ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Molecular Phylogenetics and Evolution 29 (2003) 380–395

MOLECULAR PHYLOGENETICS AND EVOLUTION

www.elsevier.com/locate/ympev

Evolution of mitochondrial gene content: gene loss and transfer to the nucleus

Keith L. Adams[a,*] and Jeffrey D. Palmer[b]

[a] *Department of Botany, Iowa State University, 353 Bessey Hall, Ames, IA 50011, USA*
[b] *Department of Biology, Indiana University, Bloomington, IN 47405, USA*

Received 24 December 2002; revised 4 April 2003

Abstract

Mitochondrial gene content is highly variable across extant eukaryotes. The number of mitochondrial protein genes varies from 3 to 67, while tRNA gene content varies from 0 to 27. Moreover, these numbers exclude the many diverse lineages of non-respiring eukaryotes that lack a mitochondrial genome yet still contain a mitochondrion, albeit one often highly derived in ultrastructure and metabolic function, such as the hydrogenosome. Diversity in tRNA gene content primarily reflects differential usage of imported tRNAs of nuclear origin. In the case of protein genes, most of this diversity reflects differential degrees of functional gene transfer to the nucleus, with more minor contributions resulting from gene loss from the cell as a consequence of either substitution via a functional nuclear homolog or the cell's dispensation of the function of the gene product. The tempo and pattern of mitochondrial gene loss is highly episodic, both across the broad sweep of eukaryotes and within such well-studied groups as angiosperms. All animals, some plants, and certain other groups of eukaryotes are mired in profound stases in mitochondrial gene content, whereas other lineages have experienced relatively frequent gene loss. Loss and transfer to the nucleus of ribosomal protein and succinate dehydrogenase genes has been especially frequent, sporadic, and episodic during angiosperm evolution. Potential mechanisms for activation of transferred genes have been inferred, and intermediate stages in the process have been identified by comparative studies. Several hypotheses have been proposed for why mitochondrial genes are transferred to the nucleus, why mitochondria retain genomes, and why functional gene transfer is almost exclusively unidirectional.
© 2003 Elsevier Science (USA). All rights reserved.

1. Introduction

By now, with many mitochondrial genomes fully sequenced, and with complete sequences also available for such crucial bacteria as *Rickettsia* (Andersson et al., 1998) and other α-proteobacteria, it is undisputable that the mitochondrion and its genome are of endosymbiotic origin, i.e., that they are the derived remnants of a once-free living bacterium, almost certainly an α-proteobacterium (Gray, 1999; Gray et al., 1999, 2001; Lang et al., 1999). What is also clear is that there has been an extraordinary degree of rewiring of the genetic circuitry of the eukaryotic cell since the mitochondrion's inception via endosymbiosis. All examined mitochondrial genomes contain vastly fewer genes (by 1–3 orders of magnitude) than the genomes of free-living α-proteobacteria, yet mitochondria are nonetheless metabolically and biochemically complex organelles requiring at least several hundred proteins to function properly. In other words, among extant eukaryotes, the great majority (up to 99%) of mitochondrial proteins are the products of nuclear genes. During the course of mitochondrial evolution, many mitochondrial genes have been functionally transferred to the nucleus, whereas others have been replaced by preexisting nuclear genes of similar function.

In this review, we first discuss the degree to which mitochondrial gene content varies among extant eukaryotes. This variation is extensive, and for both protein and tRNA genes it reflects considerable gene- and lineage-specific variation in rates of gene loss. We consider reasons why the cell can tolerate loss of a previously functional gene from the mitochondrial genome,

* Corresponding author. Fax: 1-515-294-1337.
E-mail address: KLAdams@iastate.edu (K.L. Adams).

1055-7903/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1055-7903(03)00194-5

the most common of which is transfer of the gene to the nucleus. We then review several aspects of mitochondrial gene transfer to the nucleus, focusing primarily on plants and green algae, those groups that are known to be most active in continuing gene transfer in the functional sense. Finally, we provide a more speculative treatment of some of the factors that either promote or retard gene transfer.

2. Evolution of mitochondrial gene content

2.1. Ribosomal RNA genes

Mitochondrial genes can be usefully divided into three classes—rRNA, tRNA, and protein coding genes—with respect to the frequency and underlying biology of their loss from the mitochondrial genome. The two major rRNA genes, encoding the small-subunit and large-subunit rRNAs, are present in all characterized mitochondrial genomes. Although universally retained, these genes do exhibit a remarkable range in size, especially compared to those of their eubacterial and chloroplast counterparts. In several cases, mitochondrial rRNA genes have experienced modest to extreme fragmentation into little pieces that are scattered throughout the mitochondrial genome, transcribed separately, and then reassembled at the RNA level via secondary-pairing interactions to form a functional rRNA molecule (for review, see Gray and Schnare, 1996). In the most severe cases of fragmentation, which also correspond to especially reduced and degenerate rRNAs overall, it is difficult to rule out the possibility that a small piece of rRNA sequence that is not accounted for among the rRNA fragments evident in the mitochondrial genome has been functionally transferred to the nucleus, as opposed to having simply been lost entirely from the cell. With the recent sequencing of nuclear genomes from *Plasmodium falciparum* (Gardner et al., 2002) and *Chlamydomonas reinhardtii* (soon to be released and published), two organisms with especially fragmented and degenerate mitochondrial rRNA genes, this possibility can now be rigorously tested.

In distinct contrast to the universality of SSU and LSU rRNA genes in mitochondrial genome, the gene for mitochondrial 5S rRNA has a very patchy distribution, being absent from most examined mitochondrial DNAs (mtDNAs) and present only in land plants, a subset of green algae, red algae, and brown algae (Oudot-Le Secq et al., 2001), and in the protozoan *Reclinomonas* (reviewed in Bullerwell et al., 2003a, 2003b; Gray et al., 1998). This distribution implies many separate losses of the 5S rRNA gene across the broad course of mitochondrial evolution. Alternatively, bioinformatic and/or experimental failure to recognize a highly divergent 5S rRNA could account for its apparent absence in some mitochondrial genomes, as evidenced by the recent identification of a highly divergent 5S sequence in the mitochondrial genome of *Acanthamoeba castellanii* (Bullerwell et al., 2003a, 2003b). Other possibilities to account for the missing 5S gene in many mitochondrial genomes include replacement of 5S rRNA by its cytosolic counterpart, as shown in animals (Entelis et al., 2001; Magalhaes et al., 1998) or transfer of the gene to the nucleus in some lineages. It is also possible that the 5S gene might be dispensable in the context of the somewhat to highly aberrant and degenerate ribosomes of many mitochondrial genetic systems.

2.2. Transfer RNA genes

The number of distinct tRNA genes present in mtDNA varies enormously across eukaryotes, from none in apicomplexans and the kinetoplastid *Trypanosoma brucei*, to a reasonably self-sufficient set of 22–27 tRNAs in many mitochondrial systems (Gray et al., 1998; Lang et al., 1999). To our knowledge, there is no evidence that any of these missing tRNA genes of mitochondrial origin have been functionally transferred to the nucleus. Instead, the predominant explanation for their loss is functional substitution/replacement, via nuclear tRNAs of strictly eukaryotic/nuclear/cytoplasmic origin, which occasionally do double duty, serving the needs of both cytoplasmic and mitochondrial protein synthesis. Mitochondrial import of nuclear-encoded cytoplasmic tRNAs has evolved, to a greater or lesser extent, many times over the course of mitochondrial evolution (Gray et al., 1998).

Two other, "special" cases of mitochondrial tRNA gene loss have also been described. Partial RNA editing of the anticodon of a mitochondrially encoded tRNA creates two functionally distinct tRNA species (for Gly-GCC and Asp-GUC) in opossum mitochondria (Borner et al., 1996), thereby allowing loss of a tRNA gene. In angiosperms, several missing mitochondrial tRNA genes have been functionally replaced by cognate tRNA genes of chloroplast origin that have been directly transferred *into* the mitochondrial genome (e.g., Joyce and Gray, 1989). Angiosperm mitochondria also import and use a variable number of cytoplasmic tRNAs (e.g., Glover et al., 2001; Kumar et al., 1996), and thus their mitochondrial translational apparatus is the most catholic of all, using significant numbers of tRNAs of mitochondrial, chloroplast, and nuclear/cytoplasmic origin (Small et al., 1999).

2.3. Protein genes—the big picture

As with tRNA genes, protein gene content of the mitochondrial genome is highly variable across eukaryotes. The number of mitochondrial protein genes ranges from only 3 in the malarial parasite *P. falciparum*

and all other members of phylum Apicomplexa (Feagin, 1994) to 67 in the freshwater protist *Reclinomonas americana* (Lang et al., 1997). There is little correlation within an organism between its number of mitochondrial tRNA genes and protein genes (Lang et al., 1999). The best correlations in terms of extreme reduction of both sets of genes are found in the apicomplexans, with only 3 protein genes and no tRNA genes, and the Chlamydomonad green algae, with 7 protein genes and 3 tRNA genes. However, there are plenty of counter-examples: Among fungi, both fission and fusion yeast have only half (7) the number of protein genes (13–14) typically found in fungi, yet have the usual high number (24–25) of tRNA genes. Conversely, two chytridiomycetes, *Spizellomyces* and *Hyaloraphidium*, have the usual number of fungal protein genes, but are down to only 7–8 tRNA genes (Forget et al., 2002; Lang et al., 1999). While virtually all animal mitochondrial genomes contain 13 protein and 22 tRNA genes, cnidarian mtDNAs, with the same 13 protein genes, are down to just 1 or 2 tRNA genes (reviewed in Boore, 1999).

A loose functional correlation between reduction of these two classes of mitochondrial genes can be predicted on the basis of translational demand, i.e., a mitochondrion that makes only a very few proteins can presumably get by with lower levels of mitochondrial tRNAs, and such tRNAs could be more readily supplied by "leaky" import from the cytoplasm than higher abundance tRNAs required for a more robust mitochondrial genetic system. This hypothesis might explain the correlations seen for apicomplexans and chlamydomonads. On the other hand, since tRNA loss and protein gene loss are principally the consequence of two fundamentally different mechanisms—cytoplasmic tRNA import and functional transfer of mitochondrial protein genes to the nucleus, respectively—it is not surprising to find that they are often evolutionarily uncoupled.

All mtDNAs are highly reduced (by 10–1000-fold) in protein gene content relative to the genomes of their α-proteobacterial progenitors. [*Rickettsia* has only 834 protein genes (Andersson et al., 1998), but as an obligate intracellular parasite it may be a poor exemplar (Gray et al., 2001) for the mitochondrial progenitor compared to a free-living α-proteobacterium such as *Caulobacter* (Nierman et al., 2001), with over 3600 protein genes.] Furthermore, all mtDNAs are reduced to the same sets of protein genes, involved almost exclusively in respiration and protein synthesis. Most remarkably, the 67 or so protein genes present in the most gene-rich mitochondrial genome (Lang et al., 1997) include *all* of the protein genes found variously among *all* other mitochondrial genomes characterized to date. These facts have naturally led a number of observers (Gray, 1999; Gray et al., 1999, 2001; Lang et al., 1999), including one of us (Palmer, 1997a), to conclude that most endosymbiotic/mitochondrial genes were lost or transferred to the nucleus soon after endosymbiosis, prior to the perhaps "big-bang" diversification and emergence of the major groups of eukaryotes. However, it is increasingly apparent that convergence in mitochondrial gene loss—the parallel loss of many of the same genes over and over again in different mitochondrial lineages (as reviewed below, and as analyzed in a novel context in a recent paper by Stiller et al., 2003)—has been remarkably pervasive. So pervasive, that (1) gene content similarities should no longer be considered one of the more reliable characters in support of a monophyletic origin of mitochondria, and (2) one should be open to the possibility that a larger set of mitochondrial genes persisted to the eukaryotic "big-bang," these having gone extinct in parallel in two or more descendant mitochondrial lineages. With respect to mitochondrial monophyly, we hasten to add that even though mitochondrial gene content may no longer be a particularly good character, there are still other lines of evidence (e.g., gene phylogenies and derived operon organizations) to support a monophyletic origin of mitochondria (Gray, 1999; Gray et al., 1999; Lang et al., 1999), and no good evidence to the contrary.

Looking at mitochondrial protein genes across eukaryotes, intriguing patterns emerge on both a gene-by-gene and lineage-by-lineage basis. We will focus entirely on the two major sets of mitochondrial protein genes, those involved in respiration (encoding up to 24 subunits of four major complexes of the electron transfer chain, plus the coupling factor or ATP synthase) and in protein synthesis (all ribosomal proteins, up to 27 of them). We will ignore the *tufA* gene (for elongation factor Tu) and the dozen or so genes involved in transcription and protein import and maturation (Gray et al., 1998; Lang et al., 1999) because these are so uncommon in mitochondrial genomes (many are present only in *Reclinomonas* and *Jakoba*) as to be of limited utility for inferring evolutionary patterns across eukaryotes.

On the whole, as well-illustrated in several of the tables in Gray et al. (1998) and Lang et al. (1999), ribosomal protein genes have been lost more often than respiratory genes, but there is also considerable variation within each functional class of genes as to how frequently a given gene is lost. The above-cited tables, and the text in these papers, make it clear that a "hierarchy" of gene loss exists, that certain genes are lost more frequently, more readily, than other genes. The hierarchy of gene losses is steeper and the constraints appear stronger, for respiratory genes than for ribosomal protein genes. Of the 27 ribosomal protein genes present in the *Reclinomonas* mitochondrial genome, 16 are broadly and fairly commonly found in mtDNAs of diverse eukaryotes; interestingly, 11 of 12 *Reclinomonas* small subunit ribosomal proteins fall into this commonly present class, as compared to only 5 of 15 large subunit proteins. At the same time however, four different major lineages of eukaryotes—animals, fungi

(certain fungi that is, a few retain a single ribosomal protein gene), apicomplexans, and chlamydomonad green algae—have each independently lost *all* ribosomal protein genes from their mitochondrial genomes. In sharp contrast, no respiring eukaryote examined to date has lost all mtDNA-encoded respiratory genes. All mitochondrial genomes contain at least three respiratory genes, two of which, cytochrome *b* (*cob*) and cytochrome oxidase subunit 1 (*cox1*), are "universal," i.e., present in all examined mitochondrial genomes. Moreover, this statement ignores the important fact that three other respiratory genes, *nad1*, *nad4*, and *nad5*, are what we hereby dub "de facto universal" mitochondrial genes. This is because the only two lineages lacking them (yeasts and apicomplexans) lack them not because of transfer to the nucleus, but because the entire large (>20 subunits) biochemical complex (the NADH dehydrogenase complex, or complex I of the electron transfer chain) in which they function has simply been lost from these organisms. Thus, a total of 5 respiratory genes are universally present in mtDNA except when functionally unnecessary, while most other respiratory genes are lost only rarely. As discussed in the following section, compared to the overall eukaryotic pattern, plants exhibit an even sharper distinction between relatively frequent loss of ribosomal genes and infrequent loss of respiratory genes.

Looking at mitochondrial protein gene content across eukaryotic lineages, we see a bit of a paradox. On the one hand, there is over 20-fold variation in protein gene content among examined mitochondrial genomes (from 3 to 67 genes), and many lineages have *independently* experienced not only massive if not wholesale loss of ribosomal proteins but also major loss of respiratory genes. On the other hand, relatively few differences in gene content are evident in comparing diverse members from *within* several major phyla or kingdoms (again, see tables in Gray et al., 2001, and Lang et al., 1999). Few differences are present within the following multiply sampled groups: animals, fungi, apicomplexans, ciliates, and heterokonts (also known as stramenopiles). An important caveat is that of these 5 groups, only animals and plants have been well sampled for mitochondrial gene content. Sampling is not an important issue for apicomplexans because the few that have been examined represent the diversity of the group well and all have but 3 protein genes (including, of course, cob and cox1), so there is not much opportunity for further loss. But it is entirely possible that the picture of little gene content variation noted among the 5 diverse heterokonts or 2 diverse ciliates currently examined will change once these groups are better sampled. A dramatic example of this is seen among the slime molds: the compilations in Gray et al. (1998) and Lang et al. (1999) show that the two very diverse slime molds then sequenced, *Acanthamoeba* and *Dictyostelium*, contain the same 18 respiratory genes and share 15 of 16 ribosomal protein genes. However, the subsequent sequencing (Takano et al., 2001) of a third member of the group, *Physarum*, which is actually more closely related to *Dictyostelium* than either is to *Acanthamoeba*, changes one's view of mitochondrial gene constancy within slime molds. *Physarum* mtDNA contains only 11 respiratory genes and 1 ribosomal protein gene, although the requirement for extensive and substitutional editing of mitochondrial transcripts in *Physarum* could hinder identification of some genes and thus the mitochondrial genome may contain additional yet unidentified genes.

Animals and fungi, however, are very well sampled and continue to show a remarkable constancy of mitochondrial protein gene content. About 20, highly diverse fungal mitochondrial genomes have been sequenced (e.g., Bullerwell et al., 2003a, 2003b; Forget et al., 2002; and references therein), and excepting the above noted biochemical loss of the entire NADH dehydrogenase complex in yeasts, only two of the 15 protein genes inferred to be present in the ancestral fungal mitochondrial genome have been found missing: *atp9* has been lost (and transferred to the nucleus) once among euascomycetes, and *rps3* has been lost several times among a diversity of fungal lineages. Animals are even more striking: over 100 diverse animal mitochondrial genomes have been sequenced, and the great majority contain the same 13 genes (all respiratory). The only losses detected are of *atp8*, in nematodes and in mollusks, and whether this is due to gene transfer or merely reflects a dispensable protein is unclear because a clear homolog of *atp8* has not been reported from the sequenced nuclear genome of the nematode *Caenorhabditis elegans*.

The paradox noted above, of great variation in mitochondrial protein gene content among most major groups of eukaryotes, but only limited variation within large and ancient groups, suggests a very episodic, punctuated pattern of mitochondrial gene loss over the broad sweep of eukaryotic evolution. So too does the pattern noted above within slime molds, in which *Physarum* has lost two-thirds of the mitochondrial genes that have been retained in two other deep lineages of slime molds, while a similar pattern has already been emphasized in green algae (Turmel et al., 1999, 2002). All of this implies periods and lineages, some relatively ancient, some more recent, of relatively intense loss of mitochondrial genes, and other periods and lineages of relative stasis, animals being most remarkably static in gene content. As the next section will show, plants are exceptionally well sampled and dramatically illustrate both stasis and rapid change in protein gene content.

2.4. Protein genes in plants—both stasis and lots of action

In this section, we focus on plant mitochondrial genomes. This is because, like animals, plants have been

extensively sampled for mitochondrial protein gene content, and also, unlike animals, they exhibit extensive and provocative variation in gene content. Insofar as examined, mitochondrial gene content has remained relatively stable over the broad course of land plant evolution. The only sequenced mitochondrial genome from a non-flowering plant—the liverwort *Marchantia polymorpha* (Oda et al., 1992)—contains only two protein genes not found in any angiosperm (the ribosomal proteins *rpl6* and *rps8*). *Rpl6* was transferred to the nucleus (*Arabidopsis* Genome Initiative, 2000), and *rps8* was replaced by a duplicate copy of its cytosolic counterpart, *rps15A* (Adams et al., 2002a). There are no characterized protein genes in angiosperms that are not present in *Marchantia*, although *nad7* has been functionally transferred to the nucleus, with only a *nad7* pseudogene remaining in the *Marchantia* mitochondrial genome (Kobayashi et al., 1997). Thus, the repertoire of protein genes was almost identical in the ancestral land plant and in the ancestral angiosperm, indicating near stasis in mitochondrial gene content across the ~450 million years of land plant evolution represented by these lineages. This stasis has been broken in some but not all lineages of angiosperms, as we will now discuss; the extent to which it persists in the many unsampled lineages of bryophytes and vascular plants remains to be determined.

Mitochondrial gene content in flowering plants has been studied extensively. The mitochondrial genome has been completely sequenced from *Arabidopsis thaliana* (Unseld et al., 1997), sugar beet (Kubo et al., 2000b), rice (Notsu et al., 2002), and maize (S. Clifton et al., in preparation) revealing differential loss of several protein genes in each plant. As summarized in Table 1, a comprehensive survey by Southern blot hybridization of mitochondrial protein gene content in 280 diverse angiosperms has revealed remarkably frequent loss of 16 mitochondrial genes—all 14 ribosomal protein and both succinate dehydrogenase (sdh) genes (Adams et al., 2002b). In stark contrast, only two losses were inferred among the other 24 protein genes, most of which function in respiration. As discussed in Adams et al. (2002b), the bimodal pattern of gene loss summarized in Table 1 for angiosperms largely holds up across the broad sweep of eukaryotic evolution, although the distinction between frequently and infrequently lost genes seems sharper for angiosperms. There are also a few exceptional genes, such as *atp1* and *nad7*, and to a lesser extent *nad4L* and *nad9*, that are invariably present in angiosperm mtDNAs but relatively frequently lost from other mitochondrial genomes.

Mitochondrial gene content is tremendously variable across angiosperms. Some flowering plants, in particular most of the basal angiosperms, have not lost any of the 40 known mitochondrial protein genes that were present in the common ancestor of angiosperms. In strong contrast, several flowering plants, such as *Allium* (onion), *Lachnocaulon* (see Fig. 1), and *Erodium* have lost all or almost all of their 16 ribosomal protein and succinate dehydrogenase (sdh) genes (Adams et al., 2002b). Such plants are converging on a significantly reduced mitochondrial gene content, one more like that of an animal than a typical plant. Even more remarkable is the punctuated nature of the gene losses in some plants (such as *Lachnocaulon*, Fig. 1), with most losses occurring along a single branch, suggestive of many losses taking place during a very short evolutionary time span.

Table 1
Estimated numbers of mitochondrial gene losses among 280 angiosperms (from Adams et al., 2002b)

Gene	# of losses	Gene	# of losses
rpl2	41	atp1	0
rpl5	19	atp6	0
rpl16	15	atp8	1
rps1	33	atp9	0
rps2	8*	ccb2	0
rps3	7	ccb3	0
rps4	7	ccb6c	0
rps7	42	ccb6n	0
rps10	26	cob	0
rps11	14*	cox1	0
rps12	6	cox2	1
rps13	30	cox3	0
rps14	27	nad1	0
rps19	39	nad2	0
		nad3	0
sdh3	40	nad4	0
sdh4	19	nad4L	0
		nad5	0
matR	0	nad6	0
mtt2	0	nad7	0
orf25	0	nad9	0

*Includes 1 deep loss (see text). Gene names: rpl, ribosomal protein large subunit; rps, ribosomal protein small subunit; sdh, succinate dehydrogenase; matR, maturase; mtt, protein transport subunit (name comes from mtt operon in bacteria); orf, conserved open reading frame of unknown function; atp, atp synthase; ccb, cytochrome *c* biogenesis; cob, cytochrome *b*, cox, cytochrome oxidase; nad, NADH dehydrogenase.

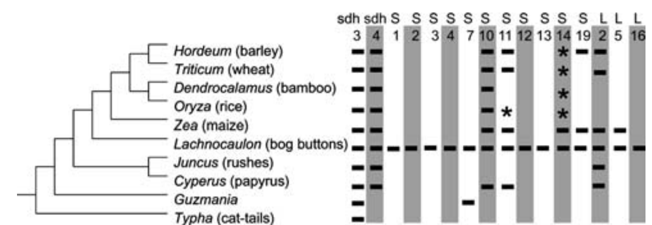

Fig. 1. Gene losses in the Poales order of monocots. Solid rectangles indicate inferred absence of a gene from mitochondrial DNA, as determined by Southern blot hybridizations (Adams et al., 2002b) and sequencing of the rice and maize mitochondrial genomes (Notsu et al., 2002; S. Clifton et al., in preparation). Asterisks indicate sequenced pseudogenes (Notsu et al., 2002; H. Ong and J. D. Palmer, unpublished data). Genes are for succinate dehydrogenase (sdh), and the proteins of the small (S) and large (L) subunits of the mitochondrial ribosome. The phylogeny is from Adams et al. (2002b).

The phylogenetic depth and patterns of gene losses in angiosperms are rather variable. Most of the roughly 370 inferred gene losses inferred by Adams et al. (2002b) are quite restricted in phylogenetic distribution. Only two losses (and transfers), one each of *rps2* and *rps11*, are very deep and ancient within angiosperms, both occurring about 100 million years ago near the base of eudicots, in the common ancestor of clades comprising nearly 180 of the 280 diverse angiosperms studied by Adams et al., 2002b. Some of the general patterns of gene loss are well illustrated in the Poales order of monocots that includes the grasses (Fig. 1). A deep loss of *sdh3* encompasses all species surveyed, whereas loss of several genes has occurred only recently in but a single species. *Lachnocaulon* has lost all 14 ribosomal protein genes and both *sdh* genes, whereas *Typha* has retained all but one ribosomal protein gene. Although *Lachnocaulon* has not yet been examined for gene transfer, many of the other losses in the Poales, including all 8 genes missing from *Zea* (maize) mtDNA, have already been shown to reflect transfer of the gene to the nucleus in one or more species (Adams et al., 2000, 2001a, 2001b, 2002b; Bonen et al., 1998; Figueroa et al., 1999; Kadowaki et al., 1996; Kubo et al., 1999; Kubo et al., 2000a).

2.5. Mitochondria without a genome

In recent years, it has become clear that a number of eukaryotes have independently, convergently lost *all* of their mitochondrial genes, and indeed their entire mitochondrial genome. This reflects the almost singular biochemical/metabolic role of the mitochondrial genome (in aerobic respiration), and the not uncommon tendency to evolve an anaerobic or specialized parasitic lifestyle in which respiration no longer occurs. Most protein genes encoded by mitochondrial genomes across eukaryotes are geared entirely towards respiration (i.e., electron transport and oxidative phosphorylation), either directly, by specifying components of these systems or their assembly, or indirectly, by maintaining the mitochondrial translational apparatus (and, rarely, the transcriptional apparatus too). Moreover, in those eukaryotes with the most reduced mitochondrial gene content, including animals, apicomplexans, and certain fungi and green algae, *all* mitochondrial protein genes function *directly* in respiration.

Given this situation, it is not surprising that the many groups of eukaryotes which have evolved non-respiratory lifestyles have also almost invariably dispensed with their mitochondrial genomes. In many cases, these events are sufficiently ancient and the resulting mitochondrion has diverged sufficiently in ultrastructure and metabolism that for a number of years it was thought that these organisms entirely lacked mitochondria. The lack of mitochondria was thought to be an ancestral feature, due to grouping of many of these organisms (for the most part artefactually), near the base of the eukaryotic tree in rRNA phylogenies. In recent years, however, a combination of approaches both molecular phylogenetic (isolating nuclear genes of mitochondrial phylogenetic origin from these organisms) and cytological (showing that the products of these genes are localized to either a previously undetected or uncharacterized organelle, or to an organelle recognized by an entirely different name) has provided moderate to strong evidence that these non-respiring, mostly anaerobic eukaryotes nonetheless contain a highly derived organelle of mitochondrial descent. We very briefly cover these cases below; for a recent and comprehensive review of this topic, see Williams and Keeling (2003).

The most common type of non-respiring, agenomic mitochondrion is the hydrogenosome, a double-membrane-bound organelle that produces ATP and evolves hydrogen. Hydrogenosomes are found in a broad range of diverse, microaerophilic to anaerobic protists, including trichomonads, certain amoeboflagellates, multiple independent lineages of free-living and rumen-dwelling ciliates, and rumen-dwelling chytrid fungi (reviewed in Dyall and Johnson, 2000; Embley et al., 1997; Hackstein et al., 1999, 2001; Palmer, 1997b; Roger, 1999). Most attempts to detect a genome in the hydrogenosome have proved fruitless (e.g., Clemens and Johnson, 2000; Van der Giezen et al., 1997). There is one claim for a hydrogenosomal genome, in the ciliate *Nyctotherus ovalis* (Akhmanova et al., 1998), but more definite evidence is needed (Dyall and Johnson, 2000).

Mitochondrial genomes have disappeared in a few other cases of non-hydrogenosomal metabolism. Microsporidia, obligate intracellular parasites (primarily of animals) once thought to be basal, primitively amitochondrial eukaryotes, are now known to be highly degenerate and reduced fungi, possessing a relict of a mitochondrion of uncertain function (see Katinka et al., 2001; Williams et al., 2002 and references therein). Likewise, *Entamoeba histolytica*, an anaerobic amoeba, contains a highly derived mitochondrion of unknown function called either a "mitosome" (Tovar et al., 1999) or a "crypton" (Mai et al., 1999).

3. Why mitochondrial gene loss?

There are three fundamental reasons why the cell can tolerate loss of a previously functional gene from the mitochondrial genome. First, the function of that gene may no longer be necessary, and so the gene is lost entirely from the cell. This was undoubtedly a major factor in early eukaryotic evolution, when the protomitochondrion probably rapidly lost many of the genes for metabolic and other functions (e.g., bacterial cell wall synthesis) that were needed for life as a free-living bacterium but not as a permanent, genetic endosymbiont.

The preceding section discussed the many instances in which loss of aerobic respiration has led to loss of the entire mitochondrial genome. Among modern-day, respiring eukaryotes with a still-active mitochondrial genome, loss of gene product function seems to occurs only rarely. For example, this is probably the case for loss of respiratory complex I and all associated NADH dehydrogenase genes in yeasts (see above) and in apicomplexans (Gardner et al., 2002). A perhaps related surprise from the *Plasmodium* genome sequence is that not only does it lack any nuclear genes for complex I, but it also lacks any genes for the mitochondrial coupling factor, or ATP synthase (Gardner et al., 2002). Apicomplexan parasites would thus appear to contain only a very rudimentary respiratory system, with this reflected in their exceptionally reduced mitochondrial genome (encoding only cytochrome *b* and two subunits of cytochrome *c* oxidase).

A second reason why mitochondrial gene loss is tolerable has been termed "gene substitution" or "gene replacement" and also involves complete loss of a gene from the cell. The distinction here is that the function of the missing mitochondrial gene is still needed and is directly replaced by some preexisting gene whose product can play the same role in the mitochondrion. Analysis of the entire yeast mitochondrial proteome (Karlberg et al., 2000) indicates that gene substitution has played a major role in mitochondrial evolution. For example, many mitochondrial ribosomal proteins, aminoacyl-tRNA synthetases, and TCA cycle components are of nuclear origin (Karlberg et al., 2000; reviewed in Martin and Schnarrenberger, 1997; Small et al., 1998). Most commonly, gene substitution occurs via duplication of an ancestrally nuclear gene whose product functions in the cytoplasm, with one gene copy acquiring a mitochondrial targeting sequence and supplanting the original mitochondrial gene. In plants, a few nuclear genes that code for mitochondrial proteins are derived by duplication from nuclear genes of ancestrally chloroplast origin (e.g., *rps13* in some angiosperms; Adams et al., 2002a; Mollier et al., 2002). Not infrequently, though, a *single* nuclear gene (of either nuclear origin, or, in the case of plants, of chloroplast origin) supplies two compartments with the same protein, through use of alternative promoters or splicing to produce products either with and without targeting elements or with two different targeting elements (Hedtke et al., 2000; Peeters and Small, 2001; Small et al., 1998).

Finally, a gene may be lost from the mitochondrion because it has been functionally transferred to the nucleus. This is pretty much the only viable pathway allowing mitochondrial gene loss for functions such as electron transfer and oxidative phosphorylation that were uniquely donated to the eukaryotic cell through mitochondrial endosymbiosis, and for which the nucleus (and chloroplast) lack functional counterparts. Ribosomal proteins are the one major class of genes still present in many mitochondrial genomes for which both transfer and substitution are reasonably viable options to drive mitochondrial gene loss, although as discussed below, accumulating data suggest that (for plants at least), gene transfer is the predominant pathway used.

4. Transfer of mitochondrial genes to the nucleus

4.1. Mitochondrial pseudogenes in the nucleus

The transfer of nucleic acids from the mitochondrion to the nucleus is an ongoing process in most eukaryotes. Transfers can be divided into two types: those that result in the transferred gene becoming functional, and those that result in pseudogenes. Numerous examples of mitochondrial pseudogenes have been documented in the nucleus of humans (e.g., Hazkani-Covo et al., 2003; Mourier et al., 2002; Woischnik and Moraes, 2002) and other animals (reviewed in Bensasson et al., 2001), and in plants (e.g., Blanchard and Schmidt, 1995; Knoop and Brennicke, 1994; Kubo et al., 2001). Such pseudogenes range considerably in size from gene fragments, to small multi-gene segments of mtDNA, up to the remarkable transfer of most of a mitochondrial genome to the centromere of chromosome 2 in *A. thaliana* (Arabidopsis Genome Initiative, 2000). Some mitochondrial pseudogenes are probably dead on arrival, particularly in groups such as animals where the mitochondrial genetic code differs from the standard code (Wolstenholme, 1992). Other transferred sequences probably had the potential to become functional upon transfer but did not acquire the necessary elements for expression and targeting before they become pseudogenes. Mitochondrial pseudogenes in the nucleus have attracted considerable attention in animal systems because their sometimes preferential amplification by PCR can complicate phylogenetic studies (e.g., Sorenson and Fleischer, 1996). Such pseudogenes have been used in evolutionary and phylogenetic studies to infer ancestral states, to root phylogenies, and to perform studies of spontaneous mutations (reviewed in Bensasson et al., 2001).

4.2. Gene escape, transfer, and integration into nuclear DNA

Nucleic acids could escape from the mitochondrion by several mechanisms, including escape during disruptions of mitochondrial membranes that might occur during digestion of organelles by lysosomes or vacuoles, during mitochondrial fusion, or during cellular stress that damages mitochondrial membranes (discussed in Berg and Kurland, 2000; Brennicke et al., 1993; Kurland and Andersson, 2000; Thorsness and Weber, 1996). Another possible escape mechanism is the illicit use of

RNA transport systems (Thorsness and Weber, 1996). The rate of organellar gene escape and uptake by the nucleus has been estimated to be relatively high from experiments with yeast mitochondria (Thorsness and Fox, 1990) and chloroplast genes in tobacco (Huang et al., 2003). DNA can migrate to the nucleus either as genomic DNA or by a cDNA intermediate (reverse transcribed RNA). Once inside the nucleus, mitochondrial DNA can integrate into the nuclear genome by double-strand-break repair, as shown in yeast (Ricchetti et al., 1999), by related chromosome end-joining processes (Blanchard and Schmidt, 1996), and perhaps by other mechanisms.

4.3. Activation and expression of a transferred gene

Transferred genes that become functional in the nucleus must gain regulatory elements for proper expression and, in most cases, a mitochondrial targeting signal sequence. Functional gene transfers have been documented almost exclusively in plants (e.g., Nugent and Palmer, 1991; for a summary, see supporting table 4 in Adams et al., 2002b—http://www.pnas.org/cgi/content/full/042694899/DC1/3) and green algae (Funes et al., 2002a; Pérez-Martínez et al., 2000, 2001). How do transferred genes gain regulatory and targeting elements? One way is by associating with pre-existing genes for mitochondrial proteins, sometimes in complex ways (Fig. 2). Two intriguing cases have been described of insertion of a transferred gene into a pre-existing gene for a mitochondrial protein. The *rps10* gene in carrot was inserted into a duplicate copy of the gene for mitochondrial hsp22, and the S10 protein utilizes the hsp22-derived presequence for targeting and import into mitochondria (Fig. 2; Adams et al., 2000). A remarkable case of acquisition of a targeting presequence by insertion into another gene and alternative splicing with the host gene's presequence has been characterized. The *rps14* gene in maize and rice was inserted into an intron of the *sdh2* gene, and *rps14* is co-expressed with *sdh2* by alternative splicing, utilizing the presequence of *sdh2* (Fig. 2; Figueroa et al., 1999, 2000; Kubo et al., 1999). A transferred gene could gain a mitochondrial presequence and regulatory elements from a tandem duplicate of a gene for a mitochondrial protein, as has been inferred for *sdh3* in *Arabidopsis* (Fig. 2; Adams et al., 2001b). In other cases, such as cotton *sdh3* (Adams et al., 2001b), *rice rps11* (Kadowaki et al., 1996), and *Fuchsia rps10* (Adams et al., 2000), a targeting presequence has been acquired from another gene for a mitochondrial protein, although no clues to the mechanism exist. Another potential source of mitochondrial targeting presequences is genes for non-mitochondrial proteins, as discussed in Adams et al. (2001b).

For other transferred genes, flanking sequences give no clues about the origin of regulatory and targeting elements, although exon-shuffling type processes have been invoked in some cases because of strategically located introns (Fig. 2; Daley et al., 2002; Nugent and Palmer, 1991; Wischmann and Schuster, 1995). Given the considerable flexibility in primary sequence, it is not difficult to imagine that some mitochondrial targeting presequences may form de novo. Considering that the *Arabidopsis* genome has been sequenced (Arabidopsis Genome Initiative, 2000), it is tempting to speculate that the presequences of unknown origin found on 3 recently transferred genes in *Arabidopsis* were derived de novo upon association with the newly transferred genes. Finally, a few transferred genes in angiosperms, including four independently transferred *rps10* genes and 5′ *rpl2* in legumes (Fig. 3), have become activated in the nucleus without gaining a mitochondrial targeting presequence and rely on internal targeting signals (Adams et al., 2000, 2001a; Kubo et al., 2000a).

Once a transferred gene becomes expressed in the nucleus, there is a period of time (however short) when both the mitochondrial and nuclear copy are simultaneously expressed. Two examples of such dual expression

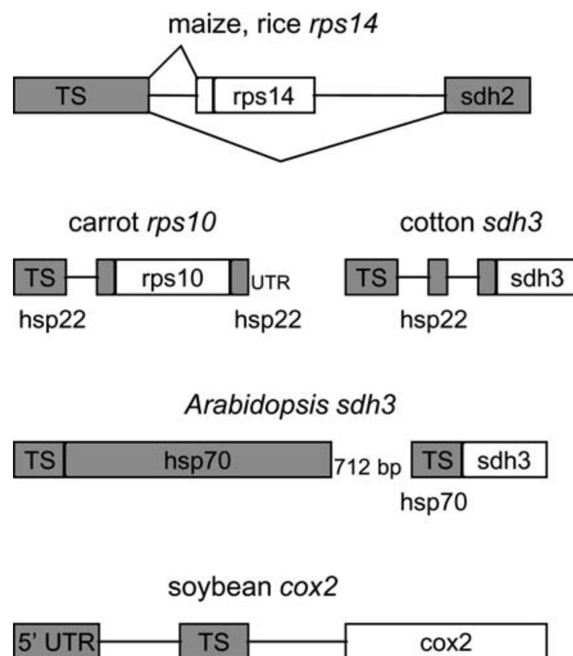

Fig. 2. Acquisition of mitochondrial targeting presequences by transferred genes. Shown are diagrams of four transferred genes whose structures suggest mechanisms for acquisition of a mitochondrial targeting presequence, plus cotton *sdh3* for comparison. Exons are indicated by boxes, introns are indicated by horizontal lines, and alternative splicing is indicated by diagonal lines. Non-shaded boxes indicate exons transferred from the mitochondrion; shading indicates exons derived from the nucleus, in some cases from pre-existing genes. Note that *sdh3* and *hsp70* in *Arabidopsis* contain multiple introns, but are shown here as cDNA sequences. Genes not drawn to scale. Abbreviations: TS, mitochondrial targeting presequence; UTR, untranslated region. See "Section 4.3" of the text for additional explanation and sources of sequences.

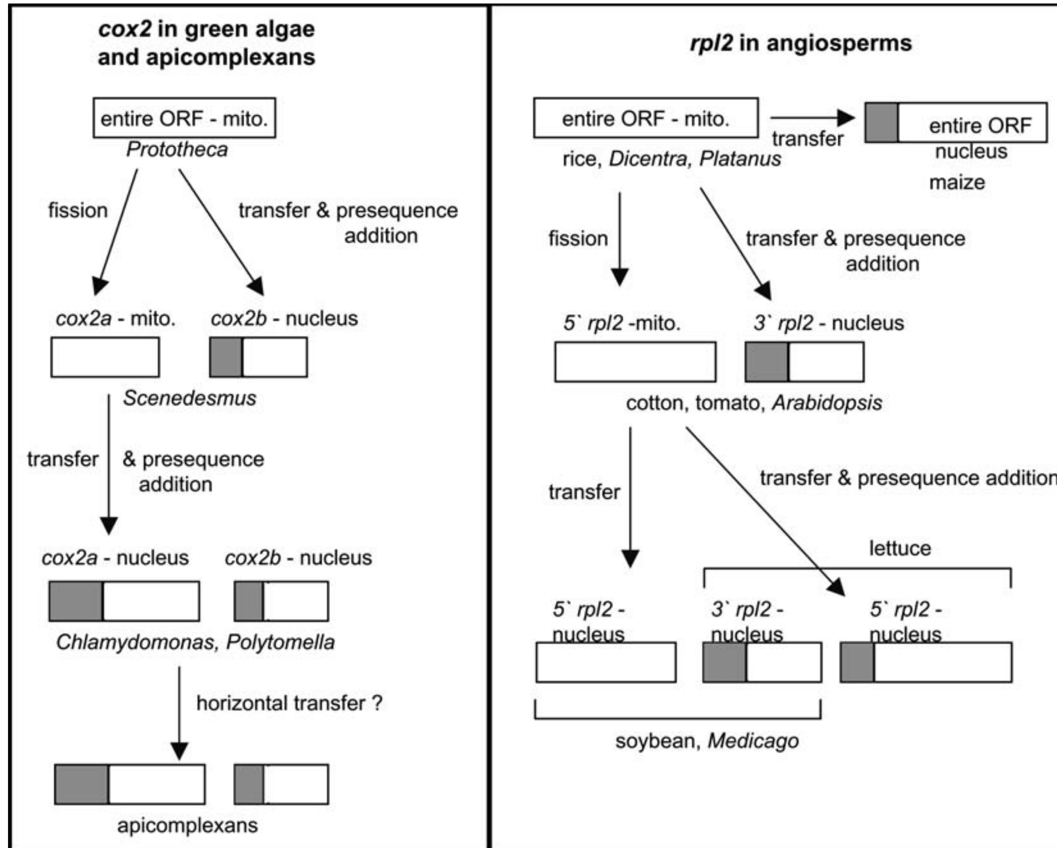

Fig. 3. Split genes in the mitochondrion and nucleus. Gene structure diagrams of *cox2* in green algae and apicomplexans, and *rpl2* in angiosperms. Boxes indicate genes, with shading indicating mitochondrial targeting presequences. Abbreviations: ORF, open reading frame, mito., mitochondrion. Sources of *cox2* data: *Chlamydomonas* and *Polytomella* (Pérez-Martínez et al., 2001), *Scenedesmus* mitochondrial *cox2* (Kuck et al., 2000; Nedelcu et al., 2000), and *Prototheca* (Wolff et al., 1994). *Rpl2* data are from Adams et al. (2001a), except for the lettuce 5′ *rpl2* sequence (GenBank ESTs: BQ869003, BU009168). The lettuce *rpl2* sequence is predicted by MitoProt II to be targeted to mitochondria (Claros and Vincens, 1996) and to contain a 22 amino acid targeting presequence. Predotar (http://www.inra.fr/Internet/Produits/Predotar) also predicts that the sequence is for a mitochondrial product, but the program does not predict presequence cleavage sites. BLAST searches of the NCBI sequence databases indicated that the lettuce 5′ *rpl2* is homologous to mitochondrial *rpl2* genes.

have been reported. Through extensive study of nuclear and mitochondrial *cox2* genes in a group of related legumes, Adams et al. (1999) determined that both *cox2* genes are transcribed in several legumes at varying levels. Following the initial co-expression, it is possible that both the nuclear and mitochondrial copies of a gene may become fixed, as is probably the case in *Neurospora* whose two copies of *atp9* are expressed at different stages of the life cycle (Bittner-Eddy et al., 1994).

Following co-expression of a transferred nuclear gene and its mitochondrial counterpart, gene silencing and loss appears to be a common, perhaps the most common, outcome. Hundreds of mitochondrial gene losses have been documented in plants and other eukaryotes, as discussed above, and most of these are likely to result from functional transfer of the gene to the nucleus. Inactivation of the previously functional nuclear copy can also occur. To prove that a transferred gene has been inactivated, it is essential to show that the gene was active at some point in the evolutionary history of the lineage. Thus, a comparative phylogenetic approach is needed to study several species that are closely related to a species containing a functional transferred gene. Such an approach was used to infer that the transferred nuclear *cox2* gene has been inactivated more than once during the evolution of a group of legumes (Adams et al., 1999). So far *cox2* in legumes is the only reported case of inactivation of a transferred gene that was previously functional in the nucleus, although other examples will probably be documented as additional cases of gene transfer are studied in detail using a comparative phylogenetic approach.

4.4. Split gene transfers

Most cases of functional gene transfer result in an entire gene becoming active in the nucleus. However, there are two genes for mitochondrial proteins that are present as partial genes in the nucleus. Intriguing

scenarios have been developed to explain the presence of these partial genes. In chlorophycean green algae, the gene *cox2* has a variable location in the mitochondrion and nucleus, depending on the species. During the evolution of these green algae the 3′ section of *cox2* was transferred to the nucleus, while the 5′ section remained in the mitochondrion; this situation is represented by *Scenedesmus* (Fig. 3; Kuck et al., 2000; Nedelcu et al., 2000). The 5′ section was transferred to the nucleus in the lineage containing *Chlamydomonas* and *Polytomella* (Pérez-Martínez et al., 2001), so that these green algae contain the two sections of *cox2* as separate genes in the nucleus. Surprisingly, apicomplexans contain a strikingly similar split pair of nuclear *cox2* genes as in *Chlamydomonas* and *Polytomella* (Funes et al., 2002b). Because of this structural similarity and because apicomplexan and green algal *cox2* genes cluster in phylogenetic analyses (Funes et al., 2002b), the apicomplexan *cox2* genes are thought to have been horizontally acquired from the nucleus of a green alga, perhaps from the green alga that is controversially suggested to have given rise by secondary symbiosis to the non-photosynthetic plastid of apicomplexans [also see Palmer (2003) and Williams and Keeling (2003) for alternative scenarios for the origin of the apicomplexan *cox2* genes].

A similar split-gene situation exists for the *rpl2* gene in angiosperms. The 3′ end of *rpl2* was transferred to the nucleus in a common ancestor of core eudicots, leaving the 5′ end as a functional gene in the mitochondria (Fig. 3; Adams et al., 2001a). The 5′ *rpl2* gene was then transferred to the nucleus at least twice, in legumes and in lettuce, and transfers probably occurred in several other lineages that lack the gene in the mitochondrion. The transfer of 5′ *rpl2* in lettuce involved the gain of an amino terminal extension of the coding region that is probably a mitochondrial targeting presequence, but the transfer in legumes did not. Presumably the 5′ *rpl2* gene in legumes has internal signals for targeting, as inferred for recently transferred *rps10* genes in grasses, spinach, and lettuce (Adams et al., 2000; Kubo et al., 2000a).

How did *cox2* and *rpl2* become split genes? To explain *rpl2* fission and transfers in eudicots, Adams et al. (2001a) proposed that the 3′ section was initially transferred to the nucleus to become a functional gene. This event enabled the mitochondrial copy to become truncated so that only the 5′ section remained expressed in the mitochondrion. Finally, the 5′ section was lost several times independently in various lineages of eudicots, including some legumes and lettuce, where the gene has been identified in the nucleus. The *cox2* split gene transfer could be explained in a similar way to *rpl2*: first the 3′ section was transferred, allowing formation of the truncated 5′ section in the mitochondrion. The 5′ section was then transferred in a common ancestor of *Chlamydomonas* and *Polytomella*. An alternative explanation is there was fission of *cox2* in the mitochondrion, followed by transfers of each section to the nucleus at different times during green algal evolution (Pérez-Martínez et al., 2001). Evidence for this possibility would be the existence of both sections of *cox2* as separate genes in the mitochondrion, an arrangement that has not been found for this gene in any mitochondrial genome. However, a precedent for this situation does exist in the case of the split *nad1* gene in the mitochondrion of two ciliate protozoa, *Tetrahymena pyriformis* and *Paramecium Aurelia* (Burger et al., 2000; Edqvist et al., 2000).

4.5. Frequency of functional mitochondrial gene transfer

Mitochondrial genome sequencing projects from a diverse array of eukaryotes, together with Southern blot surveys in angiosperms, have revealed numerous losses of respiratory and, especially, ribosomal protein genes (Adams et al., 2002b; Gray et al., 1998; Lang et al., 1999). In many cases there is evidence for many parallel losses of the same gene in different lineages. While it is likely that many (or even most) of these parallel losses reflect separate transfers to the nucleus, relatively few recently transferred nuclear genes have yet been identified across the broad sweep of eukaryotes. However, recently completed and ongoing nuclear genome sequencing projects in various protists are likely to reveal numerous transferred genes, and in some cases parallel transfers in multiple lineages.

Almost all characterized cases of functional gene transfer where there is evidence that the transfer occurred within even a *relatively* recent context, such as within a phylum, are from plants and green algae. The frequency of functional mitochondrial gene transfer has been most extensively studied in flowering plants. Transfers of 13 mitochondrial genes have been reported, including the respiratory genes *cox2*, *sdh3*, and *sdh4*, along with genes for 10 ribosomal proteins (see Adams et al., 2002b for a summary table). Transferred nuclear copies of 8 genes have been reported in lineages representing two or more parallel losses, as judged by comprehensive Southern blot hybridization surveys, and transfers in four or more parallel loss lineages have been shown for *rps10* (Adams et al., 2000), *sdh3* (Adams et al., 2001b), and *rps19* (Adams et al., 2002b). Inferences of separate transfers of *rps10*, *sdh3*, *rps19*, and other genes are based on a combination of separate parallel losses from mitochondrion, sometimes in rather distantly related lineages; mitochondrial targeting presequences that are completely different from each other, sometimes derived from other genes; and for *rps10*, phylogenetic analyses of nuclear and mitochondrial sequences. See Adams et al. (2001b) for further discussion and evaluation of evidence bearing on the frequency of gene transfers. Overall it appears that many mitochondrial genes in angiosperms have been transferred to the

nucleus many times independently. In some lineages of angiosperms, gene transfer appears to be occurring at an astonishingly high rate, i.e., several times higher than the rate of synonymous substitutions in plant mitochondrial genes (Adams et al., 2002b).

5. Factors promoting or retarding gene transfer

5.1. Is gene transfer a selective or neutral process?

Several hypotheses have been proposed regarding the role of selection in promoting gene transfer to the nucleus. One of the most widely discussed hypotheses, invoking Muller's ratchet, states that the deleterious mutational buildup in asexual mitochondrial genomes is relieved when a gene is transferred to the nucleus (see Berg and Kurland, 2000; Blanchard and Lynch, 2000; Kurland and Andersson, 2000; Martin and Herrmann, 1998). There is evidence for Muller's ratchet in some animal mitochondrial genes (Lynch, 1996) and in the genomes of endosymbiotic bacteria (Moran, 1996). Muller's ratchet could be a factor promoting transfer in eukaryotes where the mitochondrion has a relatively high nucleotide substitution rate. However, plant mitochondrial genes have a very low nucleotide substitution rate, much lower than in the nucleus (Wolfe et al., 1987), perhaps due to factors such as efficient DNA repair, high-fidelity DNA replication, or recombination between genomes. Thus, the effects of Muller's ratchet should be minimal in plant mitochondria. A related hypothesis proposes that beneficial mutations are a major selective factor (Blanchard and Lynch, 2000). When a gene is relocated to the nucleus, recombination between homologous chromosomes can occur during meiosis and help to fix beneficial mutations, a process that does not occur in asexual mitochondria. Thus, the product of a transferred gene might function better than its mitochondrial counterpart and therefore cause loss of the mitochondrial copy. This hypothesis seems plausible for most eukaryotes, provided that beneficial mutations are fixed before deleterious mutations.

Mitochondrial genome streamlining has been proposed to promote transfer of genes to the nucleus (Selosse et al., 2001). According to this hypothesis, a mitochondrial genome that is missing a gene would have an advantage in intra-organellar competition and thus be selectively favored. Genome streamlining would be a plausible factor promoting transfer from small mitochondrial genomes that already show evidence of streamlining. However, plant mitochondrial genomes tend to be very large (e.g., Unseld et al., 1997; Ward et al., 1981) and frequently take up foreign DNA (e.g., Cho et al., 1998; Nugent and Palmer, 1988; Unseld et al., 1997; Watanabe et al., 1994), suggesting that they are under no pressure to streamline. A final selection-based hypothesis is that transferring genes to the nucleus relieves them from the effects of toxic free radicals present in mitochondria that might damage DNA (Allen and Raven, 1996). As with Muller's ratchet, free radicals are unlikely to be much of a factor in plants because of their low nucleotide substitution rate.

A neutral model of gene transfer has been developed by Berg and Kurland (2000). Central to neutral views is that gene transfer is driven by the high rate of DNA escape from the mitochondrion and uptake by the nucleus, shown experimentally in yeast (Thorsness and Fox, 1990; Thorsness and Weber, 1996). A small fraction of the genes that are transferred become activated by gaining regulatory and targeting elements. Once both the nuclear and mitochondrial copies of a gene are functional, the roles of selective and neutral evolution are likely to be variable in different lineages. For lineages with small mitochondrial genomes and high nucleotide substitution rates, Muller's ratchet, streamlining, and avoidance of free radicals might favor loss of the mitochondrial copy. In lineages with large mitochondrial genomes and low nucleotide substitution rates, such as plants, easier fixation of beneficial mutations is perhaps the only major selective factor operating to allow the nuclear copy to out-compete its mitochondrial counterpart. In this regard, it is notable that in plants the nucleotide substitution rate of a gene soars after transfer to the nucleus (see, for example, Adams et al., 2000). The higher nucleotide substitution rate in the nucleus can also be a selective factor against the transferred nuclear copy if deleterious mutations are fixed before beneficial mutations. Thus, there may be a precarious balance between beneficial and deleterious mutations (including recombination) that determines the ultimate fate of the nuclear vs. the mitochondrial copy. Alternatively, it may be largely chance as to which copy becomes silenced and lost. If the nuclear copy is lost or becomes a pseudogene in a particular lineage, then transfer to the nucleus has failed, but it is possible that transfer might be repeated; such a process can be envisioned as a gene transfer ratchet (Doolittle, 1998).

5.2. Why do mitochondria retain genomes?

Maintenance of a separate genetic system in the mitochondrion requires numerous proteins involved in DNA replication, repair, recombination, transcription, RNA processing, translation, and gene regulation. If all mitochondrial genes were transferred to the nucleus, the mitochondrial genome and the many nuclear genes for proteins involved in its maintenance and expression could be eliminated. Furthermore, finely tuned nuclear-cytoplasmic interactions that control gene expression, particularly genes for proteins in multi-subunit complexes that contain subunits synthesized in both the mitochondrion and cytosol, could be eliminated.

Considering the potential advantages of eliminating mitochondrial genetic systems, why do mitochondria retain genomes? Several hypotheses have been proposed to answer this question.

A widely discussed hypothesis for why all genes have not been transferred to the nucleus is that some highly hydrophobic proteins are difficult to import across the mitochondrial membranes and sort to the correct location (e.g., Popot and de Vitry, 1990). Very hydrophobic proteins might also be mis-routed to the endoplasmic reticulum (von Heijne, 1986). It is notable that the only two protein genes contained in every completely sequenced mitochondrial genome—cox1 and cob—are by some criteria the two most hydrophobic proteins present in mitochondria (Claros et al., 1995). There is experimental evidence to support the hydrophobicity hypothesis. When synthesized in the cytosol, cob could not be imported into mitochondria in its entirety—only sections of the protein containing 3 or 4 out of the 8 transmembrane regions could be successfully imported (Claros et al., 1995). Experiments with cytochrome oxidase subunit 2 revealed that the mitochondrial COX2 protein from soybean could not be imported into mitochondria in vitro unless the first transmembrane domain was removed or if two critical amino acids within this domain were changed (Daley et al., 2002). Furthermore it was shown that changing the same amino acids in the transferred nuclear copy to what they are in the mitochondrial copy prevented import. Thus, amino acid changes in a hydrophobic transmembrane domain were necessary for successful transfer of *cox2* to the nucleus in legumes. Reductions in hydrophobicity have been noted for other cases of gene transfer. Transfers of the genes *cox2*, *cox3*, and *atp6* in *Chlamydomonas* all occurred in concert with reduced hydrophobicity in transmembrane regions (Funes et al., 2002a; Pérez-Martínez et al., 2000, 2001). Although there is both circumstantial and experimental evidence to support the hydrophobicity hypothesis, it is very unlikely to account for the current distribution of *all* genes in *every* mitochondrial genome.

A second hypothesis for retention of certain genes in the mitochondrion is that their products are toxic if present in the cytosol (Martin and Schnarrenberger, 1997). If indeed a factor, such toxicity might account for the two genes found in all sequenced mitochondrial genomes—cob and cox1. However, this hypothesis is unlikely to account for the retention of other genes in mitochondria—those that have been transferred to the nucleus in only certain species—unless amino acid modifications can take place after gene transfer to the nucleus to eliminate toxicity but without disrupting normal protein function. It would be interesting to test the toxicity hypothesis in the future with experiments using cytoplasmically synthesized versions of these two proteins. A third hypothesis for retention of certain genes in the mitochondrion whose products have key roles in electron transport and energy coupling is that expression of their products must be quickly and directly regulated by the redox state of the mitochondrion (Allen, 1993; reviewed in Race et al., 1999). So far, to our knowledge, experimental evidence for this hypothesis (for mitochondria) is lacking.

A final hypothesis for why some genes remain in the mitochondrion is that the non-standard genetic code used by the mitochondrion in many eukaryotes, including animals, prevents further gene transfer to the nucleus. In this regard, it is noteworthy that animal mitochondrial gene content is almost completely constant at 13 genes (Boore, 1999). If mitochondrial genes in animals are transferred to the nucleus and become expressed, the resulting proteins would not have the correct amino acid sequence and could contain premature stop codons, as well as many missense mutations. Thus, the non-standard genetic code is probably a major factor preventing further functional gene transfer in animals. Also noteworthy is that the genes most often lost from plant mitochondria and transferred to the nucleus—ribosomal protein genes and succinate dehydrogenase genes—have all been transferred in animals.

5.3. Are nuclear genes transferred back to mitochondria?

The complete mitochondrial genome sequences of *Arabidopsis* and rice revealed numerous remnants of retrotransposons (Notsu et al., 2002; Unseld et al., 1997), with about 5% of the Arabidopsis mitochondrial genome being composed of such elements, showing that these mobile elements can be up taken by mitochondria. However, there is only one report of a potentially functional gene in the mitochondrion that may have been transferred from the nucleus—a *mutS* gene in corals (Pont-Kingdon et al., 1998). Why are genes for mitochondrial proteins so rarely transferred from the nucleus back to the mitochondrion? There are several possibilities to account for the essentially unidirectional functional transfer of mitochondrial genes to the nucleus. Experimental studies in yeast have shown a high rate of gene transfer to the nucleus, but no reverse transfer was detected (Thorsness and Fox, 1990; reviewed in Thorsness and Weber, 1996). One suggested reason for unidirectional gene transfer is because there is vacuole-mediated release of nucleic acids from mitochondria allowing them to escape, but no comparable process for the nucleus (Berg and Kurland, 2000). To become a functional mitochondrial gene, a transferred nuclear gene would have to be originally intron-less or be moved by a cDNA-intermediate because there is no spliceosomal machinery in mitochondria. Finally, the vast majority of nuclear genes do not code for mitochondrial proteins and thus transfer to the mitochondria would not serve a biological function.

Acknowledgments

We thank two anonymous reviewers for helpful comments on the manuscript. K.L.A. is supported by a postdoctoral fellowship from the Jane Coffin Childs Memorial Fund for Medical Research. Research in the lab of J.D.P. has been supported by a grant from the National Institutes of Health.

Note in proof

Although, as discussed in the last section of the text, there is scant evidence for reacquisition of missing mitochondrial genes via reverse, intracellular transfer (i.e., from the nucleus back to the mitochondrion), a recent study has shown that mitochondrial genomes in plants sometimes reacquire missing genes by horizontal transfer, via plant-to-plant, mitochondrial-to-mitochondrial transfer from donor lineages that still have the gene in their mitochondrial genome (Bergthorsson, U., Adams, K.L., Thomason, B., Palmer, J.D. 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature, in press.).

References

Adams, K.L., Song, K., Roessler, P.G., Nugent, J.M., Doyle, J.L., Doyle, J.J., Palmer, J.D., 1999. Intracellular gene transfer in action: dual transcription and multiple silencings of nuclear and mitochondrial *cox2* genes in legumes. Proc. Natl. Acad. Sci. USA 96, 13863–13868.

Adams, K.L., Daley, D.O., Qiu, Y.-L., Whelan, J., Palmer, J.D., 2000. Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. Nature 408, 354–357.

Adams, K.L., Ong, H.C., Palmer, J.D., 2001a. Mitochondrial gene transfer in pieces: fission of the ribosomal protein gene rpl2 and partial or complete gene transfer to the nucleus. Mol. Biol. Evol. 18, 2289–2297.

Adams, K.L., Rosenblueth, M., Qiu, Y.-L., Palmer, J.D., 2001b. Multiple losses and transfers to the nucleus of two mitochondrial succinate dehydrogenase genes during angiosperm evolution. Genetics 158, 1289–1300.

Adams, K.L., Daley, D.O., Whelan, J., Palmer, J.D., 2002a. Genes for two mitochondrial ribosomal proteins in flowering plants are derived from their chloroplast or cytosolic counterparts. Plant Cell 14, 931–943.

Adams, K.L., Qiu, Y.-L., Stoutemyer, M., Palmer, J.D., 2002b. Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc. Natl. Acad. Sci. USA 99, 9905–9912.

Akhmanova, A., Voncken, F., van Alen, T., van Hoek, A., Boxma, B., Vogels, G., Veenhuis, M., Hackstein, J.H.P., 1998. A hydrogenosome with a genome. Nature 396, 527–528.

Allen, J.F., 1993. Redox control of transcription: sensors, response regulators, activators and repressors. FEBS Lett. 332, 203–207.

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

Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C.M., Podowski, R.M., et al., 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396, 133–140.

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–815.

Bensasson, D., Zhang, D., Hartl, D.L., Hewitt, G.M., 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends Ecol. Evol. 16, 314–321.

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

Bittner-Eddy, P., Monroy, A.F., Brambl, R., 1994. Expression of mitochondrial genes in the germinating conidia of *N. eurospora crassa*. J. Mol. Biol. 235, 881–897.

Blanchard, J.L., Schmidt, G.W., 1995. Pervasive migration of organellar DNA to the nucleus in plants. J. Mol. Evol. 41, 397–406.

Blanchard, J.L., Schmidt, G.W., 1996. Mitochondrial DNA migration events in yeast and humans: integration by a common end-joining mechanism and alternative perspectives on nucleotide substitution patterns. Mol. Biol. Evol. 13, 537–548.

Blanchard, J.L., Lynch, M., 2000. Organellar genes: Why do they end up in the nucleus? Trends Genet. 16, 315–320.

Bonen, L., Carrillo, C., Subramanian, S., 1998. Wheat mitochondrial gene expression: splicing, editing, and gene transfer to the nucleus. In: Møller, M., Gardeström, P., Glimelius, K., Glaser, E. (Eds.), Plant Mitochondria: From Gene to Function. Backhuys, Leiden, The Netherlands, pp. 153–158.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27, 1767–1780.

Borner, G.V., Morl, M., Janke, A., Paabo, S., 1996. RNA editing changes the identity of a mitochondrial tRNA in marsupials. EMBO J. 15, 5949–5957.

Brennicke, A., Grohmann, L., Hiesel, R., Knoop, V., Schuster, W., 1993. The mitochondrial genome on its way to the nucleus: different stages of gene transfer in higher plants. FEBS Lett. 325, 140–145.

Bullerwell, C.E., Leigh, J., Forget, L., Lang, B.F., 2003a. A comparison of three fission yeast mitochondrial genomes. Nucleic Acids Res. 31, 759–768.

Bullerwell, C.E., Schnare, M.N., Gray, M.W., 2003b. Discovery and characterization of *Acanthamoeba castellanii* mitochondrial 5S rRNA. RNA 9, 287–292.

Burger, G., Zhu, Y., Littlejohn, T.G., Greenwood, S.J., Schnare, M.N., Lang, B.F., Gray, M.W., 2000. Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. J. Mol. Biol. 297, 365–380.

Cho, Y., Qiu, Y.L., Kuhlman, P., Palmer, J.D., 1998. Explosive invasion of plant mitochondria by a group I intron. Proc. Natl. Acad. Sci. USA 95, 14244–14249.

Claros, M.G., Perea, J., Shu, Y., Samatey, F.A., Popot, J.L., Jacq, C., 1995. Limitations to in vivo import of hydrophobic proteins into yeast mitochondria. The case of a cytoplasmically synthesized apocytochrome *b*. Eur. J. Biochem. 228, 762–771.

Claros, M.G., Vincens, P., 1996. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. 241, 779–786.

Clemens, D.L., Johnson, P., 2000. Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy. Mol. Biochem. Parasitol. 106, 307–313.

Daley, D.O., Clifton, R., Whelan, J., 2002. Intracellular gene transfer: reduced hydrophobicity facilitates gene transfer for subunit 2 of cytochrome *c* oxidase. Proc. Natl. Acad. Sci. USA 99, 10510–10515.

Doolittle, W.F., 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet. 14, 307–311.

Dyall, S.D., Johnson, P.J., 2000. Origins of hydrogenosomes and mitochondria: evolution and organelle biogenesis. Curr. Opin. Microbiol. 3, 404–411.

Edqvist, J., Burger, G., Gray, M.W., 2000. Expression of mitochondrial protein-coding genes in *Tetrahymena pyriformis*. J. Mol. Biol. 297, 381–393.

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

Entelis, N.S., Kolesnikova, O.A., Dogan, S., Martin, R.P., Tarassov, I.A., 2001. 5 S rRNA and tRNA import into human mitochondria. Comparison of in vitro requirements. J. Biol. Chem. 276, 45642–45653.

Feagin, J.E., 1994. The extrachromosomal DNAs of apicomplexan parasites. Annu. Rev. Microbiol. 48, 81–104.

Figueroa, P., Gómez, I., Holuigue, L., Araya, A., Jordana, X., 1999. Transfer of *rps14* from the mitochondrion to the nucleus in maize implied integration within a gene encoding the iron–sulphur subunit of succinate dehydrogenase and expression by alternative splicing. Plant J. 18, 601–609.

Figueroa, P., Holuigue, L., Araya, A., Jordana, X., 2000. The nuclear-encoded SDH2–RPS14 precursor is proteolytically processed between SDH2 and RPS14 to generate maize mitochondrial RPS14. Biochem. Biophys. Res. Commun. 271, 380–385.

Forget, L., Ustinova, J., Wang, Z., Huss, V.A.R., Lang, B.F., 2002. *Hyaloraphidium curvatum*: a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. Mol. Biol. Evol. 19, 310–319.

Funes, S., Davidson, E., Claros, M.G., van Lis, R., Perez-Martinez, X., Vazquez-Acevedo, M., King, M.P., Gonzalez-Halphen, D., 2002a. The typically mitochondrial DNA-encoded ATP6 subunit of the F1F0-ATPase is encoded by a nuclear gene in *Chlamydomonas reinhardtii*. J. Biol. Chem. 277, 6051–6058.

Funes, S., Davidson, E., Reyes-Prieto, A., Magallon, S., Herion, P., King, M.P., Gonzalez-Halphen, D., 2002b. A green algal apicoplast ancestor. Science 298, 2155.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Paln, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.-S, Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M.A, Fairlamb, A.H., Fraunholz, M.J., Ross, D.S., Ralph, S.A., McFadden, G.I., Cummings, L., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., Barrell, B., 2002. Genome sequence of the human malarial parasite *Plasmodium falciparum*. Nature 419, 498–511.

Glover, K.E., Spencer, D.F., Gray, M.W., 2001. Identification and structural characterization of nucleus-encoded transfer RNAs imported into wheat mitochondria. J. Biol. Chem. 276, 639–648.

Gray, M.W., 1999. Evolution of organellar genomes. Curr. Opin. Genet. Dev. 9, 678–687.

Gray, M.W., Schnare, M.N., 1996. Evolution of rRNA gene organization. In: Zimmermann, R.A., Dahlberg, A.E. (Eds.), Ribosomal RNA: Structure, Evolution, Processing, and Function in Protein Biosynthesis, CRC Press, Boca Raton, FL, pp. 49–69 (Chapter 3).

Gray, M.W., Lang, B.F., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T.G., Plante, I., Rioux, P., Saint-Louis, D., Zhu, M., Burger, G., 1998. Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res. 26, 865–878.

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

Gray, M.W., Burger, G., Lang, B.F., 2001. The origin and early evolution of mitochondria. Genome Biol. 2: reviews, 1018.1–1018.5.

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

Hackstein, J.H., Akhmanova, A., Voncken, F.G., van Hoek, A., van Alen, T., Boxma, B., Moon-van der Staay, S.Y., van der Staay, G., Leunissen, J., Huynen, M., Rosenberg, J., Veenhuis, M., 2001. Hydrogenosomes: convergent adaptations of mitochondria to anaerobic environments. Zoology 104, 290–302.

Hazkani-Covo, E., Sorek, R., Graur, D., 2003. Evolutionary dynamics of large *Numts* in the human genome: rarity of independent insertions and abundance of post-insertion duplications. J. Mol. Evol. 56, 169–174.

Hedtke, B., Borner, T., Weihe, A., 2000. One RNA polymerase serving two genomes. EMBO Rep. 2000 1, 435–440.

Huang, C.Y., Ayliffe, M.A., Timmis, J.N., 2003. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature 422, 72–76.

Joyce, P.B.M., Gray, M.W., 1989. Chloroplast-like transfer RNA genes expressed in wheat mitochondria. Nucleic Acids Res. 17, 5461–5475.

Kadowaki, K., Kubo, N., Ozawa, K., Hirai, A., 1996. Targeting presequence acquisition after mitochondrial gene transfer to the nucleus occurs by duplication of existing targeting signals. EMBO J. 15, 6652–6661.

Karlberg, O., Canback, B., Kurland, C.G., Andersson, S.G., 2000. The dual origin of the yeast mitochondrial proteome. Yeast 17, 170–187.

Katinka, M.D., Duprat, S., Cornillot, E., Metenier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., Vivares, C.P., 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414, 450–453.

Knoop, V., Brennicke, A., 1994. Promiscuous mitochondrial group II intron sequences in plant nuclear genomes. J. Mol. Evol. 39, 144–150.

Kobayashi, Y., Knoop, V., Fukuzawa, H., Brennicke, A., Ohyama, K., 1997. Interorganellar gene transfer in bryophytes: the functional *nad7* gene is nuclear encoded in *Marchantia polymorpha*. Mol. Gen. Genet. 256, 589–592.

Kubo, N., Harada, K., Hirai, A., Kadowaki, K.-I., 1999. A single nuclear transcript encoding mitochondrial RPS14 and SDHB of rice is processed by alternative splicing: common use of the same mitochondrial targeting signal for different proteins. Proc. Natl. Acad. Sci. USA 96, 9207–9211.

Kubo, N., Jordana, X., Ozawa, K., Zanlungo, S., Harada, K., Sasaki, T., Kadowaki, K., 2000a. Transfer of the mitochondrial *rps10* gene to the nucleus in rice: acquisition of the 5′ untranslated region followed by gene duplication. Mol. Gen. Genet. 263, 733–739.

Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A., Mikami, T., 2000b. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA Cys (GCA). Nucleic Acids Res. 28, 2571–2576.

Kubo, T., Takano, M., Nishiguchi, M., Kadowaki, K., 2001. Mitochondrial sequence migrated downstream to a nuclear V-ATPase B gene is transcribed but non-functional. Gene 271, 193–201.

Kuck, U., Jekosch, K., Holzamer, P., 2000. DNA sequence analysis of the complete mitochondrial genome of the green alga *Scenedesmus obliquus*: evidence for UAG being a leucine and UCA being a non-sense codon. Gene 253, 13–18.

Kumar, R., Marechal-Drouard Akama, L., Akama, K., Small, I., 1996. Striking differences in mitochondrial tRNA import between different plant species. Mol. Gen. Genet. 252, 404–411.

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

Lang, B.F., Burger, G., O'Kelly, C.J., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., Gray, M.W., 1997. An

ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387, 493–497.

Lang, B.F., Gray, M.W., 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.

Magalhaes, P.J., Andreu, A.L., Schon, E.A., 1998. Evidence for the presence of 5S rRNA in mammalian mitochondria. Mol. Biol. Cell 9, 2375–2382.

Mai, Z.M., Ghosh, S., Frisardi, M., Rosenthal, B., Rogers, R., Samuelson, J., 1999. Hsp60 is targeted to a cryptic mitochondrion-derived organelle ("crypton") in the microaerophilic protozoan parasite *Entamoeba histolytica*. Mol. Cell. Biol. 19, 2198–2205.

Martin, W., Herrmann, R.G., 1998. Gene transfer from organelles to the nucleus: How much, what happens, and why? Plant Physiol. 118, 9–17.

Martin, W., Schnarrenberger, C., 1997. The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr. Genet. 32, 1–18.

Mollier, P., Hoffmann, B., Debast, C., Small, I., 2002. The gene encoding *Arabidopsis thaliana* mitochondrial ribosomal protein S13 is a recent duplication of the gene encoding plastid S13. Curr. Genet. 40, 405–409.

Moran, N.A., 1996. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc. Natl. Acad. Sci. USA 93, 2873–2878.

Mourier, T., Hansen, A.J., Willerslev, E., Arctander, P., 2002. The human genome project reveals a continuous transfer of large mitochondrial fragments to the nucleus. Mol. Biol. Evol. 18, 1833–1837.

Nedelcu, A.M., Lee, R.W., Lemieux, C., Gray, M.W., Burger, G., 2000. The complete mitochondrial DNA sequence of *Scenedesmus obliquus* reflects an intermediate stage in the evolution of the green algal mitochondrial genome. Genome Res. 10, 819–831.

Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J.A., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., Potocka, I., Nelson, W.C., Newton, A., Kolonay, J.F., Smit, J., Craven, M.B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S.L., Venter, J.C., Shapiro, L., Fraser, C.M., Eisen, J., 2001. Complete genome sequence of *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA 98, 4136–4141.

Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A., Kadowaki, K., 2002. The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. Mol. Genet. Genomics 268, 434–445.

Nugent, J.M., Palmer, J.D., 1988. Location, identity, amount and serial entry of chloroplast DNA sequences in crucifer mitochondrial DNAs. Curr. Genet. 14, 501–509.

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

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., Kohchi, T., Ohyama, K., 1992. Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. J. Mol. Biol. 223, 1–7.

Oudot-Le Secq, M.P., Fontaine, J.M., Rousvoal, S., Kloareg, B., Loiseaux-De Goer, S., 2001. The complete sequence of a brown algal mitochondrial genome, the ectocarpale *Pylaiella littoralis* (L.) Kjellm. J. Mol. Evol. 53, 80–88.

Palmer, J.D., 1997a. The mitochondrion that time forgot. Nature 387, 454–455.

Palmer, J.D., 1997b. Organelle genomes: going, going, gone!. Science 275, 790–791.

Palmer, J.D., 2003. The symbiotic birth and spread of plastids: How many times and Whodunit? J. Phycol. 39, 4–11.

Peeters, N., Small, I., 2001. Dual targeting to mitochondria and chloroplasts. Biochim. Biophys. Acta 154, 54–63.

Pérez-Martínez, X., Vázquez-Acevdeo, M., Tolkunova, E., Funes, S., Claros, M.G., Davidson, E., King, M.P., González-Halphen, D., 2000. Unusual location of a mitochondrial gene: subunit III of cytochrome *c* oxidase is encoded in the nucleus of chlamydomonad algae. J. Biol. Chem. 275, 30144–30152.

Pérez-Martínez, X., Antaramian, A., Vázquez-Acevdeo, M., Funes, S., Tolkunova, E., d'Alayer, J., Claros, M., Davidson, E., King, M.P., González-Halphen, D., 2001. Subunit II of cytochrome *c* oxidase in Chlamydomonad algae is a heterodimer encoded by two independent nuclear genes. J. Biol. Chem. 276, 11302–11309.

Pont-Kingdon, G.A., Okada, N.A., Macfarlane, J.L., Beagley, C.T., Watkins-Sims, C.D., Cavalier-Smith, T., Clark-Walker, G.D., Wolstenholme, D.R., 1998. Mitochondrial DNA of the coral *Sarcophyton glaucum* contains a gene for a homologue of bacterial MutS: a possible case of gene transfer from the nucleus to the mitochondrion. J. Mol. Evol. 46, 419–431.

Popot, J.L., de Vitry, C., 1990. On the microassembly of integral membrane proteins. Annu. Rev. Biophys. Biophys. Chem. 19, 369–403.

Race, H.L., Herrmann, R.G., Martin, W., 1999. Why have organelles retained genomes? Trends Genet. 15, 364–370.

Ricchetti, M., Fairhead, C., Dujon, B., 1999. Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. Nature 402, 96–100.

Roger, A.J., 1999. Reconstructing early events in eukaryotic evolution. Am. Nat. 154, S146–S163.

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.

Small, I., Wintz, H., Akashi, K., Mireau, H., 1998. Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol. Biol. 38, 265–277.

Small, I., Akashi, K., Chapron, A., Dietrich, A., Duchene, A.-M., Lancelin, D., Marechal-Drouard, L., Menand, B., Mireau, H., Moudden, Y., Ovesna, J., Peeters, N., Sakamoto, W., Souciet, G., Wintz, H., 1999. The strange evolutionary history of plant mitochondrial tRNAs and their aminoacyl-tRNA synthetases. J. Hered. 90, 333–337.

Sorenson, M.D., Fleischer, R.C., 1996. Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. Proc. Natl. Acad. Sci. USA 93, 15239–15243.

Stiller, J.W., Reel, D.C., Johnson, J.C., 2003. A single origin of plastids revisited: convergent evolution in organellar genome content. J. Phycol. 39, 95–105.

Takano, H., Abe, T., Sakurai, R., Moriyama, Y., Miyazawa, Y., Nosaki, H., Kawano, S., Sasaki, N., Kuroiwa, T., 2001. The complete DNA sequence of the mitochondrial genome of *Physarum polycephalum*. Mol. Gen. Genet. 264, 539–545.

Thorsness, P.E., Fox, T.D., 1990. Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae*. Nature 346, 376–379.

Thorsness, P.E., Weber, E.R., 1996. Escape and migration of nucleic acids between chloroplasts, mitochondria, and the nucleus. Int. Rev. Cytol. 165, 207–233.

Tovar, J., Fischer, A., Clark, C.G., 1999. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. Mol. Microbiol. 32, 1013–1021.

Turmel, M., Lemieux, C., Burger, G., Lang, B.F., Otis, C., Plante, I., Gray, M.W., 1999. The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*: two radically different evolutionary patterns within Green Algae. Plant Cell 11, 1717–1730.

Turmel, M., Otis, C., Lemieux, C., 2002. The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. Proc. Natl. Acad. Sci. USA 17, 11275–11280.

Unseld, M., Marienfeld, J.R., Brandt, P., Brennicke, A., 1997. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366, 924 nucleotides. Nat. Genet. 15, 57–61.

van der Giezen, M., Sjollema, K.A., Artz, R.R.E., Alkema, W., Prins, R.A., 1997. Hydrogenosomes in the anaerobic fungus *Neocallimastix frontalis* have a double membrane but lack an associated organelle genome. FEBS Lett. 408, 147–150.

von Heijne, G., 1986. Why mitochondria need a genome. FEBS Lett. 198, 1–4.

Watanabe, N., Nakazono, M., Kanno, A., Tsutsumi, N., Hirai, A., 1994. Evolutionary variations in DNA sequences transferred from chloroplast genomes to mitochondrial genomes in the Gramineae. Curr. Genet. 26, 512–518.

Ward, B.L., Anderson, R.S., Bendich, A.J., 1981. The mitochondrial genome is large and variable in a family of plants (cucurbitaceae). Cell 25, 793–803.

Williams, B.A.P., Keeling, P.J., 2003. Cryptic organelles in parasitic protists and fungi. Adv. Protistol. (in press).

Williams, B.A.P., Hirt, R.P., Lucocq, J.M., Embley, T.M., 2002. A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. Nature 418, 865–869.

Wischmann, C., Schuster, W., 1995. Transfer of *rps10* from the mitochondrion to the nucleus in *Arabidopsis thaliana*: evidence for RNA-mediated transfer and exon shuffling at the integration site. FEBS Lett. 375, 152–156.

Woischnik, M., Moraes, C.T., 2002. Pattern of organization of human mitochondrial pseudogenes in the nuclear genome. Genome Res. 12, 885–893.

Wolfe, K.H., Li, W.H., Sharp, P.M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. USA 84, 9054–9058.

Wolff, G., Plante, I., Lang, B.F., Kuck, U., Burger, G., 1994. Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii*. Gene content and genome organization. J. Mol. Biol. 237, 75–86.

Wolstenholme, D.R., 1992. Genetic novelties in mitochondrial genomes of multicellular animals. Curr. Opin. Genet. Dev. 2, 918–925.