Mitochondrial Genome Evolution in the Social Amoebae

Andrew J. Heidel and Gernot Glöckner

Genome Analysis Group, Leibniz Institute for Age Research—Fritz Lipmann Institute, Jena, Germany

Most mitochondria contain a core set of genes required for mitochondrial function, but beyond this base there are variable genomic features. The mitochondrial genome of the model species *Dictyostelium discoideum* demonstrated that the social amoebae mitochondrial genomes have a size between those of metazoans and plants, but no comparative study of social amoebae mitochondria has been performed. Here, we present a comparative analysis of social amoebae mitochondrial genomes using *D. discoideum*, *Dictyostelium citrinum*, *Dictyostelium fasciculatum*, and *Polysphondylium pallidum*. The social amoebae mitochondria have similar sizes, AT content, gene content and have a high level of synteny except for one segmental rearrangement and extensive displacement of tRNAs. From the species that contain the rearrangement, it can be concluded that the event occurred late in the evolution of social amoebas. A phylogeny using 36 mitochondrial genes produced a well-supported tree suggesting that the pairs of *D. discoideum*/*D. citrinum* and *D. fasciculatum*/*P. pallidum* are sister species although the position of the root is not certain. Group I introns and endonucleases are variable in number and location in the social amoebae. Phylogenies of the introns and endonucleases suggest that there have been multiple recent duplications or extinctions and confirm that endonucleases have the ability to insert into new areas. An analysis of *d*N/*d*S ratios in mitochondrial genes revealed that among groups of genes, adenosine triphosphate synthase complex genes have the highest ratio, whereas cytochrome oxidase and nicotinamide adenine dinucleotide (NADH) dehydrogenase genes had the lowest ratio. The genetic codes of *D. citrinum*, *P. pallidum*, and *D. fasciculatum* are the universal code although *D. fasciculatum* does not use the TGA stop codon. In *D. fasciculatum*, we demonstrate for the first time that a mitochondrial genome without the TGA stop codon still uses the release factor *RF2* that recognizes TGA. Theories of how the genetic code can change and why *RF2* may be a constraint against switching codes are discussed.

Introduction

Mitochondrial genomes vary considerably across eukaryotes providing the opportunity for numerous comparisons and investigation of the selective forces acting on genome evolution and the mechanisms of the changes. Many features of the mitochondrial genome display variation, including genome architecture and size, gene content and order, introns and intronic endonucleases, and the genetic code (Burger, Gray, and Lang 2003; Gray et al. 2004; Santos et al. 2004). Furthermore, the effect of selection and mutation pressure on these genomic features has been investigated (Lynch et al. 2006; Meiklejohn et al. 2007).

Most mitochondrial genome comparative studies have been performed in animal taxa due to the large number of animal mitochondrial sequences available. However, a number of mitochondrial genomes from unicellular organisms have recently been sequenced and display large differences from animals in many mitochondrial genomic features. In comparison to animals, the genome size in protists is more variable, ranging from 6 to 77 kb, the number of genes is generally larger, and the gene content often resembles that of plants (Burger, Forget, et al. 2003; Gray et al. 2004). How mitochondrial genomes have rearranged and the mechanisms involved have also been investigated heavily in animals (Dowton and Campbell 2001) with fewer studies in other taxa.

The social amoebae inhabit an area between unicellular and multicellular organisms; the social amoebae can not only grow indefinitely as solitary amoebae but can also aggregate to form fruiting bodies under resource deprivation. This unique developmental midway point is useful to examine steps on the pathway to multicellularity, and the nuclear and mitochondrial genomes of the model species *Dictyostelium discoideum* have recently been sequenced allowing genome-wide analysis (Ogawa et al. 2000; Eichinger et al. 2005) of both genomes. The mitochondrial genome of *D. discoideum* contains many differences in introns, intronic endonucleases, and the genetic code from its nearest solitary amoebae relatives such as *Acanthamoeba castellanii*.

Group I introns are self-splicing introns present in nuclear and mitochondrial genomes that sometimes contain homing endonucleases. The endonucleases can cleave an intronless allele leading to insertion of the intron and endonuclease into the previously intronless allele (Chevalier and Stoddard 2001). Both introns and endonucleases are present in *D. discoideum* (Ogawa, Matsuo, et al. 1997; Ogawa, Naito, et al. 1997), but their ancestry and evolution have not been investigated.

Alternative genetic codes are a hallmark of mitochondrial genomes, occurring much more often than in nuclear genomes. Indeed, all metazoans and fungi and some additional taxa including ciliates and kinetoplastids have at least the TGA stop codon switched to coding for tryptophan (Knight et al. 2001). There are different theories explaining how the genetic code changes (Santos et al. 2004) and each may apply to different situations and organisms. The mitochondria of *D. discoideum* have the universal genetic code (Cole and Williams 1994), whereas the mitochondria of the solitary amoeba *A. castellanii* have an alternative genetic code whereby the TGA stop codon codes for tryptophan (Burger et al. 1995). The reasons why a genetic code change did not occur in *D. discoideum* and did occur in *A. castellanii* are not known, but a comparative approach using more related genomes can be a useful mode of inquiry because related genomes may represent intermediate steps.

These variable parts of mitochondrial genomes discussed above have not been comparatively studied in the social amoebae with the exception of tRNA and genetic

Key words: *Dictyostelium discoideum*, social amoebae, mitochondrial genomes, genetic codes.

E-mail: aheidel@fli-leibniz.de.

Mol. Biol. Evol. 25(7):1440–1450. 2008
doi:10.1093/molbev/msn088
Advance Access publication April 15, 2008

code comparisons in *D. discoideum* and *Polysphondylium pallidum* (Gray et al. 2004). Here we perform comparative studies on the social amoebae's mitochondrial genome using a selection of social amoebae widely spaced throughout the lineage plus one close relative of *D. discoideum*. The social amoebae can be divided into 4 main phylogenic groups (Schaap et al. 2006). We use 1 species each in the most basal 2 groups: *Dictyostelium fasciculatum* (SH3) and *P. pallidum* (PN500) in groups 1 and 2, respectively. We additionally use the model species *D. discoideum* (AX-3/X22) in the most derived group 4 and a closely related species also in group 4, *Dictyostelium citrinum*. In these comparisons, we focus on synteny and a segmental rearrangement, variation in introns and intronic endonucleases, and selection and changes in the genetic code.

Methods

DNA Extraction and Sequencing

DNA from *D. citrinum*, *D. fasciculatum*, and *D. purpureum* was extracted from cells using the Blood and Cell Culture DNA Midi Kit (QIAGEN, Hilden, Germany). DNA extraction from *P. pallidum* was performed using the following summarized protocol: 1) lysis with detergent, 2) treatment with proteinase K, 3) extraction with phenol/chloroform and then chloroform again, and 4) precipitation with ethanol followed by reprecipitation with sodium acetate and ethanol. DNA was cloned and sequenced as follows: DNA was sonicated to produce fragments of 1.5–5 kb and ligated into the pUC18 vector and transformed into DH10B electrocompetent cells (Invitrogen, Carlsbad, CA). Individual clones were grown in 96-well racks, and plasmid DNA was extracted using the MagAttract 96 Miniprep Core Kit (QIAGEN). Sequencing was then performed using Applied Biosystems (ABI, Foster City, CA) BigDye Terminator v3.1 Sequencing Reagents on an ABI PRISM 3730XL 96 capillary DNA analyzer equipped with 3730xl SeqA, SeqScape, and Data Collection Software. Additionally, some sequence of *P. pallidum* was obtained with a GS20 sequencer (Roche, Basel, Switzerland).

Assembly of Mitochondrial Genomes

Mitochondrial DNA is a significant part of the total DNA of the amoeba cells. Therefore, survey sequencing and assembly of several thousand clones yielded some larger contigs for each species, several of which could be assigned to the mitochondrial genomes. Gaps or regions of poor quality sequence that remained in the mitochondrial genomes were closed or improved, respectively, by designing primers flanking the area, generating polymerase chain reaction (PCR) fragments, and sequencing the fragments. Complete mitochondrial genomes were assembled for all species except *D. purpureum*. The already sequenced mitochondrial genomes of *D. discoideum* and *A. castellanii* in GenBank were also used for analysis. There is an additional *P. pallidum* strain (CK8) sequence available at GenBank (AY700145). This strain is distinct on a sequence level from the strain (PN500) used here. For example, *nad4* and *cox3* had 86.4% and 84.7% Stretcher identity scores (Rice et al. 2000), respectively, between CK8 and PN500. This is on the same order as the difference between *D. discoideum* and *D. citrinum* that have 86.4% and 87.8% Stretcher identity scores for the same 2 genes. The sequences from CK8 were not included in analyses because we already had one member of group 2.

In *D. purpureum*, sequences from the genes on the edges of the segmental rearrangement were identified and then 3 additional sequences were generated from PCR fragments using primers in supplementary table 1 (Supplementary Material online). The GenBank accession numbers for the *D. purpureum atp1-rns*, *atp9-rnl*, and *nad3-nad9* segments are EU339292, EU339291, and EU339293, respectively.

Genomic features were identified using a combination of GeneMarkS (Besemer et al. 2001), Blast searches, tRNAscan (Lowe and Eddy 1997), and EMBOSS applications. Introns in *D. fasciculatum* were additionally confirmed by sequences derived from RNA templates. All open reading frames (ORFs) equal or greater than 200 nt were annotated. The GenBank accession numbers for the *D. citrinum*, *D. fasciculatum*, and *P. pallidum* mitochondrial genomes are DQ336395, EU275727, and EU275726, respectively.

Nuclear Genes in *D. fasciculatum*

The assembled contigs from clones of total DNA of *D. fasciculatum* also contained nuclear DNA and from these sequences plus additional sequence generated with a FLX sequencer (Roche), the nuclear genes *RF1* and *RF2* were identified. The GenBank accession numbers for *RF1* and *RF2* are EU307938 and EU307939, respectively.

RNA Extraction from *D. fasciculatum* and Sequencing

PolyA RNA was extracted with the MicroPoly(A)Purist kit (Ambion, Austin, TX). The resulting RNA was reverse transcribed and then double-stranded DNA was generated using the SMART PCR cDNA Synthesis Kit using 15 cycles of amplification (Clontech, Mountain View, CA). The DNA was then sequenced on a Roche GS20 machine. The genes *RF1* and *RF2* were additionally amplified and sequenced from first-strand synthesis cDNA using primers in supplementary table 1 (Supplementary Material online).

Alignments

The nucleic acid alignments of mitochondrial genes and introns and protein alignments of endonucleases were initially performed using ClustalW with default settings. For the 34 mitochondrial protein-coding genes and the 2 ribosomal RNAs, alignments were adjusted by hand and sites where there was a gap in more than 1 species were excluded in addition to areas not unambiguously aligned. For the endonucleases and introns, alignments were adjusted by hand and sites where there was a gap in more than 4 endonucleases or introns were excluded in addition to areas not unambiguously aligned.

Table 1
Genome Summary

	Dd	Dc	Pp	Df	Ac
Genome size (kb)	55.6	57.8	48.0	54.6	41.6
%AT	72.6	73.0	75.8	74.5	70.6
Genes	42	43	40	40	34
Total tRNAs	18	19	20	17	15
aa coverage	15	15	15	15	13
TAA	32	36	35	38	23
TAG	8	6	3	2	11
TGA	2	1	2	0	0

NOTE.—Dd, *Dictyostelium discoideum*; Dc, *Dictyostelium citrinum*; Pp, *Polysphondylium pallidum*; Df, *Dictyostelium fasciculatum*; Ac, *Acanthamoeba castellanii*; genes, the number of protein-coding genes including the 4 conserved unknown ORFs and endonucleases but not unknown unconserved ORFs; aa coverage, the number of amino acids that are covered by the tRNAs in the mitochondrial genome; TAA, TAG, and TGA, the number of times the respective stop codon is used in genes.

Tree Reconstruction

All trees were constructed using MrBayes (version 3.1.1) with the default number of runs for 5×10^5, 1×10^6, and 1×10^6 generations for the concatenated mitochondrial gene tree, the intron tree, and the endonuclease tree, respectively, and sampled every 100 generations (Huelsenbeck and Ronquist 2001). A General Time Reversible (GTR) evolutionary model was used with gamma-distributed rate variation for the concatenated mitochondrial genes and intron trees, and a mixed amino acid model was used for the endonuclease tree. All 3 analyses were run past the point when the average standard deviation of split frequencies was <0.01.

Molecular Evolution Analysis

Before analysis, all sites with the TGA codon in *A. castellanii* (encoding tryptophan) were removed from the alignments to prevent the different genetic codes from biasing the analysis. The *d*N/*d*S ratios were determined and compared with the program HyPhy using the pairwise relative ratio standard analysis (this analysis compares all pairs of *d*N/*d*S ratios) with the likelihood-based Muse-Gaut 94 codon global model (Pond et al. 2005). Comparisons between genes or categories were corrected for multiple comparisons by the sequential Bonferroni method.

Results

Genome Summary

The genomes of the 5 mitochondria have similar sizes and AT content ranging from 41.6 to 57.8 kb and 70.6 to 75.8%, respectively (table 1), and all genes are encoded on the same strand of DNA. The 5 genomes share 34 known mitochondrial protein-coding genes and 2 rRNA genes encoding the large and small rRNA subunits. In addition to the genes identified in *D. discoideum* (Ogawa et al. 2000), *atp4* can be identified based on homology and position to the *atp4* in *P. pallidum* (GenBank accession number YP_209574). The 4 social amoebae also share 4 unknown genes (orthologs to DidioMp05, DidioMp21, DidioMp26, and DidioMp36 in *D. discoideum*). DidioMp05 and DidioMp21 have a low level of similarity at the sequence level to the orthologs in *P. pallidum* and *D. fasciculatum*, but the perfectly conserved location of the genes suggests that they are indeed orthologs. DidioMp26 has partial homology to part of the *rps3* gene, so it is likely a part of the ribosomal complex. The 3 remaining conserved unknown genes have no strong similarity to any other known gene, so their function cannot be ascertained.

Dictyostelium citrinum contains an ortholog of the small RNA identified in *D. discoideum* (Pi et al. 1998) and although *D. fasciculatum* and *P. pallidum* do not have a homologous region, they do have a noncoding area in the same region 3′ to the large rRNA. The 5 genomes contain a variable number of tRNAs that in no case covers the complete set of amino acids (table 1), indicating that all must import tRNAs from the cytosol.

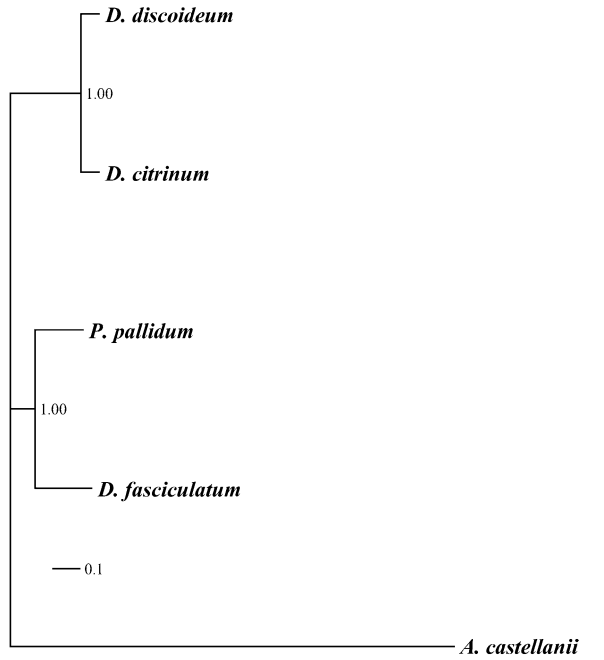

FIG. 1.—Bayesian phylogeny using the 34 mitochondrial protein-coding genes and the 2 rRNAs in common between the 5 species with *Acanthamoeba castellanii* used as the outgroup. Numbers at branches indicate Bayesian inference posterior probabilities. Branch lengths correspond to the substitutions per site given by the scale bar.

Mitochondrial Phylogeny

The 34 known protein-coding genes and 2 rRNA genes in common between the 5 species were concatenated to construct a phylogeny by Bayesian inference (fig. 1). The tree indicates that the pairs of *D. discoideum*/*D. citrinum* and *D. fasciculatum*/*P. pallidum* are both sister species and is consistent with a tree generated from small subunit (SSU) rDNA and tubulin sequences (Schaap et al. 2006).

Synteny and Segmental Rearrangement

The order and presence of tRNAs between the 4 social amoebae are similar but include multiple duplications,

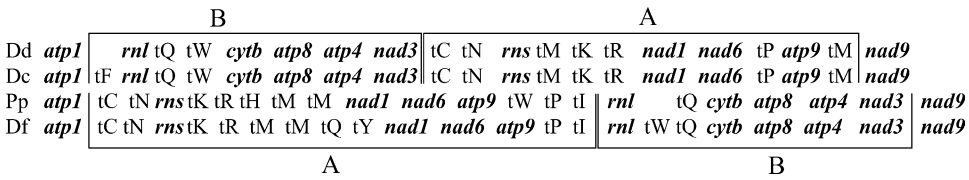

FIG. 2.—Segmental rearrangement. tF, tRNA for amino acid given by capital letter, in this case phenylalanine. Segments A and B are switched between the pair of Dd/Dc and the pair Pp/Df.

deletions, and rearrangements of individual tRNAs. All 4 social amoebae contain a single tRNA for the amino acids Q, C, N, K, R, P, Y, A, and E and contain 2 tRNAs for the amino acids M and L. However, the tRNAs are not always located in the same place in different species. For example, the tRNA for Y is located between *atp6* and *rpl11* in *D. discoideum* and *D. citrinum* but is in a different location in *D. fasciculatum* and in a different location again in *P. pallidum*. There are 1 or 2 tRNAs for W, I, H, and F depending on the species indicating that either deletions or duplications have occurred. Additionally, all amino acids that have only one tRNA use the same anticodon in all 4 species. The amino acids with multiple tRNAs also only use one anticodon for both tRNAs in all species with one exception. The exception is for the amino acid L, where in all species 1 tRNA uses the anticodon TAA and the other tRNA uses the anticodon TAG.

In contrast, the 4 social amoebae have the identical gene order in the conserved 34 protein-coding genes and rRNAs with one exception. There is much less synteny between 4 social amoebae and *A. castellanii* or the slime mold, *Physarum polycephalum* with mostly only pairs of genes in the same order.

The one exception in protein-coding and rRNA gene order in the social amoebae is a segmental rearrangement between the pairs of *D. discoideum*/*D. citrinum* and *P. pallidum*/*D. fasciculatum* (fig. 2) The rearrangement can be dated to a point after the split between the 2 pairs, however, it is difficult to determine if the rearrangement occurred in *D. discoideum*/*D. citrinum* or *P. pallidum*/*D. fasciculatum* because the mitochondrial genome of *A. castellanii*, the species used as the root, has only limited synteny to the social amoebae mitochondrial genomes. Despite this fact, *A. castellanii* does have the 2 genes *nad3* and *nad9* next to each other suggesting that the arrangement in *D. fasciculatum* and *P. pallidum* is ancestral. Segment A codes for tRNAs at both ends suggesting that recombination could have been aided by the reordering of tRNA genes. Although the protein-coding and rRNA genes in the segmental rearrangement have the same order within the rearranged sections, the tRNAs have further rearranged. For example, *P. pallidum* contains an extra W tRNA after *rnl*, which is absent in the other 3 species.

To place the rearrangement at a more specific point in the social amoebae phylogeny, we checked for the rearrangement in one additional group 4 species, *D. purpureum* (WS321). The presence or absence of the rearrangement was determined by generating and sequencing PCR fragments across the edges of segments A and B (fig. 2) (i.e., to test for the *D. fasciculatum* order, 3 fragments were generated and sequenced from 1) *atp1* to *rns*, 2) *atp9* to *rnl*, and 3) *nad3* to *nad9*). *Dictyostelium purpureum* contains the same arrangement as *P. pallidum* and *D. fasciculatum* indicating that the rearrangement occurred after the divergence of group 4.

Introns and Endonucleases

Group I introns exist in the genes for *cox1/2* and *rnl* in most species and many contain endonucleases (fig. 3). All endonucleases are contained within introns except in *P. pallidum* where both endonucleases are outside coding regions, although sequence upstream of *P. pallidum*'s first endonuclease, *en1*, has homology to a *D. fasciculatum* intron. All endonucleases contain a version of the LAGLIDADG motif, the identifying feature of a class of endonucleases; *D. discoideum's ai2a* and *D. citrinum's ai2b* contain the exact LAGLIDADG motif, and the rest usually contains functionally similar amino acids at locations that are not exact matches. The number of introns in *cox1/2* and *rnl* ranges from 0 to 4, whereas the total number of endonucleases ranges from 2 to 5. All endonucleases in *A. castellanii* are in the *rnl*, and *D. fasciculatum* retains an endonuclease, *en2*, in the rRNA, an ancestral condition (Angata et al. 1995). Endonucleases in *cox1/2* are also plesiomorphic as it has been noted that *Saccharomyces cerevisiae* also has endonucleases in *cox1* (Ogawa, Matsuo, et al. 1997).

To determine the evolutionary history of the endonucleases and introns, phylogenies were constructed using the amino acid sequences of all endonucleases (fig. 4) and the nucleotide sequence of all introns (fig. 5). For both phylogenies, terminal nodes are generally well supported whereas earlier nodes are less supported, so close relationships are clear but the relationship of these clades to each other is not. However, the distinction of *rnl* introns and *cox1/2* introns is clear with the exception of the second and third introns of *rnl* in *A. castellanii*.

The intron phylogeny is consistent with the physical location of the introns (fig. 3) for the closely related species *D. discoideum* and *D. citrinum*. The closely related first and second introns of *cox1/2* and the only intron of *rnl* are at homologous positions, respectively, in the 2 species. Similarly, in *cox1/2*, the closely related third intron of *D. discoideum* and the fourth intron of *D. citrinum* are in homologous positions. Conversely, the fourth intron of *D. discoideum* and the third intron of *D. citrinum* that have no close relation in the other species are in physical locations where no intron exists in the other species. In *D. fasciculatum* and *A. castellanii*, the physical location of the introns does not exactly match the intron phylogeny in

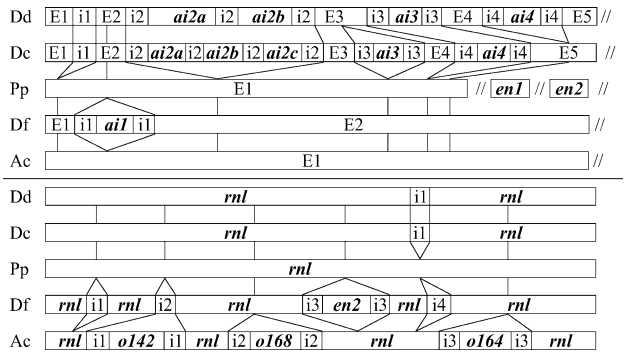

FIG. 3.—The figure is not to scale. Lines between species indicate homologous positions of beginning and end points of introns in the genes (e.g., the first introns in *cox1/2* in Dd and Dc are in homologous positions, but their location is within the first exon of Pp, Df, and Ac). Df's fourth rnl intron does not exactly match the position of the rnl intron in Dd and Dc but is 7 nt away. E(1–5) *cox1/2* exon number; *en1* and *en2*, endonucleases in *Polysphondylium pallidum* and *Dictyostelium fasciculatum* not located in *cox1/2* introns are named based on their order in the genome; o142, o168, and o164, endonucleases in *Acanthamoeba castellanii* (ORF142, ORF168, and ORF164); rnl, large ribosomal subunit RNA.

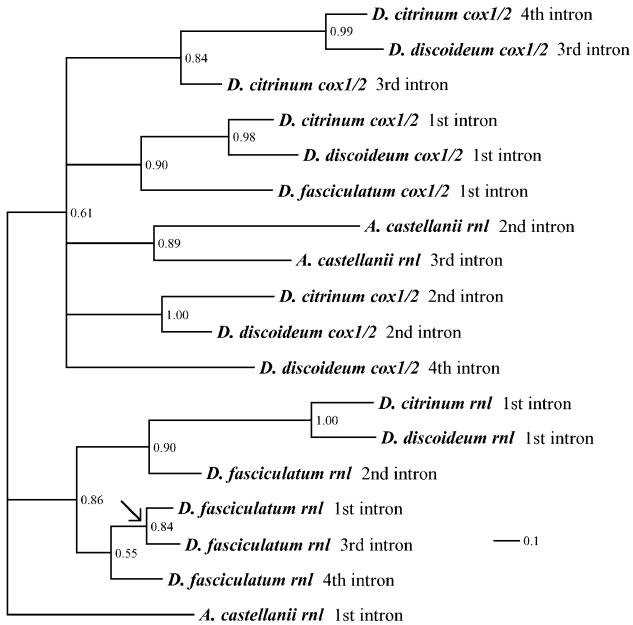

FIG. 5.—Bayesian phylogeny of introns. The arrow indicates a likely recent duplication event. The first intron from *Acanthamoeba castellanii's rnl* gene was used as the outgroup based on preliminary analysis. Numbers at branches indicate Bayesian inference posterior probabilities. Branch lengths correspond to substitutions per site given by the scale bar.

agreement with the increased evolutionary distance between these species and *D. discoideum* and *D. citrinum*. The only *cox1/2* intron in *D. fasciculatum* is related to the first introns in *D. discoideum* and *D. citrinum*, but *D. fasciculatum*'s intron is at a position homologous to a point in the second exon of the other species. A closer correlation is seen in *D. fasciculatum*'s fourth *rnl* intron that is only 7 nt away from the position of the only *rnl* intron in *D. discoideum* and *D. citrinum*, but the *D. fasciculatum* intron is only distantly related to the other introns. There is one conserved position between *D. fasciculatum* and *A. castellanii*; in *rnl*, the position of the first intron in *A. castellanii* is homologous to the position of the second intron in *D. fasciculatum*.

There is evidence that both some endonucleases and introns have duplicated recently. It is clear that in the ancestor of *D. discoideum* and *D. citrinum*, an endonuclease was duplicated in the second exon, producing *ai2a*/*ai2b* (*D. discoideum*/*D. citrinum*) and *ai2b*/*ai2c* (fig. 4, arrow). Similarly, it is likely that either the first or third intron of *rnl* in *D. fasciculatum* is a duplicate of the other (fig. 5, arrow).

The *D. fasciculatum* and *P. pallidum* endonucleases and the *D. fasciculatum* introns are differentiated from the endonucleases and introns of *D. discoideum* and *D. citrinum* in several ways congruent with the phylogenetic placement of the 2 pairs of species as distant from each other. The endonucleases of *D. discoideum* and *D. citrinum* are all located in *cox1/2* introns, whereas only 1 of the 4 endonucleases in *D. fasciculatum* and *P. pallidum* is located in *cox1/2*. This one gene, *ai1* in *D. fasciculatum*, is located in a *cox1/2* intron, and its closest phylogenic relatives are other *cox1/2* intronic endonucleases, but *D. fasciculatum's ai1* is distant from other endonucleases in its group. The *D. fasciculatum cox1/2* intron is orthologous to the first introns in *D. discoideum* and *D. citrinum* but is distant from them. The other 3 endonucleases of *D. fasciculatum* and *P. pallidum* also have long-branch lengths separating them from each other and from other endonucleases.

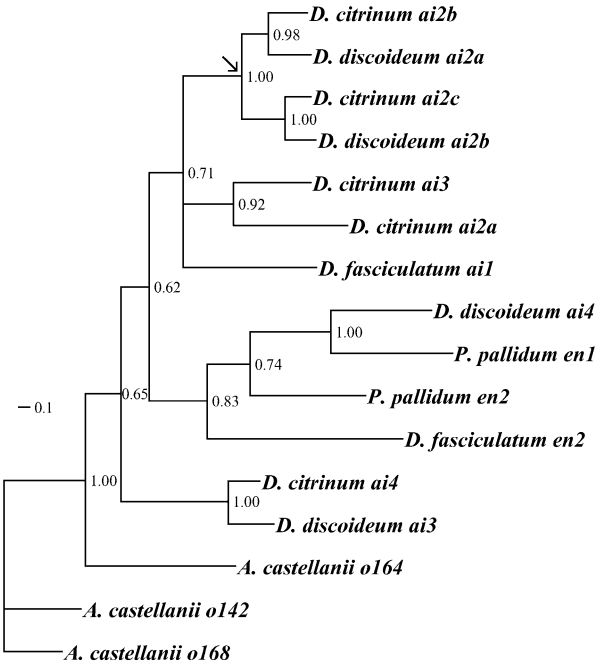

FIG. 4.—Bayesian phylogeny of endonucleases. The arrow indicates a likely recent duplication event in the Dd/Dc lineage. ORF142 from *Acanthamoeba castellanii* was used as the outgroup. Numbers at branches indicate Bayesian inference posterior probabilities. Branch lengths correspond to the substitutions per site given by the scale bar.

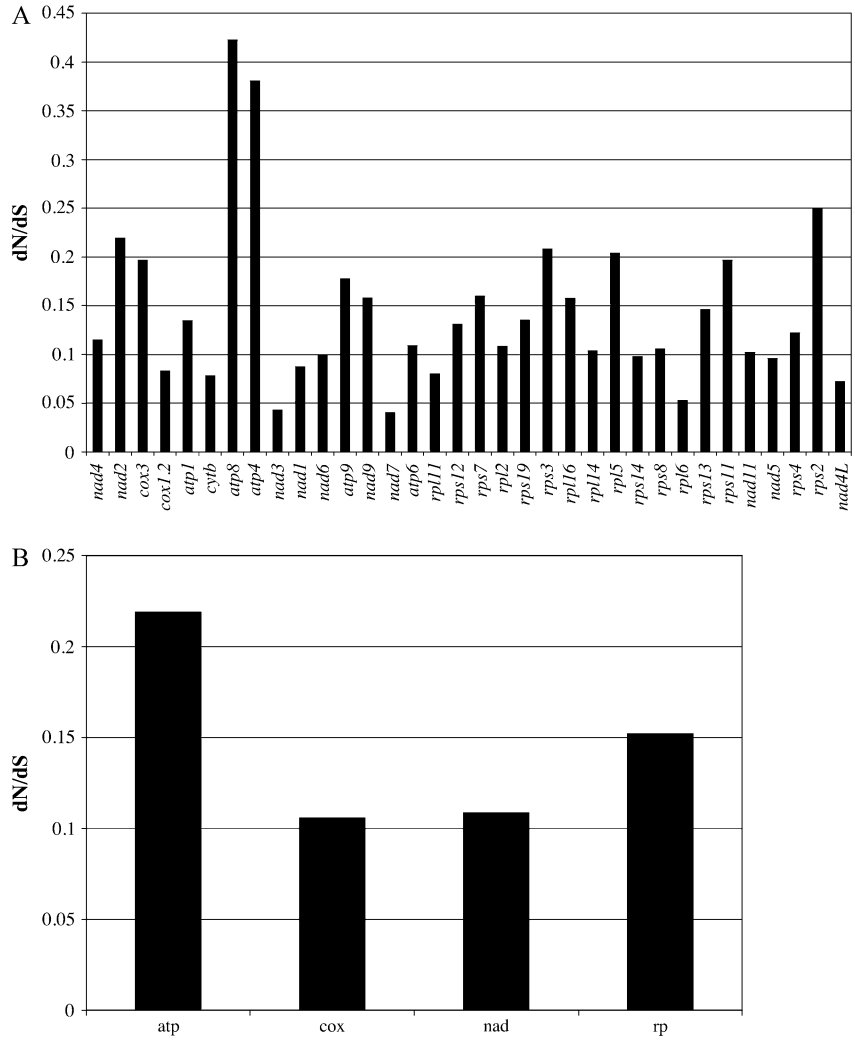

FIG. 6.—Protein evolution. (*A*) *d*N/*d*S ratios for the 34 known protein-coding genes shared between the 5 species. Significant differences between genes are listed in supplementary table 2 (Supplementary Material online). (*B*) *d*N/*d*S ratios for 4 categories of protein-coding genes. atp, all ATP synthase complex genes; cox, cytochrome oxidase 1/2 and 3 and cytochrome b; nad, all NADH dehydrogenase genes; and rp, all ribosomal proteins from the small and large subunit. All categories are significantly different from all others.

Protein Evolution

To determine the rates of change in proteins, we calculated the *d*N/*d*S ratios for all protein-coding genes shared between the 5 species. The values range more than 10-fold from a high of 0.423 for *atp8* to a low of 0.040 for *nad7* (fig. 6*A*), and significant comparisons between genes are given in supplementary table 2 (Supplementary Material online). To determine if certain classes of genes had different rates of change, we additionally calculated *d*N/*d*S ratios for 4 categories of protein-coding genes 1) all adenosine triphosphate (ATP) synthase complex genes, 2) cytochrome oxidase 1/2 and 3 and cytochrome b, 3) all nicotinamide adenine dinucleotide (NADH) dehydrogenase genes, and 4) all ribosomal proteins from the small and large subunits (fig. 6*B*). Cytochrome oxidase and NADH dehydrogenase genes had the lowest ratio followed by ribosomal proteins and the ATP synthase complex genes that had the highest ratio.

Genetic Code

We examined codon usage in the common protein-coding genes between the species to determine the mitochondrial genetic codes. Similar to *D. discoideum*, *D. citrinum*, *P. pallidum*, and *D. fasciculatum* use the universal code. There are no indications that any codons encoding amino acids are switched between codons. Additionally, *D. citrinum* and *P. pallidum* use all 3 universal stop codons, similarly to *D. discoideum*, although TGA is only used once and twice, respectively (table 1), consistent with the universal code. Interestingly, in the total of 4 TGA stop codons found in *D. discoideum*, *D. citrinum*, and *P. pallidum*, the TAA codon appears in frame as the first or second codon after the TGA.

In contrast, *D. fasciculatum* does not use TGA as a stop codon, only using the TAA stop codon in the 38 protein-coding genes shared between the 4 species, additionally using the TAG codon in both endonucleases and using TAA

and TAG in the unknown ORFs. Furthermore, the TGA codon does not appear in protein-coding sequences, so the codon has not switched to coding tryptophan or another amino acid. To determine if *D. fasciculatum* still has the ability to recognize the TGA codons, we searched in the nuclear genome for the mitochondrial-specific termination factors that bind to stop codons. The nuclear genome of *D. discoideum* has both *RF1* that recognizes TAA and TAG and *RF2* that recognizes TAA and TGA. It has been noted that bacterial release factors are distinct from eukaryotic release factors (Moreira et al. 2002), and bacterial release factors closely resemble mitochondrial release factors, so it is extremely unlikely that *RF1* and *RF2* from social amoebae can function in the cytosolic ribosomes. We identified both *RF1* and *RF2* in the nuclear genome of *D. fasciculatum*, and both are transcribed because we were able to amplify both genes from cDNA.

In both *D. discoideum* and *D. fasciculatum*, both mitochondrial release factors retain the signature tripeptide motifs that recognize the stop codons (Ito et al. 2000). *RF1* has the Pro-x-Thr (PxT) motif that recognizes the TAA and TAG stop codons, and *RF2* has the Ser-Pro-Phe (SPF) amino acid motif that recognizes the TAA and TGA stop codons.

Discussion

Genome Summary

The newly sequenced mitochondrial genomes of *D. citrinum*, *P. pallidum*, and *D. fasciculatum* are similar to *D. discoideum* and *A. castellanii* in size, AT%, and gene content, consistent with differences within protist mitochondria (Gray et al. 1998). However, the gene order within the social amoebae is different from that of *A. castellanii*, indicating that the gene order may change quicker than genome size or AT% over longer evolutionary periods. The quantity and location of tRNAs vary within the social amoebae giving more evidence that tRNAs can rearrange much more easily than rRNA or protein-coding genes at least within certain taxa. It has been shown that tRNAs can be more mobile than rRNA and protein-coding genes in taxa such as insects (Shao et al. 2001) and marsupials (Paabo et al. 1991) but not necessarily because other lineages such as fish are invariant.

The conservation of a gene across species suggests its necessity and conservation of function. The 3 unknown genes conserved in the social amoebae mitochondria (homologs to DidioMp05, DidioMp21, and DidioMp36 in *D. discoideum*) have little homology to anything else, but their conservation suggests that they have a conserved required function in mitochondria. Determination of the function may require further experimental studies.

Mitochondrial Phylogeny

It has been noted that concatenation of genes can be used to make phylogenies more robust and accurate (Baldauf et al. 2000), and this technique has been used in diverse taxa (Nakao et al. 2007; Signorovitch et al. 2007; Terasawa et al. 2007). The use of all protein-coding genes and the 2 rRNAs in common between the 5 amoebae species does produce a well-supported tree (fig. 1). The Schaap et al. (2006) study found 2 possible locations for the root of the social amoebae, and the data here support one of those locations. The social amoebae have a large evolutionary distance between them and any outgroup, and the corresponding long-branch length makes the correct root location difficult to determine definitively. Even without a definitive root, the tree is useful for placing mitochondrial genomic changes such as gene rearrangements.

Segmental Rearrangement

Several models have been proposed for how mitochondrial genomes rearrange in animals and whereas slipped-strand mispairing producing tandem duplication was often seen as the predominant mechanism, other mechanisms have more recently been proposed (Dowton and Campbell 2001). In one proposed mechanism, a section of the mitochondrial genome is removed, becomes circular, and then reinserts at a new location, and this or other types of intramolecular recombination has been potentially observed in diverse animal taxa including Plethodontid salamanders (Mueller and Boore 2005), humans (Holt et al. 1997; Kajander et al. 2000), mussels (Ladoukakis and Zouros 2001), and others (Tsaousis et al. 2005).

The rearrangement seen here cannot be explained by tandem duplication because the stretches of genes in segments A and B (fig. 2) are too long for them to likely maintain the exact gene order within the segments, and no pseudogenes resulting from unmaintained duplicates are seen within the rearrangement. Instead, a recombination event such as seen in animals seems required for the rearrangement. Such an event occurs rarely as only one rearrangement has occurred in the entire social amoebae lineage. The fact that all genes are encoded on one strand may hinder rearrangements because inversions that reverse the orientation of genes will cause genes to not be transcribed, and this in turn would be lethal.

The rearrangement may also help with phylogeny reconstruction. Because the rearrangement occurred within group 4, the presence or absence of the rearrangement can be used together with sequences for phylogeny reconstruction within group 4.

Introns and Endonucleases

The introns of *cox1/2* and *rnl* and endonucleases contained within are variable in their number and positions in the analyzed species. The extent of the variability was hinted by the previous finding that *D. discoideum* (AC4), *D. mucoroides* (Dm7), and *P. pallidum* (PN600) do not have the same second intron in *cox1/2* as *D. discoideum* (AX-3 and NC4) (Ogawa, Naito, et al. 1997). The variability is consistent with the endonucleases and introns likely ability to assist their own movement. A functional analysis of all the social amoebae endonucleases has not yet been performed, but analysis of one endonuclease indicates that

although the endonucleases were identified based on homology only, they have retained their function; *D. discoideum's ai2a* protein has endonuclease activity in vivo and was shown to cut one strand near the exon–intron boundary (Ogawa, Naito, et al. 1997).

A previous study investigated whether the endonucleases move independently of the introns and whether the introns move laterally between species and found that endonucleases do move independently albeit rarely but that introns do not move laterally (i.e., between species) at least in the cases they studied (Haugen et al. 2004). A comparison of the 2 phylogenies (figs. 4 and 5) generally supports the previous finding. In the social amoebae, the introns and endonucleases generally stay together as evidenced by the closely related endonucleases also being in closely related introns. The exceptions are the *P. pallidum* endonucleases that have moved out of introns and *D. citrinum*'s *ai2a* and *ai3* endonucleases, where one is a duplicate of the other and one moved into a new intron. Also, there is no evidence of lateral transfer of introns or endonucleases.

There is evidence for frequent duplications or extinctions of both introns and endonucleases. Without the closest relative to the last common ancestor available, it is difficult to determine the ancestral number of introns and endonucleases. It is therefore difficult to know whether the differing numbers of these elements in the social amoebae are due to extinctions or duplications, but they clearly have a more dynamic presence than the protein-coding genes in the mitochondria. There are likely some constraints on their duplication, foremost perhaps the cost of synthesizing extra DNA and RNA sequence, however, their overall effect on mitochondrial fitness remains unknown.

Selection in the Mitochondrial Genome?

It has been debated whether mitochondrial genes change mostly through neutral drift or selection (Meiklejohn et al. 2007). It has been shown that mitochondrial variation can have an effect on fitness in animals (reviewed by Ballard and Rand [2005]), although there remains a lack of knowledge about mitochondrial natural history and the basis of selection (Ballard and Whitlock 2004). To identify targets of selection, it may be helpful to know the genes that are evolving more quickly. We have determined which genes and gene classes have higher and lower *d*N/*d*S ratios. The gene *atp8* has the highest ratio in agreement with a study of mammal mitochondria (Pesole et al. 1999) that found *atp8* had the highest nonsynonomous substitution rate. The genes *cox1/2* and *cytb* that are often used in phylogenetic analyses have among the lowest rates. Selection or reduced constraint could both be responsible for elevated *d*N/*d*S ratios, and further study of within-species variation in addition to fitness assays is required for a more complete understanding of selection in social amoebae mitochondria.

We for the first time can report that the mitochondrial ribosomal protein genes have an intermediate rate of change in between the ATP synthase complex genes on one side and the cytochrome oxidase and NADH dehydrogenase genes on the other side. The lack of analysis of mitochondrial ribosomal protein genes is undoubtedly due to the fact that metazoan and fungal mitochondrial genomes with the exception of 1 fungal species do not carry these genes (Gray et al. 1998). A priori, the conservation of translational function in ribosomes would seem to severely constrain change in ribosomal genes although this may only apply to certain critical domains. However, resistance against streptomycin in bacteria demonstrates that changes in a ribosomal protein can occur and can be adaptive (Finken et al. 1993) although whether mitochondria are commonly exposed to environmental stresses such as antibiotics from the social amoebae's food source bacteria is unknown.

Genetic Code Alterations

Nonuniversal genetic codes in mitochondria are common, and only a small group of organisms including green plants and the social amoebae have retained the universal code. Two main theories have been proposed for how genetic codes could switch (Santos et al. 2004). In the codon capture theory, 1) a codon disappears, 2) the tRNA or release factors recognizing the codon becomes nonfunctional without purifying selection, 3) a mutation in another tRNA's anticodon allows a new amino acid to be coded by the extinct codon, and finally (4) the codon gradually reappears encoding a new amino acid (Osawa and Jukes 1989). Alternatively, in the ambiguous intermediate theory, the codon does not completely disappear before the mutation in the tRNA so that during the ambiguous period, 2 different amino acids (or a stop to translation) can potentially be inserted for the codon (Schultz and Yarus 1994). The social amoebae potentially occupy intermediate points in both of these switching code theories. The 4 species investigated here are all potentially close to switching the TGA codon from a stop signal to an amino acid. *Dictyostelium discoideum*, *D. citrinum*, and *P. pallidum* are at a point where the TGA codon has nearly disappeared so the ambiguous intermediate theory could apply, whereas *D. fasciculatum* has lost the TGA codon entirely so the codon capture theory could apply.

In the *D. discoideum*, *D. citrinum*, and *P. pallidum* mitochondrial genomes, there is a suggestion how the ambiguous intermediate theory could apply. In these genomes when the TGA stop codon is used, it is quickly followed by the TAA stop codon, hinting that the TAA could be used as an alternative stop codon, if a tRNA existed that detects TGA. In this scenario, a suppressor tRNA places an amino acid at the TGA codon, and the ribosome continues translating until it reaches the TAA stop codon. The suppressor tRNA phenomena has been found in other systems (Feng et al. 1990; Chittum et al. 1998), and in the most relevant example, it was demonstrated in a bacterium that the normal tryptophan tRNA could weakly suppress the TGA stop codon and incorporate tryptophan instead (Lovett et al. 1991). This is logical because although normal wobble rules do not allow UGA (stop codon) to pair with ACC (tryptophan anticodon), these 2 codons pair at least at the first 2 bases. To explain the remaining problem with the third base, it has been demonstrated that modifications outside the anticodon in tRNAs can increase the chance of this nonstandard

pairing (Beier and Grimm 2001). It has been suggested that suppressor tRNAs could be involved in genetic code changes (Massey and Garey 2007), so *D. discoideum*, *D. citrinum*, and *P. pallidum* could be in the first step toward a genetic code change. To complete the change, however, a tRNA that strongly recognizes TGA is required plus the loss of the release factor *RF2*.

Although the both genetic code change theories are plausible to explain the mechanism of a genetic code alteration, how selection operates for or against the change is less certain. The change could be neutral especially when a codon could disappear due to AT bias, followed by a mutation in a randomly chosen tRNA. However, the frequent reassignment of TGA to tryptophan suggests that a tRNA with weak binding to the codon, as discussed above for suppressor tRNAs, may be the most likely replacement. Conversely, it has also been suggested that amino acid-to-amino acid changes might be selected positively (Swire et al. 2005).

There are also reasons for the genetic code not to change from the universal code due to its believed optimal state. Maintaining 3 stop codons, as opposed to fewer stop codons, makes it more likely that a stop codon is encountered when a ribosome makes frameshift errors, thereby saving resources and energy (Itzkovitz and Alon 2007). Our results suggest that there may be another constraint on changing the genetic code: selection may maintain both release factors, *RF1* and *RF2*, so that the stop codons cannot be reassigned. *Dictyostelium fasciculatum* has no TGA codons so *RF2* is not needed; yet, both release factors are maintained and transcribed. One potential explanation is that *RF2* is better than *RF1* at recognizing the TAA stop codon (Osawa et al. 1992), the most common stop codon in the social amoebae mitochondria. The ability to recognize TAA better could prevent *RF2* from becoming redundant and nonfunctional.

This constraint is not impossible to overcome, obviously as many mitochondria have TGA encode tryptophan. To our knowledge, organisms that have a code where TGA encodes tryptophan in mitochondria have lost *RF2* while maintaining *RF1*. For example, *S. cerevisiae* lost *RF2* (Askarian-Amiri et al. 2000), a biochemical study could not find it in rat mitochondria (Lee et al. 1987) and a pBlast search of the human genome could not identify a *RF2* gene (personal observation, AJH). At least one other mitochondrial genome, the green alga *Mesostigma viride*, has lost the TGA stop codon without switching it to code for an amino acid (Turmel et al. 2002), but it is not known whether the alga has retained *RF2* in its nuclear genome. It has been postulated that a reduced proteomic constraint may be a factor in permitting genetic code changes (Massey and Garey 2007), and consistent with this idea, the mitochondrial genomes and associated proteomes of yeast, rats, and humans that have all switched TGA to tryptophan are all smaller than those of social amoebae. Furthermore, the theory previously mentioned, that having more stop codons saves energy by increasing the likelihood of termination after ribosome errors, will have less effect on a smaller proteome. In contrast, the large mitochondrial genomes of land plants such as *Arabidopsis thaliana*, *Zea mays* (strain NB), and *Oryza sativa* (indica cultivar) have retained the universal code. However, size is likely not the sole determining factor as *A. castellanii* only has a marginally smaller genome than the social amoebae and has an altered code.

Supplementary Material

Supplementary tables 1 and 2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Silke Förste, Nadine Zeise, Susanne Fabisch, and Christin Heinrich for technical assistance and Karol Szafranski for bioinformatic assistance and 2 anonymous reviewers for helpful comments. We are grateful to Pauline Schaap and Thomas Winkler for supplying all strains and DNA of *P. pallidum* (PN500). A.J.H. is supported by a grant from the Deutsche Forschungsgemeinschaft (G1235/1-1).

Literature Cited

Angata K, Ogawa S, Yanagisawa K, Tanaka Y. 1995. A group-I intron in the mitochondrial large-subunit ribosomal RNA-encoding gene of *Dictyostelium discoideum*—same site localization in alga and in vitro self-splicing. Gene. 153: 49–55.

Askarian-Amiri ME, Pel HJ, Guevremont D, McCaughan KK, Poole ES, Sumpter VG, Tate WP. 2000. Functional characterization of yeast mitochondrial release factor 1. J Biol Chem. 275:17241–17248.

Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF. 2000. A kingdom-level phylogeny of eukaryotes based on combined protein data. Science. 290:972–977.

Ballard JWO, Rand DM. 2005. The population biology of mitochondrial DNA and its phylogenetic implications. Ann Rev Ecol Evol Syst. 36:621–642.

Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. Mol Ecol. 13:729–744.

Beier H, Grimm M. 2001. Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. Nucleic Acids Res. 29:4767–4782.

Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 29:2607–2618.

Burger G, Forget L, Zhu Y, Gray MW, Lang BF. 2003. Unique mitochondrial genome architecture in unicellular relatives of animals. Proc Natl Acad Sci USA. 100:892–897.

Burger G, Gray MW, Lang BF. 2003. Mitochondrial genomes: anything goes. Trends Genet. 19:709–716.

Burger G, Plante I, Lonergan KM, Gray MW. 1995. The mitochondrial-DNA of the amoeboid protozoan, Acanthamoeba castellanii—complete sequence, gene content and genome organization. J Mol Biol. 245:522–537.

Chevalier BS, Stoddard BL. 2001. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. Nucleic Acids Res. 29:3757–3774.

Chittum HS, Lane WS, Carlson BA, Roller PP, Lung FD, Lee BJ, Hatfield DL. 1998. Rabbit beta-globin is extended beyond its

UGA stop codon by multiple suppressions and translational reading gaps. Biochemistry. 37:10866–10870.
Cole RA, Williams KL. 1994. The Dictyostelium discoideum mitochondrial genome—a primordial system using the universal code and encoding hydrophilic proteins atypical of metazoan mitochondrial DNA. J Mol Evol. 39:579–588.
Dowton M, Campbell NJH. 2001. Intramitochondrial recombination—is it why some mitochondrial genes sleep around? Trends Ecol Evol. 16:269–271.
Eichinger L, Pachebat JA, Glockner G, et al. (97 co-authors). 2005. The genome of the social amoeba Dictyostelium discoideum. Nature. 435:43–57.
Feng YX, Copeland TD, Oroszlan S, Rein A, Levin JG. 1990. Identification of amino-acids inserted during suppression of UAA and UGA termination codons at the gag-pol junction of Moloney murine leukemia virus. Proc Natl Acad Sci USA. 87:8860–8863.
Finken M, Kirschner P, Meier A, Wrede A, Bottger EC. 1993. Molecular basis of streptomycin resistance in Mycobacterium tuberculosis—alterations of the ribosomal protein S12 gene and point mutations within a functional 16s ribosomal RNA pseudoknot. Mol Microbiol. 9:1239–1246.
Gray MW, Lang BF, Burger G. 2004. Mitochondria of protists. Annu Rev Genet. 38:477–524.
Gray MW, Lang BF, Cedergren R, et al. (15 co-authors). 1998. Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res. 26:865–878.
Haugen P, Reeb V, Lutzoni F, Bhattacharya D. 2004. The evolution of homing endonuclease genes and group I introns in nuclear rDNA. Mol Biol Evol. 21:129–140.
Holt IJ, Dunbar DR, Jacobs HT. 1997. Behaviour of a population of partially duplicated mitochondrial DNA molecules in cell culture: segregation, maintenance and recombination dependent upon nuclear background. Hum Mol Genet. 6: 1251–1260.
Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics. 17:754–755.
Ito K, Uno M, Nakamura Y. 2000. A tripeptide 'anticodon' deciphers stop codons in messenger RNA. Nature. 403:680–684.
Itzkovitz S, Alon U. 2007. The genetic code is nearly optimal for allowing additional information within protein-coding sequences. Genome Res. 17:405–412.
Kajander OA, Rovio AT, Majamaa K, Poulton J, Spelbrink JN, Holt IJ, Karhunen PJ, Jacobs HT. 2000. Human mtDNA sublimons resemble rearranged mitochondrial genomes found in pathological states. Hum Mol Genet. 9:2821–2835.
Knight RD, Freeland SJ, Landweber LF. 2001. Rewiring the keyboard evolvability of the genetic code. Nat Rev Genet. 2: 49–58.
Ladoukakis ED, Zouros E. 2001. Direct evidence for homologous recombination in mussel (Mytilus galloprovincialis) mitochondrial DNA. Mol Biol Evol. 18:1168–1175.
Lee CC, Timms KM, Trotman CNA, Tate WP. 1987. Isolation of a rat mitochondrial release factor. Accommodation of the changed genetic code for termination. J Biol Chem. 262: 3548–3552.
Lovett PS, Ambulos NP, Mulbry W, Noguchi N, Rogers EJ. 1991. UGA can be decoded as tryptophan at low efficiency in *Bacillus subtilis*. J Bacteriol. 173:1810–1812.
Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:955–964.
Lynch M, Koskella B, Schaack S. 2006. Mutation pressure and the evolution of organelle genomic architecture. Science. 311: 1727–1730.
Massey SE, Garey JR. 2007. A comparative genomics analysis of codon reassignments reveals a link with mitochondrial proteome size and a mechanism of genetic code change via suppressor tRNAs. J Mol Evol. 64:399–410.
Meiklejohn CD, Montooth KL, Rand DM. 2007. Positive and negative selection on the mitochondrial genome. Trends Genet. 23:259–263.
Moreira D, Kervestin S, Jean-Jean O, Philippe H. 2002. Evolution of eukaryotic translation elongation and termination factors: variations of evolutionary rate and genetic code deviations. Mol Biol Evol. 19:189–200.
Mueller RL, Boore JL. 2005. Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. Mol Biol Evol. 22:2104–2112.
Nakao M, McManus DP, Schantz PM, Craig PS, Ito A. 2007. A molecular phylogeny of the genus Echinococcus inferred from complete mitochondrial genomes. Parasitology. 134:713–722.
Ogawa S, Matsuo K, Angata K, Yanagisawa K, Tanaka Y. 1997. Group-I introns in the cytochrome c oxidase genes of Dictyostelium discoideum: two related ORFs in one loop of a group-I intron, a cox1/2 hybrid gene and an unusually large cox3 gene. Curr Genet. 31:80–88.
Ogawa S, Naito K, Angata K, Morio T, Urushihara H, Tanaka Y. 1997. A site-specific DNA endonuclease specified by one of two ORFs encoded by a group I intron in Dictyostelium discoideum mitochondrial DNA. Gene. 191:115–121.
Ogawa S, Yoshino R, Angata K, et al. (11 co-authors). 2000. The mitochondrial DNA of Dictyostelium discoideum: complete sequence, gene content and genome organization. Mol Gen Genet. 263:514–519.
Osawa S, Jukes TH. 1989. Codon reassignment (codon capture) in evolution. J Mol Evol. 28:271–278.
Osawa S, Jukes TH, Watanabe K, Muto A. 1992. Recent evidence for evolution of the genetic code. Microbiol Rev. 56:229–264.
Paabo S, Thomas WK, Whitfield KM, Kumazawa Y, Wilson AC. 1991. Rearrangements of mitochondrial transfer RNA genes in marsupials. J Mol Evol. 33:426–430.
Pesole G, Gissi C, De Chirico A, Saccone C. 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. J Mol Evol. 48:427–434.
Pi M, Morio T, Urushihara H, Tanaka Y. 1998. Characterization of a novel small RNA encoded by Dictyostelium discoideum mitochondrial DNA. Mol Gen Genet. 257:124–131.
Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 21:676–679.
Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16:276–277.
Santos MAS, Moura G, Massey SE, Tuite MF. 2004. Driving change: the evolution of alternative genetic codes. Trends Genet. 20:95–102.
Schaap P, Winckler T, Nelson M, et al. (12 co-authors). 2006. Molecular phylogeny and evolution of morphology in the social amoebas. Science. 314:661–663.
Schultz DW, Yarus M. 1994. Transfer RNA mutation and the malleability of the genetic code. J Mol Biol. 235: 1377–1380.
Shao R, Campbell NJH, Barker SC. 2001. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, Heterodoxus macropus (Phthiraptera). Mol Biol Evol. 18:858–865.
Signorovitch AY, Buss LW, Dellaporta SL. 2007. Comparative genomics of large mitochondria in placozoans. PLoS Genet. 3(1):e13.

Swire J, Judson OP, Burt A. 2005. Mitochondrial genetic codes evolve to match amino acid requirements of proteins. J Mol Evol. 60:128–139.

Terasawa K, Odahara M, Kabeya Y, Kikugawa T, Sekine Y, Fujiwara M, Sato N. 2007. The mitochondrial genome of the moss Physcomitrella patens sheds new light on mitochondrial evolution in land plants. Mol Biol Evol. 24:699–709.

Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, Zouros E. 2005. Widespread recombination in published animal mtDNA sequences. Mol Biol Evol. 22:925–933.

Turmel M, Otis C, Lemieux C. 2002. The complete mitochondrial DNA sequence of *Mesostigma viride* identifies this green alga as the earliest green plant divergence and predicts a highly compact mitochondrial genome in the ancestor of all green plants. Mol Biol Evol. 19:24–38.

Geoffrey McFadden, Associate Editor

Accepted April 7, 2008