



HAL
open science

On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes.

Elie Desmond-Le Quéméner, Celine Brochier-Armanet, Patrick Forterre, Simonetta Gribaldo

► To cite this version:

Elie Desmond-Le Quéméner, Celine Brochier-Armanet, Patrick Forterre, Simonetta Gribaldo. On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes.. *Research in Microbiology*, 2011, 162 (1), pp.53-70. 10.1016/j.resmic.2010.10.004 . hal-00658502

HAL Id: hal-00658502

<https://hal.science/hal-00658502v1>

Submitted on 13 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes

Elie Desmond^a, Celine Brochier-Armanet^{b,c}, Patrick Forterre^{a,d}, Simonetta Gribaldo^{a,*}

^a Institut Pasteur, Unité BMGE, Department of Microbiology, 28 rue du Dr Roux, 75015 Paris, France

^b Université de Provence, Aix-Marseille I, place Victor Hugo, 13331 Marseille cedex 3, France

^c Laboratoire de chimie bactérienne, CNRS UPR9043, IFR88, 31 chemin Joseph Aiguier, 13402 Marseille, France

^d Biologie Moléculaire du Gène chez les Extrémophiles (BMGE), Institut de Génétique et Microbiologie (IGM) Bat 409, Université Paris-Sud, Centre d'Orsay, 91405 Orsay Cedex, France

Received 10 September 2010; accepted 4 October 2010

Available online 27 October 2010

Abstract

Understanding early evolution is a major challenge for the post-genomic era. A promising way to tackle this issue is to analyze the evolutionary history of key cellular systems through phylogenomic approaches. The current availability of genomic data from representatives of diverse lineages (especially eukaryotes), together with the ever growing number of proteomic characterizations now provides ample material to apply this type of analyses to trace back the origin and evolution of the three domains of life. Here, we have reconstructed the composition of the ancestral mitochondrial ribosome in the Last Eukaryotic Common Ancestor (LECA) and investigated its subsequent evolution in six major eukaryotic supergroups. We infer that LECA possessed a mitochondrial ribosome that was already much larger than its bacterial ancestor, with 19 additional specific proteins, indicating that a certain amount of time occurred between initial endosymbiosis at the origin of the mitochondrion and the diversification of present-day eukaryotic supergroups. Subsequently, mitochondrial ribosomes appear to have undergone a very dynamic evolutionary history in the different eukaryotic lineages, involving the loss of different sets of ribosomal protein-coding genes, their transfer to the host genome, as well as the acquisition of many novel components. This chaotic history for a such fundamental cellular machinery is puzzling, especially when compared to cytosolic, bacterial or chloroplastic ribosomes, which are much more stable. Intriguingly, archaeal ribosomes also show a very dynamic nature, with multiple independent losses among lineages.

© 2010 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.

Keywords: Phylogenomics; Ribosome; Mitochondria; LECA; Archaea; Eukaryotes

1. Introduction

The place of eukaryotes in the tree of life and the very origin of this domain remains one of the most intriguing issues in evolution and is subject of lively debate (Gribaldo et al., 2010). In particular, their sistership with Archaea in rooted universal trees has led to the common assumption that Eukaryotes arose from a common ancestor with Archaea (Brown and Doolittle, 1995; Gogarten et al., 1989; Gribaldo and Cammarano, 1998; Iwabe et al., 1989). Alternatively, it

has been proposed that Eukaryotes derive from a *bona fide* archaeon and should therefore emerge from within the Archaea in universal trees (see (Martin, 2005) and references therein). Much effort is made to elucidate the relationships between Archaea and Eucarya in order to test these two hypotheses using classical phylogenetic approaches (Ciccarelli et al., 2006; Cox et al., 2008; Foster et al., 2009; Gribaldo et al., 2010; Harris et al., 2003; Rivera and Lake, 2004; Yutin et al., 2008). However, an alternative way to tackle the issue is to try tracing back the nature of the Last Eukaryotic Common Ancestor (also called LECA), as this can provide important insights into the early evolution of this domain of life and on its origin (Koonin, 2010). Over the last 15 years,

* Corresponding author.

E-mail address: simonetta.gribaldo@pasteur.fr (S. Gribaldo).

the explosive increase of genomic data from eukaryotes, especially protists, coupled to the improvement of molecular phylogeny tools and approaches such as multigene analyses (Burki et al., 2008; Delsuc et al., 2005; Hackett et al., 2007; Hampl et al., 2009; Rodriguez-Ezpeleta et al., 2005; Rodriguez-Ezpeleta et al., 2007) have profoundly changed our vision of the evolution of the domain of life to which we belong. Taken together, these results suggest a rapid divergence of present-day eukaryotic lineages into seven major supergroups (Fig. 1): Opisthokonta (including Metazoa, Choanoflagellata and Fungi), Archaeplastida (including Glaucophyta, red algae, green algae and land plants), Rhizaria (including Radiolaria, Foraminifera, amoeboid flagellates), Alveolata (including Apicomplexa and Ciliata), Heterokonta (including glaucophyta, diatoms, golden and brown algae), Amoebozoa (including amoebas and slime molds), Excavata (including Diplomonads and Parabasalia, Jacobids, Heterolobosea, Euglenozoa/Kinetoplastids). Additionally, the position of major algal lineages (Haptophyta and Cryptophyta) remains uncertain, but a grouping with Rhizaria, Alveolata and Heterokonta has been recently suggested (Burki et al., 2009). Regarding the order of emergence of these major eukaryotic groups, it has been proposed that the first dichotomy from LECA occurred between a lineage of uniciliates that gave rise to Opisthokonta and Amoebozoa (Unikonts) and a lineage of biciliate cells that gave rise to the five other supergroups (Bikonts) (Stechmann and Cavalier-Smith, 2002, 2003).

Even if important progresses on the phylogeny of eukaryotes have been made, the early steps in the evolution of this domain of life remain still largely unknown. Fortunately, the current number of complete genomes from nearly all major supergroups now allows the application of phylogenomics approaches to study the early evolution of eukaryotic cells (i.e. their metabolisms, protein complexes, cellular processes, etc.). Briefly, these approaches consist in investigating the taxonomic distribution and the phylogeny of homologues of the components of a particular cellular system in main representatives of the studied group for which complete genome sequences are available (Eisen and Fraser, 2003). Taken together, these data allow deciphering the evolutionary history of the system and therefore a better understanding of the evolutionary history of the organisms carrying it. Such strategy was successfully applied for instance to study the origin and evolution of nuclear pore complexes (Baptiste et al., 2005; Mans et al., 2004), cell division structures such as the midbody (Eme et al., 2009), membrane-trafficking systems (Dacks and Field, 2007), and unique metabolisms such as the pathway for sterol biosynthesis (Desmond and Gribaldo, 2009). As a starting point, phylogenomic approaches rely on the knowledge of the composition of the cellular system under investigation in at least one representative. These data are becoming increasingly available following the development of proteomics approaches, especially for eukaryotes. To illustrate the power of phylogenomic approaches, we present here a case study with the mitochondrial ribosome.

It is known that the evolution of mitochondria has been and is still very dynamic (Bullerwell and Gray, 2004; Burger et al.,

2003; Lang et al., 1999). For example, a number of genes coding for ancestral mitochondrial proteins of alphaproteobacterial origin have been lost, while others have been transferred to the host nucleus, their products being expressed in the cytoplasm and imported into the mitochondrion by a dedicated transport system (Lang et al., 1999). These losses/translocations have been very different in the different eukaryotic lineages. For example, three protein-coding genes only have been retained in the mitochondrial genome of *Plasmodium falciparum* while nearly 100 still reside in the mitochondrial genome of the Jacobid *Reclinomonas americana* (Gray et al., 2004). However, even extremely reduced mitochondrial genomes encode at least a few genes such as those coding for the components of respiration and coupled oxidative phosphorylation, or for transfer and ribosomal RNAs (Gray et al., 2004). The protein-coding genes that still reside in the mitochondrion are transcribed, and the resulting messenger RNAs are translated by the ribosomes (hereafter referred to as mitoribosomes) of alphaproteobacterial ancestry.

Mitoribosomes have undergone important changes during the evolution of the alphaproteobacterial symbiont into an established organelle (Burger and Lang, 2003; Gray et al., 2001). In particular, regions of ribosomal RNA components have been lost/deleted and have been compensated by the recruitment of eukaryotic-specific ribosomal proteins (O'Brien, 2003; Suzuki et al., 2001). These proteins are not homologous to cytosolic ribosomal proteins and represent *de novo* eukaryotic innovations, and will be hereafter referred to as Ek-MRPs (for Eukaryotic Mitochondrial Ribosomal Proteins). As shown by the few available proteomic characterizations, mammals and yeast mitoribosomes comprise approximately 70–80 MRPs (Gan et al., 2002; Graack and Wittmann-Liebold, 1998; Koc et al., 2001a; Koc et al., 2001b; O'Brien et al., 1999; O'Brien et al., 2000; Towpik, 2005), of which around 20 are eukaryotic-specific (Koc et al., 2001b). By contrast, only 57 ribosomal proteins have been identified in Bacteria (Lecompte et al., 2002), 54 of which are present in alphaproteobacteria and were therefore also likely present in the very first ancestor of mitochondria (these 54 mitochondrial ribosomal proteins of alphaproteobacterial ancestry will be hereafter referred to as Bt-MRPs, for Bacterial Mitochondrial Ribosomal Proteins). More specifically, a proteomic analysis of the mammalian mitoribosome has identified 48 MRPs in the large subunit (LSU) and 29 MRPs in the small subunit (SSU) (Koc et al., 2001a,b). A proteomic analysis of the mitoribosome of the yeast *Saccharomyces cerevisiae* has identified 44 MRPs in the LSU and 33 MRPs in the SSU, including 24 new components not present in the mammalian mitoribosome (Graack and Wittmann-Liebold, 1998; Saveanu et al., 2001). A similar analysis of the LSU of mitoribosome from a second fungus (*Neurospora crassa*) has confirmed the presence of most of the MRPs found in *S. cerevisiae* but failed to identify new ones (Gan et al., 2002). These analyses highlight great differences in the mitoribosome composition of these organisms despite the fact that they all belong to the supergroup of Opisthokonta. Two analyses performed on members of a second eukaryotic major group, the

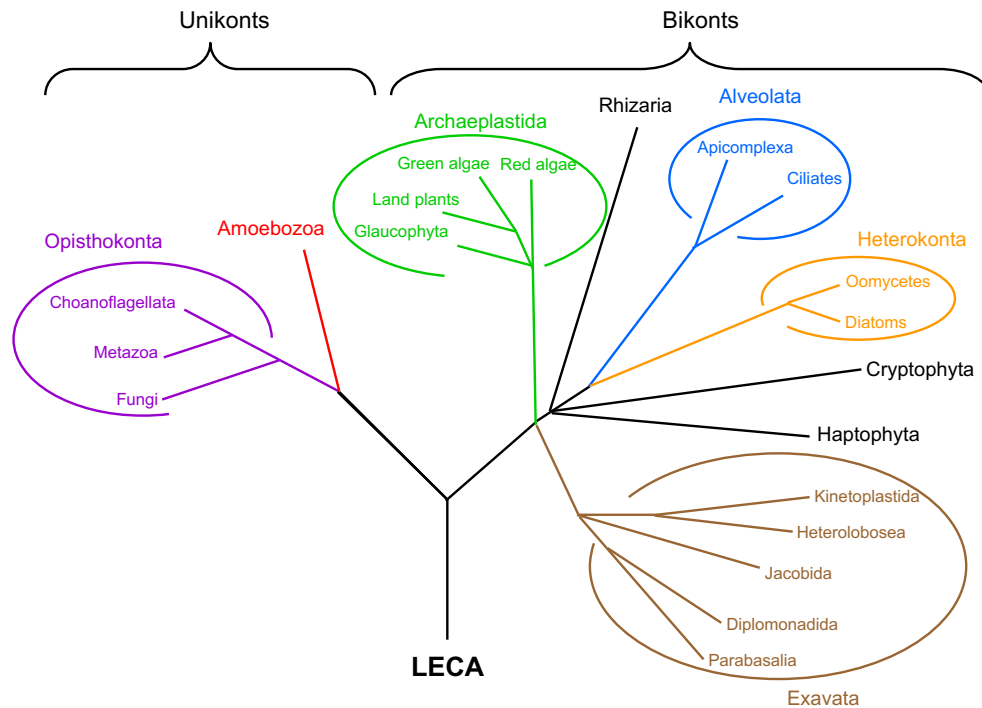


Fig. 1. Reference phylogenetic tree of Eukaryotes used to infer the evolution of the sets of MRPs during the evolutionary history of this domain.

Excavates, have shown that the inclusion of novel MRPs appears to be a general tendency in the evolution of mitochondria. In fact, a partial proteomic analysis of the mitochondria of the kinetoplastid *Leishmania tarentolae* identified 13 LSU MRPs, and 39 SSU MRPs of which 17 are unique to kinetoplastids (Maslov et al., 2006, 2007). More recently, a proteomic characterization of a mitochondria was obtained from a second representative of kinetoplastids, *Trypanosoma brucei* (Zikova et al., 2008). All MRPs characterized in *L. tarentolae* were confirmed, as well as 81 additional ones, reaching the impressive number of 133 MRPs (77 in LSU, 56 in SSU) identified with high confidence. Importantly only 38 of these have homologues to other eukaryotic MRPs, while 95 (51 LSU, 44 SSU) appear unique to Kinetoplastids (i.e. no homologues exist in other eukaryotic lineages).

Smits et al. (Smits et al., 2007) have recently investigated the distribution of MRP homologues in a sample of 17 eukaryotic genomes representatives of five supergroups. They inferred a mitochondria in LECA that contained already 16 novel Ek-MRPs with respect to its alphaproteobacterial ancestor, and reconstructed its subsequent evolution that involved large variations of protein content (both acquisitions and losses) in different lineages. Based on bioinformatic predictions, they also identified a number of previously unidentified MRPs homologues that may be responsible for mitochondria-related disorders in humans (Smits et al., 2007). Here, we have extended this study by analyzing a much larger taxonomic sampling consisting of 38 complete nuclear and mitochondria genomes including in particular representatives from a sixth major eukaryotic phylum (Heterokonts). This allowed to identify 22 MRPs previously undetected in a number of eukaryotic genomes and therefore to define with more precision the evolutionary history of the

mitochondria over eukaryotic evolution. This has been characterized by a very dynamic process, which appears to be still ongoing, and which contrasts with the overall stasis of cytosolic ribosomes. Curiously, archaeal ribosomes also display a similar chaotic history, involving mainly loss of proteins.

2. Material and methods

2.1. Dataset construction

Homologues of each of the 336 MRPs identified in the proteomic characterizations of mitochondria ribosomes of mammals (Koc et al., 2001a,b), yeast (Graack and Wittmann-Liebold, 1998; Saveanu et al., 2001), and kinetoplastids (Graack and Wittmann-Liebold, 1998; Koc et al., 2001a,b; Maslov et al., 2006, 2007; Saveanu et al., 2001) as well as homologues of the 54 alphaproteobacterial ribosomal proteins (Bt-MRP) were searched using Blastp (Altschul et al., 1997) and default parameters on the RefSeq database at the NCBI. At this preliminary stage, no global e-value cutoff was used to filter Blastp results, but we carefully examined these to select to select unambiguous homologues. These were aligned using MUSCLE 3.6 (Edgar, 2004). The MUST package (Philippe, 1993) was used to refine multiple alignments and to build preliminary neighbor-joining trees in order to clearly identify mitochondria sequences from their cytoplasmic or chloroplast homologues. MRP alignments containing only the mitochondria homologues were then used to build HMM profiles using the HMMER 2.3.2 tools (<http://hmmer.janelia.org/>). These profiles were then used to identify additional homologues (in particular very divergent ones) in a local sequence databank containing 38 complete nuclear and

mitochondrial genomes representing a broad taxonomic sampling of eukaryotes, 50 bacteria and 6 archaea (a complete list is provided in Supplementary Table 1). Genomes were retrieved from the NCBI (<ftp://ftp.ncbi.nih.gov/genomes/>), except for *Ostreococcus tauri* (Palenik et al., 2007), *Phytophthora ramorum* (Tyler et al., 2006), *Thalassiosira pseudonana* (Armbrust et al., 2004) and *Naegleria gruberi* which were retrieved from the JGI (http://genome.jgi-psf.org/euk_home.html), and *Cyanidioschyzon merolae* (Matsuzaki et al., 2004) which was retrieved from a dedicated website (<http://merolae.biol.s.u-tokyo.ac.jp>). The sequences of *Naegleria* were produced by the US Department of Energy Joint Genome Institute in collaboration with the user community (<http://www.jgi.doe.gov/>). The complete procedure (i.e. alignment reconstruction and refinement, reconstruction of preliminary neighbor-joining trees, construction of new HMM profiles, and search for additional mitochondrial homologues) was repeated until no additional homologues could be retrieved in the local database. Gene absences were systematically and carefully verified by additional searches on the corresponding complete genome sequences by tBLASTn.

2.2. Phylogenetic reconstructions

For final phylogenetic analyses, ambiguously aligned positions were manually removed using MUST. Bayesian phylogenetic trees were constructed using MrBayes 3.2 (Ronquist and Huelsenbeck, 2003). We used a mixed model of amino acid substitution and a Γ law (four discrete categories plus a proportion of invariant sites) to take into account among site rate variation. MrBayes was run with four chains for 1 million generations and trees were sampled every 100 generations. To construct the consensus tree, the first 1500 trees were discarded as “burnin”.

2.3. Inference of ancestral mitoribosome composition

Based on homologues present in analyzed eukaryotic genomes we inferred the pattern of gains and losses of MRPs along eukaryotic evolution, using a reference eukaryotic phylogenetic tree rooted between Bikonts and Unikonts. Our inference is based on a parsimony criterion minimizing the number of convergent translocation of Bt-MRPs coding genes from mitochondria to host nucleus and the number of convergent losses. We also assumed (i) that no convergent recruitment of Ek-MRPs occurred during the evolution of eukaryotes, (ii) that no gene reverse-translocation occurred (i.e. from the nucleus to mitochondria) and (iii) that no HGT of MRPs coding genes occurred between eukaryotic lineages. In practice, given that it is known that 54 Bt-MRPs were present in the ancestor of mitochondria, to infer the presence of a given Bt-MRP in the ancestor of a eukaryotic lineage it sufficed that the protein be present in at least one representative of its offspring. By contrast, we could infer the presence of a given Ek-MRP in the ancestor of a eukaryotic group only when homologues of the protein were present in at least one present-day representative from both sister lineages arising

from this ancestor. Finally, the absence of a given protein in a eukaryotic phylum was inferred only when no homologues were found in any of its present-day representatives. To avoid biases linked to our taxonomic sampling, we performed additional similarity searches in public databases to verify that no homologue of the protein was known in representatives of the phylum that were not included in our taxonomic sampling.

3. Results and discussion

We identified 22 MRPs in various eukaryotes that were previously unreported (Smits et al., 2007), both of alphaproteobacterial ancestry (Bt-MRPs) and eukaryotic-specific (Ek-MRPs) (Table 1, indicated by red boxes with X symbols). This probably results on the larger spectrum of homologues that was used to build the HMM model, likely improving detection (especially for small proteins). Moreover, our phylogenetic analysis indicates that a number of proteins previously proposed to be MRPs (Smits et al., 2007) are in fact cytoplasmic copies (see below). It should be noted that Ppe1 was reported to function as a nuclear or cytosolic phosphatase methylesterase in yeast and human whereas it was designated previously as an MRP of the small subunit (SSU) of the yeast mitoribosome (see (Smits et al., 2007) and reference therein). In addition, whereas MRPs are systematically absent in amitochondriate protists, a homologue of this protein is present in the genomes of two of them (*Giardia* and *Trichomonas*, Table 1b). Either this is a unique case of recruitment of a former MRP in amitochondriates, or else this protein was mis-annotated and does not function in the mitoribosome, as already suggested (Smits et al., 2007). In the absence of confirmation by proteomic analysis, we considered Ppe1 as a *bona fide* MRP in our analysis. Other than this exception, there was no ambiguity on the function of all other proteins analyzed.

3.1. Early events following endosymbiosis (before LECA)

Alphaproteobacteria harbor 54 ribosomal proteins (Bt-MRPs, 33 in the LSU and 21 in the SSU, Table 1) that were therefore also likely present in the alphaproteobacterial ancestor of mitochondria. We analyzed the fate of these ribosomal proteins from the time of the endosymbiosis event up to LECA. Despite a large eukaryotic taxonomic sampling, we confirm the absence of S20 in all the eukaryotic genomes (both nuclear and mitochondrial) analyzed (Table 1b), suggesting that this protein was lost very early in eukaryotic evolution (Fig. 2).

We also sought to reconstruct the events of gene translocation to the host nucleus prior to LECA. Based on the hypothesis that gene translocation occurred only in one direction (from the mitochondrion to the host nucleus), the 27 Bt-MRPs (15 LSU and 12 SSU) that are encoded in the mitochondrial genome of at least one species (Table 1) can be inferred to have been also encoded in the mitochondrial genome of LECA, i.e. before the diversification of present-day eukaryotic lineages (Fig. 2). It is noteworthy that all these 27 Bt-MRPs are still encoded in the

Table 1 (continued)

		Opisthokonta										Amoebozoa		Archaeplastida				Alveolata				Heterokonta		Excavata																											
		<i>M. brevicollis</i>	<i>D. melanogaster</i>	<i>T. castaneum</i>	<i>H. sapiens</i>	<i>M. musculus</i>	<i>D. rerio</i>	<i>C. elegans</i>	<i>A. fumigatus</i>	<i>N. crassa</i>	<i>S. cerevisiae</i>	<i>S. pombe</i>	<i>C. neoformans</i>	<i>U. maydis</i>	<i>E. curvicuti</i>	<i>E. histolytica</i>	<i>D. discoideum</i>	<i>A. thaliana</i>	<i>O. sativa</i>	<i>C. reinhardtii</i>	<i>O. tauri</i>	<i>C. merolae</i>	<i>P. tetraurelia</i>	<i>T. thermophila</i>	<i>C. parvum</i>	<i>P. falciparum</i>	<i>P. yeolii</i>	<i>T. annulata</i>	<i>P. ramorum</i>	<i>T. pseudonana</i>	<i>G. lamblia</i>	<i>T. vaginalis</i>	<i>T. cruzi</i>	<i>T. brucei</i>	<i>L. infantum</i>	<i>L. major</i>	<i>L. braziliensis</i>	<i>N. gruberi</i>	<i>mito. Reclinomonas</i>												
Tb927.6.4560	LmjF30.3220																																																		
Tb927.6.4580																																																			
Tb927.6.4930	LmjF30.3530																																																		
Tb927.7.2620	LmjF22.0830																																																		
Tb927.7.3050	LmjF22.0990																																																		
Tb927.7.3240	LmjF22.1220																																																		
Tb927.8.1430																																																			
Tb927.8.4550																																																			
Tb927.8.4860																																																			
Tb927.8.5200	LmjF16.1230																																																		
Tb09.160.3800	LmjF15.1400																																																		
Tb09.160.4750	LmjF15.0780																																																		
Tb09.211.2430	LmjF35.3940																																																		
Tb09.211.3000	LmjF35.3490																																																		
Tb09.211.4510	LmjF35.2230																																																		
Tb10.26.0050																																																			
Tb10.26.0750																																																			
Tb10.61.0190																																																			
Tb10.61.0690																																																			
Tb10.61.3110																																																			
Tb10.70.3840	LmjF03.0630																																																		
Tb10.70.4220	LmjF03.0220																																																		
Tb11.01.1910																																																			
Tb11.01.2170	LmjF36.4120																																																		
Tb11.01.3300	LmjF28.2180																																																		
Tb11.01.3710																																																			
Tb11.01.7510																																																			
Tb11.02.0130	LmjF33.2510																																																		
Tb11.02.3180	LmjF24.0830																																																		
Tb11.47.0024	LmjF27.0630																																																		

mitochondrial genome of the Jacobid *Reclinomonas americana*, of which nine are not found in any other mitochondrial genome, underscoring the uniqueness of the organelle in this protist. Indeed, if the mitochondrial genome sequence of *Reclinomonas* were not included in the analysis, only 18 Bt-MRPs (seven LSU and 11 SSU) instead of 27 would have been inferred as encoded in the mitochondrial genome of LECA. The presence of 27 Bt-MRPs in the mitochondrial genome of LECA is therefore a minimum estimate and it cannot be excluded that it will increase when additional mitochondrial genome sequences from a wider range of protists become available.

On the contrary, the 26 Bt-MRPs that are not encoded in any currently available mitochondrial genome (including that of *Reclinomonas*) might have been transferred to the host nuclear genome prior to the divergence of major eukaryotic supergroups (Table 1 and Fig. 2). This observation may be explained by two hypotheses: (i) considerable time elapsed between the endosymbiosis and the diversification of present-day eukaryotic lineages or (ii) the translocation of Bt-MRPs to the host nucleus occurred rapidly after endosymbiosis. A third possibility that cannot be totally ruled out is the independent and convergent translocation of the genes coding for these Bt-MRPs to the nucleus after the divergence of present-day eukaryotic lineages. We therefore sought to gather additional clues on the order and the timing of transfer from the ancestral mitochondrial genome

to the host nucleus by looking at bacterial endosymbionts involved in genomic reduction processes than might be similar to that undergone by the ancestor of mitochondria (Tamames et al., 2007). This is the case for four organisms belonging to distantly related lineages: the gammaproteobacteria *Candidatus Carsonella ruddii* PV (endosymbiont of psyllids; 0.16 Mb and 182 protein-coding genes (Nakabachi et al., 2006)) and *Buchnera aphidicola* str. Cc (endosymbiont of the cedar aphid, *Cinara cedri*; 0.42 Mb and 357 coding genes (Perez-Brocail et al., 2006)), the alphaproteobacterium *Candidatus Hodgkiania cicadicola* Dsem (symbiont of the cicada *Diceroprocta semicineta*; 0.14 Mb and 169 coding genes (McCutcheon et al., 2009)) and the flavobacterium *Candidatus Sulcia muelleri* DMIN (gut symbiont of the Blue-Green Sharpshooter and several other leafhopper species; 0.24 Mb and 226 coding genes, unpublished). We found that, except for *B. aphidicola*, these endosymbiont genomes lack a large number of ribosomal protein-coding genes (Table 2). Interestingly, the number of ribosomal proteins in these symbionts correlates well with the size of their genomes. More precisely, the two smallest genomes (*Candidatus C. ruddii* and in *Candidatus H. cicadicola*) encode only 40 ribosomal proteins, whereas 48 are present in *Candidatus S. muelleri* (Table 2). It should be noted that, despite an important reduction of its genome (0.42 Mb and 357 protein-coding genes), no losses of ribosomal proteins are observed in

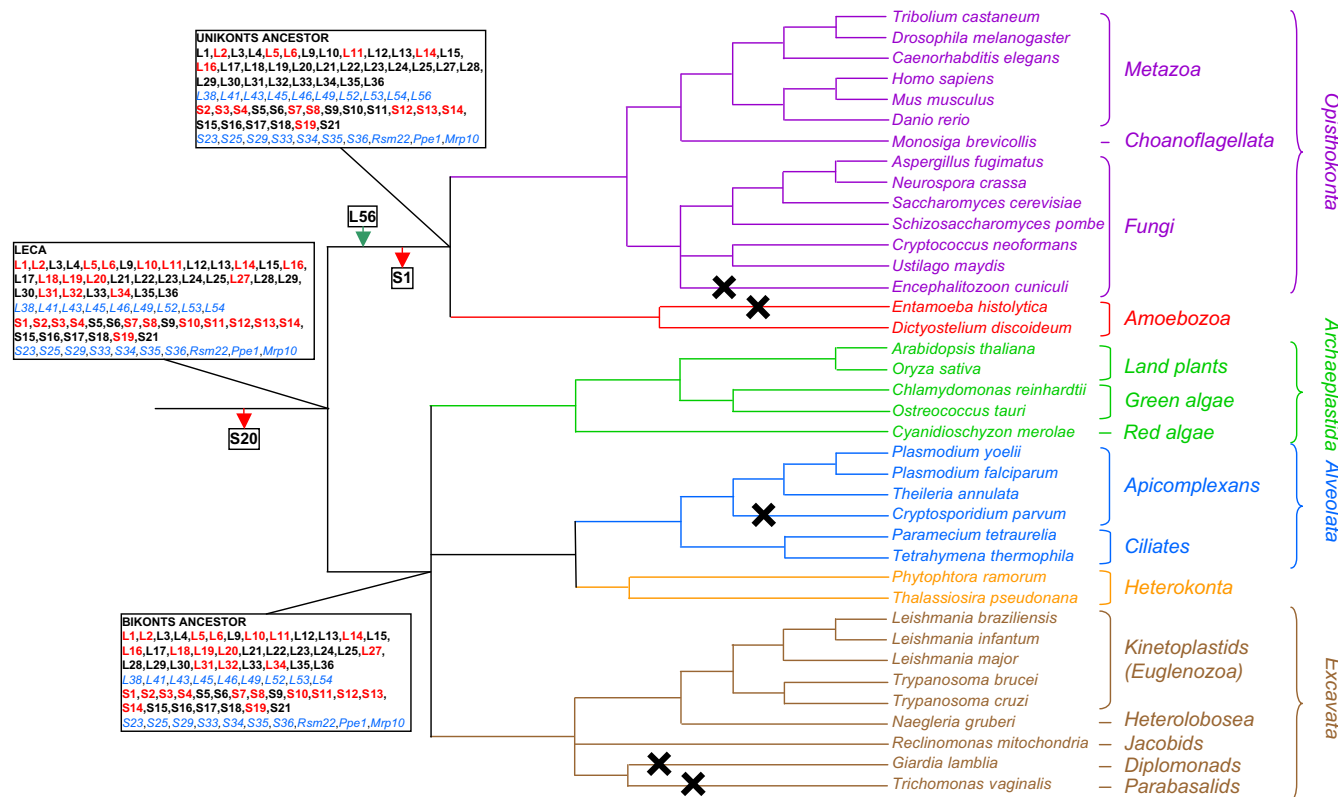


Fig. 2. Inference of the set of MRPs present in the last eukaryotic common ancestor and its subsequent evolution in Unikonts and Bikonts. MRP acquisitions are indicated by green arrows, whereas losses are highlighted by red arrows. Proteins of alphaproteobacterial origin encoded in mitochondria are indicated in bold. Bt-MRPs predicted to be encoded in mitochondria are indicated in red, whereas those encoded in the nucleus are indicated in black. EK-MRPs are indicated in blue italic font. Lineages that have lost their mitochondria are marked with crosses; in these cases no ribosomal proteins are detected in their genomes, to the exception of Ppe1 (but see text for discussion).

B. aphidicola (Table 2). This strongly suggests that ribosomal proteins are likely among the last to be lost during the process of genome reduction after the endosymbiotic event. The flavobacterium *Candidatus Sulcia muelleri* has lost six ribosomal proteins, and all of these are also missing in *Candidatus Carsonella ruddii* and *Candidatus Hodgkinia cicadicola*. The absence of ribosomal proteins in the three extremely reduced endosymbiont genomes is puzzling and has been discussed previously (Tamames et al., 2007). In fact, since these losses are not compatible with a viable ribosome, it has been put forward that the genes coding for the missing components (as other missing genes involved in important functions) have been transferred to the host genomes, and their products, like in the case of mitochondria, are now imported into the symbionts (Nakabachi et al., 2006). However, a recent analysis of the genome of *Cinara cedri* has shown that the genome reduction of its symbiont, *B. aphidicola*, has not been accompanied by gene translocation to the host nuclear genome (Nikoh et al., 2010). Alternatively, it has been proposed that key missing proteins in reduced endosymbionts might be complemented by the mitochondrial proteins of the host (Moya et al., 2008; Tamames et al., 2007). In agreement with this hypothesis, mitochondrial homologues of all the missing ribosomal proteins in these endosymbionts are encoded in the current available genomes of insects (Tables 1 and 2, and Fig. 3) and thus probably also in the genome of their hosts. However, the two Bt-MRPs (S20 and

L25) that have been lost early in metazoan evolution (see below) should be considered as definitely missing from the ribosomes of two of these three symbionts (i.e. *Candidatus Carsonella ruddii* and *Candidatus Hodgkinia cicadicola*), because these cannot be re-imported from the host genome. It has been also proposed that the loss of key proteins may be complemented by proteins coming from co-endosymbionts (Moya et al., 2008). This is a very interesting hypothesis to test because it would require a dedicated export/import system between two endosymbionts.

Interestingly, five of the ribosomal proteins that have been lost from the genomes of these endosymbionts are also missing from the inferred mitochondrial genome of LECA (Table 2). Moreover, three additional protein losses are shared between the two smallest genomes (*Candidatus Carsonella ruddii* and *Candidatus Hodgkinia cicadicola*) and the mitochondrion (Table 2). These results suggest that these proteins may be among the first to be lost/translocated during the establishment of endosymbiosis. It is therefore possible that the same proteins had already been translocated to the host genome of LECA after mitochondrial endosymbiosis. Moreover, given that the loss/translocation of ribosomal proteins appears to be a late event in endosymbiosis (as testified by their presence in the least reduced genome of *Buchnera*), it is possible that in LECA also the mitochondrial endosymbiosis was already at a late stage.

Table 2

Comparison of the 54 Bt-MRPs with the set of proteins inferred to have been encoded in the mitochondrial genome of LECA (indicated by an M) and the ribosomal proteins present in extremely reduced genomes of four bacterial endosymbionts of insects: the gammaproteobacteria *Candidatus Carsonella ruddii* PV (0.16 Mb, endosymbiont of psyllids) and *Buchnera aphidicola* str. Cc (0.42 Mb, endosymbiont of the cedar aphid, *Cinara cedri*), the alphaproteobacterium *Candidatus Hodgkinia cicadicola* Dsem (0.14 Mb, symbiont of the cicada *Diceroprocta semicincta*) and the flavobacterium *Candidatus Sulcia muelleri* DMIN (0.24 Mb, gut symbiont of the Blue-Green Sharpshooter and several other leafhopper species).

	<i>Candidatus carsonella ruddii</i>	<i>Buchnera aphidicola</i> Cc	<i>Candidatus Hodgkinia cicadicola</i> Dsem	<i>Candidatus Sulcia muelleri</i> DMIN	LECA
LSU					
L1	YP_802562	YP_802600		YP_003543227	M
L2	YP_802550	YP_802885	YP_003038714	YP_003543364	M
L3	YP_802552	YP_802888	YP_003038712	YP_003543361	
L4	YP_802551	YP_802887	YP_003038713	YP_003543362	
L5	YP_802543	YP_802876	YP_003038722	YP_003543371	M
L6	YP_802540	YP_802873	YP_003038725	YP_003543374	M
L7/L12	YP_802560	YP_802598	YP_003038704	YP_003543226	
L9		YP_802912			
L10	YP_802561	YP_802599	YP_003038703	YP_003543226	M
L11	YP_802563	YP_802601	YP_003038702	YP_003543228	M
L13		YP_802804	YP_003038663	YP_003543245	
L14	YP_802544	YP_802878	YP_003038720	YP_003543370	M
L15	YP_802537	YP_802869	YP_003038728	YP_003543377	
L16	YP_802546	YP_802881	YP_003038718	YP_003543368	M
L17	YP_802531	YP_802862	YP_003543385	YP_003038732	
L18	YP_802539	YP_802872	YP_003038726	YP_003543375	M
L19		YP_802810	YP_003038743	YP_003543287	M
L20	YP_802487	YP_802651	YP_003038794	YP_003543249	M
L21	YP_802530	YP_802800		YP_003543312	
L22	YP_802548	YP_802883	YP_003038716	YP_003543366	
L23		YP_802886			
L24		YP_802877			
L25		YP_802658		YP_003543201	
L27	YP_802529	YP_802801	YP_003038762	YP_003543313	M
L28	YP_802513	YP_802624	YP_003038786	YP_003543219	
L29		YP_802880			
L30		YP_802870			
L31	YP_802423	YP_802918		YP_003543356	M
L32	YP_802492	YP_802777		YP_003543349	M
L33	YP_802512	YP_802623	YP_003038697	YP_003543220	
L34		YP_802584			M
L35		YP_802650	YP_003038793	YP_003543248	
L36	YP_802536	YP_802867	YP_003038811	YP_003543380	
SSU					
S1	YP_802447	YP_802755	YP_003038779	YP_003543184	M
S2	YP_802417	YP_802711	YP_003038781	YP_003543243	M
S3	YP_802547	YP_802882	YP_003038717	YP_003543367	M
S4	YP_802533	YP_802864	YP_003038803	YP_003543383	M
S5	YP_802538	YP_802871	YP_003038727	YP_003543376	
S6		YP_802914	YP_003038787	YP_003543314	
S7	YP_802556	YP_802892	YP_003038708	YP_003543358	M
S8	YP_802541	YP_802874	YP_003038724	YP_003543373	M
S9	YP_802443	YP_802803	YP_003038664	YP_003543244	
S10	YP_802553	YP_802889	YP_003038711	YP_003543360	M
S11	YP_802534	YP_802865	YP_003038730	YP_003543382	M
S12	YP_802557	YP_802893	YP_003038707	YP_003543357	M
S13	YP_802535	YP_802866	YP_003038729	YP_003543381	M
S14	YP_802542	YP_802875	YP_003038723	YP_003543372	M
S15	YP_802466	YP_802790		YP_003543354	
S16	YP_802521	YP_802807	YP_003038742	YP_003543345	
S17	YP_802545	YP_802879	YP_003038719	YP_003543369	
S18		YP_802913	YP_003038788	YP_003543315	
S19	YP_802549	YP_802884	YP_003038715	YP_003543365	M
S20		YP_802664		YP_003543343	
S21		YP_802610		YP_003543321	

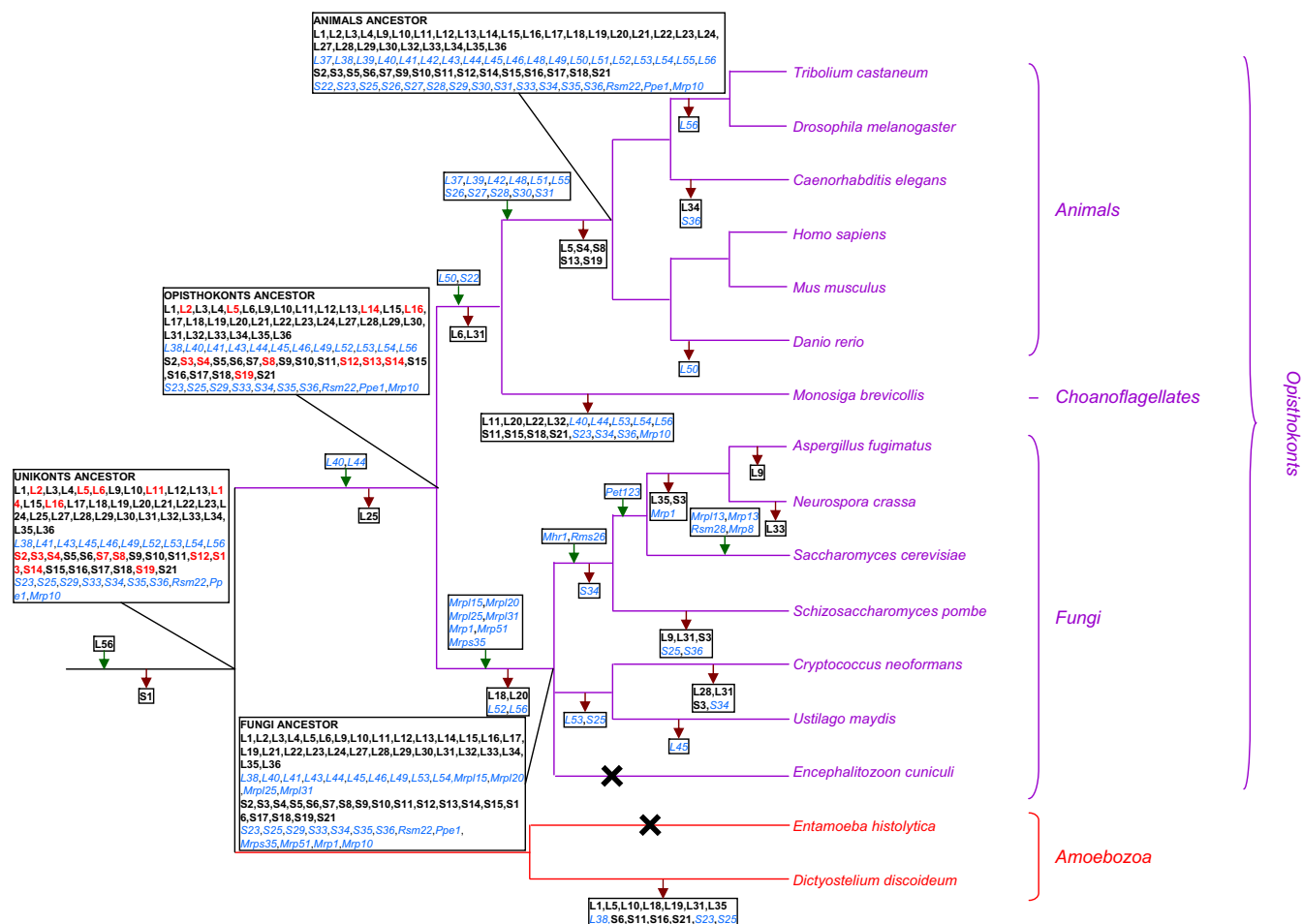


Fig. 3. Inference of the set of MRP genes in the last common ancestor of Unikonts and its subsequent evolution in Metazoa, Fungi and Amoebozoa. For details see legend to Fig. 2.

The loss of genes is a frequently observed phenomenon in endosymbiotic events (Moya et al., 2008), but the mitochondrial endosymbiosis appears to be special in having involved both losses and gains. Concerning the mitoribosome, 19 eukaryotic-specific proteins (Ek-MRPs, 9 in the LSU and 10 in the SSU) are present in at least one representative of a Unikont and a Bikont and thus must have appeared before LECA. These proteins were therefore added to the complement of initial Bt-MRPs (Table 1), meaning that the LECA mitoribosome already contained at least 72 proteins (42 in the LSU and 30 in the SSU), 18 proteins more than the original alphaproteobacterial ribosome. These 72 MRPs include 53 Bt-MRPs (26 of which may have been already translocated to the nuclear genome) and 19 Ek-MRPs (Fig. 2). Taken together, these results imply an already largely modified mitoribosome in the LECA and a fairly mature process in genome reduction after endosymbiosis, pointing to a consistent time frame between the endosymbiosis and the diversification of present-day eukaryotic lineages.

3.2. Evolution of the mitoribosome during eukaryotic diversification (after LECA)

The mitoribosomal proteomes of the ancestors of Unikonts and Bikonts are remarkably similar (Fig. 2), with only one

gain (L56) and one loss (S1) in the lineage leading to Unikonts. This may suggest that the split between the lineages leading to these two ancestors occurred very rapidly or that mitoribosomes entered in a stationary phase in these two lineages after the acquisition of 19 Ek-MRPs. The loss and gain involving two proteins in the ancestor of Unikonts appear to constitute two important synapomorphies for this group. Conversely, none of the Ek-MRPs characterized in the two proteomes of Kinetoplastids are exclusively present in Bikonts (Table 1), preventing for the time being the identification of synapomorphies for this major eukaryotic supergroup.

Except in the case of Ppe1 (see above), we could not find any homologue of MRPs in any genome of amitochondriates (Table 1), indicating that the loss of mitochondria is virtually always accompanied by the loss of all MRPs, even those encoded in the nucleus of the host. This implies multiple independent gene losses from the host genome, because at least 19 Ek-MRPs and 26 Bt-MRPs were already encoded in the nucleus of LECA. This phenomenon would have repeated multiple times during the evolution of eukaryotes because the transition to an amitochondriate state occurred independently in various unrelated lineages. Our observation indicates that no MRPs were recruited for other functions, differently from other proteins of mitochondrial origin that were retained in amitochondriate

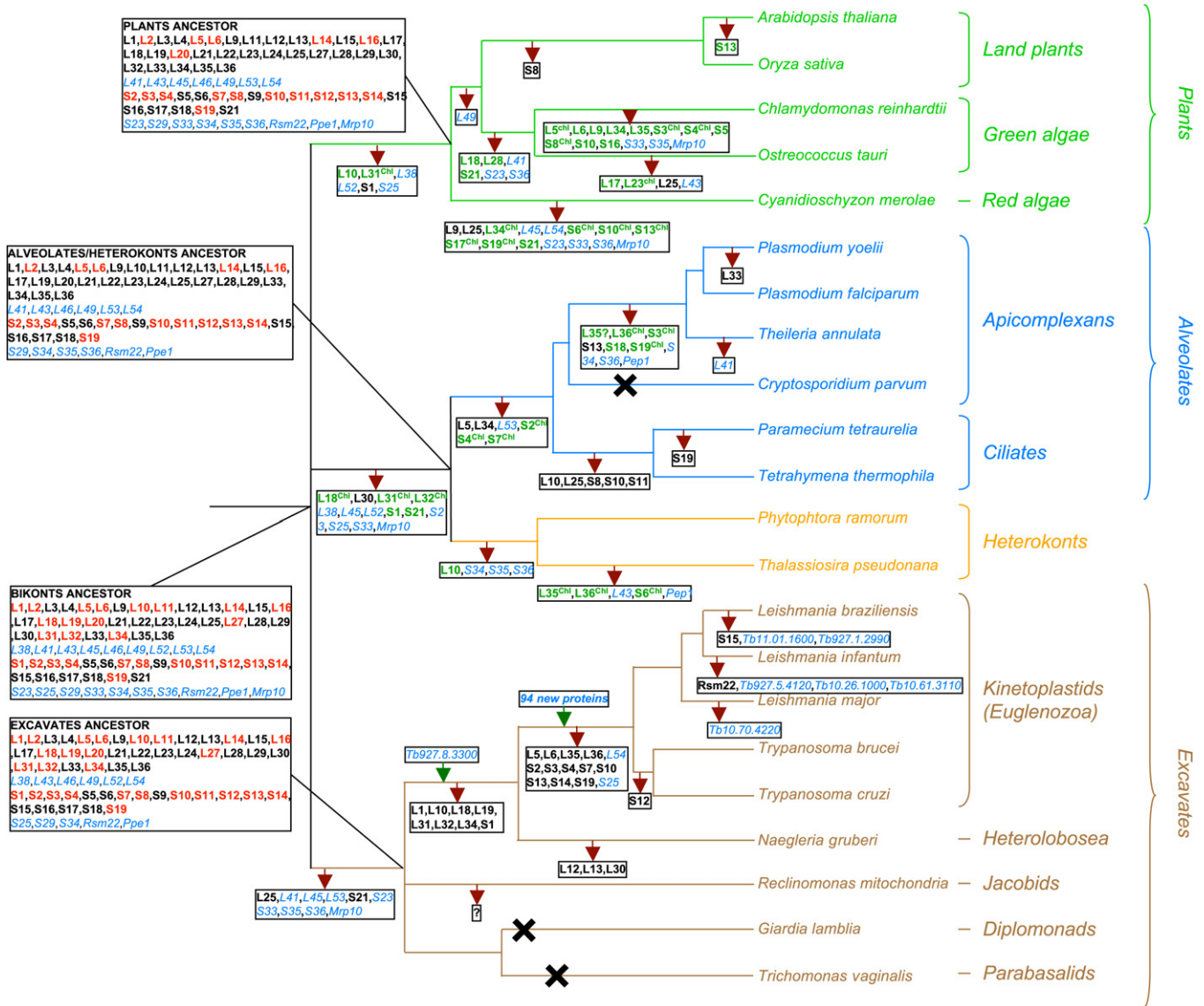


Fig. 4. Inference of the set of MRPs present in the last common ancestor of Bikonts and its subsequent evolution in Archaeplastida, Alveolata, Heterokonta and Excavata. For details see legend to Fig. 2. The missing MRPs that might have been replaced by chloroplastic homologues are indicated in green. Chl indicates that the corresponding chloroplastic protein is encoded in the chloroplast genome.

eukaryotes (Clark and Roger, 1995; Germot et al., 1996; Soltys and Gupta, 1994; Sutak et al., 2004; Tachezy et al., 2001; Tovar et al., 2003).

By assuming the most parsimonious scenarios to take into account MRP losses and gains, we then reconstructed the evolution of the mitoribosome in the six present-day eukaryotic supergroups for which complete genome sequences are currently available (Figs. 3 and 4, respectively). This was inferred according to a reference eukaryotic tree rooted between Unikonts and Bikonts, because no gains or losses of MRP are exclusive of Bikonts, assuming alternative rootings of the eukaryotic tree (e.g., within bikonts (Roger and Simpson, 2009)) does not change our inference.

3.2.1. Opisthokonta

We inferred 73 MRPs in the ancestor of Opisthokonta (44 LSU and 29 SSU), including two gains (L40, L44) that appear to

be specific synapomorphies of Opisthokonta, thanks to the proteomic characterizations of mitoribosomes from representatives of Metazoa and Fungi (Table 1 and Fig. 3). The inclusion in our sampling of the choanoflagellate *Monosiga brevicollis* changes sensibly the picture of the evolution of the mitochondrial ribosome in Opisthokonta that may be gathered by analyzing Metazoa and Fungi only (Fig. 3). In fact, *M. brevicollis* has retained 11 Bt-MRP coding genes (L2, L5, L14, L16, S3, S4, S8, S12, S13, S14, S19) in its mitochondrial genome while these have moved to the nuclear genome in all other metazoans and fungi analyzed except for S3 in some fungi (Table 1), including *Allomyces macrogynus*, a deeply branching chytrid (Bullerwell and Gray, 2004) (not shown). The sistership of Metazoa and Choanoflagellata implies that the transfer of these genes to the nuclear genome of Metazoa and Fungi occurred at least twice independently (Fig. 3) and is thus more recent than it was inferred from the analysis of mammals and fungi only (Smits et al., 2007).

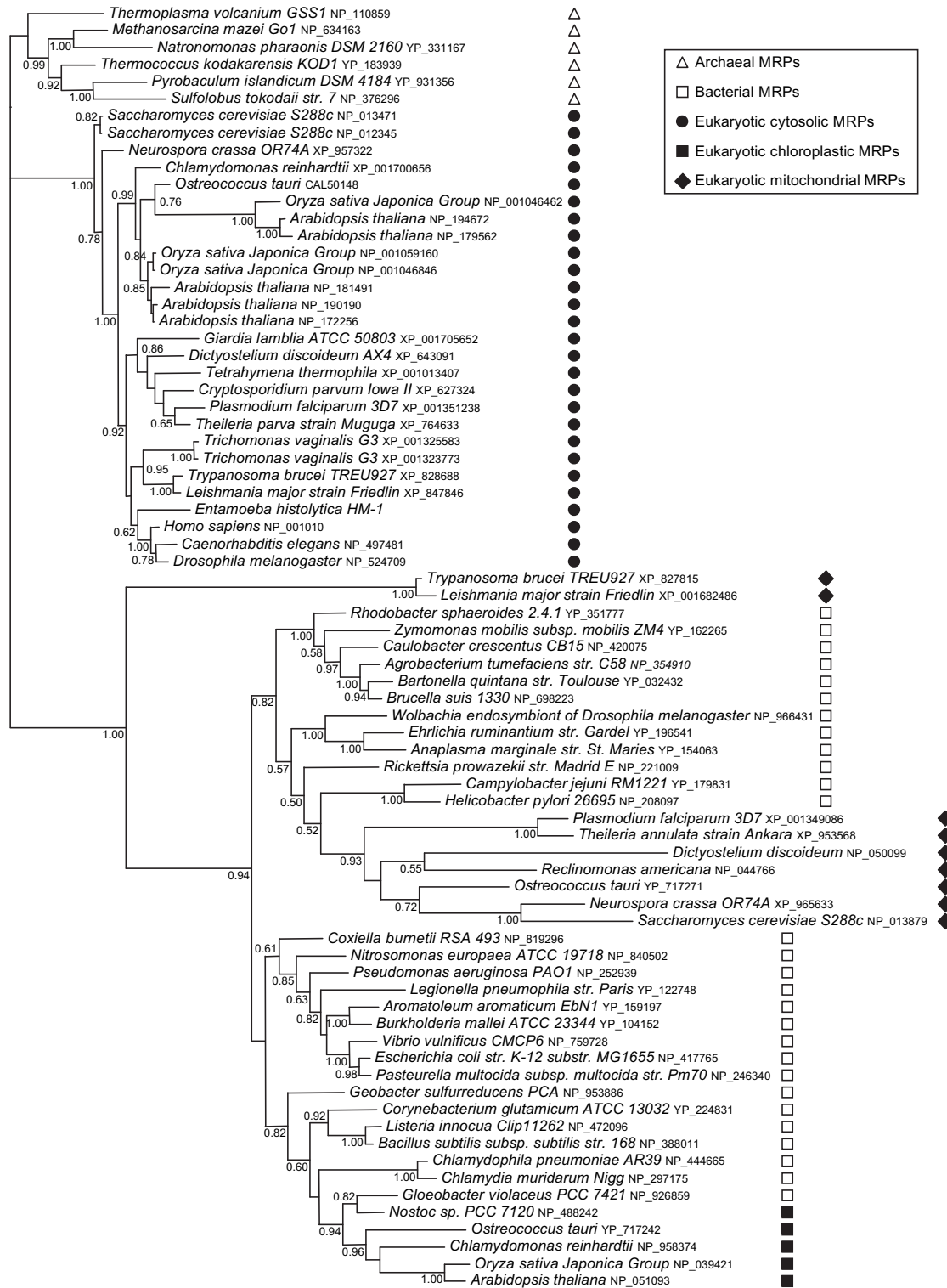


Fig. 5. Bayesian phylogenetic tree of mitochondrial ribosomal protein S8. Numbers at nodes represent posterior probabilities (PP). For clarity, only PP greater than 0.5 are shown. The scale bar represents the average number of substitutions per site that occurred during the evolution of this protein. Note that the mitochondrial sequences from *Leishmania* and *Trichomonas* do not cluster with the other mitochondrial homologues. However, based on typical sequence signatures and additional phylogenetic analyses, they are to be considered as genuine mitochondrial versions, their placement at the base of bacteria being likely due to an artefact caused by their fast evolutionary rate.

Interestingly, these genes are still encoded in the mitochondrial genomes of free-living unicellular representatives of other major eukaryotic lineages (e.g., *Phytophthora*, *Thalassiosira*, green and red algae) (Table 2). A tempting hypothesis is that the

transfer to the host nucleus of these MRPs might be linked to the passage to a multicellular lifestyle in the ancestor of Metazoans and that of Fungi. However, this does not appear to be the case for Archaeplastida.

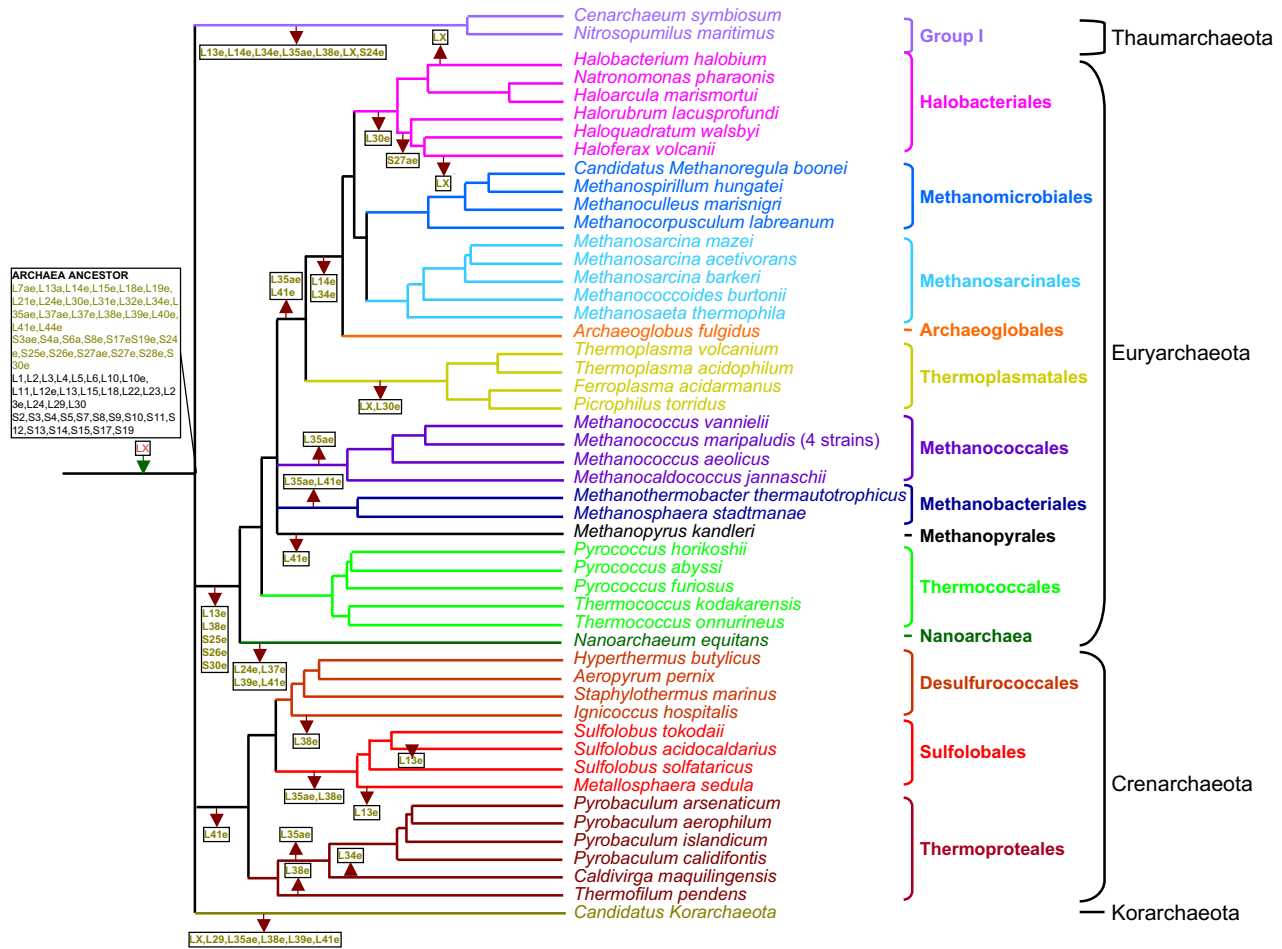


Fig. 6. Inference of the set of ribosomal proteins that was present in the last archaeal common ancestor and its subsequent evolution along the archaeal phyla: Crenarchaeota and Euryarchaeota (Woese et al., 1990), and the two additional proposed phyla: Korarchaeota (Barns et al., 1996) and Thaumarchaeota (Brochier-Armanet et al., 2008). Universal proteins are shown in black, whereas proteins found only in archaea and eukaryotes are in brown. Finally, proteins that are found only in archaea are shown in red. Protein acquisitions are indicated by dark green arrows, whereas losses are highlighted by dark red arrows. The proteins inferred to have been present in the last archaeal common ancestor are those MRPs shared by at least one present-day archaeon and one present-day eukaryote. These losses have been inferred by positing a sistership of Archaea and Eukaryotes (3-domains hypothesis). Under the alternative hypothesis that Eukaryotes derive from within the Euryarchaeota, the five losses inferred at the base of this phylum in Fig. 6 would become multiple convergent losses throughout euryarchaeal evolution. Conversely, if Eukarya derive from Crenarchaeota, the five losses at the base of Euryarchaeota would become five gains in the lineage leading to Crenarchaeota and Eukaryotes.

Compared to Metazoa and Fungi, the impressive number of 18 MRPs (of which eight Bt-MRPs) appears to have been lost in the lineage leading to *M. brevicollis* (Fig. 3). In order to confirm these gene losses and to verify whether these have been compensated by the recruitment of novel proteins it will be essential to have sequence data from a second choanoflagellate and/or to characterize the mitoribosomal proteome of a choanoflagellate representative.

After the divergence of *Monosiga* the mitochondrial ribosomal proteome of Metazoa appears to have undergone five losses and 11 gains. In the absence of proteomic characterizations other than mammals, it cannot be possible to infer the eventual gain of proteins in other metazoan lineages. Nevertheless, only few losses appear to have occurred during metazoan diversification. In contrast, our increased taxonomic sampling allows to infer a more dynamic evolution of the mitochondrial ribosome in Fungi than previously reported by

Smits et al. (Smits et al., 2007), with many losses and gains in both Basidiomycetes and Ascomycetes (Fig. 3). For example, we found a previously undetected S34 homologue in one fungal genome of our taxonomic sampling (*Ustilago maydis*). Moreover, an S34 homologue is present also in two other basidiomycetes not considered in our taxonomic sampling (*Laccaria bicolor* and *Coprinopsis cinerea*) (not shown). This means that the ancestor of Fungi likely harboured S34, which was thus lost many times independently during the evolution of this group (Fig. 3).

Four of the proteins characterized in the mitoribosomal proteome of *S. cerevisiae* appear to be specific novel recruitments in the branch of Saccharomycotina, since these have homologues in other members of this lineage (not shown) but in no other fungal genome. We also inferred five losses in the branch leading to *Schizosaccharomyces pombe*. It would be extremely interesting to characterize the mitochondrial

ribosomal proteome of additional representatives of Fungi, and in particular other lineages such as Basidiomycetes, Zygomycetes or Chytridiomycetes, in order to verify whether the recruitment of new MRPs is a widespread phenomenon in Fungi.

3.2.2. Amoebozoa

We found a previously unreported homologue of L34 in the genome of *Dictyostelium discoideum*, while we could not find any homologues of S6 and S16, contrary to previous reports (Smits et al., 2007). As in the case of *Monosiga*, an impressive number of 14 losses, 11 of which involving Bt-MRPs, appears to have occurred in *D. discoideum* (Fig. 3). It will be essential to obtain additional genomes from representatives of Amoebozoa (especially from mitochondriate relatives of *Entamoeba*), as well as proteomic characterizations to verify whether these losses are a general feature of this phylum and have been compensated by the inclusion of new proteins.

3.2.3. Archaeplastida

We inferred 66 MRPs in the ancestor of Archaeplastida (38 LSU and 28 SSU). Then, subsequent evolution of mitochondrial ribosomes appears to have been very dynamic (Fig. 4). Experimental studies on land plants have shown that mitochondrial genes are regularly transferred to the nucleus (Adams et al., 2000; Berghthorsson et al., 2003; Fallahi et al., 2005), and there are examples of shared usage of ribosomal proteins between mitochondrial, chloroplastic and the cytosolic ribosomes (e.g., S13 and L21 (Adams et al., 2002; Gallois et al., 2001)).

S23 was previously identified in Metazoans and Fungi only and therefore considered as a specific acquisition of Opisthokonta (Smits et al., 2007). However, we found previously undetected homologues of S23 in the two land plants genomes (Table 1b). Its presence in land plants pushes the origin of this protein back to LECA and increases the number of independent losses in all other major lineages (Figs. 2 and 4). S8 homologues were previously reported from the two land plants genomes (Smits et al., 2007). However, phylogenetic analysis shows that these are in fact cytoplasmic ribosomal homologues (Fig. 5) and it was experimentally shown that the cytoplasmic S8 copy works also in the mitochondrion (Adams et al., 2002). Intriguingly, a *bona fide* mitoribosomal S8 homologue is present in the mitochondrial genomes of *Ostreococcus tauri* and the red algae *Cyanidioschyzon merolae*, meaning that this protein was lost (and/or replaced by its cytosolic counterpart) at least twice independently, i.e. in *Chlamydomonas reinhardtii* and land plants. Finally, phylogenetic analysis indicates that the S5 and S13 homologues reported by Smits et al. (Smits et al., 2007) in *C. reinhardtii* are in fact chloroplastic ribosomal copies (Supplementary Figs. 1 and 2).

We found only one difference between the mitochondrial ribosomal proteome of *Oryza sativa* and that of *Arabidopsis thaliana*. In fact, S13 has been lost in *A. thaliana*, consistent with a recent duplication of the gene coding for the chloroplastic copy of S13 that was recruited to work in the

mitochondrion (Adams et al., 2002). On the contrary, we inferred a high number of losses in green and red algae (up to 15 in *C. reinhardtii* and *C. merolae* (Fig. 4)), with few overlaps, indicating that most of these losses occurred independently in these algal lineages. It would be interesting to investigate these lineages by a proteomic approach to see if some of these losses have been compensated by the recruitment of MRPs of chloroplastic origin encoded in the nucleus, when these are present (indicated in green in Fig. 3), or even of MRPs that are still encoded in the chloroplast genome (indicated by chl in Fig. 3) since this would require transport from one organelle to the other.

3.2.4. Heterokonta-Alveolata

Phylogenetic analysis indicates that Heterokonta and Alveolata share a common ancestor that possessed a chloroplast issued from a secondary endosymbiosis with a red alga (see (Simpson and Roger, 2004) and (Archibald, 2009) and references therein). We inferred for the first time the mitoribosome of this ancestor, which contained 59 MRPs (35 LSU and 24 SSU). Subsequent evolution of the mitoribosome appears to have been very different in Heterokonts and Alveolates. We inferred the loss of one Bt-MRP and of three Ek-MRPs in the branch leading to *Phytophthora ramorum* and *Thalassiosira pseudonana* (Fig. 4). No further losses occurred in *P. ramorum*, whereas six nuclear encoded MRPs (L35, L36, L43, S6, Ppe1) appear to have been specifically lost in *T. pseudonana* (Fig. 4). By contrast, more losses occurred during evolution of Alveolata (Fig. 4). We highlighted five previously unidentified MRPs (L6, L10, L25, L54, S10, (Smits et al., 2007)) in the three apicomplexan genomes analyzed (Table 1). These apicomplexan homologues are often very divergent (not shown) and we probably detected them thanks to a better HMM model based on our wide taxonomic sampling. Importantly, two (L6 and L25) have no homologues in humans (Table 1) and may thus represent potential targets for antimalarial treatment. We found also previously unreported homologues of L35, L41, and Ppe1 in the two genomes from Ciliates (Table 1). This allows us to infer the presence of L35 in the ancestor of Bikonts and the presence of Ppe1 in the ancestor of Alveolata (Fig. 4).

3.2.5. Excavata

We inferred 62 MRPs in the ancestor of Excavata (38 LSU and 24 SSU). With respect to previous analyses (Smits et al., 2007), we included a free-living representative (*Naegleria gruberi*). Moreover, we took advantage of the recent proteomic characterization of the mitoribosome from a second kinetoplastid representative (*Trypanosoma brucei*) (Zikova et al., 2008). This allowed us to identify previously unreported MRP homologues (Table 1 in red with an X) and better reconstruction of the evolution of the mitoribosome in this important eukaryotic supergroup, as well as the events that have taken place during the emergence of kinetoplastid parasites. In fact, we infer that 14 losses have occurred specifically in the branch leading to the two parasitic kinetoplastids (Fig. 4); some of these losses may have been compensated by

the specific appearance of a very large number of new MRPs (Fig. 4), as shown by the proteomic characterization of *Leishmania* and *Trypanosoma*. Indeed only one of the 95 novel MRPs that have been characterized in the two kinetoplastids is present in *N. gruberi* (Table 1), which indicates that these are genuine recent acquisitions in the ancestor of these parasites (Fig. 4). It will be crucial to obtain the mitoribosome proteome of another member of Excavata to verify if massive acquisitions of novel MRPs are a peculiarity of kinetoplastids or of the whole phylum.

3.3. Evolutionary considerations

Ribosomes are frequently considered as “frozen” cellular machines, because of their key function. However, our phylogenomic investigation shows that this does not apply to mitochondrial ribosomes, which have undergone a very dynamic history, both before and after LECA. It has been speculated that the mitochondrial endosymbiosis coincided with the origin of Eukaryotes, where the ancestor of mitochondria would have been engulfed by an archaeon, the association starting a burst of evolution giving rise to eukaryotes (Martin and Müller, 1998). However, the mitoribosomal complex has gained at least 19 novel MRPs in the time frame going from the endosymbiosis of the mitochondrial ancestor up to LECA. Finally under a scenario of eukaryotic origins coinciding with the engulfment of an alphaproteobacterium by an archaeon, it might have been expected that the increase in size of the mitochondrial ribosome would have been accomplished through recruitment of proteins already present in the archaeal host instead of new ones. It is perhaps more likely that the enrichment of the mitoribosome with novel proteins occurred in a host that was rather a proto-eukaryote (either derived from an archaeon or representing a separate lineage), with a genome containing already a large fraction of eukaryotic-specific genes. Consistently, comparison with different endosymbionts shows that the loss/translocation of ribosomal proteins is a late event in the reductive process, indicating that the endosymbiosis at the origin of the mitochondrion did not coincide with LECA but occurred some time before in an ancestor. Finally, it has been put forward that the three-domains structure of universal phylogenetic trees is the fruit of a reconstruction artefact caused by an acceleration of evolutionary rates in eukaryotes, whereas the true tree should show eukaryotes as deriving from within archaea (Cox et al., 2008; Foster et al., 2009; Lake, 1988; Rivera and Lake, 1992, 2004; Tourasse and Gouy, 1999). However, despite the very fast evolutionary rate of ribosomal mitochondrial proteins, their bacterial origin could always be easily recovered by phylogenetic analysis. This suggests either that the hypothesis of an artefact affecting universal trees does not hold, or that another type of artefact other than accelerated evolutionary rates is at work.

It should be noted that the loss of only one Bt-MRP occurred before LECA, indicating that, initially, the ancestral alphaproteobacterial ribosome experienced accretion of novel proteins uniquely, whose role was therefore not to compensate for the loss of Bt-MRPs. Such increase in size of the mitoribosome is puzzling, as it has been neither observed in the ribosomes of

other bacterial endosymbionts nor in other organelles such as plastids. Interestingly, the addition of novel proteins appears to have continued during the diversification of present-day eukaryotic lineages, and is still ongoing. Only one loss and one gain of MRPs could be inferred after the split between Unikonts and Bikonts, but a large number of such events occurred subsequently, during the diversification of present-day eukaryotic lineages, and this process is still ongoing, as exemplified by great differences between mitoribosome of closely related organisms. Interestingly, these gains and losses have occurred repeatedly and independently in different lineages. It is worth noting that both Bt-MRPs and Ek-MRPs have been equally affected by this dynamic of gene losses, suggesting that there is no general trend in protein dispensability.

In contrast to the chaotic history of mitoribosomes, their cytosolic ribosomal counterparts have remained largely stable during eukaryotic evolution (data not shown). Therefore, there is definitely some unique and still unclear factor driving the instability in composition of mitochondrial ribosomes. Intriguingly, the archaeal ribosome also appears to have experienced a dynamic history. It has in fact experienced multiple independent losses during the history of this domain (Fig. 6), in clear contrast with its bacterial counterpart, whose composition is much more stable (Lecompte et al., 2002). However, there is other evidence of streamlining in additional informational archaeal systems (see for example (Gribaldo and Brochier, 2006) and (de Crecy-Lagard et al., 2010)). This indicates that the archaeal ancestor was more complex (at least for some cellular systems) than its present-day descendants. Interestingly, and differently from mitoribosomes, losses have involved uniquely those MRPs shared between Archaea and Eukarya (in brown in Fig. 6) but not universal MRPs (in black in Fig. 6). It will be interesting to investigate further the reasons for such instability of the archaeal ribosome.

4. Conclusions

The combination of genomic and proteomic data in phylogenomic analyses promises to become a major tool of the post-genomic era to dissect early evolutionary events. Too little is known about the composition of the majority of cellular systems, especially in representatives of major eukaryotic phyla other than mammals or yeasts. Proteomics of mitoribosomes from protist lineages should be relatively easy to accomplish and would provide essential insights for the reconstruction of a complete picture of early and more recent events in its evolution, which may shed light on LECA and therefore eukaryotic origins.

Acknowledgement

We wish to thank the Fondation Des Treilles for support, as well as Celine Petitjean and Alexis Criscuolo for help with some analyses. C.B-A. is supported by an Action Thématique et Incitative sur Programme (ATIP) of the French Centre National de la Recherche Scientifique (CNRS).

Appendix. Supplementary data

Supplementary data related to this article can be found online at doi:10.1016/j.resmic.2010.10.004.

References

- Adams, K.L., Daley, D.O., Qiu, Y.L., Whelan, J., Palmer, J.D., 2000. Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. *Nature* 408, 354–357.
- Adams, K.L., Daley, D.O., Whelan, J., Palmer, J.D., 2002. Genes for two mitochondrial ribosomal proteins in flowering plants are derived from their chloroplast or cytosolic counterparts. *Plant Cell* 14, 931–943.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402.
- Archibald, J.M., 2009. The puzzle of plastid evolution. *Curr. Biol.* 19, R81–R88.
- Armbrust, E.V., Berges, J.A., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., Zhou, S., Allen, A.E., et al., 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306, 79–86.
- Bapteste, E., Charlebois, R.L., MacLeod, D., Brochier, C., 2005. The two tempos of nuclear pore complex evolution: highly adapting proteins in an ancient frozen structure. *Genome Biol.* 6, R85.
- Barns, S.M., Delwiche, C.F., Palmer, J.D., Pace, N.R., 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. U S A* 93, 9188–9193.
- Berghthorsson, U., Adams, K.L., Thomason, B., Palmer, J.D., 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* 424, 197–201.
- Brochier-Armanet, C., Boussau, B., Gribaldo, S., Forterre, P., 2008. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252.
- Brown, J.R., Doolittle, W.F., 1995. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proc. Natl. Acad. Sci. USA* 92, 2441–2445.
- Bullerwell, C.E., Gray, M.W., 2004. Evolution of the mitochondrial genome: protist connections to animals, fungi and plants. *Curr. Opin. Microbiol.* 7, 528–534.
- Burger, G., Gray, M.W., Lang, B.F., 2003. Mitochondrial genomes: anything goes. *Trends Genet.* 19, 709–716.
- Burger, G., Lang, B.F., 2003. Parallels in genome evolution in mitochondria and bacterial symbionts. *IUBMB Life* 55, 205–212.
- Burki, F., Inagaki, Y., Brate, J., Archibald, J.M., Keeling, P.J., Cavalier-Smith, T., Sakaguchi, M., Hashimoto, T., Horak, A., Kumar, S., Klaveness, D., Jakobsen, K.S., Pawlowski, J., Shalchian-Tabrizi, K., 2009. Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, telonemia and centroheliozoa, are related to photosynthetic chromalveolates. *Genome Biol. Evol.* 1, 231–238.
- Burki, F., Shalchian-Tabrizi, K., Pawlowski, J., 2008. Phylogenomics reveals a new ‘megagroup’ including most photosynthetic eukaryotes. *Biol. Lett.* 4, 366–369.
- Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B., Bork, P., 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science* 311, 1283–1287.
- 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.
- Cox, C.J., Foster, P.G., Hirt, R.P., Harris, S.R., Embley, T.M., 2008. The archaeobacterial origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* 105, 20356–20361.
- Dacks, J.B., Field, M.C., 2007. Evolution of the eukaryotic membrane-traffic system: origin, tempo and mode. *J. Cell Sci.* 120, 2977–2985.
- de Crecy-Lagard, V., Brochier-Armanet, C., Urbonavicius, J., Fernandez, B., Phillips, G., Lyons, B., Noma, A., Alvarez, S., Droogmans, L., Armengaud, J., Grosjean, H., 2010. Biosynthesis of wyosine derivatives in tRNA: an ancient and highly diverse pathway in Archaea. *Mol. Biol. Evol.* 27 (9), 2062–2077.
- Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–375.
- Desmond, E., Gribaldo, S., 2009. Phylogenomics of sterol synthesis: insights into the origin, evolution, and diversity of a key eukaryotic feature. *Genome Biol. Evol.* 2009, 364–381.
- Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinforma.* 5, 113.
- Eisen, J.A., Fraser, C.M., 2003. Phylogenomics: intersection of evolution and genomics. *Science* 300, 1706–1707.
- Eme, L., Moreira, D., Talla, E., Brochier-Armanet, C., 2009. A complex cell division machinery was present in the last common ancestor of eukaryotes. *PLoS One* 4, e5021.
- Fallahi, M., Crosthwait, J., Calixte, S., Bonen, L., 2005. Fate of mitochondrially located S19 ribosomal protein genes after transfer of a functional copy to the nucleus in cereals. *Mol. Genet. Genomics* 273, 76–83.
- Foster, P.G., Cox, C.J., Embley, T.M., 2009. The primary divisions of life: a phylogenomic approach employing composition-heterogeneous methods. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 2197–2207.
- Gallois, J., Achard, P., Green, G., Mache, R., 2001. The Arabidopsis chloroplast ribosomal protein L21 is encoded by a nuclear gene of mitochondrial origin. *Gene* 274, 179–185.
- Gan, X., Kitakawa, M., Yoshino, K., Oshiro, N., Yonezawa, K., Isono, K., 2002. Tag-mediated isolation of yeast mitochondrial ribosome and mass spectrometric identification of its new components. *Eur. J. Biochem.* 269, 5203–5214.
- Germot, A., Philippe, H., Le Guyader, H., 1996. Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. *Proc. Natl. Acad. Sci. USA* 93, 14614–14617.
- Gogarten, J.P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E.J., Bowman, B.J., Manolson, M.F., Poole, R.J., Date, T., Oshima, T., Konishi, J., Denda, K., Yoshida, M., 1989. Evolution of the vacuolar H⁺-ATPase: implications for the origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* 86, 6661–6665.
- Graack, H.R., Wittmann-Liebold, B., 1998. Mitochondrial ribosomal proteins (MRPs) of yeast. *Biochem. J.* 329 (Pt 3), 433–448.
- Gray, M.W., Burger, G., Lang, B.F., 2001. The origin and early evolution of mitochondria. *Genome Biol.* 2 REVIEWS1018.
- Gray, M.W., Lang, B.F., Burger, G., 2004. Mitochondria of protists. *Annu. Rev. Genet.* 38, 477–524.
- Gribaldo, S., Brochier, C., 2006. The origin and evolution of the Archaea: a state of the art. *Proc. Royal Acad. Sci.* 361 (1470), 1007–1022.
- Gribaldo, S., Cammarano, P., 1998. The root of the universal tree of life inferred from anciently duplicated genes encoding components of the protein-targeting machinery. *J. Mol. Evol.* 47, 508–516.
- Gribaldo, S., Poole, A., Daubin, V., Forterre, P., Brochier-Armanet, C., 2010. The origin of Eukaryotes and their evolutionary relationship with Archaea: have we reached a phylogenomic impasse? *Nat. Rev. Microbiol.* 8 (10), 743–752.
- Hackett, J.D., Yoon, H.S., Li, S., Reyes-Prieto, A., Rummele, S.E., Bhattacharya, D., 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol. Biol. Evol.* 24, 1702–1713.
- Hampl, V., Hug, L., Leigh, J.W., Dacks, J.B., Lang, B.F., Simpson, A.G., Roger, A.J., 2009. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic “supergroups”. *Proc. Natl. Acad. Sci. USA* 106, 3859–3864.
- Harris, J.K., Kelley, S.T., Spiegelman, G.B., Pace, N.R., 2003. The genetic core of the universal ancestor. *Genome Res.* 13, 407–412.
- Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., Miyata, T., 1989. Evolutionary relationship of archaeobacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA* 86, 9355–9359.
- Koc, E.C., Burkhardt, W., Blackburn, K., Moseley, A., Spemulli, L.L., 2001a. The small subunit of the mammalian mitochondrial ribosome. Identification of the full complement of ribosomal proteins present. *J. Biol. Chem.* 276, 19363–19374.
- Koc, E.C., Burkhardt, W., Blackburn, K., Moyer, M.B., Schlatzer, D.M., Moseley, A., Spemulli, L.L., 2001b. The large subunit of the mammalian mitochondrial ribosome. Analysis of the complement of ribosomal proteins present. *J. Biol. Chem.* 276, 43958–43969.

- Koonin, E.V., 2010. The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biol.* 11, 209.
- Lake, J.A., 1988. Origin of the eukaryotic nucleus determined by rate-invariant analysis of rRNA sequences. *Nature* 331, 184–186.
- Lang, B.F., Gray, M.W., Burger, G., 1999. Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* 33, 351–397.
- Lecompte, O., Ripp, R., Thierry, J.C., Moras, D., Poch, O., 2002. Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acids Res.* 30, 5382–5390.
- Mans, B.J., Anantharaman, V., Aravind, L., Koonin, E.V., 2004. Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* 3, 1612–1637.
- Martin, W., 2005. Archaeobacteria (Archaea) and the origin of the eukaryotic nucleus. *Curr. Opin. Microbiol.* 8, 630–637.
- Martin, W., Müller, M., 1998. The hydrogen hypothesis for the first eukaryote. *Nature* 392, 37–41.
- Maslov, D.A., Sharma, M.R., Butler, E., Falick, A.M., Gingery, M., Agrawal, R.K., Spremulli, L.L., Simpson, L., 2006. Isolation and characterization of mitochondrial ribosomes and ribosomal subunits from *Leishmania tarentolae*. *Mol. Biochem. Parasitol.* 148, 69–78.
- Maslov, D.A., Spremulli, L.L., Sharma, M.R., Bhargava, K., Grasso, D., Falick, A.M., Agrawal, R.K., Parker, C.E., Simpson, L., 2007. Proteomics and electron microscopic characterization of the unusual mitochondrial ribosome-related 45S complex in *Leishmania tarentolae*. *Mol. Biochem. Parasitol.* 152, 203–212.
- Matsuzaki, M., Misumi, O., Shin, I.T., Maruyama, S., Takahara, M., Miyagishima, S.Y., Mori, T., Nishida, et al., 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428, 653–657.
- McCutcheon, J.P., McDonald, B.R., Moran, N.A., 2009. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.* 5, e1000565.
- Moya, A., Pereto, J., Gil, R., Latorre, A., 2008. Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat. Rev. Genet.* 9, 218–229.
- Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H.E., Moran, N.A., Hattori, M., 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 314, 267.
- Nikoh, N., McCutcheon, J.P., Kudo, T., Miyagishima, S.Y., Moran, N.A., Nakabachi, A., 2010. Bacterial genes in the aphid genome: absence of functional gene transfer from *Buchnera* to its host. *PLoS Genet.* 6, e1000827.
- O'Brien, T.W., 2003. Properties of human mitochondrial ribosomes. *IUBMB Life* 55, 505–513.
- O'Brien, T.W., Fiesler, S.E., Denslow, N.D., Thiede, B., Wittmann-Liebold, B., Mougey, E.B., Sylvester, J.E., Graack, H.R., 1999. Mammalian mitochondrial ribosomal proteins (2). Amino acid sequencing, characterization, and identification of corresponding gene sequences. *J. Biol. Chem.* 274, 36043–36051.
- O'Brien, T.W., Liu, J., Sylvester, J.E., Mougey, E.B., Fischel-Ghodsian, N., Thiede, B., Wittmann-Liebold, B., Graack, H.R., 2000. Mammalian mitochondrial ribosomal proteins (4). Amino acid sequencing, characterization, and identification of corresponding gene sequences. *J. Biol. Chem.* 275, 18153–18159.
- Palenik, B., Grimwood, J., Aerts, A., Rouze, P., Salamov, A., Putnam, N., Dupont, C., Jorgensen, R., et al., 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. USA* 104, 7705–7710.
- Perez-Brocail, V., Gil, R., Ramos, S., Lamelas, A., Postigo, M., Michelena, J.M., Silva, F.J., Moya, A., Latorre, A., 2006. A small microbial genome: the end of a long symbiotic relationship? *Science* 314, 312–313.
- Philippe, H., 1993. MUST, a computer package of management utilities for sequences and trees. *Nucleic Acids Res.* 21, 5264–5272.
- Rivera, M.C., Lake, J.A., 1992. Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* 257, 74–76.
- Rivera, M.C., Lake, J.A., 2004. The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* 431, 152–155.
- Rodriguez-Ezpeleta, N., Brinkmann, H., Burey, S.C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H.J., Philippe, H., Lang, B.F., 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* 15, 1325–1330.
- Rodriguez-Ezpeleta, N., Brinkmann, H., Burger, G., Roger, A.J., Gray, M.W., Philippe, H., Lang, B.F., 2007. Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. *Curr. Biol.* 17, 1420–1425.
- Roger, A.J., Simpson, A.G., 2009. Evolution: revisiting the root of the eukaryote tree. *Curr. Biol.* 19, R165–R167.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Saveanu, C., Fromont-Racine, M., Harington, A., Ricard, F., Namane, A., Jacquier, A., 2001. Identification of 12 new yeast mitochondrial ribosomal proteins including 6 that have no prokaryotic homologues. *J. Biol. Chem.* 276, 15861–15867.
- Simpson, A.G., Roger, A.J., 2004. The real 'kingdoms' of eukaryotes. *Curr. Biol.* 14, R693–R696.
- Smits, P., Smeitink, J.A., van den Heuvel, L.P., Huynen, M.A., Ettema, T.J., 2007. Reconstructing the evolution of the mitochondrial ribosomal proteome. *Nucleic Acids Res.* 35, 4686–4703.
- Soltys, B.J., Gupta, R.S., 1994. Presence and cellular distribution of a 60-kDa protein related to mitochondrial hsp60 in *Giardia lamblia*. *J. Parasitol.* 80, 580–590.
- Stechmann, A., Cavalier-Smith, T., 2002. Rooting the eukaryote tree by using a derived gene fusion. *Science* 297, 89–91.
- Stechmann, A., Cavalier-Smith, T., 2003. The root of the eukaryote tree pinpointed. *Curr. Biol.* 13, R665–R666.
- Sutak, R., Dolezal, P., Fiumera, H.L., Hrdy, I., Dancis, A., Delgadillo-Correa, M., Johnson, P.J., Muller, M., Tachezy, J., 2004. Mitochondrial-type assembly of FeS centers in the hydrogenosomes of the amitochondriate eukaryote *Trichomonas vaginalis*. *Proc. Natl. Acad. Sci. USA* 101, 10368–10373.
- Suzuki, T., Terasaki, M., Takemoto-Hori, C., Hanada, T., Ueda, T., Wada, A., Watanabe, K., 2001. Structural compensation for the deficit of rRNA with proteins in the mammalian mitochondrial ribosome. Systematic analysis of protein components of the large ribosomal subunit from mammalian mitochondria. *J. Biol. Chem.* 276, 21724–21736.
- Tachezy, J., Sanchez, L.B., Muller, M., 2001. Mitochondrial type iron-sulfur cluster assembly in the amitochondriate eukaryotes *Trichomonas vaginalis* and *Giardia intestinalis*, as indicated by the phylogeny of IscS. *Mol. Biol. Evol.* 18, 1919–1928.
- Tamames, J., Gil, R., Latorre, A., Pereto, J., Silva, F.J., Moya, A., 2007. The frontier between cell and organelle: genome analysis of *Candidatus Carsonella ruddii*. *BMC Evol. Biol.* 7, 181.
- Tourasse, N.J., Gouy, M., 1999. Accounting for evolutionary rate variation among sequence sites consistently changes universal phylogenies deduced from rRNA and protein-coding genes. *Mol. Phylogenet. Evol.* 13, 159–168.
- Tovar, J., Leon-Avila, G., Sanchez, L.B., Sutak, R., Tachezy, J., van der Giezen, M., Hernandez, M., Muller, M., Lucocq, J.M., 2003. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature* 426, 172–176.
- Towpik, J., 2005. Regulation of mitochondrial translation in yeast. *Cell Mol. Biol. Lett.* 10, 571–594.
- Tyler, B.M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R.H., Aerts, A., Arredondo, F.D., Baxter, L., et al., 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261–1266.
- Woese, C.R., Kandler, O., Wheelis, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* 87, 4576–4579.
- Yutin, N., Makarova, K.S., Mekhedov, S.L., Wolf, Y.I., Koonin, E.V., 2008. The deep archaeal roots of eukaryotes. *Mol. Biol. Evol.* 25, 1619–1630.
- Zikova, A., Panigrahi, A.K., Dalley, R.A., Acestor, N., Anupama, A., Ogata, Y., Myler, P.J., Stuart, K., 2008. *Trypanosoma brucei* mitochondrial ribosomes: affinity purification and component identification by mass spectrometry. *Mol. Cell Proteomics* 7, 1286–1296.