia* is not permanently maintained in any one host, but alternates between vectors and hosts, successful gene transfers may not be as simple as for mitochondrial genes. Concerning your question about regulation, we know that some genes, such as those coding for citrate synthase and ATP–ADP translocases, are regulated in *Rickettsia* according to the levels of ATP in the host's cytoplasm. Whether this and other regulatory circuits would work as efficiently if the genes had been downloaded to the nuclear genomes, I do not know.

W. Martin. You mentioned Mereschkowsky in the context of nineteenth-century ideas on the endosymbiotic origin of mitochondria, but Mereschkowsky (Martin & Kowallik 1999; Mereschkowsky 1905) suggested that only plastids and the nucleus were bacterial endosymbionts, not mitochondria. Altmann (Altmann 1890) is often credited with suggesting an endosymbiotic origin of mitochondria in '*Die Elementarorganismen*', but as far as I am familiar with his paper he never really suggested that; rather he just argued that there may be a basic organizational unit of living things below the level of the cell (the Bioblasten), no word of endosymbiosis.

S. G. E. Andersson. Yes, of course you are absolutely right. Thanks for reminding me.

Additional references

Allen, J. F. 1993 Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. *J. Theor. Biol.* **165**, 609–631.

Allen, J. F. & Raven, J. A. 1996 Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. *J. Mol. Evol.* **42**, 482–492.

Altmann, R. 1890 *Die Elementarorganismen und ihre Beziehungen zu den Zellen*. Leipzig: Verlag von Veit & Comp.

Gray, M. W. 1998 *Rickettsia*, typhus and the mitochondrial connection. *Nature* **396**, 109–110.

Martin, W. & Kowallik, K. V. 1999 Annotated English translation of Mereschkowsky's 1905 paper 'Über Natur und Ursprung der Chromatophoren im Pflanzenreiche'. *Eur. J. Phycol.* **34**, 287–295.

Mereschkowsky, C. 1905 Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. Centralbl.* **25**, 593–604.

Tielens, A. G. M., Rotte, C., Van Hellemond, J. J. & Martin, W. 2002 Mitochondria as we don't know them. *Trends Biochem. Sci.* **27**, 564–572.

GLOSSARY

AAC: ADP–ATP carrier protein
FRD: fumurate reductase
ME: minimum evolution
NiR: nitrate reductase
PDH: pyruvate dehydrogenase
PFO: pyruvate : ferredoxin oxodireductase
SDH: succinate dehydrogenase
TIM: translocase of inner membrane
TOM: translocase of outer membrane