42

Organization of Genetic Material
3. Organellar Genomes
(Mitochondrion, Chloroplast and Apicoplast)

The genomes the three eukaryotic organelles including mitochondria, chloroplasts and apicoplast have been studied in great detail ever since their discovery in 1960s and thereafter. Starting in 1995, now in 2017, whole genomw sequences have become available for mitochondrial genomes of ~8000 animal/plant species, chloroplast genomes of >800 plant species and apicoplast genomes in >20 parasitic apicomplexan species (as in 2017). In this chapter, we discuss the organization of mitochondrial, chloroplast and apicoplast genomes.

SIZE, SHAPE AND NUMBER (PER CELL) OF ORGANELLAR GENOMES

Organellar genomes in mitochondria and chloroplasts vary in their size, shape and number per cell. Same level of variation has not been reported yet for apicoplasts, perhaps because large number of apicoplast genomes have not been examined so far.

Size and Organization of Genomes

The organeller genomes are relatively simple in organization and small in size (Table 42.1). Size of mitochondrial genome differs greatly among different organisms, ranging from 6 kb in *Plasmodium* (malarial parasite) to 2500 kb in melon (*Cucumis melo*). The variation in size of mitochondrial genomes is, however, unrelated to the variation in complexity of organisms (Fig. 42.1). For instance, most multicellular animals have small mitochondrial genomes with a compact genetic organization, the genes being close together with little space between them. This is exemplified by human mitochondrial genome (Fig. 42.2), which is only 16,569 bp (~16.6 kb) in length. In contrast, lower eukaryotes such as *S. cerevisiae* (Fig. 42.3), as well as flowering plants, have larger and less compact mitochondrial genomes, with a number of genes, which contain introns. In humans, the two strands of double stranded mtDNA are also distinguished as **heavy strand** (H-strand) and **light strand** (L-strand), since they differ in density, when measured on CsCl density gradient.

Table 42.1. **Sizes of mitochondrial and chloroplast genomes.**

Species	Genome size (kb)
Mitochondrial genomes	
Plasmodium falciparum	6
Chlamydomonas reinhardtii	16
Mus musculus	16
Homo sapiens	17
Metridium senile	17
Drosophila melanogaster	19
Chondrus crispus	26
Aspergillus nidulans	33
Reclinomonas americana	69
Saccharomyces cerevisiae	75
Suillus grisellus	121
Brassica oleracea	160
Arabidopsis thaliana	367
Zea mays	570
Cucumis melo	2500
Chloroplast genomes	
Pisum sativum	120
Marchantia polymorpha	121
Oryza sativa	136
Nicotiana tabacum	156
Chlamydomonas reinhardtii	204
Apicoplast genomes	
Leucocytozoon caulleryi	34.8
Plasmodium falciparum	34.7
P. chabuadi chabaudi	29.6
Toxoplasma gondii	35.0
Eimeria tenella	34.8
Theileria parva	39.6
Babesia bovis	33.4
Babesia microti	28.7
Cyclospora cayetanensis	33.2

In comparison of mitochondrial genomes, chloroplast and apicoplast genomes are less variable in size. The chloroplast genome ranges from 120 kb in pea to 204 kb in *Chlamydomonas* and the genomes of apicoplast range from ~27 kb in *Babesia microti* to ~39.6 kb in *Theileria parva* (Table 42.1). Most chloroplast genomes have a structure similar to that shown in Figure 42.4 for the rice chloroplast genome. Similarly, most apicoplast genomes have a structure shown in Figure 42.5 for *Plasmodium falciparum.* However, the size of different DNA molecules within a chloroplast may vary, so that in many eukaryotes, larger molecules coexist with smaller molecules that contain subcomponents of the larger multigenome molecules. The variation in size of

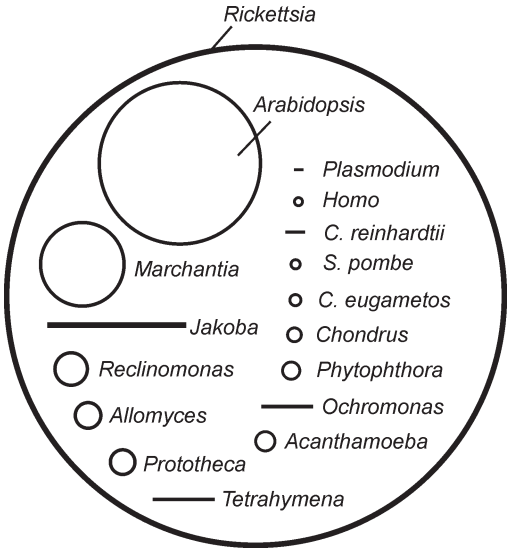

Fig. 42.1. Mitochondrial genomes, showing variation in size. For comparison, genome of mitochondria-like proteobacteria of the Rickettsia group (*Reclinomonas americana*) is also shown.

apicoplast genome within a species and within the same cell has not been reported, but may become known, when many more genomes are sequenced.

Shape of Genome (DNA Molecules)

The mtDNA and cpDNA were initially believed to be circular (like bacterial chromosomes) and any linear DNA molecules observed, were actually attributed to broken circles ('broken circle theory'). Both in bacteria and the cell organelles, the proposal of circular shape was actually also based on the circular nature of their genetic maps, without realizing that circular genetic maps could also result from linear head-to-tail concatamers. For almost 30 years (1970-2000), the 'broken circle theory' remained a strong barrier to the discovery of linear mtDNA/cpDNA. The apicoplast DNA is also circular in nature with some exceptions.

During 1990s and in the early years of the present century (2000-2004), it was demonstrated that a large fraction of *in vivo* mtDNA and cpDNA in many (and perhaps all) eukaryotes are

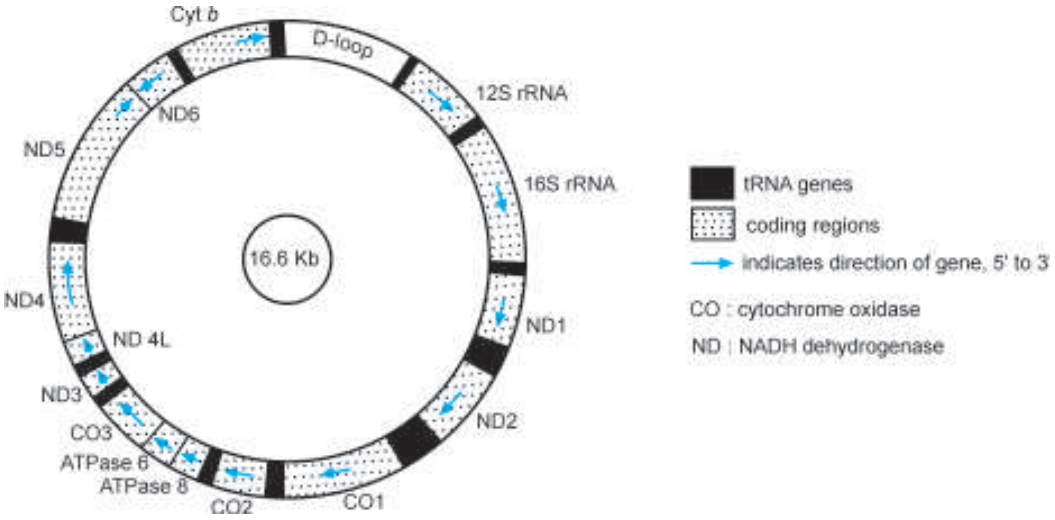

Fig. 42.2. The human mitochondrial genome showing (i) 2 rRNA genes, (ii) 13 protein coding regions and (iii) regions containing genes for 22 tRNAs (14 transcribed in the clockwise direction and 8 transcribed in the anticlockwise direction).

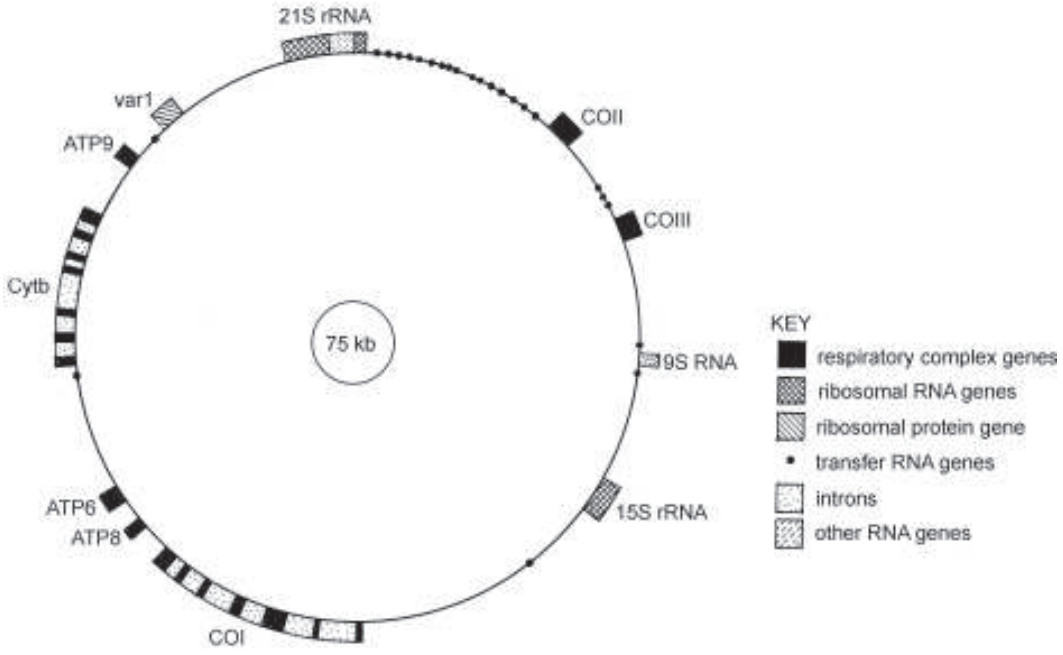

Fig. 42.3. The *Saccharomyces cervisiae* (budding yeast) mitochondrial genome. In the yeast genome, the genes are more spaced out than in the human mitochondrial genome and some of the genes have introns. This type of organization is typical of many lower eukaryotes and plants. (Abbreviations : ATP6, ATP8, ATP9, genes for ATPase subunits, 6, 8 and 9, respectively; COI, COII, COIII, genes for cytochrome C oxidase subunits I, II and III respectively; Cyt b, gene for apocytochrome b; var I, gene for a ribosome-associated protein. Ribosomal RNA and transfer RNA are two types of non-coding RNA. The 9S RNA gene specifies the RNA component of the enzyme ribonuclease P.

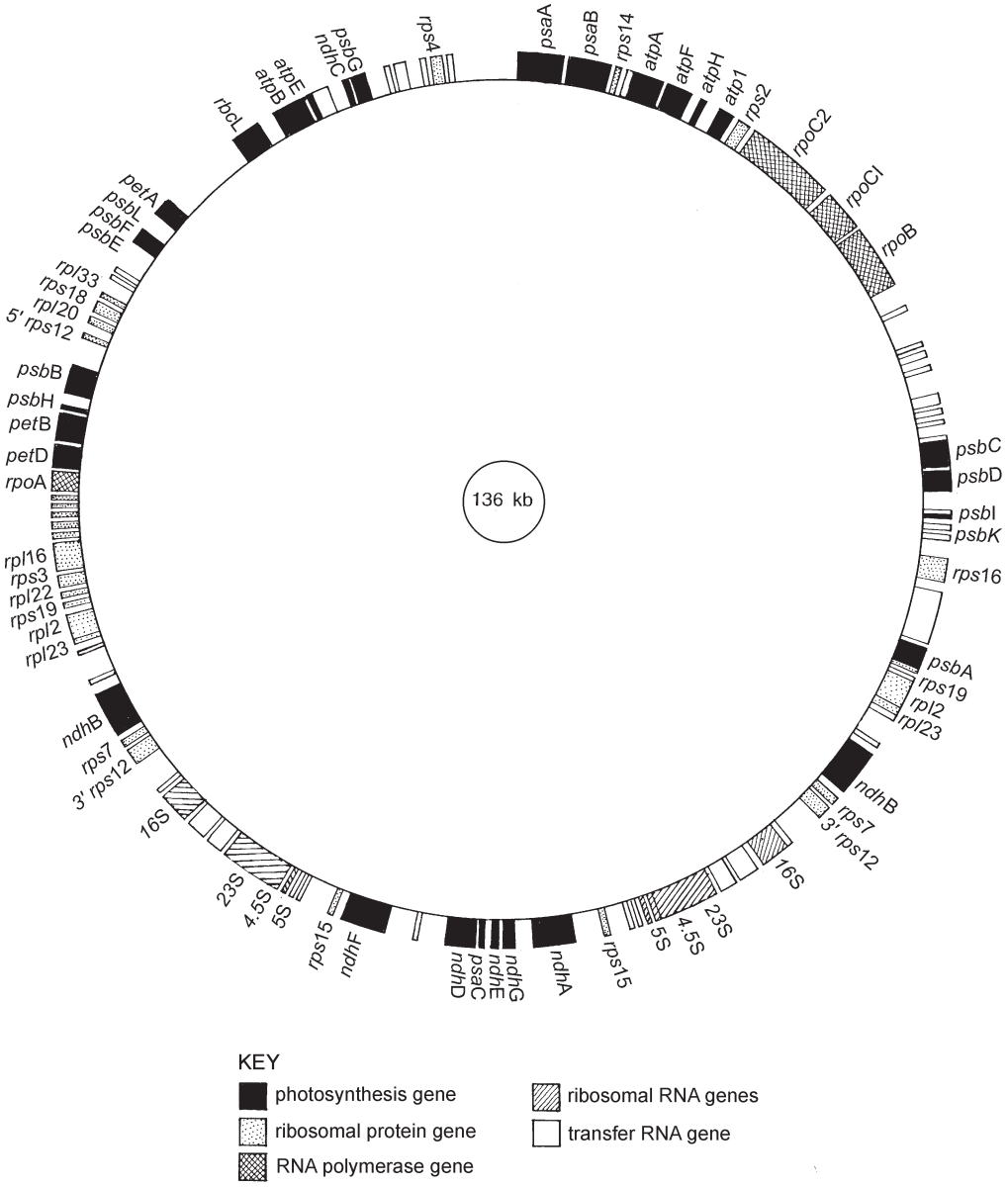

Fig. 42.4. The rice chloroplast genome. Only genes with known functions are shown. A number of genes contain introns which are not indicated on the map. These discontinuous genes include several of those for tRNAs; this is why the tRNA genes are of different lengths although the tRNAs that they specify are all of similar size.

essentially linear in shape, with only a small fraction occurring as circular DNA. Linear mtDNA has been reported in some algae (*e.g.*, *Chlamydomonas*), several fungi (including yeasts), plants (*e.g.*, *Chenopodium*), some protozoa (*e.g.*, *Paramecium*), and some apicomplexans (*e.g.*, *Plasmodium*). Similarly, linear cpDNA was reported in maize. For some apicoplast genomes also, linear DNA has been reported..

Table 42.2. **Relative amounts of organellar DNA in some cells and tissues.**

Organelle and organism	Tissue or cell type	DNA molecules (genomes) per organelle	Organelle per cell	Organellar DNA (as % total cell DNA)
I Mitochondrial DNA				
Rat	Liver	5-10	1000	1
Yeast	Vegetative	2-50	1-50	15
Frog	Egg	5-10	10^7	99
II Chloroplast DNA				
Chlamydomonas	Vegetative	80	1	7
Maize	Leaves	20-40	20-40	15

Number of Genomes per Organelle

The number of genomes in an individual organelle also varies. A large number of mitochondria or chloroplasts are also available within an individual cell, thus increasing many-fold the number of organellar genomes within a cell (Table 42.2). For instance, each human cell contains 800 mitochondria, and each mitochondrion contains as many as 10 identical genomes, so that each cell contains about 8000 mitochondrial genomes. In yeast, although the number of genomes per mitochondrion is about 100, but there are fewer mitochondria per cell, thus making the number

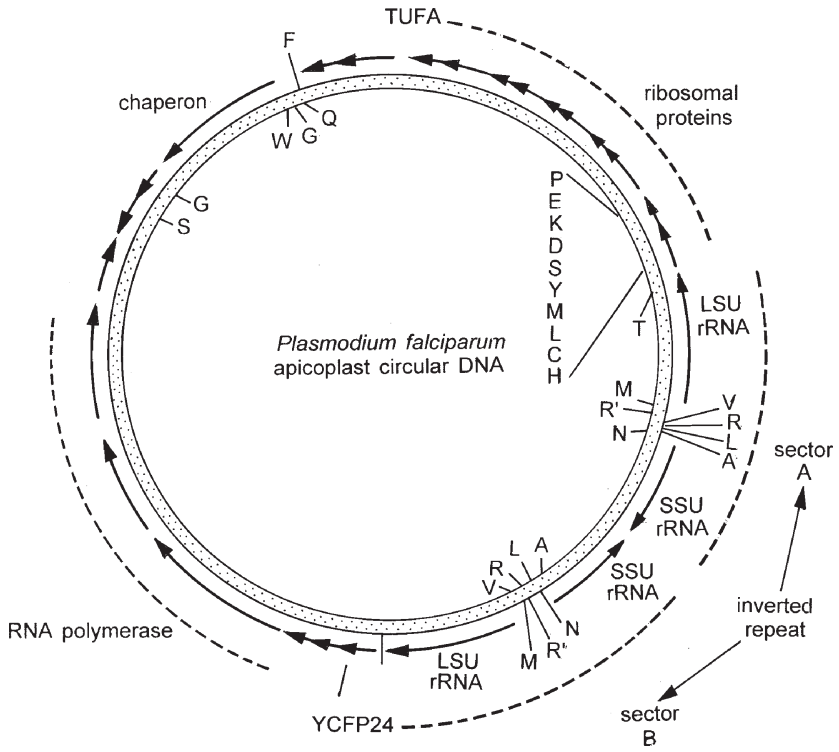

Fig. 42.5. Apicoplast genome from *Plasmodium falciparum*; genes are indicated as arrows indicating the direction of transcription; broken lines indicate functional regions; ribosomal RNA genes are labeled as LSU (large subunit) and SSU (small subunit); tRNA are shown by single letters indicated their corresponding amino acids.

of genomes per cell closer to that in humans (~5000). In case of plants, the number of chloroplast genomes per cell is approximately 5000, but that in photosynthetic microorganisms like *Chlamydomonas*, it is only 1000 per cell. Apicoplasts appear to contain, each a single genome, although future work may show presence of more than one genome per apicoplast, and more than one apicoplasts per cell.

REPLICATION OF ORGANELLAR DNA

Several mechanisms for the replication of organellar DNA have been proposed in the past. The most popular of these models is the model, involving initial formation of D-loops followed by the formation of intermediate θ shaped structures and finally rolling circles (see Chapter 39). This model was later questioned, and **recombination-dependent replication** (**RDR**) has been proposed for both mtDNA and cpDNA (for details, consult Chapter 30, and explore at the internet). Similar mechanism of replication may operate in apicoplast DNA also.

GENES IN ORGANELLAR GENOMES

Organellar genomes are much smaller than their nuclear counterparts. Therefore, it is anticipated that the gene contents in these organellar genomes are much more limited, which is indeed the case.

Mitochondrial genomes exhibit greater variability in gene content. Mitochondrial genomes display greater variability, the number of genes ranging from 5 for the malaria parasite (*P. falciparum*) to 92 for the protozoan, *Reclinomonas americana* (Table 42.3). All mitochondrial genomes contain genes for the non-coding rRNAs and for at least some of the protein components of the respiratory chain, the latter being the main biochemical feature of the mitochondrion. Some gene-rich mitochondrial genomes also code for tRNAs, ribosomal proteins and proteins involved in transcription and translation; they also code for proteins involved in transport of other proteins into mitochondrion from the surrounding cytoplasm (Table 42.3; Fig. 42.6). There are also a number of unidentified reading frames (URFs) representing protein coding genes with unknown

Table 42.3. **Features of mitochondrial genomes in six different organisms.**

Feature/types of genes	*Plasmodium falciparum*	*Chlamydomonas reinhardii*	*Homo sapiens*	*Saccharomyces cerevisiae*	*Arabidopsis thaliana*	*Reclinomonas americana*
Genome size (kb)	6	16	17	75	367	69
Total number of genes	5	12	37	35	52	92
Protein-coding genes	5	7	13	8	27	62
Respiratory complex	0	7	13	7	17	24
Ribosomal proteins	0	0	0	1	7	27
Transport proteins	0	0	0	0	3	6
RNA polymerase	0	0	0	0	0	4
Transcription factors	0	0	0	0	0	1
Non-coding RNA genes	0	5	24	26	25	30
Ribosomal RNA genes	0	2	2	2	3	3
Transfer RNA genes	0	3	22	24	22	26
Other RNA genes	0	0	0	1	0	1
Number of introns	0	1	0	8	23	1

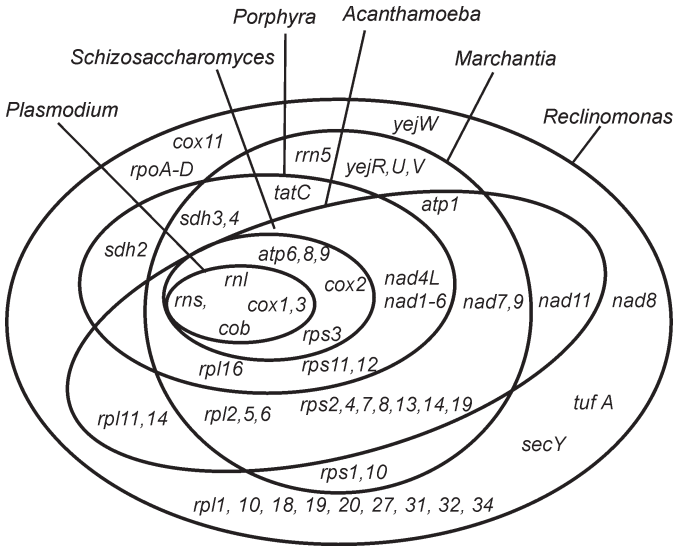

Fig. 42.6. Mitochondrial genomes of a variety of organisms showing variation in gene content.

products. The organization of human mitochondrial genome is shown in Figure 42.7.

Why are mitochondrial genomes much larger in plants? Mitochondrial genomes are much larger in plants (150 kb to 2500 kb) than in animals (6 kb to 70 kb). It has been shown that perhaps much extra DNA in plant mitochondria is 'junk DNA'. In several plant species (*e.g., Zea mays* = maize), cytoplasmic male sterility is also determined by mtDNA.

Mitochondrial genome has several surprising features. The mitochondrial genome also has several surprising features, which include the following: (i) there are very few regulatory DNA sequences, most regulatory sequences making a part of one or the other coding sequence; (ii) only 22-26 tRNA genes encoding 22 tRNAs are found in mitochondria as against 30 tRNAs found each in the cytosol and the chloroplasts; (iii) for at least four of the 64

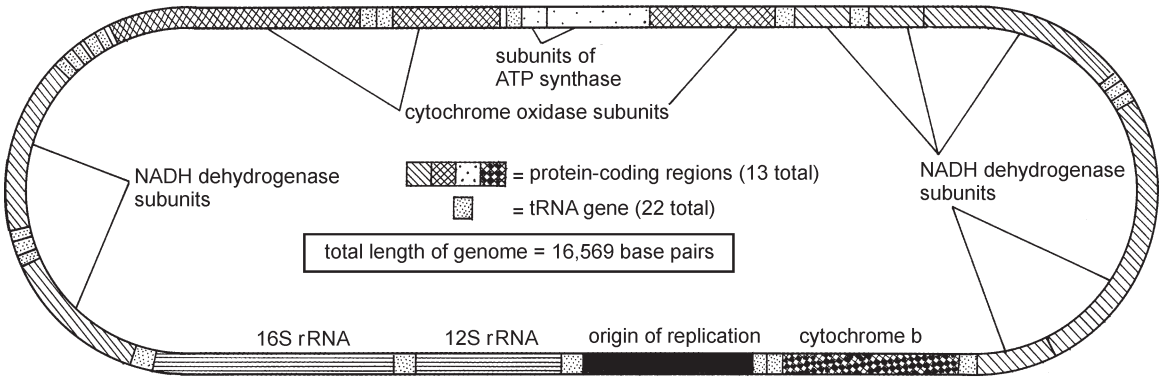

Fig. 42.7. The organization of human mitochondrial genome.

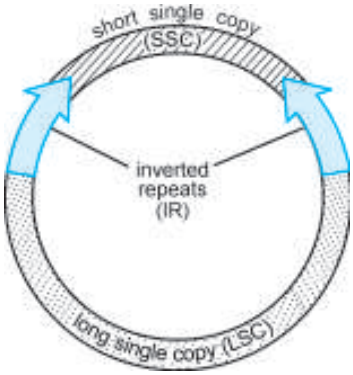

Fig. 42.8. Three characteristic regions of a chloroplast genome.

codons, the genetic code in mitochondria differs not only from that in the cytosol and chloroplasts, but also among themselves (for details, consult Chapter 46 on 'The Genetic Code').

Most chloroplast genomes possess same set of genes. Most chloroplast genomes appear to possess the same set of about 110-120 genes, which code for rRNAs and tRNAs, as well as ribosomal proteins and proteins involved in photosynthesis (see Table 42.4). In most plants, cpDNA is characterized by the presence of the following three regions shown in Figure 42.8: (i) two inverted repeats (IR), each 10-24 kb long and carrying ribosomal RNA genes, (ii) a short single copy (SSC) sequence, 18-20kb long, and (iii) a long single copy (LSC) sequence. A summary of 110 genes that are encoded in cpDNA is presented in Table 42.4.

***Table 42.4.* Genes encoded in chloroplast DNA.**

RNA/Protein	Genes	RNA/protein	Genes
1. Genes for protein synthesis apparatus		**2. Genes for photosynthesis apparatus**	
16S rRNA	1	Photosystem I	2
23S rRNA	1	Photosystem II	7
4·5S rRNA	1	Cytochrome b/f	3
5S rRNA	1	ATPase	6
tRNA	30	NADH dehydrogenase	6
r-proteins	19	Ferredoxin	3
RNA polymerase	3	RUBP carboxylase	1
		3. Unidentified genes	**26**
		Total	**110**

Most apicoplasts also contain the same set of genes. The apicoplast genome carries genes for at least 30 different proteins, ~25 tRNA, two rRNA (ribosomal RNAs) and for a number of ribosomal proteins. In majority of cases, both strands are coding strands, and only in some cases one strand is coding, the other being antisense strand.

Some Organellar Genes Contain Introns

Unlike human mtDNA, some plant and fungal mitochondrial genes contain introns, which need to be spliced out from RNA transcripts during processing. Introns have also been found in atleast 20 chloroplast genes. Introns may also be present in some apicoplast genes. The presence of introns in these genes is surprising because these are not common in bacteria, which are assumed to have given rise to mitochondria and chloroplasts, although introns may be present in some genes of red algae, which are believed to have given rise to apicoplasts. In yeasts, the same mitochondrial gene may have an intron in one strain, but not in the other strains. Such introns perhaps behave like transposable elements (TEs). In some cases, introns were located at the same position in the same gene in the mitochondrial genomes of yeast, *Aspergillus* and *Neurospora*, suggesting their common origin. These observations also suggest that introns have ancient origin and were lost in bacterial genomes, but retained in some organellar and nuclear genomes of eukaryotes.

ORGANELLAR PROTEIN SYNTHESIS AND SEMI-AUTONOMY

We know that the organellar genome specifies only some of the proteins found in the organelle, but not all of them. The other proteins are coded by nuclear genes, synthesized in the cytoplasm, and transported into the organelle. If the cell has mechanisms for transporting proteins into

mitochondria and chloroplasts, then why not have all the organelle proteins specified by the nuclear genome? We do not yet have a convincing answer to this question, although it has been suggested that at least some of the proteins coded by organellar genomes are extremely hydrophobic and cannot be transported through the membranes that surround mitochondria and chloroplasts, and so simply cannot be moved into the organelle from the cytoplasm. The only way in which the cell can get them into the organelle is to make them there itself.

Eukaryotic organelles resemble prokaryotes in protein synthesis. It is known now that each of the three eukaryotic organelles (mitochondrion, chloroplast and apicoplast) contains all the four components necessary for autonomy, that is, DNA, DNA polymerase, RNA polymerase and a protein synthesizing system. Most of these components and the mechanism of protein synthesis in each of these three organelles resemble those in prokaryotes. For instance, RNA polymerase in mitochondria resembles that in certain bacteriophages (single subunit), and RNA polymerase in chloroplasts and apicoplast largely resembles that in *E. coli*. Also the mRNAs in these three organelles lack a 5¢ cap (5¢ cap is characteristic of all eukaryotic mRNAs) and is synthesized as polycistronic mRNA (as in *E. coli*).

Both strands of mtDNA are transcribed. In human cells, both strands of mtDNA are transcribed at the same rate from a single promoter region on each strand, but in opposite directions. Thus two different giant RNA molecules are produced. The transcripts made on **heavy strand** (H-strand), are extensively processed to yield two rRNAs, most tRNAs and 10 poly-A containing mRNAs (no 5´ cap is found). The transcripts from the **light strand** (L-strand) is processed to produce only tRNAs and one small poly-A containing mRNA (remaining 90% of L-strand transcript is degraded).

All the three organelles contain 70S ribosomes resembling those of bacteria but differing from 80S ribosomes found in cytoplasm (consult Chapter 49). In higher plants, chloroplast ribosomes make 50% of the total cellular ribosomes. The organellar ribosomes resemble those in prokaryotes in the following additional features: (i) use of N-formyl methionine as initiating amino acid, (ii) sensitivity to chloramphenicol, lincomycin and spectinomycin, which inhibit protein synthesis, and (iii) basic structures of their RNA and proteins (consult Chapter 49).

ORGANELLAR PROTEOME

The organellar proteomes have also been studied in some detail during the last three decades (1990-2017). In each case, the organellar proteome includes following two sets of proteins: (*i*) proteins encoded in the nuclear genome, synthesized in the cytoplasm and later transported to the organelle; and (ii) the proteins encoded by the organellar genome and synthesized within the organelle itself using its own protein synthesizing machinery. In 2010, ~300,000 mitochondrial proteins and ~120,000 chloroplast proteins were available in the GOBASE database. Later, this number must have gone up substantially. However the total number of proteins available at a time in an organelle must be just a few thousand, of which 100-500 may be encoded by their own genome. Similarly the proteins in an apicoplast are estimated to be 500 only, of which about 30 proteins (excluding ribosomal proteins) are synthesized with the apicoplast.

Unlike chloroplast, apicoplast are non-photosynthetic organelles. However, this organelle retains other plastic functionalities such as fatty acid synthesis, heme synthesis, and isoprenoid

synthesis. The products of these pathways are essential to the parasite, which requires these products in order to survive.

ORGANELLAR GENOME/ PROTEOME DATABASES

Databases containing details of genomes and proteomes have also been developed. For instance, GOBASE database was developed by the University of Montreal, Canada and contained information about mitochondrial/chloroplast genomes and proteomes. In the year 2010, GOBASE release 25 (Jume 2010) contained >1,000,000,000 mitochondrial genome sequences (including ~300,000 proteins), and >300,000 chloroplast sequences (including ~120,000 proteins). This was, however the last release of this database, because this database could not continue later due to lack of funds.

A separate database for chloroplast genomes is ChloroplastDB (http://chloroplastcbio.psu.edul) that was launched in 2005. This database contains information about all fully sequenced plastid genomes, containing protein, DNA and RNA sequences, gene locations and RNA-editing sites. The information about mitochondrial and chloroplast genomes is also available at NCBI site as GenBank records. On this site, a list is available which is up-dated on day to-day basis for mitochondrial and plastid genomes that are fully sequenced. In early, 2014, this list contained 6000 mitochondrial genomes and >200 plastid genomes that were fully sequenced at that time.

A mitochondrial protein database is MITOP, which was created in 1998 at the University of Munich in Germany. This database provided comprehensive list of mitochondrial proteins of yeast, man, mouse, the nematode worm (*C. elegans*), and the fungus *Neurospora crassa*. Later, in 2004, MitoP2 database (http:/ www.mitop.de) was launched as an extension of **MITOP**, to integrate further the information on mitochondrial proteins and their molecular functions for yeast, human and mouse. **MitoProteome** is another database giving information on mitochondrial proteomes. MitoP2, MitoProteome, and HMPDb (a mitochondrial protein database only for humans) should prove extremely useful for further studies on mitochondrial proteome, particularly with the objective to find the genes and proteins that underlie common and rare mitochondrial diseases in humans. More details about mitochondrial and chloroplast genomes and proteomes are available at the NCBI site.

Mitochondrial Diseases

During the last 15 years (1990 onwards), several mitochondrial defects have been found to be associated with diseases in human beings and mouse (used as a model). About 50 pathogenic mtDNA base substitutions and rearrangements (deletions and insertions) have been associated with a variety of degenerative diseases. Other diseases involved defects in mitochondrial bioenergetics or biogenesis, which are linked with mutations in the nuclear genes. The three important functions of mitochondria (covered by oxidative phosphorylation), defects in which are related with diseases, include the following: (i) energy production, (ii) generation of reactive oxygen species (ROS) and (iii) regulation of programmed cell death or **apoptosis**.

EVOLUTIONARY ORIGIN OF MITOCHONDRIA AND PLASTIDS

Monophyletic Origin of Mitochondrion from a Free Living Proteobacterium

Mitochondria resemble bacteria in many respects (Table 42.5) and are believed to have originated from some bacteria, which belong to **rickettsial** subdivision of a-proteobacteria. During 1997-99, complete DNA sequences of bacteria-like mitochondrial genome (*e.g., Reclinomonas americana*; 69,034 bp) and mitochondria-like eubacterial genome (*e.g., Rickettsia prowazekii;*

Table 42.5. **Comparison between mammalian mitochondria and bacteria.**

Characters	Mitochondria	Bacteria
1. Size : width	0·2—1·0μ	0·2—2·0μ
length	1·0—4μ	0·3—10·0μ
2. Lipoprotein membrane	6—7 (nm)	7—8 (nm)
3. Invaginations of membrane	Cristae	Mesosomes
4. Respiratory control	Marked	Low or absent
5. Inhibitors of phosphorylation (*e.g.*, CN—, CO, azide, antimycin, DNP and oligomycin)	Effective	Less effective or ineffective
6. DNA shape	Closed circle	Closed circle
7. Inhibitor of protein synthesis	Chloramphenicol	Chloramphenicol
8. Ribosomes	70S (variable)	70S

1,111,523 bp) were worked out and compared, suggesting strongly that all extant mtDNAs had their origin in a single ancestral **protomitochondrial genome**, which in its turn, must have originated from a eubacterial genome (**monophyletic origin** rather than **polyphyletic origin**). A remarkable variation exists in the size of mtDNA and its gene content (Figs. 42.1, 42.6). Based on this information, mtDNAs have been classified in two basic types: (i) **ancestral type** with many genes, as in *R. americana* and (ii) **derived type** marked with extensive gene loss, divergence in coding and non-coding DNA sequences, a new mitochondrial code, RNA editing, etc. (*e.g.*, animal and fungal mtDNAs).

Simultaneous Origin of Eukaryotic Nuclei and Mitochondria
('Serial Endosymbiosis' vs 'Hydrogen Hypothesis')

A serial endosymbiosis model has also been proposed for the origin of nuclear genome and mitochondria. In this model, it is assumed that in the first step, the nuclear genome of the host itself resulted from the fusion of eubacterial and archaebacterial partners and that only in the second step, the mitochondrion was acquired as a symbiont. A group of eukaryotes known as **Archezoa**, do lack mitochondria, suggesting that perhaps eukaryotes without mitochondria originated first as a primitive form and that the mitochondria entered the cells later.

The above view, however, has recently (1997-2010) been challenged and it is suggested that the eukaryotes acquired the nucleus and mitochondria simultaneously, and that **Archezoa** (lacking mitochondria) must have resulted due to loss of mitochondria. This view was supported by the observation that genes for several mitochondrial proteins of bacterial origin are found in the genome of amitochondriate Archezoa and also in the organelle **hydrogenosome**, recently discovered in protists lacking mitochondria. This view is described as '**hydrogen hypothesis**' since it assumes chimeric origin of eukaryotic nucleus, due to symbiotic association between (i) a eubacterium (a proteobacterium, the symbiont), producing H2 as the end product of anaerobic metabolism and (ii) a hydrogen - requiring autotrophic archaebacterium (the host). This hypothesis thus allows the possibility of simultaneous origin of the ancestor of eukaryotic cell and its mitochondrion (Fig. 42.9).

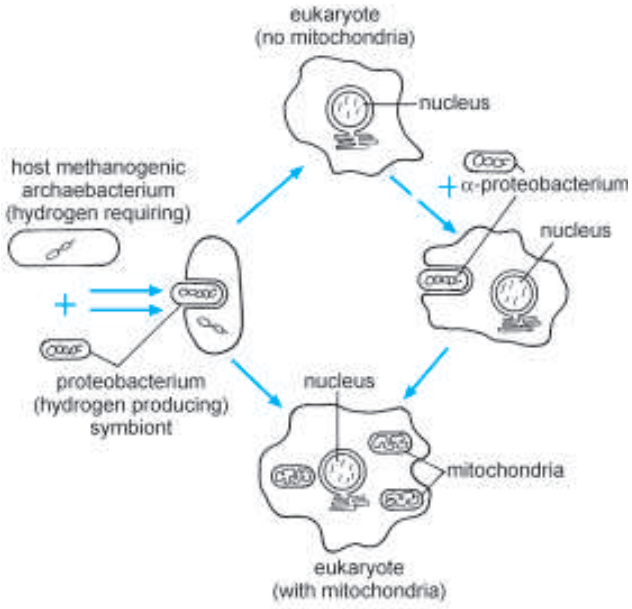

Fig. 42.9. Two alternative hypotheses for the origin of eukaryotic cell and its mitochondrion.

Evolutionary Origin of Plastids from Cyanobacteria and 'Endosymbiotic Gene Transfer'

A comparison of sequences of chloroplasts from a large number of species with those of cyanobacteria, suggested that the plastids (particularly chloroplast) originated from a cyanobacterium. This origin involved primary endosymbiosis followed by extensive **'endosymbiotic gene transfer'** from the cyanobacterium to the eukaryotic host nucleus (for details, consult Chapter 7).

A detailed study of chloroplast genome in eukaryotes, has also revealed that the chloroplast genomes share both prokaryotic and eukaryotic features. Therefore, the view that chloroplast are the remnants of entrapped prokaryotic genomes is an oversimplification. Some of the features in which chloroplasts resemble the prokaryotes and those in which they differ are as follows.

(a) Resemblances between chloroplast and prokaryotes. (i) The ribosomes are 70S in size in both the cases. (ii) The rRNA genes as well as several tRNA genes have very high degree of homology as revealed by sequence homology. (iii) In most chloroplast genes, preceding the initiation codon AUG, there are present ribosome binding sites, and the sequences, which bear strong resemblance to '**Pribnow Box**' (–10 and the -35 regions) of bacteria (consult Chapters 47) (iv) There is sequence overlapping between some genes in chloroplasts, a feature known in prokaryotes. In chloroplasts it can not be due to economy, since non-coding sequences are also available in chloroplasts. (v) There is a resemblance between RNA polymerases of chloroplasts and prokaryotes. For instance, in *Chlamydomonas reinhardii*, the enzyme is sensitive to **rifampicin** and contains a s (sigma) like polypeptide. There is also evidence of homology between b and b¢ subunits of *E. coli* RNA polymerase and those of *C. reinhardii* chloroplast RNA polymerase. (vi) Protein synthesis in chloroplasts can be inhibited by chloramphenicol as in bacteria, since only 70S ribosomes are sensitive to this drug. Cycloheximide does not inhibit protein synthesis in chloroplasts, as it does in the cytoplasm (80S ribosomes) of eukaryotes. (vii) Initiation of protein synthesis takes place by formyl methionyl tRNA both in chloroplasts and prokaryotes.

(b) Differences between chloroplast and prokaryotic genomes. (i) In chloroplasts, there are split genes (*e.g.*, rRNA genes and protein genes in chloroplasts of *C. reinhardii*, protein genes in chloroplasts of *Euglena*, and tRNA genes of several chloroplasts of higher organisms); introns of these split genes fall in three groups and differ from those in the nuclear genes of eukaryotes (consult Chapter 35). (ii) Chloroplast DNA sequences have the ability to promote autonomous replication, this feature is absent in prokaryotes. (iii) Presence of short repetitive elements interspersed throughout several, but not all chloroplast genomes is also a eukaryotic feature.

(c) Comparative genomics. In the previous section, we briefly discussed the monophyletic endosymbiotic origin of mitochondria from proteobacteria. During 1990s, due to the availability of extensive data on the nucleotide sequences of a number of chloroplast genomes and the corresponding functional genes in each case, it became possible to apply the principle of comparative genomics in case of chloroplasts also. One of the smallest genomes of cyanobacteria, sequenced in the recent past is 3,573,470 bp long and belongs to **Synechocystis**; in contrast, the biggest chloroplast genome sequenced so far (till 1998) is 191,028 bp long, which belongs to **Porphyra**. A comparison of sequences of chloroplast genomes from 12 species with those of cyanobacteria, suggested that the plastids (particularly chloroplasts) originated from a cyanobacterium, due to endosymbiosis followed by extensive '**endosymbiotic gene transfer**' from the cyanobacterium to the eukaryotic host nucleus. While a primary endosymbiosis led to the origin of simple two membraned plastids, a secondary endosymbiosis is assumed to have been involved in the origin of complex three or four membraned plastids (Fig. 42.10). As in case of mitochondria, it is also believed that during evolution

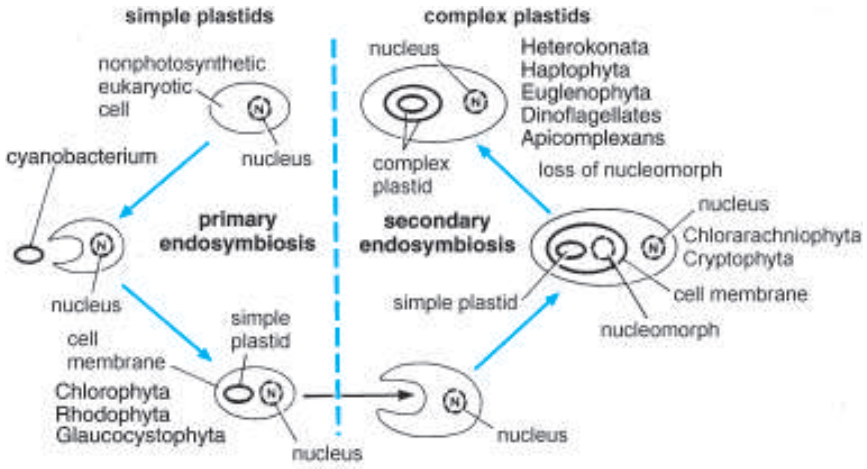

Fig. 42.10. A model for the evolutionary origin of simple and complex plastids from a cyanobacterium through endosymbiosis.

chloroplasts exported their genes to the nucleus, but reimported the products with the help of transit peptides and protein import machinery. Sufficient evidence is now available to suggest that '**endosymbiotic gene transfer**' from organelles to nucleus really took place in eukaryotic cells.

It has also been shown that the plastid genome though generally encodes many more (21 to 200) proteins, relative to mitochondrial genome (5 to 63), but they still code only 1% to 5% of the protein-coding genes found in a relatively small cyanobacterium genome. In contrast to this small number of genes, the proteins that are found in a fully functional plastid have been estimated to be up to 5000 in *Arabidopsis*, suggesting that vast majority of these proteins are imported. Recent analysis also suggested that a minimum of 50 genes, that are found in atleast one plastid genome, have functional homology with the nuclear genes of at least one higher plant, indicating large scale transfer of genes from plastid to the nucleus. There are also proteins in the chloroplasts, which are encoded in the nuclear genes, but perhaps originally belonged to mtDNA, so that some of the chloroplast proteins were originally encoded in the mtDNA.

Evolutionary Origin of Apicoplasts

The apicoplast has a secondary endosymbiotic origin (plastids had a primary endosymbiotic origin from a cyanobacterium-like prokaryote), which is evident from the fact that apicoplasts generally have four membranes (primary plastids invariably have only two membranes). However, there seems to be some disagreement about the nature of the secondary endosymbiont (green alga vs red alga), although there seems to be an agreement now that apicoplast must have originated from red alga only. This is based on the discovery of photosynthetic apicomplexan geneus *Chromera*, which clearly has a plastid (similar to apicoplast) derived from a red alga.

PROMISCUOUS DNA AND HORIZONTAL TRANSFER OF GENES

(Movement of DNA Between the Organelles and the Nucleus)

During 1982-84, evidence became available, which suggested transfer of DNA from

chloroplasts to mitochondria and from both these organelles to the nuclear genome,. For instance, it was suggested that apparent translocation of DNA sequences from chloroplasts to mitochondria has taken place in corn. This conclusion was based on the discovery that a 12 kilobase segment of mitochondrial DNA (mtDNA) was homologous to a section of chloroplast DNA (cpDNA) of the same organism, corn. During the same period, a report from Texas, U.S.A., described the same phenomenon in yeast, where a segment of mtDNA showed similarity with a segment in nuclear DNA, suggesting flow of DNA from mitochondria to nucleus. A similar report of transfer of mtDNA to nucleus in sea urchin also came from California Institute of Technology. In 1984, more reports of promiscuous DNA came from different places in plant systems like mungbean, spinach, corn and peas, suggesting that movement of DNA sequences among organelles and between organelles and nucleus must be very common and not exceptional. The movement of DNA among mitochondria, chloroplasts and nucleus are shown in Figure 42.11. Apicoplast is also believed to have transferred many genes to the the nucleus. Some of these might be encoding proteins, which are transported back to the organelle.

The evolutionary implications of promiscuous DNA are far reaching and it is speculated that chloroplasts and mitochondria, before invading eukaryotic cell were free living primitive prokaryotes. The endosymbiont hypothesis suggests that in course of time these entrapped organisms lost their independence by transfer of genetic information to nuclear genome. Thus the promiscuity of DNA gave further support to endosymbiont hypothesis, but exchange of DNA between mitochondria and chloroplasts is surprising, since they are independent of each other. This may be partly explained on the basis of great variation and bigger size of mtDNA, relative to cpDNA which is tightly constrained in size and sequence even between species. While mtDNA can tolerate foreign DNA and other disturbances, chloroplasts are sensitive to any such variation or modification. This is confirmed by evidence, where it was shown that chloroplast sequences from a variety of plants are scattered

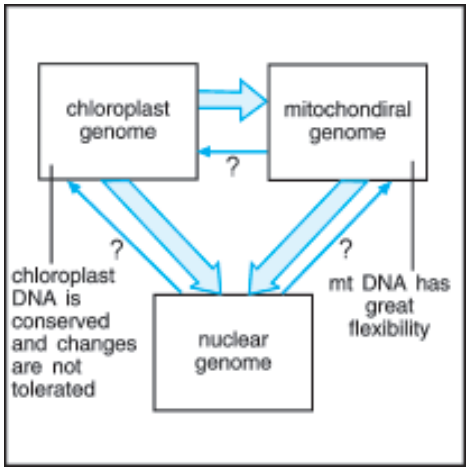

Fig. 42.11. Exchange of genetic material between nuclear genome, chloroplast genome and mitochondrial genome. Thick arrows indicate transfer of promiscuous DNA for which evidence is available.

throughout mitochondrial genomes, and some of these transfers are believed to be rather recent. The mechanism of this transfer of the so called promiscuous DNA is not known, but two possibilities were considered, (i) Some kind of a vector or transposon or transducing phage may be involved. (ii) The two organelles might have undergone fusion at some stage. This second mechanism involving fusion is preferred on the basis of electron microscopic observations.

RETENTION OF GENES IN CHLOROPLASTS AND MITOCHONDRIA : WHY?

Since there has been a large-scale horizontal transfer (HT) of genes from chloroplast, mitochondrial and apicoplast genomes to the nuclear genomes, there has also been a debate why some genes and not others have been retained in these three organelles. Three possible hypotheses, which were put forward in the past to explain this retention of organellar genes are listed in Table 42.6. Among the two of these three hypotheses that were initially suggested, one

Table 42.6. **Three different hypotheses explaining retention of genes in organelles.**

Hypothesis	Summary statement of hypothesis
Code disparity hypothesis (CDH)	Divergence of genetic code, especially the codon encoding STOP in the nucleus but not in the organelle, retains genes in the organelle, if they contain any such codons. This barrier to functional gene transfer during evolution is greater than any other.
Hydrophobicity hypothesis (HH)	Hydrophobicity (measured as 'mesohydrophobicity') of a protein quantitatively impedes its import into the mitochondrial matrix via TIM23, which must unfold its cargo fully so as to translocate it without collapsing the mitochondrial membrane potential. HH does not apply to chloroplasts, because they lack a membrane potential and so can import folded proteins.
Co-location for redox regulation (CORR)	Proximity of genes to their products' site of action facilitates rapid and microenvironment-specific response of gene expression to metabolic demands. This is especially advantageous for organellar electron transport processes and constitutes a barrier to functional transfer of either cpDNA or mtDNA genes.

is based on genetic code disparity between nuclear and organellar genes, and the other is based on hydrophobicity of proteins encoded by organellar genes. These two hypotheses assume that the organellar genes if transferred to the nucleus, will lead to synthesis of proteins containing incorrect amino acids, and that any hydrophobic organellar protein, if synthesized in the cytosol, will have difficulty in its transport from the cytosol to the organelles. In contrast to these two hypotheses, there are hypotheses which presumably provide much more powerful barriers to allow retention of genes in chloroplasts and mitochondria. One such hypothesis listed in Table 42.7, concerns role of proteins in redox processes, and received wide support, although the debate still continues.

PROBLEMS

1. Describe the variation in shape, size and number of mitochondrial and chloroplast genomes in a eukaryotic cell.
2. Discuss the structure of mitochondrial genome in different groups of plants and animals. Describe some of the diseases caused due to mutations in mitochondrial genes.
3. Compare the mitochondrial genomes of animals with those in plants. Why are mitochondrial genomes in plants much larger in size?
4. Describe the *heavy* and *light strands* (H-and L-strands) of human mtDNA. How are they transcribed and what is the fate of the two transcripts?
5. Describe briefly the genetic content of mitochondrial and chloroplast genomes.
6. Discuss the simultaneous endosymbiotic origin of mitochondria and eukaryotic cells from proteobacteria and archaebacteria; also discuss the 'serial endosymbiosis' vs 'hydrogen hypothesis' for the origin of mitochondrion and the eukaryotic nucleus.
7. Describe briefly the evidences, which support the view that plastids originated from cyanobacteria in the remote past and that there has been endosymbiotic gene transfer from plastids and mitochondria to the nucleus.
8. Write short notes : (i) human mitochondrial genome; (ii) yeast mitochondrial genome; (iii) tRNAs in mitochondria; (iv) mitochondrial genes; (v) chloroplast genes; (vi) endosymbiotic origin of mitochondria and chloroplasts.