Available online at www.sciencedirect.com

SCIENCE DIRECT®

Biochimica et Biophysica Acta 1659 (2004) 212–220

BIOCHIMICA ET BIOPHYSICA ACTA

BBA

http://www.elsevier.com/locate/bba

Review

Shaping the mitochondrial proteome

Toni Gabaldón*, Martijn A. Huynen

NCMLS, Nijmegen Center for Molecular Life Sciences, P/O: CMBI, Center for Molecular and Biomolecular Informatics, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

Received 4 June 2004; received in revised form 15 July 2004; accepted 28 July 2004
Available online 11 September 2004

Abstract

Mitochondria are eukaryotic organelles that originated from a single bacterial endosymbiosis some 2 billion years ago. The transition from the ancestral endosymbiont to the modern mitochondrion has been accompanied by major changes in its protein content, the so-called proteome. These changes included complete loss of some bacterial pathways, amelioration of others and gain of completely new complexes of eukaryotic origin such as the ATP/ADP translocase and most of the mitochondrial protein import machinery. This renewal of proteins has been so extensive that only 14–16% of modern mitochondrial proteome has an origin that can be traced back to the bacterial endosymbiont. The rest consists of proteins of diverse origin that were eventually recruited to function in the organelle. This shaping of the proteome content reflects the transformation of mitochondria into a highly specialized organelle that, besides ATP production, comprises a variety of functions within the eukaryotic metabolism. Here we review recent advances in the fields of comparative genomics and proteomics that are throwing light on the origin and evolution of the mitochondrial proteome.

Keywords: Mitochondria; Mitochondrial evolution; Alpha-proteobacteria; Proteome; Endosymbiosis

1. Introduction

According to the widely accepted endosymbiotic theory, mitochondria are eukaryotic organelles of bacterial descent. Phylogenetic data supporting this theory point to an ancient alpha-proteobacterium, most likely an ancestor of the Rickettsiales, establishing a symbiotic relationship inside a primitive eukaryotic cell circa 2 billion years ago [1–3]. Since then the mitochondrion has undergone major changes that transformed it into a highly specialized organelle that plays a key role within the metabolism of most eukaryotic cells. This process involved not only a dramatic reduction at the level of its genome but also an extensive renewal at the level of its proteome that affected more than 80% of the protein content of modern mitochondria [4].

Nowadays, mitochondria are present in a multitude of eukaryotic organisms adapted to a variety of niches [5], with modern mitochondrial proteomes reflecting a large diversity in organellar functions. Moreover, there is increasing evidence that certain organisms that lack mitochondria "sensu-stricto" actually harbor relicts of these organelles. These, together with the hydrogenosomes, most of which appear of mitochondrial descent [6], reflect extreme forms of mitochondrial adaptation to microaerophilic or anaerobic environments. The evolution of the mitochondrial genome in both its structure and gene content has been the focus of several reviews [7–9], but the scarcity of data has long prevented similar surveys on the evolution of the mitochondrial proteome. Recent advances in the fields of proteomics and comparative genomics give us a picture of the processes that shaped the mitochondrial proteome from the early stages of endosymbiosis to modern adaptations of mitochondria to diverse environments and cellular functions. Recent reviews have focused on the evolution of a specific mitochondrial system, such as the mitochondrial import machinery [10], or have analyzed more generally the processes that transformed ancestors of both mitochondria

* Corresponding author. Tel.: + 31 24 3653374; fax: +31 24 3652977.
E-mail address: T.Gabaldon@cmbi.ru.nl (T. Gabaldón).

0005-2728/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbabio.2004.07.011

and plastids into modern organelles [11]. Here we review recent data that define the proteome of the mitochondrial ancestor and the modern mitochondria. We also discuss the most relevant events that affected the mitochondrial protein content and hence its metabolic capacities.

2. The starting point: the proto-mitochondrial proteome

To unravel the evolution of the mitochondrial proteome, it is essential to get an idea of what it looked like during the first endosymbiotic stages, the common point from which all modern mitochondrial proteomes evolved. Defining the proteome of the mitochondrial ancestor would also help in understanding how the initial symbiosis was established and perpetuated, an issue that is hotly debated and for which several hypotheses have been proposed [12–16]. First approaches to infer the nature of the mitochondrial ancestor were based on the similarity, in terms of metabolic capabilities, of mitochondria with some bacterial groups. In this way the gamma-proteobacterium *Bdellovibrio* and the alpha-proteobacterium *Paracoccus* were the first proposed models for the proto-mitochondrion [17]. Later on, phylogenetic analyses of small subunit ribosomal RNAs and proteins from the respiratory complexes confirmed a monophyletic origin of mitochondria from an alpha-proteobacterial ancestor [1,3]. Moreover, the sequencing of the genome of the bacterium *Rickettsia prowazekii* [18] and subsequent phylogenetic analyses [18,19] identified members of the *Rickettsia* genus as the closest relatives of modern mitochondria. Some of the Rickettsiales are obligate intracellular parasites, a feature that suggested a similar lifestyle for the mitochondrial ancestor [18,20]. However, any attempt to establish parallels between the proto-mitochondrial proteome and those of modern alpha-proteobacteria must be cautious, bearing in mind the roughly estimated 2 billion years of evolution that separate them. Indeed, the adaptation to an intracellular lifestyle of the modern Rickettsiales is the result of a different event than that of the endosymbiosis of mitochondria [18]. Therefore, the similarities in their respective adaptations are probably the result of convergent evolution [18,21]. The identification of its phylogenetic affiliation with the alpha-proteobacteria narrows the scope of the speculations on the proto-mitochondrion's lifestyle. However, the great diversity of size and composition of modern alpha-proteobacterial genomes, ranging from 834 to more than 8000 protein-coding genes, provides enough room for many alternative models.

The problem can be reformulated as to what extent does the proto-mitochondrial proteome resemble that of modern alpha-proteobacteria. To answer this question, it is necessary to distinguish truly common features from those resulting from secondary adaptations. A valid approach is to trace the origin of the modern proteomes to determine which proteins are directly derived from the endosymbiotic event. Assuming no genetic transfer to the mitochondrial genome, the proteins encoded there constitute a 'bona-fide' subset of proteins derived from the proto-mitochondrion. But mitochondrial genomes sequenced so far encode only 3–67 proteins that are involved in few processes, mainly respiration and protein synthesis and, occasionally, transcription, RNA maturation and protein import [7]. The set has been extended by a phylogenetic analysis of the yeast mitochondrial proteome [22] that identified an additional number (~20) of nuclear-encoded proteins whose phylogenies indicated an alpha-proteobacterial origin. When combined, both sets form a reduced core of the proto-mitochondrial proteome whose deduced metabolic capabilities reflect that of a cell harboring few metabolic pathways, but able to couple electron transport to the production of ATP as well as with the capacity to synthesize the required proteins.

This picture changed considerably after a large-scale phylogenetic comparison of alpha-proteobacterial and eukaryotic genomes [4] identified 630 eukaryotic proteins that were likely derived from the alpha-proteobacterial ancestor of mitochondria. Mapping the metabolic functions of these orthologous groups, a minimal metabolism for the proto-mitochondrion ancestor could be reconstructed. Besides the abovementioned processes, other pathways such as the oxidation and synthesis of fatty acids, biotin and heme synthesis, iron-sulfur cluster assembly, and fructose and sucrose metabolism pathways emerged. Also notable was the presence of many metabolite transporters.

Altogether the accumulation of data on the proto-mitochondrial proteome point towards a (facultatively) aerobic organism living on several compounds provided by the host. Whether the proto-mitochondrion was a parasite, something that is compatible with the data available, depends on the existence of potential benefits for the host. In the case of a mutual benefit, the lack of an ATP transporter suggests that ATP was not the main currency used by the proto-mitochondrion to pay back host's services. Alternative benefits for the host have been proposed in hypotheses that consider a hydrogen-producing [12] or an oxygen-detoxifying [13] endosymbiont. The presence of the Fe–S cluster assembly pathway and the ancestor of the ABC transporter that is likely involved in the export of Fe–S clusters from mitochondria (ATM1) [23] indicate an alternative benefit to the host that could have been a key in the initial symbiotic relationship. This would be in agreement with the finding that proteins of this pathway are among the few conserved by mitochondrial remnants in the microsporidian *Enzephalitozoon cuniculi* [24], the protozoan *Giardia intestinalis* [25] and the apicomplexan *Cryptosporidium parvum* [26]; although we cannot discard a secondary loss in these organisms of another pathway that provided the original selective pressure for the symbiosis.

3. A crucial step: the origin of the mitochondrial import machinery

The proto-mitochondrial proteome soon underwent a series of transformations that shaped it. Most important to this transformation was the acquisition of a mechanism that facilitated the import of proteins from the cytosol into the mitochondrion. The abovementioned proto-mitochondrial reconstruction points to an ancient endosymbiont that was autonomous in protein synthesis, with no sophisticated system for the import of proteins synthesized in the cytosol. This contrasts with the modern situation in which most mitochondrial proteins are encoded by nuclear genes, synthesized in the cytoplasm and subsequently imported into the organelle. The latter step is carried out by a complex machinery consisting of dozens of proteins located in the inner and outer membranes of the mitochondria [27], as well as many soluble chaperones that assist in the process. This machinery recognizes specific N-terminal signals that are sufficient and necessary to direct the import of proteins into mitochondria. Such a system is a prerequisite not only for the escape from mitochondria to the nucleus of genes whose products should be targeted back but also for the recruitment of proteins of different origin to the organelle. Considering that both processes have been rampant [4,8,22], there is little doubt that the emergence of the protein import system was a crucial step in the evolution of mitochondria. Indeed, it might well be the event that marked the beginning of the transition from endosymbiont to organelle. Once genes encoding essential mitochondrial proteins were transferred to the nucleus, the host took command of the mitochondrion.

The phylogenetic analysis of the mitochondrial import machinery [10] reveals that although most of its components are of eukaryotic origin, some of the components in the inner membrane and most of the soluble chaperones that assist the translocation have bacterial homologs. This mixed origin argues in favor of the hypothesis [10] that a rudimentary system of chaperones and porins already existed in the ancestor. This system rapidly evolved into a more sophisticated machinery able to import proteins with high efficiency and specificity. Furthermore, in its initial stages the protein import system could have been tightly coupled with translation as suggested by the preferential synthesis of ancient mitochondrial proteins by polysomes attached to mitochondria [28]. The sophistication of the translocation system should have run parallel to the evolution of the addressing sequences that direct the targeting of the mitochondrial proteins. It has been suggested that these sequences have evolved from proteins with an inherent propensity to be targeted to the mitochondria [29]. Furthermore, such targeting sequences are not uncommon in prokaryotic proteins [29] and can easily be matched by random sequences [30]. Once present in a few genes, it is easy to conceive that these pre-sequences were passed to other genes by means of duplication and recombination events [31], allowing evolution to potentially test the mitochondrial localization of any nuclear-encoded protein. Such a transport system would presumably pave the way for the recruitment of proteins for service in the mitochondrion. Therefore, the protein transport system contributed a new dimension to the evolution of the mitochondrial proteome by facilitating expansion of the proteome as well as reduction of the organellar genome.

4. Turning mitochondria into cell's energy factory

Perhaps the function gain that most radically affected the role of the mitochondrion within the eukaryotic cell was the acquisition of an ATP/ADP translocase and therefore the ability to exchange ATP between the mitochondria and the cell's cytoplasm. Although the ATP-production system is derived from the bacterial ancestor, the ATP/ADP translocase has a eukaryotic origin [32]. The origin of the ATP/ADP exchanger provided mitochondria with a new function for the cell: that of an energy-converting organelle. This new role might have also favored subsequent mitochondrial transformations such us the elaborate folding pattern of the inner membrane (cristae) and the increase in complexity of the ATP-synthase [33] as well as the electron transport chain (ETC) complexes [34].

This increase in complexity has affected all mitochondrial ETC complexes with the exception of succinate dehydrogenase, notably the only ETC complex that does not pump protons across the inner membrane. For the other complexes, as much as a threefold increase in terms of protein content may be observed when the bacterial components are compared to their mammal counterparts: 3 to 11 subunits in cytochrome bc_1 complex, 4 to 13 in cytochrome *c* oxidase and 14 to 46 in NADH:ubiquinone oxidoreductase (complex I). In the latter, the addition of extra subunits has not altered the characteristic L-shaped structure of the complex [35]. Notably, not all new subunits recruited to the ETC complexes are of eukaryotic origin. The acyl carrier component (ACP) of complex I is actually present in alpha-proteobacteria, but has never been identified as part of complex I in prokaryotes [36]. It is difficult to assess the function of the so-called subsidiary subunits. Their participation in the biogenesis or stabilization of the complexes is the most common speculative explanation of their roles [35,37]. A processing peptidase activity has been confirmed for some subsidiary subunits such as the 8-kDa subunit of cytochrome bc_1 complex in plants [38]. However, different roles have been proposed for other subunits such as the participation in fatty acid synthesis of the ACP component of *Neurospora* Complex I [39].

The presence of the mitochondrial-type ATP/ADP translocase in both aerobic and anaerobic mitochondria studied so far [40] as well as in hydrogenosomes [6,41] suggests that this gain of function preceded the diversification of

mitochondria in terms of the terminal acceptor of electrons used to fuel the synthesis of ATP. It also raises doubts on whether oxygen was the original sink for the electron transport chain [5]. Besides oxygen, modern mitochondria are able to use alternative electron acceptors such as nitrate, nitrite or fumarate. Although some anaerobic mitochondria, such as the ones present in parasitic helminths and freshwater snails, seem to be the result of secondary adaptations to anerobiosis [42], current data on other anaerobic mitochondria and hydrogenosomes [5] can also be explained by a scenario in which main anaerobic traits are vertically derived from a facultative anaerobic ancestor. The recent finding of an evolutionarily related but distinct ATP/ADP translocase in the flagellate *Trichomonas gallinae* [43,44] may indicate multiple origins for ATP/ADP exchangers.

5. Modern mitochondrial proteomes

Modern mitochondria represent extremes of the mitochondrial proteome evolution. There are in fact multiple ends, as mitochondria diverged along the different eukaryotic lineages (Fig. 1). Defining the modern mitochondrial proteome could seem easier than inferring the ancestral one, since we have the extant organisms at hand. But only recently has it become technically feasible to obtain comprehensive data for the protein content of mitochondria, thanks to the application of several new proteomic techniques to the analyses of organellar proteomes [45–48]. Analysis of the mitochondrial proteome starts by a purification of the organelle by means of centrifugation on a density gradient followed by separation of the protein content to identify the individual proteins. Different groups have been fine-tuning every single step of the process to cover an increasing fraction of the mitochondrial proteome. In the separation phase, 1D or 2D polyacrylamide gel electrophoresis has been combined or used in parallel with other techniques such us liquid-chromatography or isoelectric focusing [49]. In the final step, the identification of proteins in the different gel-spots or aliquots is facilitated by a variety of mass spectrometry (MS) techniques. These include, among others, matrix-assisted laser-desorption ionization (MALDI) MS, electrospray ionization MS and tandem MS. Successive proteomic studies have progressively increased molecular coverage. Consequently, we have now reasonably good proteomic sets for the yeast [50] and human [51,52] mitochondria, while great progress has been made for other model organisms such as mouse [49], rice [53] and *Arabidopsis* [54] (Fig. 1). Alternatively, epitope-

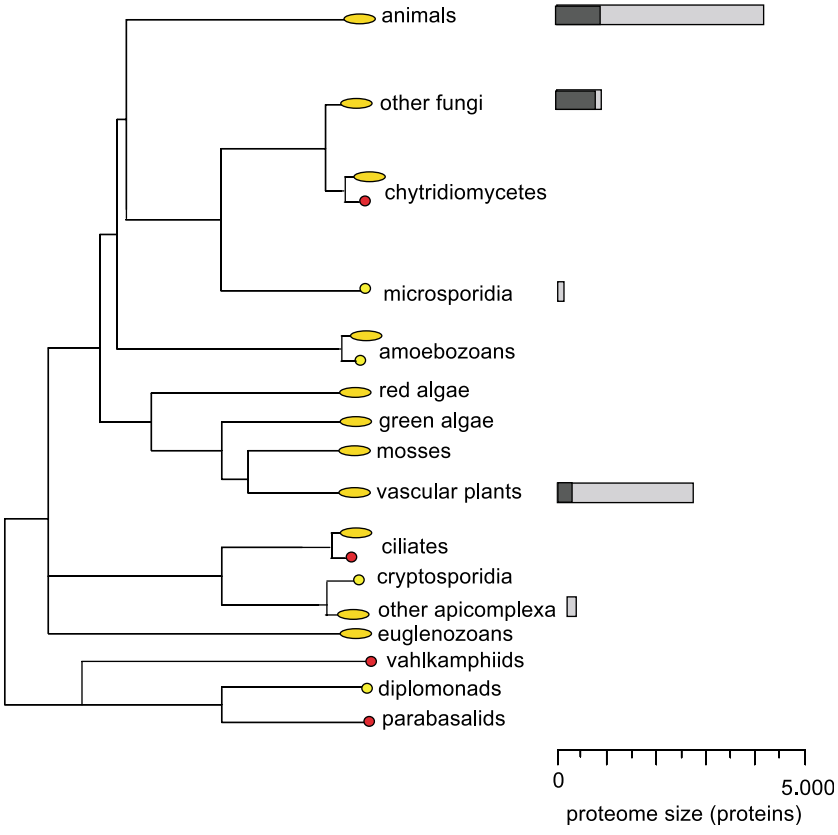

Fig. 1. Eukaryotic phylogeny according to Refs. [2,87] (left); symbols next to a phylum's name represents possession of canonic mitochondria (brown ovals), mitosomes or other mitochondrial remnants (yellow circles) and hydrogenosomes (red circles). Bars on the right represent total estimated proteome size (in grey) according to Ref. [59], and fraction of experimentally determined proteome (in black) according to Refs. [51] (human), [50] (yeast) and [54] (*Arabidopsis*).

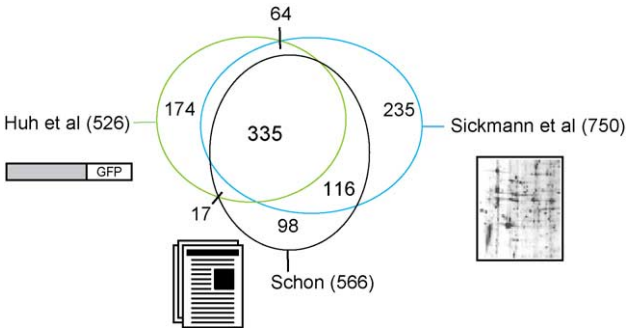

Fig. 2. Partial overlaps between three mitochondrial proteome sets: 526 proteins identified by GFP-tagging and fluorescence microscopy by Huh et al. [56], 750 proteins identified by mitochondrial isolation and proteomics analyses by Sickmann et al. [50] and 566 proteins of a manually curated list of experimentally identified mitochondrial proteins by Schon [57].

[55] and GFP- [56] protein fusions combined with (immuno)fluorescence microscopy have been used on a genome-wide scale to analyze the subcellular localization of members of the yeast proteome, identifying 332 and 526 mitochondrial proteins, respectively.

The comparison of the proteomic [50], the GFP-tagging genome-wide analyses [56] and an expert-compiled list of experimentally determined mitochondrial proteins [57] (Fig. 2) shows a higher coverage of the reference set by proteomic techniques (80%) compared to the GFP-tagging analysis (62%). Such high coverage indicates that we are getting very close to the complete identification of the yeast mitochondrial proteome. Nevertheless, there is still a large fraction of proteins only detected by one of the techniques (49% of the combined set). In particular, there are nearly hundred proteins for which a mitochondrial localization has been found by small-scale experiments but not by the high-throughput ones. These results are in agreement with a recent survey on different mitochondrial proteome analysis techniques in yeast [58], and show that there is still a part of the mitochondrial proteome resisting identification by large-scale approaches. These particular proteins tend to be present in low concentrations or only expressed under certain conditions [58]. As expected, the situation in human, mouse and other model organisms is no better in terms of coverage. In the absence of proteomic data, mitochondrial-targeting prediction tools [59,60] have also been used to obtain a rough estimate of the protein contents of mitochondria of organisms with a fully sequenced genome (Fig. 1). Although the range of specificity (28–55%) [58] of these computational methods is lower than most large-scale proteomic techniques, this can be increased by the combined use of different algorithms [59], albeit at a loss of coverage. The increasing availability of fully sequenced genomes of species for which no large-scale subcellular proteomics are available makes the use of computational methods an essential tool to study the mitochondrial proteome. Even when proteomic techniques are applied, the fact that different experimental and computational tools show specific biases makes a combination of techniques the most adequate approach [54,58,59].

A preliminary analysis of the available sets reveals that modern mitochondrial proteomes are highly variable not only in size but also in content, likely reflecting a parallel diversity in metabolic and regulatory functions among the different phyla. In addition to their function as energy producing organelles, mitochondria are pivotal in many other processes such as the synthesis of the heme group [61] and iron–sulfur clusters [62]. In addition, they harbor metabolic pathways which are specific for certain species or genera [5,61,63]. Many of these lineage-specific functions, such as the role of plant mitochondria in folate synthesis [64], have long been known from physiological studies but the availability of large-scale proteomic data is likely to uncover many more aspects of mitochondrial diversity. This functional diversity is the result of the two processes that shaped the mitochondrial proteome (Fig. 3): First, there is the continuous loss of proteins, which has reduced the ancient fraction of the proteome and therefore the bacterial identity of the organelle. Second, there is a continuous recruitment of proteins, which expands the proteome, providing it with new functions or adapting the

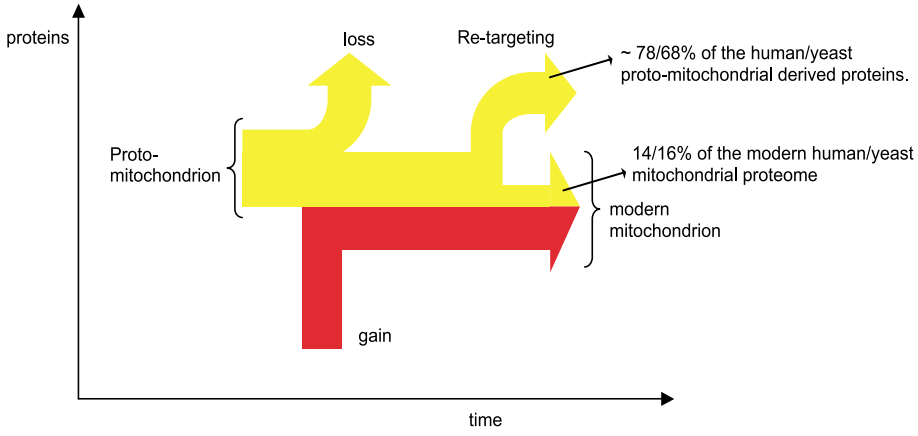

Fig. 3. Overview of the processes that transformed the proto-mitochondrial proteome into those of modern mitochondria. The estimates for the fraction of the modern proteome affected are according to Ref. [4]. Estimates for the proteomic fraction affected in human and yeast, according to Ref. [4], are indicated.

old ones to the new circumstances. We emphasize that the loss from the mitochondrial proteome does not necessary imply loss from the total cell's proteome. Indeed, ancient mitochondrial proteins can be retargeted to other locations in the cell [4] as exemplified by the peroxisomal location of part of the beta-oxidation pathway in yeast.

Gain and loss processes have acted along the various lineages in a different manner in terms of quality and intensity. Mitochondria have selectively lost proteins whose functions were no longer needed in a certain lineage, such as the specific complex I loss from yeast, while keeping required functions. In the yeast mitochondrial proteome, most of what remains from the alpha-proteobacterial ancestor (59 out of 97 proteins) is related to respiration according to a genome-wide analysis [65]. This suggests that these mitochondria are more specialized towards respiration than their ancestor.

Major expansions in the size of the proteome are observed in vascular plants and metazoans and these may reflect the adaptation to multicellularity and tissue differentiation. Indeed, many tissue-specific mitochondrial properties and proteins have been described in plants [66] and mammals [49]. In plants, comparisons between photosynthetic and non-photosynthetic tissues have demonstrated that the composition of mitochondrial proteomes varies in accordance to their different metabolic needs. For example, there is a greater demand for the mitochondrial oxidation of glycine associated with photo-respiration [67]. In mammals, mitochondria of certain tissues have been shown to participate in tissue-specific processes such as the regulation of insulin secretion in pancreatic beta-cells [68] or steroidogenesis in adrenal cortex [69].

Besides these lineage-specific expansions, the mitochondrial proteome likely gained a common set of proteins before the divergence of most eukaryotic species. These common gains include the abovementioned ATP/ADP translocase and part of the mitochondrial protein import machinery. The total overlap between yeast and human mitochondrial proteomes, 371 proteins in the E. Schon set [57], greatly exceeds the alpha-proteobacterial fraction, 79 proteins (~20%) [4], indicating that the mitochondrial proteome experienced a major expansion before the divergence of fungi and metazoa. Consistently the ETC complexes that are present in both organisms share most of the subsidiary subunits of eukaryotic origin.

It is worth recalling that the recruitment of proteins to the mitochondrial proteome was not limited to proteins of eukaryotic origin. In principle, any nuclear-encoded protein could potentially be targeted to mitochondria once the necessary targeting-sequence is attached to the gene. Therefore, it is not surprising to find mitochondrial proteins of diverse origin such as the components of the oxidative branch of the Krebs cycle, which are phylogenetically close to the Cyptophaga–Flavobacterium–Bacteroides (CFB) bacterial group [70], although a scenario in which CFB bacteria gained these enzymes from eukaryotes cannot be ruled out. Similarly the viral-like mitochondrial RNA-polymerase seems to be a clear case of non-orthologous gene displacement, because the proto-mitochondrial ancestor likely harbored the eubacterial-type RNA-polymerase that is encoded in the mitochondrion of the protist *Reclinomonas americana* [71]. The RNA-polymerase case is not the only example of a viral-like protein working in mitochondria. The so-called twinkle protein is a mitochondrial DNA-helicase with homology to phage-T7 primase helicase [72]. Phylogenetic analyses of mitochondrial proteomes have so far not identified other proteins of viral origin (results not shown).

6. Mitochondrial proteome remnants in "amitochondriate" eukaryotes

The monophyletic origin of mitochondria and their nearly ubiquitous distribution among eukaryotes suggest that the acquisition of these organelles occurred very early in the evolution of eukaryotes. How early is still a matter of discussion. The initial placing of most amitochondriate eukaryotes at the base of the reconstructed phylogenies suggested that these diversified prior to the mitochondrial endosymbiosis [73]. Recent phylogenetic reconstructions

Table 1
List of *E. cuniculi* proteins with a likely proto-mitochondrial origin [4] that have a detectable homolog (*E* value $<10^{-15}$, with BlastP search) in the genome of *C. parvum*

Gene name	Annotation	Mit.
Ecu-CU09_920	Putative ribosomal RNA methyl transferase	
Ecu-CU07_1340	Ribosomal RNA methyl transferase	
Ecu-CU11_0540*	Mitochondrial HSP70	
Ecu-CU03_0240	Similar to ABC transporter	
Ecu-CU11_1200*	Similar to ABC transporter (ATM1)	
Ecu-CU04_0480	ABC transporter 7	
Ecu-CU07_0600*	Adrenodoxin	
Ecu-CU11_1770*	Similar to NFS1	
Ecu-CU05_1380*	Similar to hypoth. Prot. (glutaredoxinlike)	
Ecu-CU01_0510*	NIFU-like protein	
Ecu-CU05_1250	CDP diacyl glycerol synthase	
Ecu-CU05_1190	ABC transporter	
Ecu-CU04_0600	Proline aminopeptidase	
Ecu-CU05_0310	AcetylCoA synthetase	
Ecu-CU10_0710	DNA mismatch repair protein	
Ecu-CU06_0730	Ribonucleoside di-P reductase chain M2	
Ecu-CU11_0300	Similar to Dephospho-CoA kinase	n.p
Ecu-CU07_1320	Similar to YAGE-SCHPO	

Proteins within the same box belong to the same orthologous group and proteins involved in Fe–S cluster assembly are marked with an asterisk. Column on the right denote the mitochondrial localization of a yeast mitochondrial ortholog according to [50,57], whenever an orthologous yeast protein exists; otherwise, absence of a yeast ortholog is indicated by "n.p" (not present).

[74,75], however, support the relocation of some of the amitochondriates, such as microsporidia, certain amoebozoa and parabasalia, higher in the tree. Their lack of mitochondria is now considered the result of secondary adaptations to anaerobic environments. Recently subcellular structures that appear to be mitochondrial remnants have been described in microsporidia [76], cryptosporidia [77], amoebozoa [78] and diplomonads [25]. In addition, there are hydrogenosomes in parabasalids, ciliates and anaerobic fungi [79] that might represent forms of highly specialized—rather than degraded—mitochondria. The existence of these subcellular structures supports the hypothesis that mitochondrial endosymbiosis predated the diversification of the analyzed eukaryotic species. This view is also supported by the presence of proteins of a likely mitochondrial origin in all fully sequenced eukaryotes so far.

A search for proto-mitochondrial derived proteins in nine fully sequenced eukaryotic genomes [4] revealed an ubiquitous but variable distribution of ancient mitochondrial proteins, including 29 proteins in the microsporidian *Encephalitozoon cuniculi.* A large fraction (22) of the proto-mitochondrial derived proteins in *E. cuniculi* have yeast orthologs targeted to mitochondria, which suggests they might be part of the mitosome. These results are similar to those of a large-scale prediction of mitochondrial targeting signals in eukaryotic genomes [59] that predicted a mitochondrial localization for 156 proteins in *E. cuniculi*, 26 of which have alpha-proteobacterial homologs. Moreover, 18 of the proto-mitochondrial derived proteins, including those involved in Fe–S cluster assembly, can also be found in the recently sequenced genome of *C. parvum* [80], another parasite with a mitochondrial remnant (Table 1). In addition, other Fe–S cluster assembly genes and mitochondrial chaperonin-like genes have been found in amoebozoans, diplomonads, amitochondriate apicomplexans and parabasalids [81–84]. Some of these are targeted to the mitochondrial remnants.

The apparently ubiquitous presence of proto-mitochondrial derived proteins in virtually all eukaryotes supports the view of an early endosymbiosis at the root of the eukaryotic tree. Alternatively it might be argued that these proteins could have been gained via subsequent horizontal transfers. In the end, the greater plasticity of the mitochondrial proteome, when compared to that of the mitochondrial genome, makes it less useful in the search for the origin of highly reduced remnants. The discovery of a potential residual genome in the hydrogenosomes of the ciliate *Nyctoterus ovalis* [85] provides a way out of this dilemma. The initial analyses of this genome strongly suggest a mitochondrial ancestry for the hydrogenosomes in this species [86].

The emerging picture is that there is no extant eukaryote whose amitochondriate state can be categorically considered ancestral. If this is proven to be true, especially when a greater diversity of early-diverging eukaryotes are sequenced, it would have tremendous implications for the role that mitochondrial endosymbiosis played in the process of eukaryogenesis.

Acknowledgements

We are grateful to Charles Kurland for critically reviewing the manuscript. This work was supported in part by a grant from the Netherlands organization for Scientific Research (NWO).

References

[1] M.W. Gray, G. Burger, B.F. Lang, Mitochondrial evolution, Science 283 (1999) 1476–1481.
[2] S.B. Hedges, J.E. Blair, M.L. Venturi, J.L. Shoe, A molecular time scale of eukaryote evolution and the rise of complex multicellular life BMC, Evol. Biol. 4 (2004) 2.
[3] D. Yang, Y. Oyaizu, H. Oyaizu, G.J. Olsen, C.R. Woese, Mitochondrial origins, Proc. Natl. Acad. Sci. U. S. A. 82 (1985) 4443–4447.
[4] T. Gabaldón, M.A. Huynen, Reconstruction of the proto-mitochondrial metabolism, Science 301 (2003) 609.
[5] A.G. Tielens, C. Rotte, J.J. van Hellemond, W. Martin, Mitochondria as we don't know them, Trends Biochem. Sci. 27 (2002) 564–572.
[6] F. Voncken, B. Boxma, J. Tjaden, A. Akhmanova, M. Huynen, F. Verbeek, A.G. Tielens, I. Haferkamp, H.E. Neuhaus, G. Vogels, M. Veenhuis, J.H. Hackstein, Multiple origins of hydrogenosomes: functional and phylogenetic evidence from the ADP/ATP carrier of the anaerobic chytrid *Neocallimastix* sp, Mol. Microbiol. 44 (2002) 1441–1454.
[7] G. Burger, M.W. Gray, B.F. Lang, Mitochondrial genomes: anything goes, Trends Genet. 19 (2003) 709–716.
[8] J.N. Timmis, M.A. Ayliffe, C.Y. Huang, W. Martin, Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes, Nat. Rev., Genet. 5 (2004) 123–135.
[9] J. Nosek, L. Tomaska, Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy, Curr. Genet. 44 (2003) 73–84.
[10] J.M. Herrmann, Converting bacteria to organelles: evolution of mitochondrial protein sorting, Trends Microbiol. 11 (2003) 74–79.
[11] S.D. Dyall, M.T. Brown, P.J. Johnson, Ancient invasions: from endosymbionts to organelles, Science 304 (2004) 253–257.
[12] W. Martin, M. Müller, The hydrogen hypothesis for the first eukaryote, Nature 392 (1998) 37–41.
[13] C.G. Kurl, S.G. Andersson, Origin and evolution of the mitochondrial proteome, Microbiol. Mol. Biol. Rev. 64 (2000) 786–820.
[14] D. Moreira, P. López-García, Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis, J. Mol. Evol. 47 (1998) 517–530.
[15] W. Martin, M. Hoffmeister, C. Rotte, K. Henze, An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondria and hydrogenosomes), and their heterotrophic lifestyle, Biol. Chem. 382 (2001) 1521–1539.
[16] P. López-Garcia, D. Moreira, Metabolic symbiosis at the origin of eukaryotes, Trends Biochem. Sci. 24 (1999) 88–93.
[17] L. Margulis, Symbioses in Cell Evolution, W.H. Freeman, San Francisco, 1981.
[18] S.G. Andersson, A. Zomorodipour, J.O. Andersson, T. Sicheritz-Ponten, U.C. Alsmark, R.M. Podowski, A.K. Naslund, A.S. Eriksson, H.H. Winkler, C.G. Kurland, The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria, Nature 396 (1998) 133–140.

[19] V.V. Emelyanov, Evolutionary relationship of *Rickettsiae* and mitochondria, FEBS Lett. 501 (2001) 11–18.
[20] V.V. Emelyanov, Rickettsiaceae, rickettsia-like endosymbionts, and the origin of mitochondria, Biosci. Rep. 21 (2001) 1–17.
[21] M. Muller, W. Martin, The genome of *Rickettsia prowazekii* and some thoughts on the origin of mitochondria and hydrogenosomes, Bioessays 21 (1999) 377–381.
[22] O. Karlberg, B. Canback, C.G. Kurl, S.G. Andersson, The dual origin of the yeast mitochondrial proteome, Yeast 17 (2000) 170–187.
[23] G. Kispal, P. Csere, C. Prohl, R. Lill, The mitochondrial proteins Atm1p and Nfs1p are essential for biogenesis of cytosolic Fe/S proteins, EMBO J. 18 (1999) 3981–3989.
[24] C.P. Vivares, M. Gouy, F. Thomarat, G. Metenier, Functional and evolutionary analysis of a eukaryotic parasitic genome, Curr. Opin. Microbiol. 5 (2002) 499–505.
[25] J. Tovar, G. Leon-Ávila, L.B. Sanchez, R. Sutak, J. Tachezy, M. van der Giezen, M. Hernandez, M. Müller, J.M. Lucocq, Mitochondrial remnant organelles of *Giardia* function in iron–sulfur protein maturation, Nature 426 (2003) 172–176.
[26] M.J. LaGier, J. Tachezy, F. Stejskal, K. Kutisova, J.S. Keithly, Mitochondrial-type iron–sulfur cluster biosynthesis genes (IscS and IscU) in the apicomplexan *Cryptosporidium parvum*, Microbiology 149 (2003) 3519–3530.
[27] N. Wiedemann, A.E. Frazier, N. Pfanner, The protein import machinery of mitochondria, J. Biol. Chem. 279 (2004) 14473–14476.
[28] P. Marc, A. Margeot, F. Devaux, C. Blugeon, M. Corral-Debrinski, C. Jacq, Genome-wide analysis of mRNAs targeted to yeast mitochondria, EMBO Rep. 3 (2002) 159–164.
[29] R. Lucattini, V.A. Likic, T. Lithgow, Bacterial proteins predisposed for targeting to mitochondria, Mol. Biol. Evol. 21 (2004) 652–658.
[30] B.D. Lemire, C. Fankhauser, A. Baker, G. Schatz, The mitochondrial targeting function of randomly generated peptide sequences correlates with predicted helical amphiphilicity, J. Biol. Chem. 264 (1989) 20206–20215.
[31] K. Kadowaki, N. Kubo, K. Ozawa, A. Hirai, Targeting presequence acquisition after mitochondrial gene transfer to the nucleus occurs by duplication of existing targeting signals, EMBO J. 15 (1996) 6652–6661.
[32] H. Amiri, O. Karlberg, S.G. Andersson, Deep origin of plastid/parasite ATP/ADP translocases, J. Mol. Evol. 56 (2003) 137–150.
[33] P.D. Boyer, The ATP synthase—a splendid molecular machine, Annu. Rev. Biochem. 66 (1997) 717–749.
[34] S. Berry, Endosymbiosis and the design of eukaryotic electron transport, Biochim. Biophys. Acta 1606 (2003) 57–72.
[35] V. Guenebaut, A. Schlitt, H. Weiss, K. Leonard, T. Friedrich, Consistent structure between bacterial and mitochondrial NADH:ubiquinone oxidoreductase (complex I), J. Mol. Biol. 276 (1998) 105–112.
[36] U. Sackmann, R. Zensen, D. Rohlen, U. Jahnke, H. Weiss, The acyl-carrier protein in *Neurospora crassa* mitochondria is a subunit of NADH:ubiquinone reductase (complex I), Eur. J. Biochem. 200 (1991) 463–469.
[37] J. Carroll, I.M. Fearnley, R.J. Shannon, J. Hirst, J.E. Walker, Analysis of the subunit composition of complex I from bovine heart mitochondria, Mol. Cell Proteomics 2 (2003) 117–126.
[38] S. Brumme, V. Kruft, U.K. Schmitz, H.P. Braun, New insights into the co-evolution of cytochrome *c* reductase and the mitochondrial processing peptidase, J. Biol. Chem. 273 (1998) 13143–13149.
[39] R. Zensen, H. Husmann, R. Schneider, T. Peine, H. Weiss, De novo synthesis and desaturation of fatty acids at the mitochondrial acyl-carrier protein, a subunit of NADH:ubiquinone oxidoreductase in *Neurospora crassa*, FEBS Lett. 310 (1992) 179–181.
[40] F. Palmieri, Mitochondrial carrier proteins, FEBS Lett. 346 (1994) 48–54.
[41] M. van der Giezen, D.J. Slotboom, D.S. Horner, P.L. Dyal, M. Harding, G.P. Xue, T.M. Embley, E.R. Kunji, Conserved properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles, EMBO J. 21 (2002) 572–579.
[42] J.J. van Hellemond, A. van der Klei, S.W. van Weelden, A.G. Tielens, Biochemical and evolutionary aspects of anaerobically functioning mitochondria, Philos. Trans. R. Soc. Lond., B Biol. Sci. 358 (2003) 205–213 (discussion 213–5).
[43] J. Tjaden, I. Haferkamp, B. Boxma, A.G. Tielens, M. Huynen, J.H. Hackstein, A divergent ADP/ATP carrier in the hydrogenosomes of *Trichomonas gallinae* argues for an independent origin of these organelles, Mol. Microbiol. 51 (2004) 1439–1446.
[44] S.D. Dyall, C.M. Koehler, M.G. Delgadillo-Correa, P.J. Bradley, E. Plumper, D. Leuenberger, C.W. Turck, P.J. Johnson, Presence of a member of the mitochondrial carrier family in hydrogenosomes: conservation of membrane-targeting pathways between hydrogenosomes and mitochondria, Mol. Cell. Biol. 20 (2000) 2488–2497.
[45] F.M. Canovas, E. Dumas-Gaudot, G. Recorbet, J. Jorrin, H.P. Mock, M. Rossignol, Plant proteome analysis, Proteomics 4 (2004) 285–298.
[46] D.E. Warnock, E. Fahy, S.W. Taylor, Identification of protein associations in organelles, using mass spectrometry-based proteomics, Mass Spectrom. Rev. 23 (2004) 259–280.
[47] S.W. Taylor, E. Fahy, S.S. Ghosh, Global organellar proteomics, Trends Biotechnol. 21 (2003) 82–88.
[48] M. Dreger, Proteome analysis at the level of subcellular structures, Eur. J. Biochem. 270 (2003) 589–599.
[49] V.K. Mootha, J. Bunkenborg, J.V. Olsen, M. Hjerrild, J.R. Wisniewski, E. Stahl, M.S. Bolouri, H.N. Ray, S. Sihag, M. Kamal, N. Patterson, E.S. Lander, M. Mann, Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria, Cell 115 (2003) 629–640.
[50] A. Sickmann, J. Reinders, Y. Wagner, C. Joppich, R. Zahedi, H.E. Meyer, B. Schonfisch, I. Perschil, A. Chacinska, B. Guiard, P. Rehling, N. Pfanner, C. Meisinger, The proteome of Saccharomyces cerevisiae mitochondria, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 13207–13212.
[51] D. Cotter, P. Guda, E. Fahy, S. Subramaniam, MitoProteome: mitochondrial protein sequence database and annotation system, Nucleic Acids Res. 32 (2004) D463–D467 (Database issue).
[52] S.W. Taylor, E. Fahy, B. Zhang, G.M. Glenn, D.E. Warnock, S. Wiley, A.N. Murphy, S.P. Gaucher, R.A. Capaldi, B.W. Gibson, S.S. Ghosh, Characterization of the human heart mitochondrial proteome, Nat. Biotechnol. 21 (2003) 281–286.
[53] N. Tanaka, M. Fujita, H. Handa, S. Murayama, M. Uemura, Y. Kawamura, T. Mitsui, S. Mikami, Y. Tozawa, T. Yoshinaga, S. Komatsu, Proteomics of the rice cell: systematic identification of the protein populations in subcellular compartments, Mol. Genet. Genomics (2004) 566–576.
[54] J.L. Heazlewood, J.S. Tonti-Filippini, A.M. Gout, D.A. Day, J. Whelan, A.H. Millar, Experimental analysis of the *Arabidopsis* mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins, Plant Cell 16 (2004) 241–256.
[55] A. Kumar, S. Agarwal, J.A. Heyman, S. Matson, M. Heidtman, S. Piccirillo, L. Umansky, A. Drawid, R. Jansen, Y. Liu, K.H. Cheung, P. Miller, M. Gerstein, G.S. Roeder, M. Snyder, Subcellular localization of the yeast proteome, Genes Dev. 16 (2002) 707–719.
[56] W.K. Huh, J.V. Falvo, L.C. Gerke, A.S. Carroll, R.W. Howson, J.S. Weissman, E.K. O'Shea, Global analysis of protein localization in budding yeast, Nature 425 (2003) 686–691.
[57] E. Schon, in: P.M. Leslie Wilson (Ed.), Methods Cell Biol., Academic Press, New York, 2001, pp. 463–482.
[58] H. Prokisch, C. Scharfe, D.G. Camp 2nd, W. Xiao, L. David, C. Andreoli, M.E. Monroe, R.J. Moore, M.A. Gritsenko, C. Kozany, K.K. Hixson, H.M. Mottaz, H. Zischka, M. Ueffing, Z.S. Herman, R.W. Davis, T. Meitinger, P.J. Oefner, R.D. Smith, L.M. Steinmetz,

Integrative analysis of the mitochondrial proteome in yeast PLoS, Biology 2 (2004) E160.
[59] E. Richly, P.F. Chinnery, D. Leister, Evolutionary diversification of mitochondrial proteomes: implications for human disease, Trends Genet. 19 (2003) 356–362.
[60] C. Guda, E. Fahy, S. Subramaniam, MITOPRED: a genome-scale method for prediction of nucleus-encoded mitochondrial proteins, Bioinformatics (2004) W372–W374.
[61] I.E. Scheffler, Mitochondria make a come back, Adv. Drug Deliv. Rev. 49 (2001) 3–26.
[62] R. Lill, G. Kispal, Maturation of cellular Fe–S proteins: an essential function of mitochondria, Trends Biochem. Sci. 25 (2000) 352–356.
[63] S. Ohta, A multi-functional organelle mitochondrion is involved in cell death, proliferation and disease, Curr. Med. Chem. 10 (2003) 2485–2494.
[64] M. Neuburger, F. Rebeille, A. Jourdain, S. Nakamura, R. Douce, Mitochondria are a major site for folate and thymidylate synthesis in plants, J. Biol. Chem. 271 (1996) 9466–9472.
[65] L.M. Steinmetz, C. Scharfe, A.M. Deutschbauer, D. Mokranjac, Z.S. Herman, T. Jones, A.M. Chu, G. Giaever, H. Prokisch, P.J. Oefner, R.W. Davis, Systematic screen for human disease genes in yeast, Nat. Genet. 31 (2002) 400–404.
[66] C.G. Bowsher, A.K. Tobin, Compartmentation of metabolism within mitochondria and plastids, J. Exp. Bot. 52 (2001) 513–527.
[67] R. Srinivasan, D.J. Oliver, Light-dependent and tissue-specific expression of the H-protein of the glycine decarboxylase complex, Plant Physiol. 109 (1995) 161–168.
[68] P. Maechler, Mitochondria as the conductor of metabolic signals for insulin exocytosis in pancreatic beta-cells, Cell. Mol. Life Sci. 59 (2002) 1803–1818.
[69] T.J. Rosol, J.T. Yarrington, J. Latendresse, C.C. Capen, Adrenal gland: structure, function, and mechanisms of toxicity, Toxicol. Pathol. 29 (2001) 41–48.
[70] A.D. Baughn, M.H. Malamy, A mitochondrial-like aconitase in the bacterium *Bacteroides fragilis*: implications for the evolution of the mitochondrial Krebs cycle, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 4662–4667.
[71] B.F. Lang, G. Burger, C.J. O'Kelly, R. Cedergren, G.B. Golding, C. Lemieux, D. Sankoff, M. Turmel, M.W. Gray, An ancestral mitochondrial DNA resembling a eubacterial genome in miniature, Nature 387 (1997) 493–497.
[72] J.A. Korhonen, M. Gaspari, M. Falkenberg, TWINKLE Has 5′ →3′ DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein, J. Biol. Chem. 278 (2003) 48627–48632.
[73] T. Cavalier-Smith, E.E. Chao, Molecular phylogeny of the free-living archezoan *Trepomonas agilis* and the nature of the first eukaryote, J. Mol. Evol. 43 (1996) 551–562.
[74] T. Cavalier-Smith, The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa, Int. J. Syst. Evol. Microbiol. 52 (2002) 297–354.
[75] A. Stechmann, T. Cavalier-Smith, Rooting the eukaryote tree by using a derived gene fusion, Science 297 (2002) 89–91.
[76] B.A. Williams, R.P. Hirt, J.M. Lucocq, T.M. Embley, A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*, Nature 418 (2002) 865–869.
[77] C.E. Riordan, J.G. Ault, S.G. Langreth, J.S. Keithly, *Cryptosporidium parvum* Cpn60 targets a relict organelle, Curr. Genet. 44 (2003) 138–147.
[78] J. Tovar, A. Fischer, C.G. Clark, The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*, Mol. Microbiol. 32 (1999) 1013–1021.
[79] T.M. Embley, M. van der Giezen, D.S. Horner, P.L. Dyal, P. Foster, Mitochondria and hydrogenosomes are two forms of the same fundamental organelle, Philos. Trans. R. Soc. Lond., B Biol. Sci. 358 (2003) 191–201 (discussion 201–2).
[80] M.S. Abrahamsen, T.J. Templeton, S. Enomoto, J.E. Abrahante, G. Zhu, C.A. Lancto, M. Deng, C. Liu, G. Widmer, S. Tzipori, G.A. Buck, P. Xu, A.T. Bankier, P.H. Dear, B.A. Konfortov, H.F. Spriggs, L. Iyer, V. Anantharaman, L. Aravind, V. Kapur, Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*, Science 304 (2004) 441–445.
[81] N. Arisue, L.B. Sanchez, L.M. Weiss, M. Müller, T. Hashimoto, Mitochondrial-type hsp70 genes of the amitochondriate protists, *Giardia intestinalis*, *Entamoeba histolytica* and two microsporidians, Parasitol. Int. 51 (2002) 9–16.
[82] A. Germot, H. Philippe, H. Le Guyader, Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 14614–14617.
[83] A. Germot, H. Philippe, Critical analysis of eukaryotic phylogeny: a case study based on the HSP70 family, J. Eukaryot. Microbiol. 46 (1999) 116–124.
[84] J. Slapeta, J.S. Keithly, *Cryptosporidium parvum* mitochondrial-type HSP70 targets homologous and heterologous mitochondria, Eukaryotic Cell 3 (2004) 483–494.
[85] A. Akhmanova, F. Voncken, T. van Alen, A. van Hoek, B. Boxma, G. Vogels, M. Veenhuis, J.H. Hackstein, A hydrogenosome with a genome, Nature 396 (1998) 527–528.
[86] A.H. van Hoek, A.S. Akhmanova, M.A. Huynen, J.H. Hackstein, A mitochondrial ancestry of the hydrogenosomes of *Nyctotherus ovalis*, Mol. Biol. Evol. 17 (2000) 202–206.
[87] S.L. Baldauf, A.J. Roger, I. Wenk-Siefert, W.F. Doolittle, A kingdom-level phylogeny of eukaryotes based on combined protein data, Science 290 (2000) 972–977.