Nucleomorph Genomes

P Mackiewicz, A Bodył, and P Gagat, University of Wrocław, Wrocław, Poland

Glossary

Chlorarachniophytes Marine photosynthetic and mixotrophic amoeboflagellates producing anastomosing reticulate pseudopods used for prey capture and formation of netlike connections between cells. They gained a plastid via a secondary endosymbiosis with a green alga.

Cryptophytes A group of unicellular algae common in freshwater, marine, and brackish environments. Most cryptophytes are photosynthetic, but some show mixotrophy. Cryptophytes contain a plastid acquired via the secondary endosymbiosis with a red alga. Their asymmetric cell is equipped with two unequal flagella and distinct extrusomes (ejectisomes).

Endosymbiotic gene transfer Gene transfer occurring during endosymbiosis in which a gene from an endosymbiont or organelle genome is transferred and integrated into the host's nuclear genome.

Microsporidia A group of obligate intracellular eukaryotic parasites. Previously they were regarded as ancient eukaryotes that diverged before the origin of mitochondria; however, later studies confirmed the presence of reduced descendents of these organelles. It also was found that microsporidia represent a highly modified fungal lineage. During their adaptation to a parasitic lifestyle, some microsporidial nuclear genomes underwent drastic reduction, evolving several features that also characterize nucleomorphs.

Presequence An N-terminal amino acid sequence of a pre-protein responsible for its targeting to a specific subcellular compartment. In the case of secondary (complex) plastids, the presequence usually consists of a signal peptide followed by a transit peptide. The signal peptide is cleaved in the endoplasmic reticulum, whereas the transit peptide is removed in the stroma of the plastid.

Secondary endosymbiosis An endosymbiosis in which a eukaryote containing a primary plastid (i.e., previously derived from primary endosymbiosis) is acquired by another eukaryotic host cell.

Signal peptide An N-terminal amino acid sequence (usually 10–40 residues) that targets pre-proteins to the endoplasmic reticulum. The signal peptide is rich in hydrophobic amino acid residues and is removed in the endoplasmic reticulum by a signal peptidase.

Spliceosomal introns The most common type of introns present in nuclear protein-coding genes. They are removed from precursor mRNA by a large complex called the spliceosome that is composed of characteristic small nuclear (sn) RNAs and proteins.

Synteny Conserved order of genes or other genetic elements in the chromosomes of different species.

Transit peptide An N-terminal amino acid sequence of a pre-protein responsible for targeting it to the mitochondrion or plastid. This peptide is often cleaved off when the organelle's matrix is reached. Plastid transit peptides are variable in length (mostly 50–70 residues) and amino acid composition but usually are characterized by a relatively high content of hydroxylated residues.

Nucleomorph Genomes

Nucleomorphs are relict nuclei found in multimembrane plastids of some eukaryotic algae, such as chlorarachniophytes and cryptophytes. Their plastids are surrounded by four membranes and evolved via secondary endosymbiosis from green and red algal endosymbionts, respectively. They still maintain the vestigial cytoplasm of the engulfed alga, called the periplastidal compartment, which is located between the inner and outer pairs of envelope membranes (Figure 1). This compartment contains the nucleomorph, eukaryotic (80S) ribosomes, and starch grains (in cryptophytes). Although secondary plastids are widespread in many protist lineages (e.g., euglenids, heterokonts, haptophytes, and even some apicomplexan parasites), reduced nuclei from their endosymbionts only have been found in cryptophytes and chlorarachniophytes. It is also possible, however, that the green alga-derived plastid in the dinoflagellate *Lepidodinium chlorophorum* possesses a nucleomorph, as suggested by microscopic analyses and the existence of two distinct ribosomal RNA (rRNA) genes, one of which appears to be derived from the 'green' plastid donor.

Nucleomorphs were first discovered by Greenwood in 1974 in cryptophytes and from the beginning were correctly interpreted as vestigial cell nuclei. In 1984, Hibberd and Norris described the nucleomorph in chlorarachniophytes. Early cytochemical tests confirmed the presence of DNA within nucleomorphs, and also revealed a nucleolus-like region rich in rRNAs. In 1991, two very different rRNA genes were isolated from cryptophytes and phylogenetic analyses confirmed that they are derived from distinct organisms. Further studies demonstated that one of these genes resides in the nucleomorph genome and that its rRNA product functions in the periplastidal compartment. Similar data also were obtained for chlorarachniophytes. Decisive evidence for the red and green algal origins of nucleomorphs was provided by complete genome sequences. To date, complete nucleomorph genomes from one chlorarachniophyte (*Bigelowiella natans*) and three cryptophytes (*Guillardia theta*, *Hemiselmis andersenii*, and *Cryptomonas paramecium*) have been sequenced (Table 1).

 doi:10.1016/B978-0-12-374984-0.01701-0

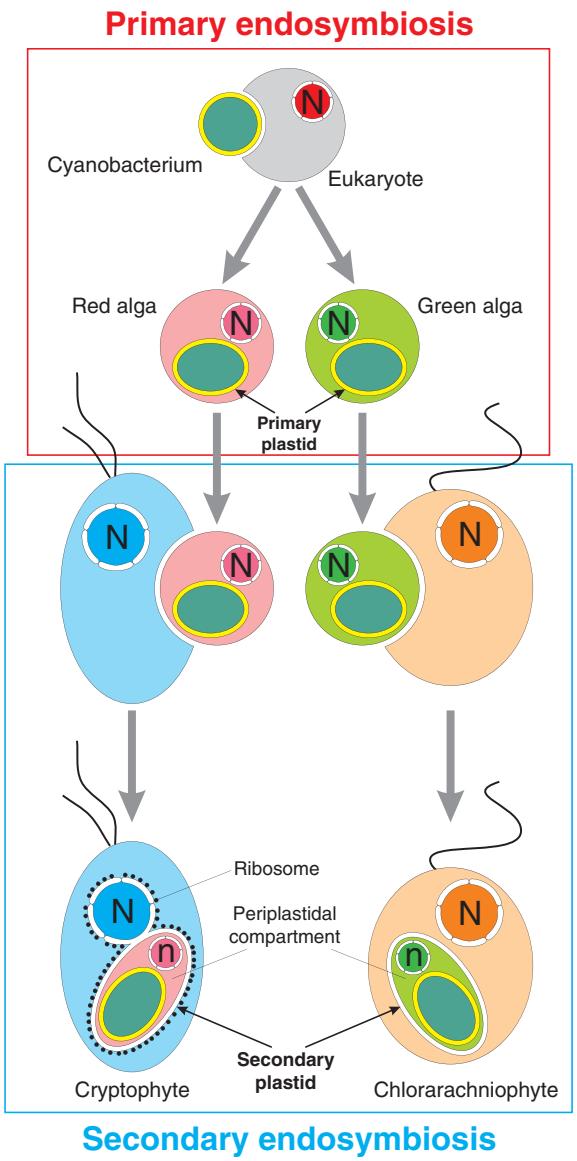

Figure 1 Complex evolutionary history of cryptophyte and chlorarachniophyte plastids. These plastids are surrounded by four envelope membranes, maintaining a residual eukaryotic cytoplasm or 'periplastidal space' that contains a vestigial nucleus called the nucleomorph (n). This peculiar organelle is enclosed by a nuclear envelope with pores and has linear chromosomes, as in typical eukaryotic nuclei (N). The complex structure of cryptophyte and chlorarachniophyte plastids, and especially the presence of nucleomorphs, indicates that their evolution proceeded in two main endosymbiotic stages. During an initial primary endosymbiosis, a phagotrophic eukaryote enslaved a cyanobacterium that was transformed into a primary plastid with two envelope membranes. Such primary plastids are found in three eukaryotic lineages: glaucophytes, red algae, and green plants, including green algae and higher plants relatives. In the next evolutionary step, red and green algal cells were engulfed independently by the heterotrophic ancestors of cryptophytes and chlorarachniophytes, respectively, in a process termed secondary endosymbiosis. These algal cells were finally reduced to secondary plastids surrounded by four envelope membranes with highly reduced nuclei. The outermost plastid membrane is ribosome-free in chlorarachniophytes, whereas it is covered with ribosomes in cryptophytes and is connected to the nuclear envelope. This suggests that the outermost membrane underwent fusion with the rough endoplasmic reticulum in the cryptophyte lineage.

Genome Structure and Organization

Nucleomorphs preserve several features characteristic of typical nuclei, such as a nuclear envelope with pores and linear eukaryotic chromosomes (Table 1); however, they contain the smallest known eukaryotic genomes. They are even several times smaller than the highly reduced genomes of parasitic microsporidia; for example, the 2.3 Mbp genome of *Encephalitozoon intestinalis*. Nucleomorph genomes of cryptophytes have a known size distribution of 485–845 kbp, whereas those from chlorarachniophytes range between 330–1035 kbp (Table 1).

Although cryptophyte and chlorarachniophyte nucleomorphs evolved independently, their genomes share several common features resulting from striking evolutionary convergence (Table 1). For example, they contain only three chromosomes (100–300 kbp), possess subtelomeric ribosomal DNA (rDNA) operons (Figure 2), and their A + T content is very high (65–77%). Nucleomorph chromosome ends with telomeric repeats are unusual in cryptophyte chromosomes, while in the chlorarachniophyte *B. natans* they are more typical (Table 1). It is hypothesized that the tiny nucleomorph chromosomes are large enough to avoid loss during mitosis, yet small enough to fit within the confines of a remnant nucleus.

Another common feature of cryptophyte and chlorarachniophyte nucleomorph genomes is their very compact structure (Table 1). Genes constitute 78–91% of their genomes, similar to what is seen in bacterial genomes. They are devoid of transposons and have lost almost all noncoding DNA, with the exception of very small or very rare spliceosomal introns (Table 1). The nucleomorph genome of the chlorarachniophyte *B. natans* contains 852 spliceosomal introns that are the smallest known (18–21 bp). A similar size range of introns has been found in the nucleomorph genome of another chlorarachniophyte *Gymnochlora stellata*, suggesting that extremely reduced intron size is characteristic of chlorarachniophyte nucleomorphs. The nucleomorph genomes of the cryptophytes *G. theta* and *C. paramecium* possess only 17 and 2 introns, respectively, but they are significantly longer than those of chlorarachniophytes and within the range characteristic of intron length in protist genomes. In accordance with the presence of spliceosomal introns, the nucleomorph genomes of *B. natans*, *G. theta*, and *C. paramecium* encode small nuclear (sn) RNAs. No introns were found in the nucleomorph genome of the cryptophyte *H. andersenii*. Although intron loss clearly has occurred in cryptophyte nucleomorphs, only intron shortening has been demonstrated in chlorarachniophytes. The drastic differences in the number of introns in the nucleomorphs of these two protist groups may reflect intron numbers in the nuclear genomes of their algal ancestors. It is known that the genomes of green algae are generally intron-rich, while those of red algae are intron-poor.

Genome compaction of cryptophyte and chlorarachniophyte nucleomorphs also has involved significant shortening of intergenic regions to an average distance of ~100 bp (Table 1). It appears that, as a result of this compaction, regulatory elements (e.g., promoters and transcription terminators) have moved from intergenic into coding regions. Moreover, some gene sequences overlap (Table 1). Interestingly, a substantial fraction of proteins encoded in the nucleomorph genomes of the cryptophytes *G. theta* and

Table 1 Characteristics of completely sequenced nucleomorph genomes

Species	Bigelowiella natans	Cryptomonas paramecium	Guillardia theta	Hemiselmis andersenii
Taxonomy	Chlorarachniophyte	Cryptophyte	Cryptophyte	Cryptophyte
Origin	Green alga	Red alga	Red alga	Red alga
Plastid photosynthetic ability	Present	Lost	Present	Present
Total genome length (bp)[a]	372 870	487 066	551 264	571 872
Length of particular chromosomes (bp)[a]	140 590	177 338	196 216	207 524
	134 144	160 189	180 915	184 755
	98 136	149 539	174 133	179 593
Guanine plus cytosine content	28	26	26	25
Telomere sequence	$[TCTAGGG]_{25–45}$	$[GA_9]_{>9}$	$[(AG)_7AAG_6A]_{11}$	$[GA_{17}]_{4–7}$
Number of genes[a]	328	519	560	525
Number of protein genes[a]	283	466	485	471
Number of plastid genes	17	18	30	30
Number of overlapping genes[b]	23	28	45	9
Mean intergenic distance (bp)[b]	114	103	93	132
Number of spliceosomal introns and their length	852 (18–21 bp)	2 (62, 100 bp)	17 (42–52 bp)	0

[a]Data taken from current annotations of these genomes in GenBank database.
[b]Calculated according to current gene locations in GenBank database.

H. andersenii have been affected by reductive evolution as well. They are shorter than their red algal and green plant homologs, which results from losses of amino acid resides at their termini as well as internal deletions, including even of whole domains. Similar trends were observed in proteins encoded in the reduced genome of the microsporidian *E. intestinalis*.

Gene Content

Although the nucleomorph gene sets were reduced to 328 and between 519-560 genes in chlorarachniophytes and cryptophytes, respectively, convergent reductive pressures led to important similarities between them (Table 1). All nucleomorphs contain a 'the core set' of housekeeping genes involved in translation, transcription, protein folding/degradation, and DNA/RNA metabolism. The presence of centromere-binding proteins suggests that centromeres play a role in nucleomorph mitosis. It is interesting that cryptophyte and chlorarachniophyte nucleomorph genomes encode exactly the same aminoacyl-tRNA synthetase gene (for serine), the only one known to be retained in nucleomorphs. Almost 82% of the 120 conserved protein genes in the chlorarachniophyte

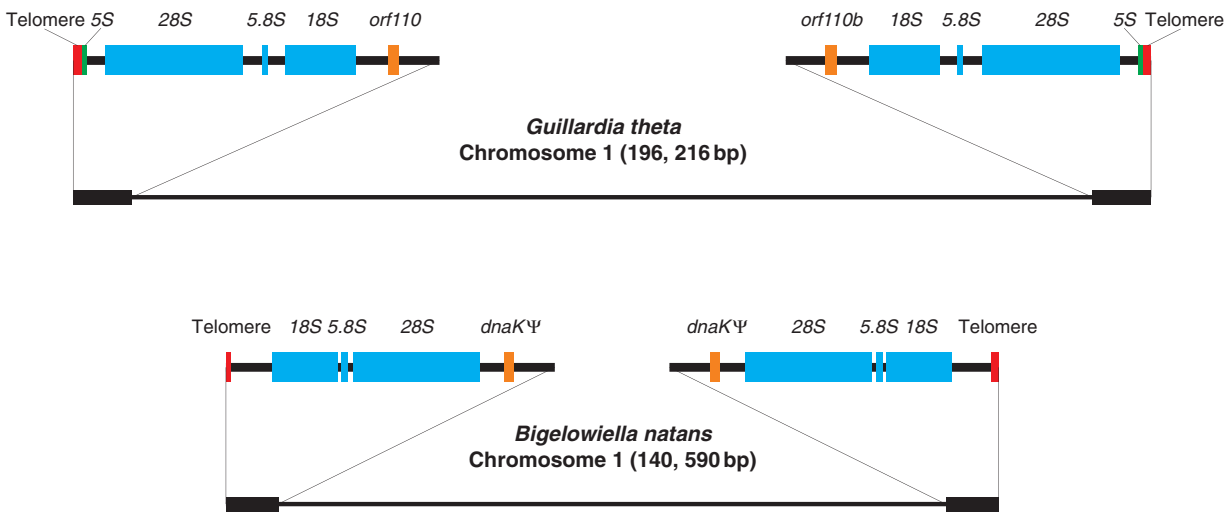

Figure 2 Structure of nucleomorph chromosome ends in the cryptophyte *G. theta* and the chlorarachniophyte *B. natans*. Chromosome 1 was chosen as an example for both species. Interestingly, each of these chromosomes has characteristic telomeric and subtelomeric rDNA repeat regions at their ends. The subtelomeric region of the cryptophyte chromosome consists of a *28S-5.8S-18S* rDNA operon and associated *5S* gene. The comparable region of the chlorarachiophyte chromosome does not contain a *5S* gene, and its rRNA genes constitute an operon oriented in the opposite direction; that is, *18S-5.8S-28S*. This chromosome end architecture is a characteristic of other chromosomes in a variety of cryptophyte and chlorarachniophyte species; however, only three of the six nucleomorph chromosome ends in the cryptophyte *H. andersenii* contain intact subtelomeric repeats, whereas the other three (both ends of chromosome 2 and one of chromosome 3) include only the *5S* rDNA locus. Similarly, in the cryptophyte *C. paramecium*, one end of both chromosomes 1 and 2 contains only a *5S* gene without a complete a rDNA operon.

nucleomorph also are present in all three cryptophyte genomes that, in turn, share 217 such genes.

About 30% of all the identified open reading frames in cryptophyte nucleomorphs show no significant primarily sequence similarity to genes with defined function. Many of them occupy evolutionarily stable positions on chromosomes (termed synteny) and are significantly enriched in amino acids encoded by A- and T-rich codons, such as phenylalanine, isoleucine, asparagine, lysine, and tyrosine. This peculiar combination of amino acids suggests that these sequences may encode membrane proteins.

Plastid-Associated Genes

Prior to their reduction during the course of secondary endosymbiosis, nucleomorphs presumably encoded ~1000 genes for plastid-targeted proteins, but this number was reduced drastically during evolution. The nucleomorph genomes of the cryptophytes *G. theta* and *H. andersenii* still retain 30 such genes, whereas there are only 18 in the cryptophyte *C. paramecium*; as a nonphotosynthetic species, it has lost genes involved in photosynthesis (Table 1). The nucleomorph genome of the chlorarachniophyte *B. natans* contains only 17 genes for plastid-targeted proteins. These proteins participate in distinct plastid functions, such as transcription (Sig2), iron–sulfur cluster formation (SufB, SufC), protein transport (Toc75, Tic20, TatC, SecY), protein folding (Hsp93 (ClpC), DnaK, Hsp60), protein degradation (ClpP), and plastid division (MurL).

Nucleomorph-encoded, plastid-targeted proteins have N-terminal transit peptides responsible for protein translocation through the two innermost plastid membranes (Figure 3). Because these membranes correspond to the envelope membranes of primary plastids of red and green algae, it is likely that protein translocation proceeds in a comparable manner across the two inner membranes of cryptophyte and chlorarachniophyte plastids. Protein import into primary plastids depends on two translocons: Toc (located in the outer membrane) and Tic (located in the inner membrane). Interestingly, the nucleomorph genome of the chlorarachniophyte *B. natans* encodes both Toc75 and Tic20, which are key protein-conducting channels. The nucleomorph genomes of the cryptophytes *G. theta* and *H. andersenii* encode other Toc and Tic proteins (representing IAP100 and Tic22 receptor subunits), which suggests that genes for Toc75 and Tic20 were transferred to their host nuclei.

Genome Rearrangements and Sequence Evolution

Although chlorarachniophyte and cryptophyte nucleomorph genomes are products of convergent evolution, they exhibit significant differences in rates of sequence evolution. In general, nucleomorph proteins from the chlorarachniophyte *B. natans* have accumulated more frequent substitutions than their homologs in higher plants or red algae, whereas cryptophyte sequences are not so divergent, especially plastid-targeted proteins. This is consistent with the possibility that cryptophyte nucleomorphs have 'stabilized' in comparison with chlorarachniophytes. In accordance with this, comparisons of gene order among cryptophyte nucleomorph genomes revealed large blocks of synteny, which suggest that genomic recombination and rearrangement is infrequent. On the other hand, considerable variation in nucleomorph genome size among closely related cryptophyte strains suggests that the reduction of nucleomorph genomes is an ongoing and rapid process in many independent cryptophyte lineages.

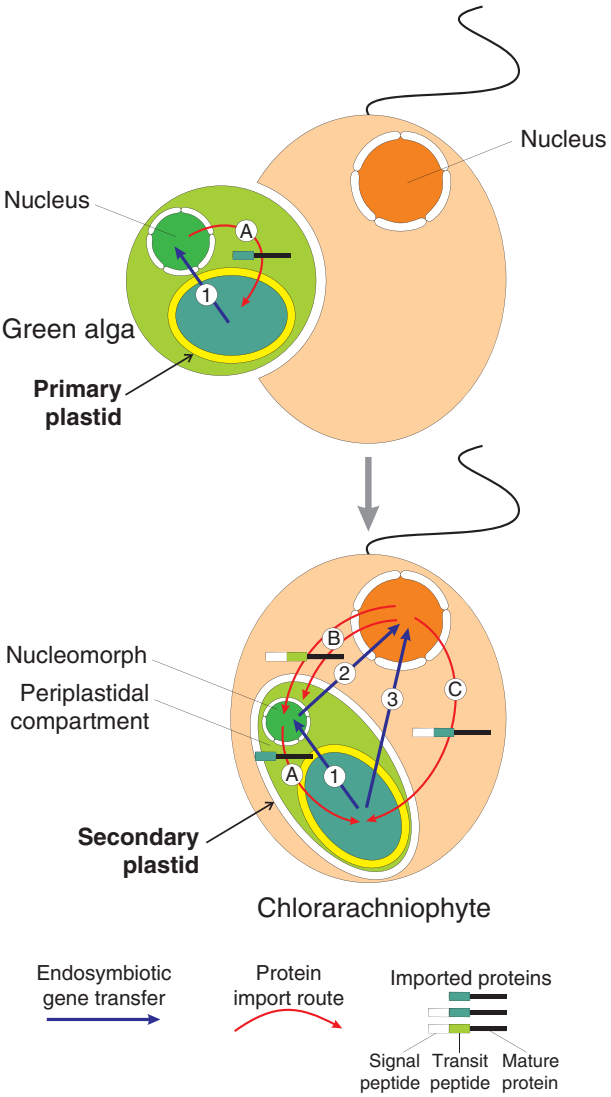

Figure 3 Endosymbiotic gene flow and plastid protein import routes within the cells of chlorarachniophyte algae containing green alga-derived plastids. During the primary cyanobacterial endosymbiosis, genes present in the primary plastid genome were transferred to the green algal nucleus (1). The transferred genes were equipped with sequences encoding N-terminal transit peptides responsible for translocation of their protein products through the two-membrane envelope of primary plastids (A). During the secondary green algal endosymbiosis, genes for the plastid proteins encoded in the green algal nucleus, which was reduced to the nucleomorph, were transferred to the nucleus of the secondary host; (2) note, however, that their direct transfer from the plastid genome to the secondary host nucleus also is feasible (3). During the reduction of the nucleomorph genome, other genes unrelated to plastid function also were moved to the host nucleus (2). All these proteins acquired bipartite presequences consisting of signal and transit peptides responsible for their translocation through two (in the case of nucleomorph- and periplastidal compartment-targeted proteins) (B) or four plastid membranes (in the case of plastid stromal-targeted proteins) (C). Because the transit peptides of plastid stroma-targeted proteins and nucleomorph- and periplastidal compartment-directed proteins have distinct physicochemical features, they are shown in turquoise and light green colors, respectively.

Endosymbiotic Transfer of Nucleomorph Genes to Host Nuclear Genomes

Nucleomorph genomes (e.g., in the chlorarachniophyte *B. natans*) encode 50 times fewer genes than modern-day algae believed to be related to their potential donors (e.g., the green alga *Chlamydomonas reinhardtii* has about 15 000 protein-coding genes). This reduction not only included loss of many plastid-associated genes, but also genes necessary for nucleomorph maintenance and expression. Nucleomorphs of cryptophytes and chlorarachniophytes do not encode genes for DNA-modifying enzymes or DNA polymerases. Additionally, genes encoding tubulins, proteasome subunits, 5S rRNA, telomerase RNA, and numerous tRNAs are missing from the chlorarachniophyte *B. natans* nucleomorph.

The 'missing' genes most likely have been transferred to the host nucleus via endosymbiotic gene transfer. Many examples of host nucleus-encoded, nucleomorph-derived genes for plastid proteins have been documented, such as those involved in the formation of light-harvesting complexes. There also are genes whose products are targeted to the periplastidal compartment (e.g., translation elongation factor-like protein) or imported directly into the nucleomorph (e.g., histones) (**Figure 3**). Interestingly, such transfers also appear to have been a source of evolutionary innovation for the host cell in the cryptophyte *Pyrenomonas helgolandii*, which acquired a novel type of actin gene from the nucleomorph of its endosymbiont. The full extent of such endosymbiont contributions to the biology of the cryptophyte and chlorarachniophyte host cells is unknown.

During the course of evolution, transferred genes encoding proteins that function within the multimembrane plastid acquired sequences encoding N-terminal targeting signals. These are more complex than the presequences of proteins targeted to primary plastids because they carry bipartite targeting signals composed of a signal peptide followed by a transit peptide (**Figure 3**). Such complex presequences are responsible for translocation of these proteins through two (in the case of nucleomorph- and periplastidal compartment-targeted proteins) or four plastid envelope membranes (in the case of plastid-targeted proteins). In the case of chlorarachniophytes, the electric charges of transit peptides could be responsible for the correct sorting of the imported proteins. Peptides of plastid-targeted proteins are positively charged, whereas those targeted to nucleomorph or periplastidal compartment have negative or near-neutral charges.

Driving Forces for Nucleomorph Reduction

Compared to their presumed algal progenitors, nucleomorph genomes are drastically reduced. This is manifested in a massive 'escape' of genes from nucleomorph to the host nucleus, which could have been driven by selection for faster genome replication. In contrast to polyploid mitochondrial and plastid genomes, however, nucleomorph chromosomes are not multicopy, which restricts intraorganellar replication rate competition between them. Nevertheless, competition favoring more rapid growth and division, either of complex plastids within a single host cell, or of whole host cells within their populations, could have resulted in the reduction of nucleomorph genomes. Selection for faster genome replication also could explain the loss of non-coding DNA, compression of intergenic spacers as well as intron loss and/or shortening described previously.

Large-scale transfers of nucleomorph genes to the host nuclear genome also are consistent with a Muller's ratchet effect, the irreversible accumulation of deleterious mutations due to absence of recombination combined with genetic drift in small, asexual nucleomorph populations. Presumably, any sexual processes in cryptophytes and chlorarachniophytes involve genetic exchange only between their nuclear genomes with nucleomorphs remaining intact within the two outer plastid envelope membranes that surround them. It also cannot be excluded that transfer of endosymbiont genes is favored by more efficient regulatory control in the host cell nucleus, thereby conferring greater integration with the whole cell's metabolism.

Why Do Nucleomorphs Still Persist?

Several hypotheses have been put forth to explain the maintenance of genomes in endosymbiont-derived organelles, such as mitochondria, plastids, and nucleomorphs. For example, the reduced genomes of mitochondria and plastids encode mainly redox proteins. According to the colocalization for redox regulation hypothesis, by sensing redox potential these organelles are able to initiate the quick synthesis of proper redox proteins to prevent deleterious consequences of oxidative stress. This hypothesis does not apply to the persistence of nucleomorph genomes, however, because redox reactions do not take place in the nucleomorph or the periplastidal compartment.

It was initially supposed that nucleomorphs are retained because they encode numerous proteins necessary for maintenance of plastid function; however, the complete sequences of nucleomorph genomes revealed that they encode only 17–30 such genes (**Table 1**). Moreover, chlorarachniophyte and cryptophyte genomes share only four genes (*cpn60*, *rpoD* (*sig2*), *clpP1*, and *clpP2*) encoding plastid-targeted proteins. These data indicate that, in principle, the nucleomorph genes for plastid proteins could be transferred to their host genomes, as has happened in many nucleomorph-lacking lineages of protists with eukaryotic alga-derived plastids (e.g., euglenids and haptophytes), but they are maintained in the nucleomorph for unknown reasons.

One hypothesis posits that the genes for plastid proteins still reside in nucleomorphs because they have very small introns that might not be recognizable by the nuclear splicing machinery. Although chlorarachniophyte nucleomorph genes do possess the smallest known spilceosomal introns, those of the cryptophyte *G. theta* have typical size for protist introns, whereas the nucleomorph of the cryptophyte *H. andersenii* is devoid of introns (**Table 1**). Thus, spliceosomal introns do not provide a good explanation for nucleomorph maintenance.

It is also possible that the rate of gene transfer is correlated with the number of cell organelles present, because frequent transfers depends on organelle lysis and DNA release. For example, plastid-to-nucleus gene transfer is undetectable in the green alga *C. reinhardtii* with one plastid, whereas it is astonishingly frequent in tobacco cells harboring several hundred plastids per cell. Interestingly, cryptophyte cells usually have a single plastid

and nucleomorph, and the outer plastid membrane is continuous with the nuclear envelope. An exception is the cryptophyte *C. paramecium*, in which two plastids and two nucleomorphs are present. In turn, each chlorarachniophyte cell harbors at most three or four plastid–nucleomorph complexes. In accordance with this hypothesis, chlorarachniophyte nucleomorphs contain fewer genes (including those for plastid proteins) than do cryptophyte nucleomorphs.

It is an unresolved question whether these miniaturized (or *bonsai*) nuclei still are undergoing reductive evolution and will eventually disappear, or whether they are moving toward a limit representing a final irreducible genome sizes. Additional studies involving more nucleomorph genomes from distantly related lineages of cryptophytes and chlorarachniophytes are required to answer this question.

See also: Base Composition; *Chlamydomonas reinhardtii*; Chromosome; Chromosome Number; Cytoplasm; Endosymbiont Theory; Evolution: Eukaryotes; Gene Number; Gene Rearrangement, Eukaryotic; Genome; Genome Organization; Genome Size; Housekeeping Gene; Horizontal Gene Transfer; Introns and Exons; Karyotype; Muller's Ratchet; Non-coding; Nucleus; Pre-mRNA Splicing; rDNA Amplification; Rearrangements; Ribosomal RNA; Ribosomes; Signal Sequence; Small Nuclear RNA; snRNPs; Splicing; Split Genes; Symbionts, Genetics of; Synteny (Syntenic Genes); Telomeres.

Further Reading

Archibald JM and Lane CE (2009) Going, going, not quite gone: Nucleomorphs as a case study in nuclear genome reduction. *The Journal of Heredity* 100: 582–590.

Cavalier-Smith T and Beaton MJ (1999) The skeletal function of non-genic nuclear DNA: New evidence from ancient cell chimaeras. *Genetica* 106: 3–13.

Douglas S, Zauner S, Fraunholz M, *et al.* (2001) The highly reduced genome of an enslaved algal nucleus. *Nature* 410: 1091–1096.

Gilson PR, Su V, Slamovits CH, *et al.* (2006) Complete nucleotide sequence of the chlorarachniophyte nucleomorph: Nature's smallest nucleus. *Proceedings of the National Academy of Sciences of the United States of America* 103: 9566–9571.

Hirakawa Y, Burki F, and Keeling PJ (2011) Nucleus- and nucleomorph-targeted histone proteins in a chlorarachniophyte alga. *Molecular Microbiology* 80: 1439–1449.

Hirakawa Y, Gile GH, Ota S, Keeling PJ, and Ishida K (2010) Characterization of periplastidal compartment-targeting signals in chlorarachniophytes. *Molecular Biology and Evolution* 27: 1538–1545.

Lane CE, van den Heuvel K, Kozera C, *et al.* (2007) Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. *Proceedings of the National Academy of Sciences of the United States of America* 104: 19908–19913.

McFadden GI (1993) Second-hand chloroplasts: Evolution of cryptomonad algae. *Advances in Botanical Research* 19: 189–230.

Moore CE and Archibald JM (2009) Nucleomorph genomes. *Annual Review of Genetics* 43: 251–264.

Patron NJ, Rogers MB, and Keeling PJ (2006) Comparative rates of evolution in endosymbiotic nuclear genomes. *BMC Evolutionary Biology* 6: 46.

Selosse M, Albert B, and Godelle B (2001) Reducing the genome size of organelles favours gene transfer to the nucleus. *Trends in Ecology & Evolution* 16: 135–141.

Tanifuji G, Onodera NT, and Hara Y (2010) Nucleomorph genome diversity and its phylogenetic implications in cryptomonad algae. *Phycological Research* 58: 230–237.

Tanifuji G, Onodera NT, Wheeler TJ, *et al.* (2011) Complete nucleomorph genome sequence of the nonphotosynthetic alga *Cryptomonas paramecium* reveals a core nucleomorph gene set. *Genome Biology and Evolution* 3: 44–54.