Covariation of Mitochondrial Genome Size with Gene Lengths: Evidence for Gene Length Reduction During Mitochondrial Evolution

André Schneider, Dieter Ebert

Department of Biology, University of Fribourg, Chemin du Musee 10, 1700 Fribourg, Switzerland

Abstract. Reduction of genome size and gene shortening have been observed in a number of parasitic and mutualistic intracellular symbionts. Reduction of coding capacity is also a unifying principle in the evolutionary history of mitochondria, but little is known about the evolution of gene length in mitochondria. The genes for cytochrome *c* oxidase subunits I–III, cytochrome *b,* and the large and small subunit rRNAs are, with very few exceptions, always found on the mitochondrial genome. These resident mitochondrial genes can therefore be used to test whether the reduction in gene lengths observed in a number of intracellular symbionts is also seen in mitochondria. Here we show that resident mitochondrial gene products are shorter than their corresponding counterparts in α-proteobacteria and, furthermore, that the reduction of mitochondrial genome size is correlated with a reduction in the length of the corresponding resident gene products. We show that relative genomic AT content, which has been identified as a factor influencing gene lengths in other systems, cannot explain gene length/genome size covariance observed in mitochondria. Our data are therefore in agreement with the idea that gene length evolves as a consequence of selection for smaller genomes, either to avoid accumulation of deleterious mutations or triggered by selection for a replication advantage.

Introduction

Gene shortening in small genomes has been observed in a number of obligate intracellular symbionts (parasitic and mutualistic) including diverse groups such as symbiotic bacteria (*Buchnera*) in aphids (Charles et al. 1999), a microsporidium infecting humans (Vivares et al. 2002), and the vestigial nuclear genomes of the nucleomorphs of cryptomonads (Cavalier-Smith 2002). It has been suggested that the reductions of gene lengths and genome size which are observed in all these organisms are a consequence of their intimate symbiotic life style (Wernegreen 2002). Mitochondria can also be considered obligate intracellular symbionts and a hallmark of their evolution is the reduction in genome size (Andersson and Kurland 1998). While genome size evolution of mitochondria has been studied intensively, little attention has been paid to gene length evolution. If reductions in gene length and genome size are general features of obligate intracellular symbionts, we expect that mitochondria should have shorter genes than their putative ancestors and, further, that this reduction in gene length covaries with the reduction in genome size.

Mitochondria are descendants of free-living α-proteobacteria (Andersson et al. 1998). Reduction of the coding capacity of their genome is a unifying

Correspondence to: André Schneider; *email:* andre.schneider@unifr.ch

principle in their evolutionary history and is mainly achieved by two means. First, genes were lost from the mitochondrial genome since they became nonessential due to the intracellular life style of mitochondria. Many genes coding for biosynthetic activities became obsolete because metabolites could be imported from the cytosol of the host. Second, most of the mitochondrial genes which are essential for organellar function were transferred to the nuclear genome. The proteins encoded by the transferred genes are then synthesized in the cytosol and subsequently imported into mitochondria (Gray et al. 1999; Kurland and Andersson 2000; Lang et al. 1999). Three main models have been proposed to explain the evolution of the mitochondrial genome (Berg and Kurland 2000; Blanchard and Lynch 2000). All state that a gene transferred from the mitochondrion to the nucleus capable of producing a functional mitochondrial protein would render the corresponding ancestral mitochondrial gene redundant. As a consequence the mitochondrial gene would be inactivated and eventually be lost by mutation. The models, however, differ in explaining why the transfer of mitochondrial genes occurred and whether it had a selective advantage. The *unidirectional transfer hypothesis* suggests that nuclear transfer of mitochondrial genes is essentially a neutral event but that since the DNA transfer mechanism is biased toward the nucleus all genes may eventually end up in the nucleus. This mechanism has been termed the "gene transfer ratchet" (Doolittle 1998). The model is supported by experimental data obtained in yeast showing that transfer of DNA from the mitochondria to the nucleus is at least 10^5 times more frequent than in the other direction (Thorness and Fox 1990). While this model explains genome size reduction, it would not explain gene length reduction. *Müller's ratchet* explains the advantage of nucleus-encoded mitochondrial proteins with a reduced rate of accumulating deleterious mutations (Lynch 1996). Many mitochondrial genomes, probably due to oxygen radicals created during respiration, have a higher mutation rate than their corresponding nuclear DNA (Allen and Raven 1996). Müller's ratchet states that in the absence of recombination, which is typical for many intracellular symbionts including mitochondria, these deleterious mutations can accumulate and lead to a mutational meltdown. Finally, the *replication advantage hypothesis* states that a smaller mitochondrial genome would be selected for in intracellular competition due to its faster replication rate (Selosse et al. 2001). Both Müller's ratchet and the replication advantage hypothesis predict that genomes and genes should evolve to become smaller. Thus, these two hypotheses predict a positive covariance between gene length and genome size when analyzed across taxa.

The extent of nuclear transfer of mitochondrial genes varies greatly among species. Interestingly there is a group of mitochondrial genes which rarely have been transferred to the nucleus and are therefore well suited to test the hypothesis on the evolution of gene length in mitochondria. These genes encode subunits I–III of cytochrome *c* oxidase (coxI–III) and cytochrome *b* (cytb), as well as the small (ssu) and large (lsu) subunit rRNAs (Gray et al. 1998, 1999: Lang et al. 1999). It has been suggested that the presence of coxI–III and cytb genes in the mitochondrial genome is essential since their protein products can regulate their expression in response to the redox potential of the mitochondrion (Allen 2003). The rRNA genes might be refractory to nuclear transfer since it may not be possible to import long RNA molecules.

Here we test whether the reduction in gene length observed in a number of obligate intracellular symbionts is also seen in the cox, cytb, and rRNA genes of mitochondria.

Materials and Methods

Data for α-proteobacteria were collected from the http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/eub_g.html Web site using annotated information for the ssu/lsu rRNA genes (see supporting table in the electronic apendix). In order to identify the closest homologues of mitochondrial coxI–III and cytb the genome of each α-proteobacteria was analyzed using the BLAST search software.

Predicted lengths of gene products, genome sizes, and AT contents for up to 278 eukaryotes (see supporting Table), whose mitochondrial genomes have been completely sequenced and annotated, were collected using sequence information deposited on the http://www.megasun.bch.umontreal.ca/ogmp/projects/other/mt_list.html Web site (Shimko et al. 2001). For species whose mitochondrial genome has only partially been analyzed, annotated sequence information was obtained from the http://oberon.rug.ac.be: 8080/rRNA/ Web site. As coxI–III function in the same protein complex, we combined their lengths for the analysis. The same was done with the ssu and lsu rRNAs, both of which are components of the mitochondrial ribosomes. We attempted to use a data set which is as complete as possible, however, as indicated below some cases in which the gene length could not be determined precisely due to special evolutionary events had to be excluded. For the coxI–III data set this concerns *Scenedesums obliquus*, since a parts of its coxIII gene had moved to the nucleus; *Pylaiella littoralis* and *Laminaria digitata*, since their coxII gene contains a large in frame insertion of unknown function (Secq et al. 2001); and *Acanthamoeba castellanii* and *Dictyostelium discoideum*, because their coxI–II genes are represented by a single open reading frame only. For the ssu/lsu rRNA data all apicomplexans were excluded since their mitochondrial rRNAs genes are fragmented and encoded by a large number of incompletely characterized transcripts (Gillespie et al. 1999).

For the statistical analysis we used Spearman rank correlations, as this test makes no assumption about the distribution of the data. For the comparative analysis we used a nonparametric method of Burt (1989), which allows us to test for correlated evolution among traits, disregarding their distribution. The method is based on the use of monophyletic groups (based on information from http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/) as inde-

pendent events in the evolution of the traits analyzed. Monophyletic groups were chosen such that within each group variation for all traits in question was present and the group contained at least three species. For the analysis of all genes combined, we weighted all three gene groups (ssu/lsu rRNAs, coxI-III, and cytb) equal. For this we standardized their means to zero and variances to one and calculated the sum of the three groups as a measure of combined gene length.

Results

Gene Length in Mitochondria and Bacteria

In order to test whether the reduction in gene length observed in a number of obligate intracellular symbionts is also seen in mitochondria, we focused on six mitochondrial genes (coxI–III, cytb, ssu and lsu rRNAs) essential for oxidative phosphorylation which in most species are refractory to nuclear transfer. Predicted lengths of gene products and genome sizes were collected for 278 eukaryotes and 11 α-proteobacteria (see supporting table).

Consistent with our hypothesis, we find that mitochondrial genes are on average shorter than the genes of α-proteobacteria, their presumed ancestors, supporting the idea of the evolution of gene length reduction in intracellular symbionts (Fig. 1; for all genes, the mean length of the mitochondrial genes is significantly shorter than the lengths of the α-proteobacteria: $p < 0.001$ for all genes). The gene homologs in *Rickettsia*, the closest known relative of mitochondria, are also longer than the mitochondrial genes, but this difference was only significant for coxII, coxIII, cytb, and lsu rRNA ($p < 0.05$), and not for coxI ($p = 0.30$) or ssu rRNA ($p = 0.09$). However, the power of this test is low, as we had data for only two *Rickettsia* species available. On average mitochondrial genes are 17.4% shorter than genes in *Rickettsia* (Fig. 1). A problem with these tests of the data in Fig. 1 is that we commit pseudoreplication, because all mitochondria are derived from one ancestor and therefore the species are not independent data points. This problem is unavoidable, as mitochondria are of monophyletic origin. However, we want to stress that all genes used in our analysis show a reduction in gene length. If genes are considered to be independent (all of them are older than the origin of mitochondria), one can use a sign test (giving a negative sign when the mitochondrial gene is shorter then the presumed ancestral gene) asking whether genes in general are shorter in mitochondria than in the α-proteobacteria and the *Rickettsia*. This is the case: sign test, $p < 0.05$ with $n = 6$ genes.

Mitochondrial gene length of all investigated genes varied strongly across species and among larger taxa. Despite this variation, the maximal observed lengths of the mitochondrial gene products are usually shorter than the mean of their bacterial homologues.
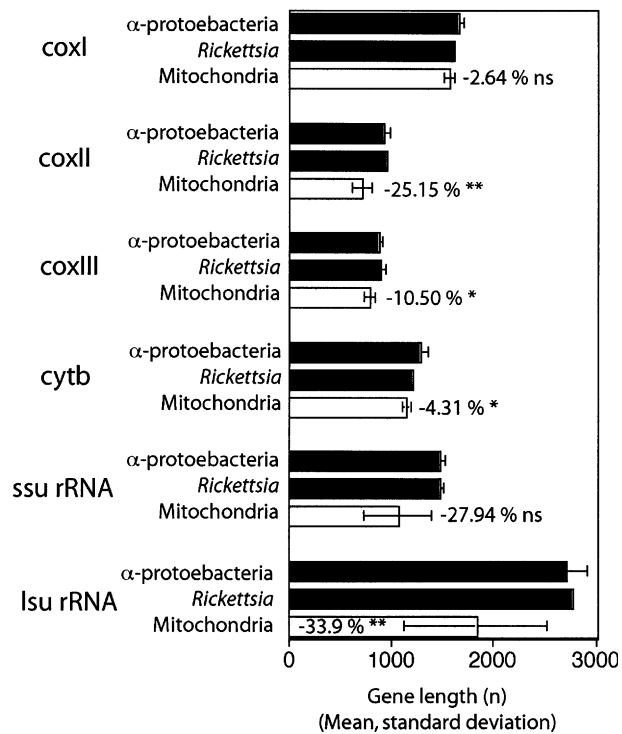

Fig. 1. Mean length (in nucleotides) of coxI, coxII, coxIII, cytb, ssu rRNA, and lsu rRNA from α-proteobacteria including two *Rickettsia* species (11 species), the two *Rickettsia* species only, and mitochondria (171 to 272 species). Error bars indicate one standard deviation below and above the mean. For the *Rickettsia* some of the standard deviations are zero. Mitochondrial protein encoding genes are on average 17.4% shorter than the their orthologs in the *Rickettsia*. The numbers after the white bars show the percentage reduction in mitochondrial gene length relative to the *Rickettsia* gene length. Significant reductions at *$p < 0.05$ and **$p < 0.01$. The comparison of the mitochondrial genes with the lengths of the α-proteobacteria is significant in all cases ($p < 0.01$). For species included in the analysis see supporting table in the electronic appendix.

The only exception to this rule are the mitochondrial ssu/lsu rRNA genes of certain algal, plant, and fungal species, which are longer than those of their bacterial counterparts (see supporting table).

Covariation of Gene Length with Genome Size

The shortest and the longest combined lengths of coxI–III, the lengths of cytb, and the combined lengths of ssu/lsu rRNAs across species vary by a factor of 1.59, 1.4, and 4.03, respectively. In contrast to this, the range of genome sizes spans approximately two orders of magnitude (5.9 to 570 kb).

Already in 1991 it was shown for a set of seven species that the length of mitochondrial rRNAs is correlated with the size of their corresponding organellar genomes (Andersson and Kurland 1991). Using a much larger data set with more species and more genes, we present here a more detailed analysis of this observation. Highly significant positive correlations are observed between mitochondrial genome size and the

Table 1. Spearman rank correlations between genome size and gene length

	coxI–III		cytb		ssu/lsu rRNAs		All genes	
All data (Fig. 2)	0.46***	(n = 253)	0.51***	(n = 272)	0.86***	(n = 163)	0.78***	(n = 146)
Species with mitochondrial genomes of 5.9 to 20 kb	0.20***	(n = 218)	0.37***	(n = 223)	0.72***	(n = 116)	0.61***	(n = 114)
Species with mitochondrial genomes of more than 20 kb	−0.20 (ns)	(n = 35)	0.30*	(n = 49)	0.61***	(n = 47)	0.37*	(n = 32)
Means of monophyletic taxa (Fig. 3)	0.80***	(n = 30)	0.71***	(n = 35)	0.94***	(n = 30)	0.92***	(n = 25)

Note. coxI–III, combined length of coxI–III; ssu/lsu rRNAs, combined length of ssu/lsu rRNAs; all genes, length of all genes combined after standardizing their means and variances; ns, nonsignificant. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. n is the sample size (number of species or number of taxa, respectively).

combined length of coxI–III, the length of cytb, and the combined length of ssu/lsu rRNAs (Table 1, Fig. 2). These positive correlations become even stronger when we use only the means of monophyletic taxa (Table 1, Fig. 3). Figure 2 and 3 reveal that in smaller genomes (6–20 kb) the slope of the gene length/genome size relationship is much steeper than for larger genomes (> 20 kb). A separate analysis of these two genome size class groups shows that the covariation between genome size and gene length remains largely unchanged (Table 1). Only the correlation between the combined length of coxI–III and genome sizes larger than 20 kb is nonsignificant (Table 1).

Figure 2 shows that the length of coxI–III and cytb is, in most species, smaller than the average gene product length of the α-proteobacteria. However, the length of the combined rRNAs of mitochondria is often longer than that of the corresponding gene products in α-proteobacteria. With few exceptions the species with such long rRNAs belong in the taxa Fungi, Stramenopiles, Chlorophyta, Rhodophyta, and Streptophyta.

To test if the observed gene length/genome size relationship is only found across the entire data set, or whether it indicates coevolution within different taxa, we tested for correlations between the combined gene length within a set of monophyletic taxa, using the comparative method for correlated traits of Burt (1989). These taxa were chosen because they had a reasonable amount of interspecies variation within the taxa and they contained more than two complete data sets, data for all gene available (three is the smallest number of data points one can use to test for a correlation among traits). For 12 of the 14 taxa we obtained positive covariances ($p < 0.05$ using a sign test; taxa are shown in Fig. 3 with $n > 2$), indicating that gene length coevolved with genome size independently within the majority of taxa. The two taxa with negative covariances were the Stramenopiles and the Nematoda.

AT Content Does Not Explain Gene Size/Genome Length Covariance

A potential explanation for the observed shrinkage of coxI–III and cytb could be related to the fact that the ATG start codon and all three translational stop codons (TAA/TAG/TGA) are rich in AT. Organellar genomes, possibly due to the nature of the DNA damage to which they are exposed, have a tendency to become AT rich (Howe et al. 2000). This is expected to increase the overall density of start and stop codons and may therefore result in shorter average gene lengths (Oliver and Marin 1996). Should the observed shortening of coxI–III and cytb mainly be caused by a high AT content, we would predict that the AT content of mitochondrial genomes should be negatively correlated with gene lengths. A weak but significant correlation is observed for cytb ($r = -0.14$, $p = 0.02$; $n = 255$), but the prediction is not supported for coxI–III ($r = -0.06$, $P = 0.34$; $n = 240$). There is also no significant correlation between mitochondrial genome size and AT content ($r = 0.03$, $p = 0.6$; $n = 258$). Thus, it seems that in mitochondria AT content is not a major factor in the evolution of gene length and genome size. Further, speaking against the AT hypothesis is our finding of a strong correlation between the length of the ssu/lsu rRNAs and genome size (Table 1, Fig. 2). As rRNAs are not translated, a higher density of start and stop codons cannot explain the observed gene length/genome size relationship.

Discussion

While it is clear that reduction of mitochondrial coding capacity during evolution is mainly caused by loss and/or transfer of genes to the nucleus, our results show that shortening of the remaining mitochondrial genes contributes to the process. In agreement with studies from obligate intracellular symbionts, we found that the lengths of mitochondrial genes are usually shorter then those of their bacterial homologs. What could cause the shortening of mitochondrial coding regions? Shorter mitochondrial genes may have a selective advantaged. A shorter gene might give its carrier a replication advantage during intraorganellar competition. Further, shorter genes may accumulate slightly deleterious mutations slower (Müller's ratchet). The hypothesis that shorter genes and smaller genomes are selectively favored is supported
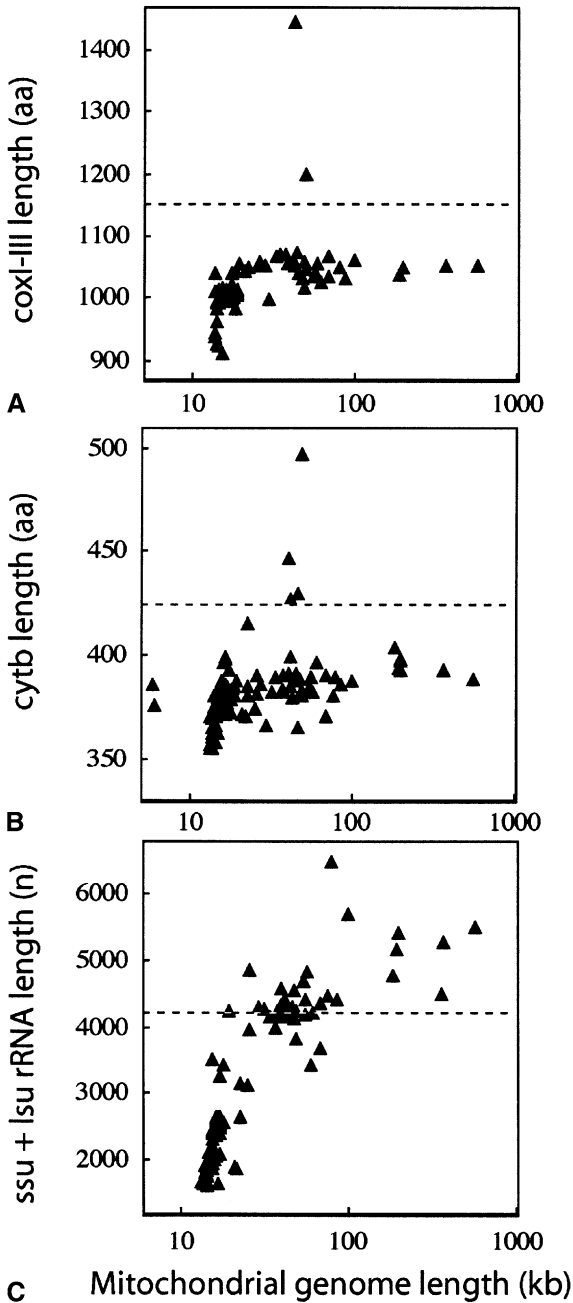
Fig. 2. Covariation of mitochondrial genome size with gene lengths. Gene lengths of coxI–III (**a**), cytb (**b**), and ssu/lsu rRNA (**c**) of different species are plotted against the corresponding mitochondrial genome size. The dashed horizontal line marks the average gene product length of the α-proteobacteria. The species having longer rRNAs than the α-proteobacteria belong mainly to the taxa Fungi, Stramenopiles, Chlorophyta, Rhodophyta, and Streptophyta. For species included in the analysis see supporting table in the electronic appendix. Sample size: coxI–III, $n = 254$; cytb, $n = 271$; ssu + lsu rRNA, $n = 162$.

Fig. 3. Mitochondrial genome sizes of all genes' lengths combined (coxI–III/cytb/ssu and lsu rRNA) after standardizing their means and variances for monophyletic taxa are plotted against the means of the corresponding mitochondrial genome sizes. Names and number of species of the monophyletic taxa are indicated. For species included in the analysis see supporting table in the electronic appendix.

smaller the genome size. Therefore, the likelihood of fixation of such deletions is higher in smaller than in larger genomes. Any of the potential adaptive advantages (escape form Müllers's ratchet, replication advantage) of shorter genes is expected to be relatively bigger in the context of small mitochondrial genomes, and becomes progressively more neutral in larger genomes. As a consequence, the covariance between genome size and gene length is expected to be strongest in smallest genomes, which is in agreement with the data (Table 1, Fig. 1). This process is further facilitated by a mutation asymmetry. In most bacterial genomes deletions are more common than insertions (Mira et al. 2001), which could lead to a net reduction in gene length.

As mentioned in the Introduction, gene shortening in obligate intracellular symbionts has been documented in a number of cases (Cavalier-Smith 2002; Charles et al. 1999; Vivares et al. 2002; Wernegreen 2002). The taxonomic distribution of these symbionts indicates that gene shortening evolved several times independently. Our results show that gene shortening has also occurred in mitochondria. Mitochondrial genes are on average shorter than those of their bacterial homologs (Fig. 1). However, the mechanisms by which gene lengths evolve may not be the same. The genes of the aphid endosymbiontic bacteria *Buchnera* were on average 0.8% shorter than their corresponding homologs in *E. coli* (Charles et al. 1999). It has been suggested that this might be caused by the high AT content of the genome and by a

by the shape of the gene length–genome size relationship, which is strongly asymptotic for all gene investigated here (Figs. 1 and 2). If a small deletion, which is neutral with respect to gene function, gives its carrier an advantage due to the smaller size, then one can expect that this advantage is relatively larger the
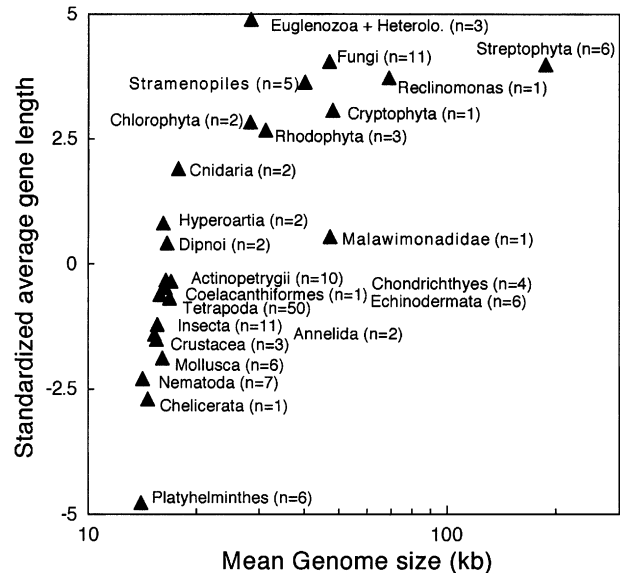
mutational bias for deletions. Dramatic shortening of coding regions has also been observed in the obligate intracellular parasite *Encephalitozoon cuniculi*: 85% of its proteins are on average 14.6% smaller than their yeast homologs (Vivares et al. 2002). Likewise, unusual short protein coding genes are found in the vestigial nuclear genomes of the nucleomorphs of cryptomonads (Cavalier-Smith 2002). Nucleomorphs are the nuclei of unusual plastids, which arose by endosymbiosis of an eukaryotic alga. In all three examples short coding sequence length is linked to a general trend for genome compactness and AT richness and the latter might in part explain the length reduction of the coding sequences. This seems not to be the case for mitochondrial genes, as we found no convincing correlation between AT richness and gene length and further found reduced lengths of genes, which are not translated.

A gene length/genome size relationship is not observed in free-living bacteria (data not shown). Why do mitochondrial genomes behave differently than genomes of free-living bacteria? First, as seen in our data, the covariance between gene length and genome size becomes progressively clouded when genome size increases. This may be because selection for reduced gene length becomes progressively weaker when genomes become bigger. Even bacteria with the smallest genomes have genomes much larger than large mitochondrial ones. For example, the *Rickettsia* genomes are larger than 1 Mb. It may therefore be that in bacteria small deletions within genes do not provide an equally strong advantage associated with the reduced genome size, as in mitochondria, and therefore rarely go to fixation.

Second, genes of bacteria may be exposed to different and variable selective pressures depending on the various ecological niches they live in, while mitochondria of different host species experience a rather similar environment, with possibly rather similar forces of selection. The diversity of selective forces acting on bacteria may render a cross-species comparative study very difficult, because factors other than selection for shorter genomes may dominate their evolution.

As mentioned above, certain algae, plants, and fungal species have longer mitochondrial ssu/lsu rRNAs genes than their bacterial counterparts. They also contain mitochondrial genomes which are bigger than those of any other eukaryotes but which are still smaller than the size of bacterial genomes. In particular, the Streptophyta (plants) have mitochondrial genomes which are often unusally large. In our data set all mitochondrial genomes larger than 120 kb are plants (Fig. 3). Their large genome sizes are, however, not correlated with information content. It has therefore been suggested that the large size of the plant mitochondrial DNA is a secondary (derived)

trait caused by a relaxation of the pressure for genome compactness (Marienfeld et al. 1999). If true, this offers an explanation for their comparatively long ssu/lsu rRNAs genes. Gene length evolution in plants may be the result of an ancestral selection process leading to strongly reduced gene lengths in ancestral plants, followed by a phase of relaxation during which genomes could expand. The genomes expanded in part by recombination events and in part by acquiring external sequences (Palmer et al. 2000). Our data seem to indicate that the length of the protein coding regions (coxI–III and cytb) did not expand, while the rRNA genes expanded and became even longer than their bacterial counterparts. Expansion of rRNAs length is not unusual and has been described for cytosolic rRNAs in some taxons (Busse and Preisfeld 2002). Variable regions of the rRNA are generally located at the surface, whereas the conserved core is situated in the center of the ribosome. Expansion of the exposed variable regions does not appear to be restriced since it is not expected to interfere with ribosomal function (Wuyts et al. 2001). According to this hypothesis, expansion of rRNA is much more likely than expansion of protein coding regions. In other words, losing small pieces of protein coding regions was by and large an irreversible event.

In summary, there are two types of evolutionary forces which reduce the coding capacity of mitochondrial genomes: transfer of coding sequences to the nucleus and loss of coding sequences from the cell. In this work we show that the latter may be responsible not only for deletion of entire genes but also for reduction of the coding sequence length of resident mitochondrial genes.

References

Allen JF (2003) The function of genomes in bioenergetic organelles. Phil Trans R Soc Lond B Biol Sci 358:19–37

Allen JF, Raven JA (1996) Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. J Mol Evol 42:482–492

Andersson GE, Kurland CG (1991) An extreme codon preference strategy: Codon reassignment. Mol Biol Evol 8:530–544

Andersson SG, Kurland CG (1998) Reductive evolution of resident genomes. Trends Microbiol 6:263–268

Andersson SGE, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UCM, Podowski RM, Näslund AK Eriksson A-S, Winkler HH, Kurland CG (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396:133–140

Berg OG, Kurland CG (2000) Why mitochondrial genes are most often found in nuclei. Mol Biol Evol 17:951–961

Blanchard JL, Lynch M (2000) Organellar genes: Why do they end up in the nucleus? Trends Genet 16:315–320

Burt A (1989) Comparative methods using phylogenetically independent contrasts. Oxford Surv Evol Biol 6:33–53

Busse I, Preisfeld A (2002) Unusually expanded SSU ribosomal DNA of primary osmotrophic euglenids: Molecular evolution and phylogenetic inference. J Mol Evol 55:757–767

Cavalier-Smith T (2002) Nucleomorphs: Enslaved algal nuclei. Curr Opin Microbiol 5:612–619

Charles H, Mouchiroud D, Lobry J, Goncalves I, Rahbe Y (1999) Gene size reduction in the bacterial aphid endosymbiont, *Buchnera*. Mol Biol Evol pp 1820–1822

Doolittle WF (1998) You are what you eat: A gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14:307–311

Gillespie DE, Salazar NA, Rehkopf DH, Feagin JE (1999) The fragmented mitochondrial ribosomal RNAs of *Plasmodium falciparum* have short A tails. Nucleic Acids Res 27:2416–2422

Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Brossard N, Delage E, Littlejohn TG, Plante I, Rioux P, Saint-Louis D, Zhu Y, Burger G (1998) Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res 26:865–878

Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481

Howe CJ, Barbrook AC, Lockhart PJ (2000) Organelle genes—Do they jump or are they pushed? Trends Genet 16:65–66

Kurland CG, Andersson SG (2000) Origin and evolution of the mitochondrial proteome. Microbiol Mol Biol Rev 64:786–820

Lang BF, Gray MW, Burger G (1999) Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33:351–397

Lynch M (1996) Mutation accumulation in transfer RNAs: Molecular evidence for Muller's ratchet in mitochondrial genomes. Mol Biol Evol 13:209–220

Marienfeld J, Unseld M, Brennicke A (1999) The mitochondrial genome of *Arabidopsis* is composed of both native and immigrant information. Trends Plant Sci 12:495–502

Mira AH, Ochman H, Moran NA (2001) Deletional bias and the evolution of bacterial genomes. Trends Genet 17:589–596

Oliver JL, Marin A (1996) A relationship between GC content and coding-sequence length. J Mol Evol 43:216–223

Palmer JD, Adams KL, Cho Y, Parkinson CL, Qiu YL, Song K (2000) Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. Proc Natl Acad Sci USA 97:6960–6966

Secq MPO-L, Fontaine JM, Rousvoal S, Kloareg B, Goer SL-D (2001) The complete sequence of a brown algal mitochondrial genome, the ectocarpale *Pylaiella littoralis* (L.) Kjellm. J Mol Evol 53:80–88

Selosse M, Albert B, Godelle B (2001) Reducing the genome size of organelles favours gene transfer to the nucleus. Trends Ecol Evol 16:135–141

Shimko N, Liu L, Lang BF, Burger G (2001) GOBASE: The organelle genome database. Nucleic Acids Res 29:128–132

Thorness PE, Fox TD (1990) Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae*. Nature 346:376–379

Vivares CP, Gouy M, Thomarat F, Metenier G (2002) Functional and evolutionary analysis of a eukaryotic parasitic genome. Curr Opin Microbiol 5:499–505

Wernegreen JJ (2002) Genome evolution in bacterial endosymbionts of insects. Nat Rev Genet 3:850–861

Wuyts J, Van de Peer Y, De Wacher R (2001) Distribution of substitution rates and location of insertion sites in the tertiary structure of ribosomal RNA. Nucleic Acids Res 29:5017–5028