Reconstructing the evolution of the mitochondrial ribosomal proteome

Paulien Smits¹, Jan A. M. Smeitink¹, Lambert P. van den Heuvel¹, Martijn A. Huynen² and Thijs J. G. Ettema²,*

¹Nijmegen Center for Mitochondrial Disorders, Nijmegen Center for Molecular Life Sciences, Radboud University Nijmegen Medical Center, Geert-Grootplein-Zuid 10 and ²Center for Molecular and Biomolecular Informatics, Nijmegen Center for Molecular Life Sciences, Radboud University Nijmegen Medical Center, Geert-Grootplein 28, 6525 GA Nijmegen, The Netherlands

Received March 20, 2007; Revised May 16, 2007; Accepted May 17, 2007

ABSTRACT

For production of proteins that are encoded by the mitochondrial genome, mitochondria rely on their own mitochondrial translation system, with the mitoribosome as its central component. Using extensive homology searches, we have reconstructed the evolutionary history of the mitoribosomal proteome that is encoded by a diverse subset of eukaryotic genomes, revealing an ancestral ribosome of alpha-proteobacterial descent that more than doubled its protein content in most eukaryotic lineages. We observe large variations in the protein content of mitoribosomes between different eukaryotes, with mammalian mitoribosomes sharing only 74 and 43% of its proteins with yeast and Leishmania mitoribosomes, respectively. We detected many previously unidentified mitochondrial ribosomal proteins (MRPs) and found that several have increased in size compared to their bacterial ancestral counterparts by addition of functional domains. Several new MRPs have originated via duplication of existing MRPs as well as by recruitment from outside of the mitoribosomal proteome. Using sensitive profile–profile homology searches, we found hitherto undetected homology between bacterial and eukaryotic ribosomal proteins, as well as between fungal and mammalian ribosomal proteins, detecting two novel human MRPs. These newly detected MRPs constitute, along with evolutionary conserved MRPs, excellent new screening targets for human patients with unresolved mitochondrial oxidative phosphorylation disorders.

INTRODUCTION

Mitochondria are primary ATP-producing organelles which originated from an ancestral alpha-proteobacterial endosymbiont (1,2). The oxidative phosphorylation (OXPHOS) pathway, consisting of the mitochondrial respiratory chain and the ATP synthase complex, is the final biochemical pathway in energy conversion (3). However, mitochondria have been found to perform pivotal roles in many other metabolic, regulatory and developmental processes. During the evolution of the mitochondria, most genes of the ancestral endosymbiont have been transferred from the mitochondrial to the nuclear genome. The genes that reside on the mitochondrial DNA (mtDNA), e.g. 13 OXPHOS-related genes in human, are translated into their respective protein products by the mitochondrial translation machinery, which resembles the prokaryotic translation system (4), comprising a mitochondrial ribosome (mitoribosome) and several translation factors. In mammals, all proteins that constitute the mitochondrial translation machinery are encoded by the nuclear genome. Also the mitochondrial genomes of fungi generally encode only very few mitochondrial ribosomal proteins (MRPs). In contrast, the mitochondrial genome of many protozoa and plants encode numerous MRPs. Most notably, the ‘primitive’ mitochondrial genome of the freshwater protist Reclinomonas americana was found to contain as many as 27 MRP-encoding genes (5).

Mitoribosomes have undergone major remodeling during their evolution. For instance, it has been found that, despite the fact that their rRNA content is only half the content of bacterial ribosomes (6), mitoribosomes generally exceed the bacterial ribosomes both in molecular mass as well as in physical dimensions (6–8). Mass spectrometry studies of the bovine and yeast
mitoribosomes, which have served as model systems during the past years, revealed that mammalian mitoribosomes comprise about 80 MRPs (9–13). In yeast, about 70 MRPs have been identified thus far, although protein–protein interaction data and mutational analysis suggest that this number might be substantially higher (14–16). Interestingly, a recent proteomics survey of mitochondrial ribosome-related complexes in the kinetoplast-mitochondria of Leishmania tarentolae revealed an additional protein complex which was found to be specifically associated with the small subunit (SSU), but not with the large subunit (LSU) of the mitoribosome (17). The biological role of the additional SSU-associated complex, which mostly consists of proteins unique to kinetoplastida, remains unknown.

Apparently, mitoribosomes have expanded their protein content in the course of evolution by acquiring numerous extra ‘supernumerary’ MRPs. Currently information regarding the functions of these supernumerary MRPs is rather limited. In addition, loss of MRPs that are normally part of the ‘bacterial core ribosome’ is also observed.

Recently, several attempts have been made to identify the MRPs in other eukaryotes, such as Neurospora crassa (18), Arabidopsis thaliana (19) and L. tarentolae (20), however, a comprehensive study of the evolution of the mitoribosomal proteome remains to be established. Such a study bears relevance for the potential identification of genes that cause mitochondrial dysfunctions in human patients. Since mitochondria perform many fundamental functions, mitochondrial dysfunction results in a wide variety of multisystemic diseases, predominantly affecting tissues with high metabolic energy rates (21). These disorders are mostly caused by the dysfunction of one or more enzyme complexes of the OXPHOS system and several mutations have been identified in mtDNA (22) as well as in nuclear DNA (23). Not only mutations in structural OXPHOS genes, but also mutations in genes involved in the mitochondrial transcription or translation, or in the assembly of OXPHOS complexes can result in OXPHOS disease. Although the vast majority of components of the mitochondrial translation system are nuclear encoded, thus far most mutations associated with mitochondrial translation defects have been reported in mtDNA-encoded tRNAs and rRNAs (22,24). Recently, mutations in four different nuclear gene products involved in mitochondrial protein synthesis have been reported in patients with mitochondrial disease: elongation factors EFG1 (25), EFTs (26) and EFTu (27), and small ribosomal subunit protein MRPS16 (28). Additional information regarding which MRPs could be implicated in disease could be obtained by studying the evolutionary conservation of the MRPs.

The aim of the present study is 2-fold, from an evolutionary and a disease point of view: (i) gain insight into the evolution of the mitoribosome and its protein content in various eukaryotic species; (ii) prioritize MRPs as candidates for their involvement in mitochondrial disease. We have performed a comprehensive comparative genomics analysis of the MRPs in 18 eukaryotic species from which both the complete nuclear and the organellar genome sequences were available. Representatives are included from different phylogenetic groups: six metazoan (Homo sapiens, Mus musculus, Tetraodon nigroviridis, Drosophila melanogaster, Anopheles gambiae and Caenorhabditis elegans), two fungi (Saccharomyces cerevisiae and N. crassa), one microsporidian (Encephalitozoon cuniculi), one mycetozoan (Dictyostelium discoideum), one plant (A. thaliana), one alga (Chlamydomonas reinhardtii), one apicomplexan (Plasmodium falciparum), one ciliophora (Tetrahymena thermophila), two kinetoplastida (Trypanosoma brucei and Leishmania major), one diplomonad (Giardia lamblia) and the mitochondrial genome of the freshwater protist Reclinomonas americana. In addition, we included a balanced, phylogenetically diverse subset of completely sequenced prokaryotic genomes in our survey.

Our analysis retrieved multiple previously unidentified MRPs in several species, two of which constitute potential novel human MRPs. Furthermore, the current study established orthology relationships between seven MRPs reported to be fungi specific and ribosomal proteins from other eukaryotes. The establishment of these homology relationships enabled us to trace the origins of some of the supernumerary MRPs and to predict their molecular functions. In addition, we have investigated all MRPs in the presence of additional, newly acquired protein domains that might point at mitochondria-specific adaptations of the mitoribosome, and we have reconstructed the evolutionary history of the mitoribosomal proteome in terms of gains and losses of the MRPs along the different eukaryotic lineages. The newly detected mitochondrial ribosomal genes constitute, in addition to the set of evolutionary conserved MRPs, excellent new screening targets for human patients with unresolved mitochondrial oxidative phosphorylation disorders.

**MATERIALS AND METHODS**

Use of genomic data

Genome data (both proteomes and genomes) of the eukaryotes described in this study were obtained from the Integr8 database (http://www.ebi.ac.uk/integr8/, Release 41) at EBI (29), except for the genomes of Tetrahymena thermophila (version 2 obtained from TIGR, ftp://ftp.tigr.org/pub/data/Eukaryotic_Projects/t_thermophila/), Trypanosoma brucei (version 4 obtained from Sanger, ftp://ftp.sanger.ac.uk/pub/databases/T.brucei_sequences/), N. crassa (version 7.0 obtained from Broad institute at MIT, http://www.broad.mit.edu/annotation/genome/neurospora/Downloads.html) and Chlamydomonas reinhardtii (version 3.0 obtained from JGI, ftp://ftp.jgi-psf.org/pub/JGI_data/Chlamy/v3.0/). For most eukaryotic species, the organelle genomes were included in the gene set per species, otherwise those were downloaded separately from GOBASE (30) (http://megason.bch.umontreal.ca/ogmp/projects/other/mt_list.html).

In addition to the eukaryotic genomes mentioned above, we have also included a balanced, phylogenetically diverse subset of completely sequenced
prokaryotic genomes. We have explicitly chosen not to include the complete set of completely sequenced prokaryotic genomes available, since the large number of prokaryotic proteins might cause an imbalance in the PSI-BLAST analysis described below. Sequence data (proteomes and genomes) of the selected prokaryotic genomes were all downloaded from the Integr8 database at EBI, and included: *Escherichia coli*, *Pseudomonas aeruginosa*, *Vibrio cholerae* (gamma-proteobacteria), *Bacillus subtilis*, *Clostridium acetobutylicum* (firmicutes), *Borrelia burgdorferi* (spirochetes), *Agrobacterium tumefaciens* WashU, *Brucella suis*, *Rickettsia prowazekii*, *Caulobacter crescentus*, *Rhizobium melliloti*, *Rhodopseudomonas palustris* (alpha-proteobacteria), *Nostoc* sp. PCC 7920, *Synechocystis* sp. PCC 6803, *Prochlorococcus marinus* SS120 (cyanobacteria), *Deinococcus radiodurans* (Deinococcus-Thermus group), *Aquifex aeolicus* (aquificae), *Sulfolobus solfataricus* (crenarchaeae), *Pyrococcus furiosus* (euryarchaeae).

Detection of MRPs in eukaryotic genomes

For each of the experimentally verified proteins of the *Leishmania* (17,20), fungal (14–16) and mammalian (6) mitochondrial ribosomal proteins we have performed iterative PSI-BLAST searches (31) against a locally assembled protein database comprised of the proteomes of the organisms indicated above. For each of these searches, an E-value inclusion threshold of 0.005 was applied, and the proteins that were retrieved after five iterations were aligned using MUSCLE (32).

For those cases where a candidate MRP, from a species from which the query MRP was still missing, appeared in the PSI-BLAST hit list, but with an E-value between 0.005 and 10, a reciprocal search was performed. If this search successfully retrieved a bona fide MRP within five iterations (E-value inclusion threshold of 0.005), it was added to the initial set of homologous proteins, and a new alignment was created. Whenever a given MRP was still not found in the predicted proteome of a eukaryotic organism, a hidden Markov model (HMM) of the respective protein was created from the above-mentioned alignments using the HMMER program (33). Subsequently, the HMM was used to screen the genomic proteome content encoded in a wide variety of eukaryotic species. To avoid confusion regarding different nomenclatures that are used for ribosomal proteins, we have adopted the nomenclature for human MRPs, approved by the Human Genome Nomenclature Committee (http://www.gene.ucl.ac.uk/nomenclature/). All proteins are designated MRPSxx or MRPLxx, where S and L stand for small and large subunit, respectively, and xx is the number of the corresponding bacterial ribosomal protein. The MRPs lacking bacterial orthologs (supernumerary proteins) are assigned higher numbers, starting at MRPS22 and MRPL37 for the small and large subunit, respectively. In yeast, unfortunately, the nomenclature is as variable as the methods employed for their characterization. Therefore, in order to prevent further confusion about proteins with the same name in human and yeast that are not part of the same orthologous group, we have indicated all human MRPs in capital letters throughout the manuscript (e.g. MRPS28), whereas for the yeast proteins we only capitalized the first letter (e.g. Mrps28, which corresponds to MRPS15 in human, also see Table 1 and Supplementary Data).

Starting from a curated list of MRPs that have been identified in various studies, we first searched for homologs using several search algorithms (see Materials and Methods). Subsequently, the list of homologs was subjected to a phylogenetic analysis in order to define orthologous groups of MRPs.

Phylogenetic analyses

Maximum-likelihood (ML) trees were created using PHYML (version 2.2.4) (38) based on alignments of the above-mentioned orthologous groups. Prior to the analyses, global and pairwise gaps were removed from the alignments. The appropriate model of protein evolution was selected using MODESELECTOR (39). Bootstrap support values were derived from 100 replicates.

Detection of newly evolved MRP domains

All MRPs homologous to bacterial core RPs that were identified in the current research were examined for additional, potentially newly evolved domains. Using HMMER (33), each MRP was analyzed with an HMM profile that was built from an alignment of its ancestral bacterial counterpart. These parts of the MRPs that were not covered by the respective HMM profile (minimum length 30 amino acids), were searched against the PFAM (40) and CDD (41) databases using HMMER (33) and RPS-BLAST (41) respectively. Hits below an E-value threshold of 0.01 were regarded significant.

RESULTS

Detection of MRPs in eukaryotic genomes

We have employed a systematic analysis of nuclear and mitochondrial genomes in order to identify the mitoribosomal proteome content encoded in a wide variety of eukaryotic species. To avoid confusion regarding different nomenclatures that are used for ribosomal proteins, we have adopted the nomenclature for human MRPs, approved by the Human Genome Nomenclature Committee (http://www.gene.ucl.ac.uk/nomenclature/). All proteins are designated MRPSxx or MRPLxx, where S and L stand for small and large subunit, respectively, and xx is the number of the corresponding bacterial ribosomal protein. The MRPs lacking bacterial orthologs (supernumerary proteins) are assigned higher numbers, starting at MRPS22 and MRPL37 for the small and large subunit, respectively. In yeast, unfortunately, the nomenclature is as variable as the methods employed for their characterization. Therefore, in order to prevent further confusion about proteins with the same name in human and yeast that are not part of the same orthologous group, we have indicated all human MRPs in capital letters throughout the manuscript (e.g. MRPS28), whereas for the yeast proteins we only capitalized the first letter (e.g. Mrps28, which corresponds to MRPS15 in human, also see Table 1 and Supplementary Data).

Starting from a curated list of MRPs that have been identified in various studies, we first searched for homologs using several search algorithms (see Materials and Methods). Subsequently, the list of homologs was subjected to a phylogenetic analysis in order to define orthologous groups of MRPs.
Table 1. Phylogenetic distribution of SSU mitochondrial ribosomal proteins and their bacterial counterparts

<table>
<thead>
<tr>
<th>MRP</th>
<th>Yeast name</th>
<th>Hsa</th>
<th>Mmu</th>
<th>Tri</th>
<th>Dme</th>
<th>Aga</th>
<th>Cel</th>
<th>Sce</th>
<th>Ecu</th>
<th>Ddi</th>
<th>Ath</th>
<th>Cre</th>
<th>Prf</th>
<th>Tth</th>
<th>Ram (mt)</th>
<th>Tbr</th>
<th>Lma</th>
<th>Eco</th>
<th>BsU</th>
<th>Rpr</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>Mrp4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3/S24</td>
<td>Var1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>Nam9/Mna6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>Mrps5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S6</td>
<td>Mrp17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>Rsm7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>Mrps8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S9</td>
<td>Mrps9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S10</td>
<td>Rsm10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S11</td>
<td>Mrps18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S12</td>
<td>Ynr036c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S13</td>
<td>Sws2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S14</td>
<td>Mrp2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>Mrps28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S16</td>
<td>Mrps16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S17</td>
<td>Mrps17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S18</td>
<td>Rsm18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S18A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S18B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S18C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S19</td>
<td>Rsm19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S21</td>
<td>Mrp21/Mrp50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S23</td>
<td>Rsm25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S25</td>
<td>Mrp49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S29</td>
<td>Rsm23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S33</td>
<td>Rsm27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S35</td>
<td>Rsm24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S36</td>
<td>Ymr31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rsm22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ppe1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mrps35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mrp51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mrp13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mrp1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
The first and second column contain, if applicable, the MRP identifiers according to the human and yeast nomenclature, respectively. MRPs and bacterial RPs that were identified in the predicted proteome are indicated in yellow. MRPs that were not found in the predicted proteome, but for which the unpredicted gene could be identified on the genomic DNA are indicated in blue. MRPs that have been experimentally verified as being associated with the mitoribosome by means of large-scale proteomics interaction studies or small-scale functional studies are indicated in green (except for the human MRPs, for which mostly the bovine counterparts were experimentally verified). MRPs that are encoded by the mitochondrial genome are indicated in red. A bracket (§) denotes a peculiar case of a split MRPS3 gene in *D. discoideum*, which is encoded by its mitochondrial genome. The 5' and 3' ends are fused to genes of unknown function (ORF425 and ORF1740, respectively) (79,80). It is unknown if the split MRPS3 genes results in a functional protein. An asterisk (*) denotes cases where a given MRP was not identified most probably as a result of incompleteness of the genome sequence, i.e. the MRP was not found in the predicted proteome, nor was it found by analyzing the genomic DNA, but could be identified in a closely related species. See Supplementary for protein identifiers.


### Table 1. Continued

<table>
<thead>
<tr>
<th></th>
<th>Pet123</th>
<th>Rsm26</th>
<th>Mrp10</th>
<th>Rsm28</th>
<th>Mrp8</th>
<th>Mhr1</th>
</tr>
</thead>
<tbody>
<tr>
<td>LmjF03.0220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF03.0630</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF15.0080</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF15.0410</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF15.0780</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF15.1400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF16.1230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF22.0830</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF22.0990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF22.1220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF24.0830</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF25.2160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF26.2270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF27.0180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF27.0630</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF27.1895</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF28.2180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF29.0430</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF30.0650</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF30.0740</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF30.3220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF30.3530</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF32.0650</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF32.0010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF33.2510</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF35.2230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF35.3490</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF35.3940</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF36.4120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Not only does our analysis distinguish MRPs from their cytosolic counterparts, it also enabled for differentiation between ‘true’ MRPs and homologous ribosomal proteins of other organelles, such as the chloroplast (in plants) or apicoplast (in the apicomplexa *P. falciparum*). We realize that this approach has its limitations, as it does not account for potential retargeting events that have been shown to occur, such as in *A. thaliana*, where a chloroplast-type S13 protein encoded by a diverged duplicated nuclear gene has been demonstrated to be targeted to the mitochondrion (42). In order to account for such cases, we included such an experimentally verified protein in our list, whenever our analysis did not suggest a candidate MRP.

Altogether, our survey resulted in the identification of highly variable numbers of MRPs across the genomes of the investigated eukaryotic species, several of which had not been identified as MRPs previously (see Table 1, Table 2 and Supplementary Data). The total number of identified MRPs ranged from 81 in most metazoan species, to 80 in yeast, to 63 in plants, down to a mere 39 MRPs in the apicomplexan *P. falciparum*.

We recognize the fact that the latter numbers are probably an underestimate of the actual numbers of MRPs in these species, as there are no experimental studies to identify potential lineage-specific supernumerary MRPs. Moreover, since protein sequences of these species have been found to evolve relatively fast (43), it is at least feasible that we failed to identify some MRPs that have evolved beyond the detection limits of our search algorithms.

Apart from mapping MRPs that are encoded by eukaryotic genomes, we were able to find hitherto undetected homology between experimentally verified supernumerary MRPs and bacterial ribosomal proteins that were presumed to have been lost from the primitive eukaryotic mitoribosome. Using sensitive profile–profile searches for detection of distant homology between protein families, we connected the MRPS24 orthologous group to the bacterial RPS3 ribosomal protein family (E-value 2.2e–5) and similarly, we connected the MRPL47 orthologous group to the bacterial RPL29 ribosomal protein family (E-value 8.6e–7) (see Figure 1). The complementary phylogenetic distribution patterns of the respective genes are in support of this suggestion. For example, MRPL47 is present in all eukaryotes analyzed (except for *E. cuniculi* and *G. lamblia*), but not in bacteria. The opposite phylogenetic pattern is observed for bacterial ribosomal protein RPL29, for which no homolog was identified in any eukaryotic genome. Apparently, these MRPs have diverged almost beyond homology detection limits during the course of evolution (see Figure 1), possibly as an adaptation to specific mitoribosomal functionality. Given the detected homology between the above-mentioned MRPs and bacterial ribosomal proteins, we have placed the MRPs in the orthologous group of their respective bacterial counterparts (see Table 1 and 2).

We also encountered some cases where we could not detect an expected MRP ortholog in some species, which might be caused by incompleteness of the respective genomes. For example, despite the fact that we could identify orthologs of MRPL50, MRPS28 and Mrp10 in all metazoan analyzed in this survey, as well as in the *Danio rerio* genome (data not shown), we were unable to detect these MRPs in the genome of *T. nigroviridis*. Likewise, MRPL56 was found in all metazoa, including *A. gambiae*, but not in *D. melanogaster* (see Table 2). In addition, in some instances our analysis proposed different proteins than have been reported in earlier reports. The discrepancies between our study and that of others are outlined in the Supplementary Data.

Not unexpectedly, MRPs were not detected in the genomes of the microsporidian *E. cuniculi* and the diplomonad *G. lamblia*. These organisms do not contain full-fledged mitochondria, but instead they contain mitosomes (*E. cuniculi*) or mitochondrial remnants (*G. lamblia*), which typically lack organellar DNA. Consequently, these organisms have no need for mitochondrial translation machineries, explaining the observed loss of MRP-encoding genes.

### Identification of novel (human) MRPs

Our analysis expanded orthologous groups that were previously reported to contain fungi-specific MRPs to include other eukaryotic proteins, resulting in the identification of two potentially novel human MRPs, namely the human orthologs of yeast SSU proteins Rsm22 and Mrp10, extending the total number of human MRPs to 81. The detection of human orthologs of yeast MRP proteins Rsm22 and Mrp10 is of particular interest, as defects within these genes might be linked to mitochondrial dysfunctions in human. Interestingly, extensive biochemical data is available for the yeast gene products, enabling to speculate about their respective molecular functions.

Yeast Rsm22 is homologous to an rRNA methyltransferase and has been shown to be part of the SSU of the yeast mitochondrial ribosome (44). The association of Rsm22 with the SSU has also been confirmed in a large-scale affinity-capture mass spectrometry study (45). We have detected an ortholog of Rsm22 in most eukaryotes, as well as in some prokaryotes. Interestingly, the STRING database (46) shows that the gene encoding the Rsm22 ortholog in some prokaryotes is clustered on the genome with genes that encode ribosomal proteins (not shown), suggesting that the Rsm22 orthologs in these species are also associated with the ribosome. In *Schizosaccharomyces pombe*, the genes encoding Rsm22 and Cox11 are fused. Cox11 is an assembly factor of the OXPHOS enzyme complex cytochrome c oxidase and in *S. cerevisiae*, Cox11 has been shown to associate with the mitoribosome and is postulated to function in co-translational insertion of copper ions into Cox1, a subunit of cytochrome c oxidase (47). Finally, Rsm22-deficient yeast strains exhibit a growth defect on non-fermentable (respiratory) carbon sources, indicative of mitochondrial respiratory dysfunction (44). All these lines of evidence suggest that Rsm22 is a protein of the mitoribosomal small subunit that is essential for proper mitochondrial function. As Rsm22 is a predicted
Table 2. Phylogenetic distribution of LSU mitochondrial ribosomal proteins and their bacterial counterparts

<table>
<thead>
<tr>
<th>MRP</th>
<th>Yeast name</th>
<th>Hsa</th>
<th>Mmu</th>
<th>Tni</th>
<th>Dme</th>
<th>Aga</th>
<th>Cel</th>
<th>Sce</th>
<th>Ncr</th>
<th>Ecu</th>
<th>Ddi</th>
<th>Ath</th>
<th>Cre</th>
<th>Pfa</th>
<th>Tfa</th>
<th>Tth</th>
<th>Ram (mt)</th>
<th>Tbr</th>
<th>Lma</th>
<th>Gia</th>
<th>Eco</th>
<th>Bsu</th>
<th>Rpr</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Mrpl1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>Mrpl2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L3</td>
<td>Mrpl9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L4</td>
<td>Yml6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L5</td>
<td>Mrpl7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L6</td>
<td>Mrpl6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L7</td>
<td>Mrpl50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L8</td>
<td>Mrpl11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L9</td>
<td>Mrpl19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L10</td>
<td>Mrpl1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L11</td>
<td>Mrpl19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L12</td>
<td>Mrpl1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L13</td>
<td>Mrpl23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L14</td>
<td>Mrpl3/Mrpl34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L15</td>
<td>Mrpl10/Mrpl18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L16</td>
<td>Mrpl16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L17</td>
<td>Mrpl6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L18</td>
<td>Mrpl1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L19</td>
<td>Mrpl49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L20</td>
<td>Mrpl22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L21</td>
<td>Mrpl41/Mrpl20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L22</td>
<td>Mrpl40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L23</td>
<td>Mrpl2/Mrpl7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L24</td>
<td>Mrpl24/Mrpl14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L25</td>
<td>Mrpl4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L26</td>
<td>Mrpl33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L27</td>
<td>Mrpl36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L28</td>
<td>Mrpl32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L29</td>
<td>Mrpl39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L30</td>
<td>Ydr115w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L31</td>
<td>Ypl183w-a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L32</td>
<td>Mrpl35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L33</td>
<td>Mrpl3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L34</td>
<td>Mrpl28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L35</td>
<td>Mrpl27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L36</td>
<td>Mrpl51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L37</td>
<td>Mrpl31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L38</td>
<td>Mrpl30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L39</td>
<td>Mrpl3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L40</td>
<td>Mrpl2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L41</td>
<td>Mrpl1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L42</td>
<td>Mrpl17/Mrpl30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 2. Continued

For explanation of used MRP protein names, color coding and species abbreviations, see comments below Table 1. Subunits that are indicated to be experimentally characterized for L. major were in fact identified to be part of the mitoribosome of the closely related species L. tarentolae (17). A bracket (\(\#\)) denotes a special case of the A. thaliana MRPL2, of which the N-terminal half is encoded by the mitochondrial genome and the C-terminal half is nuclear encoded. See Supplementary Data for protein identifiers. An asterisk (*) denotes cases where a given MRP was not identified as a result of incompleteness of the genome sequence, i.e. the MRP was not found in the predicted proteome, nor was it found by analyzing the genomic DNA, but could be identified in a closely related species. For example, MRPS28 and MRP10 were not detected in T. nigroviridis, but could be found in Danio rerio. For example, MRPL56 was not detected in D. melanogaster, but was found in the genomes of D. willistoni and A. gambiæ. Originally, MRPS32 was identified as a protein in the small subunit (9) and MRPL42 as a protein in the large subunit [MRPL31 in (89)]. However, the protein sequences of these proteins were found to be identical. Later, MRPS32 was omitted since its identification was probably caused by contamination in the preparation of small subunits, and thus was in fact MRPL42 (Thomas O'Brien, personal communication).

Figure 1. Distant homology between mitochondrial and bacterial ribosomal protein families detected in this study. Multiple sequence alignments of MRPL47 and RPL29 proteins (A) and of MRPS24 and RPS3 proteins (B). Amino acid residues are shaded according to the following physicochemical properties: \(t\): tiny (blue font on yellow shading); \(m\): amphoteric (red font on green shading); \(h\): hydrophobic (white font on black shading); \(p\): positive (red font on blue shading); \(p\): polar (black font on green shading); \(s\): small (green font on yellow shading); \(a\): aliphatic (red font on gray shading), \(r\): aromatic (blue font on gray shading); \(c\): charged (white font on blue shading). Sequences are annotated by the organism name and Swiss-Prot protein identifier and the boundaries of the segment that is used in the alignment are indicated. A 90% consensus sequence is depicted below the alignments using the same classification, with invariant residues being indicated in capitals.
rRNA methyltransferase, it might be involved in methylation of mitochondrial rRNA, and as such play a role in the regulation of protein translation efficiency.

Yeast Mrp10 was shown to be a component of the SSU of mitoribosomes as well and disruption of the gene resulted in a mitochondrial translation defect and a tendency to accumulate deletions in mtDNA (48). Jin et al. observed low sequence similarity to yeast Mrp10, a constituent of the LSU (49). However, we identified yeast Mrp10 as an ortholog of human MRPL54 and did not detect any significant sequence homology between Mrp10 and Mrpl37. Nevertheless, we did find Mrp10 orthologs in many eukaryotes, all of which were found to contain a CHCH domain (Coiled coil 1—Helix 1—Coiled coil 2—Helix 2, PFAM domain 06747). Within each helix of this CHCH domain, two invariant cysteine residues are present in a C–X₉–C motif. Besides in Mrp10, CHCH domains have also been detected in Cox19 and in the respiratory Complex I (NADH-ubiquinone oxidoreductase) 19 kDa subunit (NDUFA8) (50). In yeast, Cox19 is present in the cytoplasm as well as in the inter membrane space of the mitochondria, where it is suggested to play a posttranslational role in the assembly of cytochrome c oxidase through copper transport to mitochondria. It has been proposed that the conserved cysteine residues of the CHCH motif in Cox19 mediate as ligands for the copper ions (51,52). Most likely, yeast Mrp10 and their orthologs in other eukaryotes, including human, constitute an essential component of the mitochondrial translation machinery and we anticipate that the CHCH domains that are present in this orthologous group, as well as in Complex I subunit NDUFA8, are involved in trafficking of metal ions.

We also detected a potential third novel human MRP, Ppe1, however this case is less clear-cut than these described above. Ppe1 is a protein, which has been shown to be part of the SSU of the yeast mitoribosome (53). We detected Ppe1 orthologs in all eukaryotes analyzed, except for *P. falciparum* and *T. thermophila* (and the mtDNA lacking eukaryotes). However, the localization of Ppe1 in the mitochondrion is questionable, as the report by Kitakawa *et al.* is the only indication for this (53). Moreover, the yeast and human Ppe1 have previously been reported to function as protein phosphatase methylesterases of Protein phosphatase 2A (54,55), which are mainly localized in the nucleus and cytoplasm (56). Additionally, yeast Ppe1 mutant strains do not exhibit any growth defects (55), while mutations affecting synthesis or function of respiratory chain components generally result in growth impairment on a non-fermentable carbon source. Thus currently there is, besides the reported association with SSU of the mitoribosome in yeast, no clear indication that Ppe1 is implicated in mitochondrial translation. Therefore, additional research is needed to corroborate the cellular localization and function of yeast Ppe1 and its orthologs in other eukaryotes.

### The evolution of the mitochondrial ribosomal proteome along eukaryotic lineages

The protein composition of mitoribosomes is variable, not only in terms of the number of MRPs, but also in terms of the specific MRPs contained (see Tables 1 and 2). By mapping the distribution of the orthologous groups of MRPs onto a reconstructed tree of the eukaryotes (35) while always considering the most parsimonious scenario, we obtained an overview of the history of gains and losses of MRPs during the divergence of eukaryotes (see Figure 2). Mitoribosomes originated from the bacterial ribosome present in the alpha-proteobacterial ancestor of mitochondria, consisting of a microbial core of 54 ribosomal proteins in total. Already in the earliest stage of eukaryotic evolution, prior to the major divergence of the fungal, plant and metazoan lineages, several supernumerary proteins were recruited to the mitoribosome, while only one bacterial core protein was lost at this stage (RPS20, see Figure 2), resulting in a primitive eukaryotic mitoribosome of 68 MRPs. Subsequent to the gain of this primary set of 15 proteins, the mitoribosome diversified along the various eukaryotic lineages and a number of lineage-specific gain and loss events can be observed. Another gain of MRPs occurred at the level of the metazoa, prior to the bifurcation of nematodes and eumetazoan. At this stage, 14 proteins were recruited to the metazoan mitoribosome, whereas seven bacterial core proteins were lost. Major protein gains at these two time points have also been observed in the evolutionary history of Complex I (57). Additionally, 11 fungi-specific MRPs were recruited after the animal-fungi radiation. In the lineage towards the kinetoplastida, a major surge of 29 proteins is observed. These proteins have been shown to comprise a protein complex that is associated with the SSU of kinetoplastid-mitoribosomes (17). The role of this additional protein complex and its association with the SSU is still unclear. Apart from this major gain in the kinetoplastida, we observe loss of several proteins in the AVE (alveolates, viridiplantae and excavates) lineage, predominantly consisting of bacterial core proteins, but also including supernumerary MRPs in some lineages. Possibly, some of these proteins remained undetected due to sequence divergence. The fact that the AVE lineages are dominated by loss events of MRPs is likely due to the lack of experimental data of mitoribosomal content in these species.

Compared to the 54 ribosomal proteins present in the alpha-proteobacterial ancestor, 81 MRPs in human represent a substantial gain in complexity (see Figure 2B). However, the observed increase in complexity following the endosymbiotic event is not uncommon, as it has also been reported for electron transport chain complexes (57). It has been proposed that the recruited supernumerary proteins are mainly involved in the assembly and stabilization of the complexes (57,58). The mammalian supernumerary MRPs are shown to be mostly localized to the peripheral regions of the mitoribosome (59–61). Since all proteins synthesized by metazoan mitoribosomes are inserted into the inner mitochondrial membrane, at least...
some of the metazoa-specific supernumerary MRPs are presumed to be associated with positioning the mitoribosome during co-translational insertion of nascent polypeptides into the membrane. Interestingly, a protein that has been proposed to be essential for co-translational insertion in yeast, Mba1 (MRPL45, see ‘Origins from outside of the mitoribosomal proteome—MRP recruitment’), was already part of the primitive eukaryotic mitoribosome and has been recruited from the alpha-proteobacterial ancestor. Similar to the evolution of Complex I, we observe the recruitment of proteins of bacterial origin to the mitoribosome (MRPL45 and Rsm22), however, this number is considerably lower (only two, compared to six in Complex I), especially considering that in the evolution of the mitoribosome more proteins were added compared to Complex I (37 and 32 in the human mitoribosome and Complex I, respectively).

**Tracing the origins of supernumerary MRPs**

As outlined above, the protein content of the mitoribosome has significantly increased in most species during the evolution of the mitochondria. Our analysis indicates that in several stages of the mitoribosomal evolution various extra proteins, the supernumerary MRPs, have been added to the mitoribosomal proteome. Some of these
supernumerary MRPs were found to be present in all eukaryotes that were investigated, but not in bacteria, such as MRPS29 (see Table 1). Clearly, these MRPs have been recruited in an early stage of the evolution of the eukaryotes, before the formation of major eukaryotic kingdoms. In contrast, other MRPs have originated relatively recently, such as MRPL37, which is specific for the metazoa. However, it is not only interesting to know ‘when’ certain MRPs were added to the mitoribosome, it is also enticing to trace their origins. We found that several of the supernumerary MRPs are homologous to other proteins, some of which are also part of the mitoribosome (see Table 3). The latter finding implies that gene duplication events have played a substantial role in the expansion of the mitoribosomal proteome (see Table 3). Below we will discuss some of the most protruding cases.

Origins from within the mitoribosomal proteome—MRP duplication

Some supernumerary MRPs are homologous to MRPs that are part of the bacterial core of the mitoribosome. These proteins have seemingly been duplicated from within the mitoribosomal proteome. One of the most straightforward examples is the emergence of three variants of MRPS18 in the metazoan lineage resulting from gene duplication events (see Figure 3). In contrast with earlier reports (9), we identified three MRPS18 variants in C. elegans and we found that all three metazoan variants emerged from the same ancestral metazoan sequence, which is unlikely to be derived from chloroplast S18, as was reported by Cavdar Koc and co-workers (9) (see Figure 3). All three C. elegans MRPS18 variants seem to be essential in embryonal development, as targeted disruption of these genes result in an embryonic lethal phenotype (62). Independent from the observed MRPS18 duplications in metazoa, it is interesting to note that most sequenced actinobacterial genomes encode two variants of the S18 gene, probably the result of a gene duplication event in these species. Currently, experimental evidence regarding the functions of the different MRPS18 versions in metazoa is lacking, but it is believed that each mitoribosome contains only one copy of MRPS18, giving rise to a heterogeneous population of mitoribosomes (9).

Another example can be found in the duplication of MRPS10 in the lineage of the metazoa, which gave rise to MRPL48 (see Table 3 and Figure 4). This duplication event constitutes one of the cases where the duplicate gene product has become part of the other mitoribosomal subunit, in this case from the SSU (MRPS10) to the LSU (MRPL48). Likewise, a duplication event at the base of the metazoan radiation seems to have given rise to the supernumerary MRPs MRPS30 and MRPL37, both of which are only present in metazoa (see Figure 5).

Interestingly, MRPS30 is also known as programmed cell death protein 9 (PDCD9 or p52) (63,64). Another supernumerary protein, MRPS29, is also implicated in apoptosis, as it has been shown to be identical to death-associated protein 3 (DAP3).

Table 3. Paralogous relationships involving MRPs

<table>
<thead>
<tr>
<th>MRP</th>
<th>Mitochondrial [cytosolic] ribosome homolog</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S10</td>
<td>L48, [S20]</td>
<td></td>
</tr>
<tr>
<td>S18</td>
<td>S18A, S18B, S18C</td>
<td></td>
</tr>
<tr>
<td>S26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S30</td>
<td>L37</td>
<td></td>
</tr>
<tr>
<td>Mrp1</td>
<td>Rsm26 (yeast)</td>
<td></td>
</tr>
<tr>
<td>Pet123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mrp10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rsm28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF30.3530</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF25.2160</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LmjF32.0650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>[L8]</td>
<td></td>
</tr>
<tr>
<td>L38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L43</td>
<td>S25</td>
<td></td>
</tr>
<tr>
<td>L44</td>
<td>Mrp15, Mrp3 (yeast)</td>
<td></td>
</tr>
<tr>
<td>L45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mrp20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Names of MRPs are indicated in the first column and supernumerary MRPs and proteins from outside the ribosome are indicated in the second and third column, respectively. In addition, if a homologous cytosolic ribosomal protein is present, it is indicated between square brackets in the second column.

protein 3 (DAP3) (63,65). Given the dual roles of these MRPs, the processes of mitochondrial translation and apoptosis are apparently closely coupled (66).

**Origins from outside of the mitoribosomal proteome—MRP recruitment**

We found that some supernumerary MRPs have probably been recruited to the mitoribosome, as they display significant homology to other mitochondrial proteins, some of which are also implicated to play a role in mitochondrial translation (see Table 3). Of course, it cannot always be resolved if an MRP has been part of the mitoribosome prior to the duplication event. For some instances, we can use the phylogenetic distribution patterns of the copies in order to point out which was first. One of such examples, where a pre-existing protein has been recruited to the mammalian mitoribosome, is the acquisition of MRPL39, which is homologous to the N-terminal domain of threonyl-tRNA synthetase and which has a universal phylogenetic distribution. This homology has been noticed before and it is suggested that MRPL39 can bind tRNA in a similar manner as the *E. coli* threonyl-tRNA synthetase (67). Conceivably, the N-terminal tRNA-binding domain of a mitochondrial threonyl-tRNA synthetase was recruited to the mitoribosome in order to compensate for the loss of bacterial ribosomal proteins that are involved in tRNA binding, such as RPS13 and RPL5 (see Tables 1 and 2).

The most striking example of MRP recruitment is that of MRPL45 (see Figure 6). First we found that MRPL45 and yeast protein Mba1 belong to the same orthologous group, which in addition comprises proteins from all metazoa, the plant *A. thaliana*, the alga *C. reinhardtii* as well as from the mycetozoan *D. discoideum*. Mba1 is a mitochondrial protein associated with the matrix face of the inner membrane, which associates with the highly conserved inner membrane protein Oxa1. Mba1 and Oxa1 orchestrate the insertion of both mitochondrial- and nuclear-encoded proteins from the mitochondrial matrix into the inner membrane (68,69). Moreover, we detected homology between MRPL45 (including Mba1) and Tim44 (see Figure 6), a peripheral mitochondrial inner membrane protein associated with the matrix face of the inner membrane, which is involved in the insertion of both mitochondrial- and nuclear-encoded proteins from the mitochondrial matrix into the inner membrane (68,69).
membrane protein that is an essential component of the import motor PAM of the TIM23 complex, which is implicated in the translocation of proteins from the intermembrane space to the mitochondrial matrix [for reviews see (70,71)]. The apparent sequence homology, as well as functional homology (i.e. protein translocation), suggests that MRPL45 and Tim44 share a common alpha-proteobacterial origin. Many alpha-proteobacteria contain two distinct types of Tim44/MRPL45-like genes (see Figure 6).

A close examination of MRPL45 and Tim44 proteins reveals that the N-terminal region of Tim44 is absent in MRPL45 (not shown). This region is supposed to be located in the mitochondrial matrix where it interacts with mitochondrial Hsp70 chaperone, which is also a component of the PAM complex (72). Apart from its association with Oxa1, MRPL45 has been shown to be a constituent of the mitoribosomal LSU in multiple studies (10,73). Moreover, it has been shown that in yeast, transcription of MBA1 is tightly co-regulated with that of genes encoding MRPs (74), implying a functional link between the respective gene products. The anticipated association of the mitoribosome and membrane-bound constituents of the mitochondrial protein insertion machinery would explain the fact that up to 50% of the bovine mitoribosomes are found to associate with the inner membrane fraction of mitochondria (75). Taking all this in consideration, we postulate that MRPL45 functions as a bridging factor between the mitoribosome and Oxa1, and as such constitutes an essential component for successful co-translational insertion of proteins into the inner membrane of mitochondria in eukaryotes.

Interestingly, the MRPL45 orthologous group, which also includes the bacterial Tim44-like proteins, shows a similar phylogenetic distribution pattern as the orthologous group of Cox11, an assembly factor of cytochrome oxidase, which is involved in the co-translational insertion of copper ions into the nascent Cox1 protein (47). The observed pattern of co-presence and absence of these genes across a phylogenetically diverse set of species indicates that their gene products are involved in the insertion of copper ions into the nascent Cox1 protein (47).
might be involved in the same process (76). As Cox11 has been shown to associate transiently with the mitoribosome as well, it is feasible that MRPL45 and Cox11 act together in order to functionally insert Cox1 into the inner mitochondrial membrane (47).

Other examples of MRPs that have been recruited to the mitoribosome are the supernumerary proteins MRPL43 and MRPS25 (see Table 3), which are homologous to each other as well as to the B8 subunit (N18M) of respiratory Complex I (NADH-ubiquinone oxidoreductase) of the OXPHOS system (57,77). Presumably, MRPL43 and the B8 subunit of Complex I originated from a single gene that was present in the primitive eukaryote, as they are present in a phylogenetically wide range of organisms. Most likely, another duplication of MRPL43 then gave rise to MRPS25, which is only found in metazoa and fungi. More examples of MRP recruitment include yeast MRPs Mrpl and Rsm26, which are both part of the iron/manganese superoxide dismutase family (44), and MRPL44, which is homologous to a wide range of double-stranded RNA-binding proteins (see Table 3).

Detection of orthologs of MRPs previously reported to be fungi specific

Several comprehensive experimental studies on the yeast mitochondrial translation machinery have significantly extended its mitoribosomal proteome (14–16). However, for a large number of MRPs that have been characterized in yeast thus far, no significant or conclusive sequence similarity has been reported to ribosomal proteins from other sources. In the current study, we were able to link seven of these seemingly fungi-specific MRPs to known ribosomal proteins in other species, as will be outlined below.

Apart from the homology between yeast Mba1 and MRPL45, which has already been discussed above, we identified Mrp49 as the ortholog of MRPS25. It is, however, surprising that the orthologs of MRPS25 found in S. cerevisiae as well as in N. crassa are reported to be located in the large instead of the small subunit of the mitoribosome (18,78), in contrast with the location of bovine MRPS25 (59). It is possible that the subunit localization of the MRPS25 orthologs is different between metazoa and fungi; however, further experimental evidence is needed in order to clarify this.

We also connected Rsm27, a yeast protein of the small subunit (44), to the MRPS33 protein family. In addition, we detected MRPS33 orthologs in all metazoa as well as D. discoideum and A. thaliana. Moreover, we found that yeast MRPs Mrpl28 (53), Mrpl44 (79), Mrpl50 (53) and Mrpl40 (53) are in fact the orthologs of MRPL40, MRPL53, MRPL9 and MRPL24, respectively. Both MRPL24 and Mrpl40 were found to contain a KOW motif (PFAM domain pfam00467), which is present in a variety of ribosomal proteins as well as in the bacterial transcription elongation factor NusG (80). The homology between Mrpl50 and bacterial RPL9 has been noted before (14), but it was dismissed as being insignificant, since the sequence similarity was not conserved throughout the full length of Mrpl50. In our analysis, however, we did detect significant homology between Mrpl50 and the MRPL9 family (including bacterial RPL9). Our finding is supported by the fact that Mrpl50 and MRPL9 representatives reside within the same protein superfamily (SSF55658, ‘L9 N-domain-like’) (81).

Finally, we were able to confirm the distant homology between Var1 and Rps5, two mtDNA-encoded homologs of bacterial RPS3 in fungi, as was previously reported by Lang and colleagues (82). Var1 and Rps5, which are the only mtDNA-encoded MRPs in yeast and N. crassa, respectively, have both been identified as essential components of the SSU of the mitoribosome, showing similar phenotypes in mutant strains (82). Despite the fact that RPS3 has been shown to play an essential role in bacterial ribosome assembly, and that Var1 and Rps5 are both essential for assembly of the mitochondrial SSU in fungi, we were unable to detect candidate mitochondrial RPS3 homologs in the metazoa. In the latter organisms, this essential role in ribosome assembly could be performed by MRPS24, which displays distant homology to RPS3 (see Figure 1). Moreover, we could detect mtDNA-encoded RPS3 homologs in T. thermophila and D. discoideum (82,83), the latter of which constitutes a peculiar case where the RPS3 encoding gene is split and both the 5’ and 3’ parts are fused to an ORF of unknown function (ORF425 and ORF1740, respectively). In any case, the fact that genes encoding RPS3 homologs are present on several mitochondrial genomes implies an alpha-proteobacterial origin of this protein.

Adaptive evolution of core MRPs

Compared to bacterial ribosomes, mammalian mitoribosomes contain scarcely half as much rRNA and over twice as much protein, due to the presence of enlarged bacterial core proteins and supernumerary proteins (6). Previously, it has been hypothesized that these proteins might compensate for the loss of rRNA segments in the mitoribosome (6,84). Recently however, it was shown that many supernumerary MRPs occupy new quaternary positions in both the large and small subunits of the mitoribosome (61), inconsistent with this hypothesis.

We compared the sequences of all core MRPs with their bacterial counterparts in order to explore the possibility that core MRPs contain additional domains that perform extra, mitochondria-specific functions. Indeed, we observe that MRPs are significantly larger (on average almost twice as large) than their bacterial ancestral sequences (see Supplementary Data, Figure S8). In some cases, we could identify functional domains within these newly evolved regions, which we expect to be linked to ribosome assembly or function (see Figure 7). For instance, we identified a RNA recognition motif (RRM_L, pfam00076) and a copper-binding domain (Cu_bind_like, pfam02298) in the N-terminal regions of A. thaliana MRPS19 and MRPL22, respectively (see Figure 7). The RRM motif that is present in MRPS19 (85) is found in a wide variety of RNA-binding proteins, which are implicated in a wide variety of cellular processes. Hence, the motif is most likely
involved in increasing the binding affinity to and/or stabilization of the tertiary structure of rRNA. The role of the copper-binding domain in MRPL22 is less clear. It might be involved in the co-translational insertion of copper ions into proteins that require copper for activity, in a similar way as has been observed for the cytochrome c oxidase assembly factor in yeast, Cox11 (47) (also see above). The functions of other functional domains that are fused to MRPs are still obscure, such as a bacterial-type membrane protein (COG5373) to MRPL27 in N. crassa (see Figure 7).

DISCUSSION

In the present study, we have mapped the evolution of the mitochondrial ribosomal proteome using a comparative genomics approach that combined the most recent experimental data and computational techniques. The results of our study hint at a complex evolutionary scenario in which an ancestral ribosome of alpha-proteobacterial descent doubled its amount of protein in most eukaryotic lineages by elongation of its proteins and by recruiting new proteins of diverse origins. We observe that the mitoribosome has a complex history in terms of MRP gene gain and loss events during evolution. A major protein gain of the mitoribosome in eukaryotic evolution occurred prior to the divergence of the main eukaryotic domains, resulting in a primitive eukaryotic mitoribosome of 68 MRPs. Subsequently, the mitoribosome diversified along the different eukaryotic lineages with a continuous recruitment of new proteins, including a large gain after the bifurcation of the animal-fungi clade in each lineage. Another notable surge has taken place in the kinetoplastida lineage, where 29 new proteins were found to form a new, distinct protein complex with unknown function, which associates with the SSU of the *Leishmania* mitoribosomes (17).

In addition to the gain of MRPs, we also observe that numerous MRPs have been lost during the course of evolution, some of which have occurred independently in different lineages, as guided from ribosomal protein content that is encoded by the ‘ancestral’ mitochondrial genome of *R. americana* (5). The extensive loss of genes encoding core MRPs in some lineages contrasts with the evolutionary scenario that is observed for another mitochondrial protein complex with an alpha-proteobacterial ancestry, respiratory Complex I (NADH:Ubiquinone oxidoreductase). During the evolution of this protein complex, which has approximately tripled its size in most eukaryotes, hardly any gene loss was observed, except for these lineages in which the complete respiratory complex was lost (57). Moreover, the bacterial core of Complex I has been restricted from gene loss events in all eukaryotic genomes analyzed thus far, which contrasts with the scenario of the mitoribosome. What is the underlying reason for the observed differences in the evolutionary trajectories? Evidently, some of the core ribosomal proteins are not as essential as the core subunits of Complex I for proper functioning of the protein complex, and are therefore expendable. This is the case for RPS20 and RPL25, which are absent from most mitoribosomes, for example. RPS20-deficient *E. coli* strains were found to be viable, but showed an increased misreading ability of nonsense codons (86); RPL25-defective and wild-type *E. coli* ribosomes were found to translate at the same rate *in vitro*, albeit that the mutant type was less efficient (87). In contrast, some MRPs that are missing in most eukaryotic species, such as MRPS1 and MRPL5, are indispensable for translation in *E. coli* and result in lethal phenotypes when mutated (87,88). Moreover, MRPL5-deficient yeast strains display a respiratory-deficient phenotype, probably caused by a defect in the mitochondrial translation machinery (15). Apparently, the evolution of the mitoribosomal proteome involved several drastic events that altered the absolute necessity of some ribosomal proteins.
In addition, the present study was performed in order to prioritize MRPs for their potential involvement in mitochondrial respiratory diseases. Based on the fact that essential genes appear more evolutionarily conserved than non-essential genes, the bacterial core ribosomal proteins and translation factors that are conserved in most eukaryotes constitute excellent screening targets for human mitochondrial disorders. Mutations in non-essential MRPs could be risk factors for these diseases. Thus far, mutations in patients with mitochondrial respiratory disorders have been reported for the gene encoding small ribosomal subunit protein MRPS16 (28). But what about the supernumerary MRPs? In general, they are phylogenetically less conserved than the bacterial core proteins; however, their recruitment bears functional relevance, as is indicated by knock-out studies in yeast and *C. elegans*. As such, the two novel human MRPs, orthologs of yeast Rsm22 and Mrp10, are good candidates for unresolved mitochondrial disorders, since they show a wide phylogenetic distribution, have not been investigated before and yeast mutants show clear respiratory-deficient phenotypes. Translation-related mitochondrial disorders are not solely caused by mutations in mitoribosomal genes, but also by mutations in genes encoding factors that functionally interact with the mitoribosome, such as mitochondrial translation elongation factors EFG1 (25) and EFTs (26). In addition to the newly discovered human MRPs, these interactors constitute excellent screening targets for human patients with unresolved mitochondrial OXPHOS system disorders.

Altogether, our study reveals that the mitoribosomal proteome has, after initially having been acquired via the alpha-proteobacterial endosymbiont, been subjected to a complex evolution that involved numerous gain and loss events, resulting in a large variety in mitoribosomal protein content between different eukaryotic lineages. Potentially, the increased amount of proteins that seem to be associated with all modern eukaryotic mitoribosomes might reflect specific adaptations to mitochondria-specific functions, or, it might even reflect the increased complexity of eukaryotes themselves. Finally, we expect that additional experimental studies of mitoribosomal proteomes, such as the one that was recently performed on *Leishmania* mitoribosomes, will reveal an even more dramatic picture of the evolutionary trajectory of this protein complex. It will be a major challenge to gain more insight in the functional roles of the newly acquired MRPs.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

This work was supported by a Rubicon grant from The Netherlands Organization for scientific research (NWO) and by the European Union’s 6th Framework Program, contract number LSHM-CT-2004-503116 (EUMITOCOMBAT) and contract number LSHM-CT-2004-005260 (MITOCIRCLE). We thank Siv Andersson (Evolutionary Biology Center, Uppsala University, Sweden) for reading the manuscript and for helpful suggestions prior to submission. Funding to pay the Open Access publication charges for this article was provided by the European Union’s 6th framework program EUMITOCOMBAT.

**Conflict of interest statement.** None declared.

**REFERENCES**


