Size isn't everything: lessons in genetic miniaturisation from nucleomorphs

Paul R Gilson*, Uwe-G Maier† and Geoffrey I McFadden‡

Nucleomorphs are the vestigial nuclear genomes of eukaryotic algal cells now existing as endosymbionts within a host cell. Molecular investigation of the endosymbiont genomes has allowed important insights into the process of eukaryote/eukaryote cell endosymbiosis and has also disclosed a plethora of interesting genetic phenomena. Although nucleomorph genomes retain classic eukaryotic traits such as linear chromosomes, telomeres, and introns, they are highly reduced and modified. Nucleomorph chromosomes are extremely small and encode compacted genes which are disrupted by the tiniest spliceosomal introns found in any eukaryote. Mechanisms of gene expression within nucleomorphs have apparently accommodated increasingly parsimonious DNA usage by permitting genes to become co-transcribed or, in select cases, to overlap.

Addresses
*‡Plant Cell Biology Research Centre, School of Botany, University of Melbourne, Parkville 3052, Victoria, Australia
*e-mail: pgilson@rubens.its.unimelb.edu.au
†Universität Bayreuth, Bayreuth D-95440, Germany;
e-mail: maier@mailer.uni-marburg.de

Current Opinion in Genetics & Development 1997, 7:800–806

http://biomednet.com/elecref/0959437X0070800

Abbreviations

CCMP	culture collection of marine phytoplankton
ER	endoplasmic reticulum
ORF	open reading frame
rRNA	ribosomal RNA

Introduction

Eukaryotes are a promiscuous lot, readily striking up intimate relationships with other organisms in an attempt to flourish within nature's jungle. As with most opportunists, eukaryotes are not always original — often exploiting the inventions of other organisms by detaining them as endosymbionts. No greater example of this trend exists than the continuous quest by eukaryotic organisms to acquire the extremely useful process of photosynthesis.

A large number of photosynthetic eukaryotes — red algae, green algae (and their descendants, the land plants) and the glaucocystophytes — harbour plastids bound by two membranes [1]. These algae almost certainly acquired their plastids by engulfing and retaining cyanobacterial-like endosymbionts [2,3,4••,5]. Over time, the cyanobacterial-like endosymbionts were reduced to plastids and many of their genes were either lost or transferred to the eukaryotic nucleus (Figure 1). We call this process primary endosymbiosis and it seems likely that all plastids ultimately derive from a single successful instance of this process [5,6].

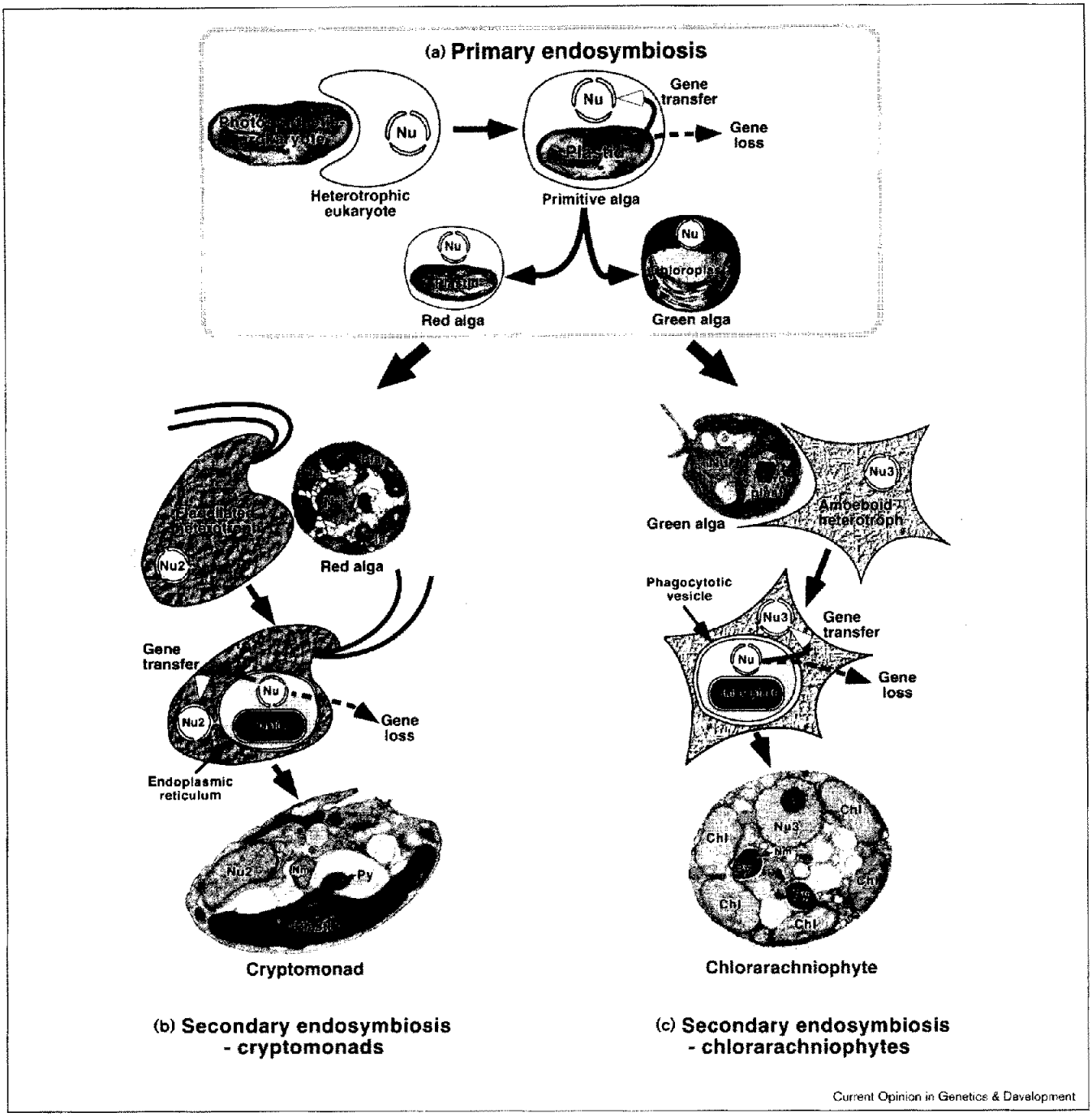
In addition to the algae mentioned above, there exist several major eukaryotic groups — in terms of diversity and biomass — that possess plastids surrounded by more than two membranes. Dinophyceae (dinoflagellates) and Euglenophyceae (e.g. *Euglena sp.*) contain plastids enveloped by three membranes, whereas heterokonts (e.g. brown algae, chrysophytes, and diatoms) and haptophytes (e.g. coccolithophorids) contain plastids bound by four membranes [6]. These multimembrane-bound plastids are believed to have been acquired second-hand from other unrelated eukaryotic algae [4••,6] rather than by direct inheritance through filial radiation after the primary cyanobacterial endosymbiosis. Referred to as secondary endosymbiosis because a eukaryotic alga is engulfed by a eukaryotic heterotroph and is retained as a photosynthetic endosymbiont, this process has apparently been a potent driver of eukaryotic cell diversity. Secondary endosymbiotic origins are invoked for six or more algal phyla [4••,7], comprising at least 37,600 species [8•]. It is believed that one of the extra membranes enveloping the original two-membrane-bound plastid is the residual plasma membrane of the captured algae. Other than the plastid, however, this extra membrane is the only physical remnant of the captured algae, making it difficult to substantiate these hypothetical scenarios of plastid acquisition [4••,7].

Fortunately, nature has provided us with examples of algae that retain unequivocal evidence of secondary endosymbioses. Cryptomonads and the chlorarachniophytes have plastids that not only have the extra membranes of secondary endosymbiosis but also retain a small volume of endosymbiont cytoplasm and a vestigial endosymbiont nucleus [7]. These diminutive nuclei, called nucleomorphs, contain the smallest known eukaryotic genomes, being only ~0.5 megabases in size. In partnership with collaborators (S Douglas, T Cavalier-Smith) we have commenced studies of these so-called 'bonsai genomes' [9•] to help unravel the secrets of eukaryote/eukaryote endosymbioses and illuminate some fundamental aspects of eukaryote genome organisation and composition.

Nucleomorphs

Early evidence that cryptomonad and chlorarachniophyte nucleomorphs were vestigial endosymbiont nuclei was provided by phylogenetic analyses of their ribosomal RNA (rRNA) genes. Analysis of the rRNA genes of the cryptomonad nucleomorph indicated that it was derived from a red alga (Figure 1) [10,11,12•,13•]. The host cell,

Figure 1



(a) Plastids are the descendants of a photosynthetic prokaryote which was engulfed by a heterotrophic eukaryote. **(b,c)** Primitive eukaryotic algae later diverged into different groups which were, in turn, engulfed by other eukaryotes. (b) The plastids of cryptomonads are descendants of red algal endosymbionts and reside inside the ER of a flagellated heterotroph. (c) Chlorarachniophytes, on the other hand, were created when an amoeboid heterotroph engulfed a green alga and enveloped it inside a modified phagocytic vesicle. Both the red and green algal endosymbionts are drastically reduced and contain vestigial nuclei that house tiny genomes – the nucleomorphs. Many non-essential genes have been lost from these genomes or have been transferred to the nuclei of the host cells (Nu2 and Nu3). Nu, nucleus of primitive alga and its descendants which is later reduced to a nucleomorph (Nm); Chl, plastid; Py, pyrenoid.

however, was shown to be related to *Goniomonas truncata*, a plastid-lacking heterotroph (Figure 1) [11,13•]. Similar analyses showed the chlorarachniophyte endosymbiont to be a green alga (Figure 1) [12•,13•] and the host cell to be a relative of the filose amoebae (Figure 1) [13•,14]. Thus, different hosts acquired different endosymbionts in two independent secondary endosymbioses to create cryptomonads and chlorarachniophytes. The two systems,

therefore, constitute excellent parallels for exploring the processes involved.

The organisation of nucleomorph DNA was first examined by isolating cryptomonad nucleomorphs [15]. Pulsed-field gel electrophoresis of DNA from isolated nucleomorphs identified their genomes as comprising three remarkably small linear chromosomes, each being only a few hundred kilobases in length (Figure 2) [16]. Curiously, each chromosome was found to carry nucleomorph-type rRNA genes [17]. Even more curiously, pulsed-field gel electrophoresis of total chlorarachniophyte DNA also showed that nucleomorphs of these algae contained three tiny chromosomes which, again, each carried rRNA genes (Figure 2) [18]. Subsequent analyses of a number of species have shown the haploid sizes of cryptomonad and chlorarachniophyte nucleomorph genomes range between 550–660 kb [19] and 380–455 kb respectively (P Gilson, unpublished data).

We believe that nucleomorph genomes are tiny because many redundant genes were lost post-endosymbiosis. As an endosymbiont, the engulfed alga has no need for most of the structures and biosynthetic pathways it originally possessed. In addition, some genes required to maintain the endosymbiont and its plastid appear to have been transferred to the host cell nucleus (Figure 1). Intriguingly, both cryptomonad and chlorarachniophyte nucleomorphs have remarkably similar karyotypes, despite the fact that they are derived from two completely unrelated endosymbioses (Figure 2) [12•,13•]. This similarity suggests that genome reduction processes have honed nucleomorphs towards a common minimal viable genetic architecture [20•,21••]. The minimal chromosome length of ~100 kb is similar to the minimum viable size defined for yeast chromosomes [22]. Convergence of the two nucleomorphs on three chromosomes is difficult to rationalise, however, as eukaryotes with only two [23] or even one [24] haploid chromosome are known.

What do nucleomorphs encode?

The plastid genomes of plants and algae typically encode a mere 10% of the proteins present in these organelles with the remainder being encoded by the cell's nucleus [4••]. It follows then that the nucleomorph — once the nucleus of a eukaryotic alga — should encode many of the proteins required by the endosymbiont's plastid. In addition, nucleomorphs would need to possess the many genetic housekeeping functions required to express these plastid proteins and replicate nucleomorph DNA [20•]. To date, sequence analyses of cryptomonad and chlorarachniophyte nucleomorph DNA have identified a plethora of housekeeping genes (Table 1) the functions of which include transcription, transcript processing, translation, protein processing, and cell cycle control [19,20•,21••,25]. Puzzlingly, precious few plastid protein genes have been found to date (possible candidates being the protease ClpP [21••] and the putative plastid division protein FtsZ [U-G Maier, unpublished data]). We already know of several plastid proteins encoded by the host cell nucleus (e.g. GAPDH [26•]; chlorophyll-binding proteins [J Deane, GI McFadden, unpublished data]; and the α subunit of phycoerythrin [27]) and it is likely that extensive transfer of nucleomorph genes to the host nucleus has occurred (Figure 1).

Many open reading frames (ORFs) have been identified within nucleomorph DNA. Some ORFs are novel and may perform nucleomorph-specific functions. Other ORFs exhibit extensive similarity to those from other eukaryotes. At present, neither cryptomonads nor chlorarachniophytes are amenable to genetic analyses and we anxiously await the elucidation of ORF function from more genetically tractable organisms such as *Saccharomyces cerevisiae*.

Nucleomorph genome architecture and expression

Nucleomorph genome architecture, like that of many other organellar genomes, is highly parsimonious with regard to DNA organisation and expression.

Telomeres and terminal inverted repeats

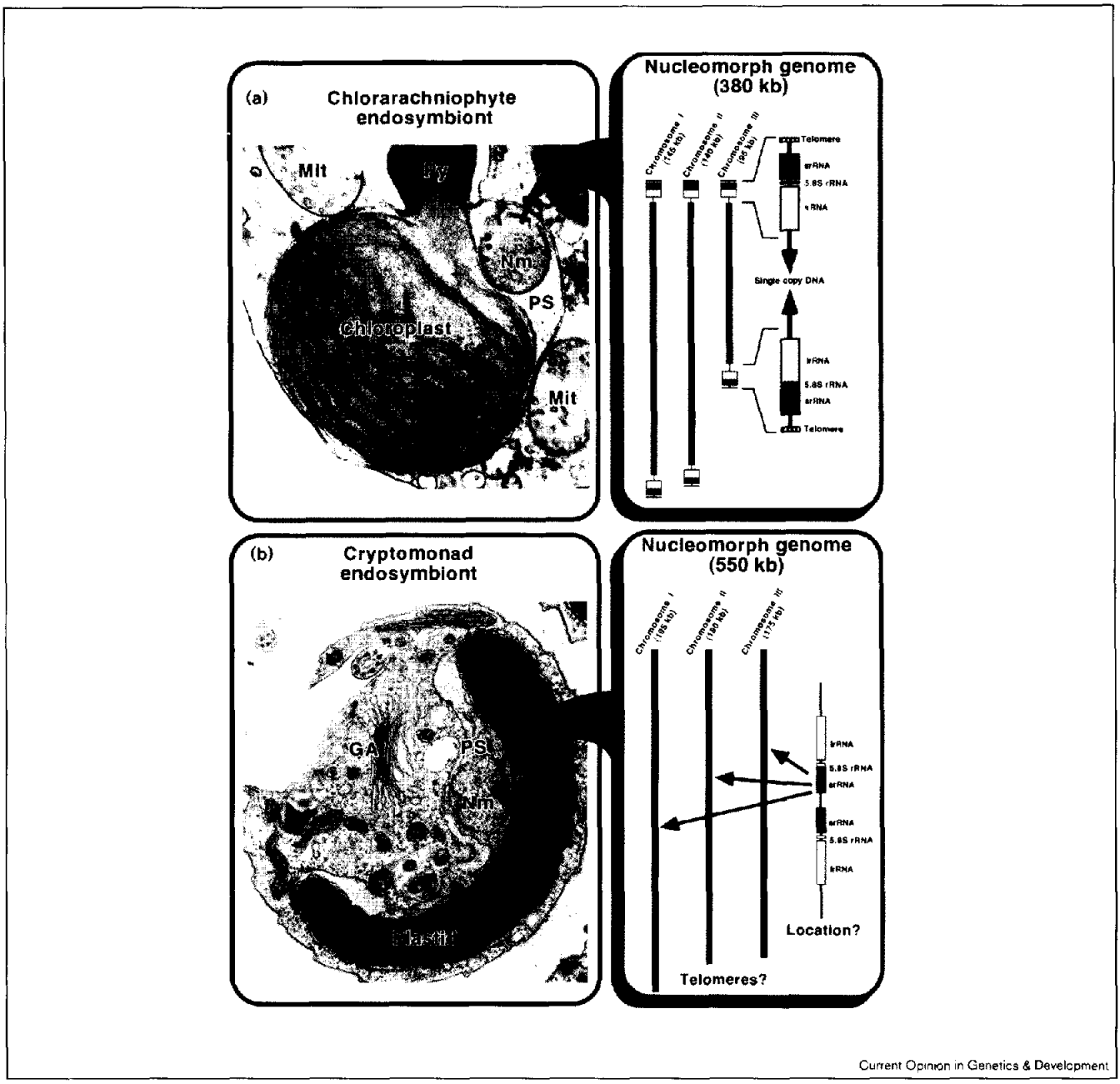
Chromosomes within the chlorarachniophyte nucleomorph are capped with telomeres comprising TCTAGGG repeats [20•,28], rather like those of plants and green algae [29,30]. Linked to these relatively short telomeres (of 25–45 repeats) are single rRNA gene operons (Figure 2). Together, the telomere and rRNA gene operon form an identical 8.5 kb cap on the termini of each nucleomorph chromosome (Figure 2) [20•,28]. We believe that the identity of nucleomorph rRNA genes is maintained by intra- and intermolecular recombination between the terminal repeats followed by gene conversion processes [28]. It is interesting that the chromosomal termini of some pathogenic eukaryotes such as *Plasmodium sp.* and *Trypanosoma sp.* are also recombinatorial hot-spots [31,32]; however, whereas recombination apparently keeps nucleomorph rRNA genes identical, terminus recombination in these pathogen chromosomes creates novel surface antigens to confuse the host immune system.

All three cryptomonad nucleomorph chromosomes encode rRNA gene operons but it is not known if they are terminally positioned and nothing is yet known about cryptomonad nucleomorph telomeres (Figure 2) [9•].

Compact genomes: operons and overlapping genes

Very little spacer DNA separates nucleomorph genes. The average spacer length within the chlorarachniophyte nucleomorph is a mere 100 bp and is only marginally greater in cryptomonads [9•,20•]. Reduction has apparently been so pervasive that the individual transcription of select genes has been abandoned in favour of contraction to an operon-like organisation. Northern blot and cDNA analyses indicate that three adjacent chlorarachniophyte nucleomorph genes are co-transcribed [21••] and many

Figure 2



(a) Chlorarachniophyte (CCMP 621) and **(b)** cryptomonad (*Guillardia theta*) endosymbionts with their nucleomorph genomes. (a) The chlorarachniophyte nucleomorph contains three small linear chromosomes which terminate in identical inverted repeats. Each repeat comprises a telomere linked to a single rRNA operon. (b) Cryptomonad nucleomorphs also contain three small rRNA-encoding chromosomes. Although the rRNA operons form inverted repeats, their position on the nucleomorph chromosome is unknown. Nm, nucleomorph; PS, periplastidal space; Py, pyrenoid; GA, Golgi apparatus, Mit, mitochondria.

more operons have been hypothesised [20•]. As yet, it has not been established how these polycistronic messengers are translated. There is no evidence of *trans*-spliced leader sequences ligated to the front of any nucleomorph mRNAs (P Gilson, unpublished data), as has been observed in other eukaryotes such as nematodes and trypanosomes which also have operons [33–35]. In addition to the compression of genes into operons, we have also found examples of overlap where a gene encoding a spliceosomal RNA lies within the 3′ untranslated region of a ribosomal protein gene [21••]. Whether this minimization of DNA content is caused by some active

Table 1

Comparison of chlorarachniophyte and cryptomonad nucleomorph genomes.

Features	Chlorarachniophytes	Cryptomonads
Location of endosymbiont	Inside a cytoplasmic vesicle	Inside lumen of ER
Number of chromosomes	Three (95, 140 and 145 kb)	Three (160, 170 and 185 kb)
Haploid nucleomorph genome size	380–455 kb	480–660 kb
Average AT content	70%	60%
Telomeres	$(TCTAGGG)_n$ repeats	Unknown
Spliceosomal introns	Numerous; 18, 19 or 20 nucleotides long	None detected to date
Operons	Present	None detected to date
Overlapping genes	Present	None detected to date
Functions common to both nucleomorphs		
Transcription	RNA polymerase subunit RPB8 [45]	TATA-box binding protein TFIIB, RNA polymerase subunit RPB1 [45]
RNA processing	snRNP E, Prp6, and U6	rRNA helicase, mRNA-capping protein
Translation	Ribosomal proteins S4 and S13; Eif 4α	Ribosomal proteins S4, L8, L27 and L9; tRNA Leu; tRNA His, tRNA Lys
Protein processing	Hsp70	Hsp70, ubiquitin-conjugating enzyme, proteosome α subunit, TCP1 subunit
Plastid functions	Protease ClpP	Plastid division protein FtsZ
Functions unique to one nucleomorph		
Cell cycle control	None detected to date	cdc48-like kinase
Signal transduction	None detected to date	G proteins
Cytoskeletal	None detected to date	α-tubulin

pressure to reduce genome size *per se* [36,37] or whether nucleomorph genomes are simply prone to DNA loss remains undetermined.

Nucleomorph coding sequences are highly enriched for AT nucleotides (60–75%) and gene spacers and introns are even further enriched (80–90%) [20•]. Such nucleotide bias occurs in many organelle genomes but its basis remains enigmatic [38•].

Pygmy introns and spliceosomes

Chlorarachniophyte nucleomorph genes contain a surprising surfeit of spliceosomal introns [21••]; but nucleomorph introns are the 'pygmies' of the spliceosomal intron world being smaller (only 18–20 nucleotides in length) than those found in any other eukaryote gene [4••,21••]. The occurrence of so many introns, albeit very small ones, in an otherwise compact genome is surprising. We suspect that the nucleomorph introns were not easily lost in evolution but they have became smaller in size; perhaps as small as an intron can get and still be recognised and removed by a splicesome.

Chlorarachniophyte nucleomorph introns possess the classic spliceosomal 5'GT and AG3' dinucleotide borders and a 'loose' consensus sequence [20•] and several components of the spliceosomal machinery are known to be encoded by the nucleomorph [21••]. We have utilised *in situ* hybridisation and immunolocalisation techniques to verify the presence of spliceosomes within the chlorarachniophyte nucleomorph and its seems that many chlorarachniophyte nucleomorph genes (Table 1) may be devoted to processing RNA transcripts [21••]. To date, no introns have been discovered within the cryptomonad nucleomorph [9•]. As cryptomonad nucleomorphs are derived from red algal endosymbionts and the nuclear genes of red algae are generally intron-poor [39], the cryptomonad nucleomorph genome will probably also be intron-poor.

Given the apparently inexorable trend towards reduction, one is tempted to ask why the nucleomorphs of cryptomonads and chlorarachniophytes have not disappeared entirely, as is proposed for most cases of secondary endosymbiosis [4••,7,40••]. Such loss of the nucleomorph

presumably requires transfer of all essential genes into the host nucleus and it is possible that there has not yet been sufficient time for this to occur in cryptomonads and chlorarachniophytes. Alternatively, it is possible that some time post-endosymbiosis the transfer of nucleomorph genes was halted as they became increasingly weird by developing aberrant regulatory sequences, biased nucleotide compositions and diminutive intronic sequences. It is hard to imagine that such aberrant genes could have easily been accommodated by host-cell genomes and the nucleomorph may have been fixed in its present form for some time now.

Host–endosymbiont relations: protein transport

Cryptomonad endosymbionts are located within the endoplasmic reticulum (ER) of the host cell whereas chlorarachniophyte endosymbionts are situated inside cytoplasmic vesicles (Figure 1). Analyses of genes for nuclear-encoded plastid proteins in cryptomonads suggests the presence of a bipartite leader which comprises a typical ER signal peptide followed by a plastid transit peptide [26•,27]. This implies that nuclear-encoded plastid proteins are cotranslationally inserted into ER from where they migrate into the plastid courtesy of their transit peptides. How these proteins actually transverse the residual plasma membrane of the endosymbiont to arrive at the plastid is unknown. Nuclear genes encoding chlorarachniophyte plastid proteins also possess amino-terminal extensions (J Deane, G McFadden, unpublished data) and their transport mechanisms are being investigated. Unravelling the complex mechanisms by which four genomes — nucleus, nucleomorph, plastid, and mitochondrion — act in concert to form the blueprint of a eukaryotic cell remains a major challenge.

Conclusions

At present, we are sequencing the nucleomorph genomes via a 'shotgun cloning' strategy and eventually predict that each genome could encode ~300 genes. At first glance, this number is remarkably similar to the estimates of the minimal genome size [41•,42] and we were initially hopeful that sequence analysis of nucleomorphs could reveal a distillation of basic eukaryotic functions. Nucleomorphs are unlikely to be autonomous, however, and they probably rely on host cell sources for many building blocks and even complete macromolecules. Indeed, just as plastid genomes cannot tell us everything about photosynthesis, or mitochondrial genomes everything about respiration, nucleomorph genomes cannot tell us everything about nuclei. Nevertheless, nucleomorphs provide a fascinating window into the genome-moulding forces of endosymbiosis. Comparative genomics with plastids [6] and mitochondria [43•,44•] have revealed some marvellous principles. Fortunately, we know of at least two nucleomorphs and the comparison of their residual gene complements will also be fascinating.

Of great interest also is the means by which proteins synthesised in the host cell are transported to the endosymbiont and its chloroplast. We are presently identifying such proteins and are endeavouring to dissect the intracellular route they travel. It will be interesting to compare and contrast the amino-terminal targeting sequences of host cell synthesised proteins destined to reside in the endosymbiont's cytoplasm and nucleus compared to those that are further despatched on to the chloroplast.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Cavalier-Smith T: **Membrane heredity, symbiogenesis, and the multiple origins of algae.** In *Biodiversity and Evolution.* Edited by Arai R, Kato M, Doi Y. Tokyo: National Science Museum; 1995:75-114.
2. Howe C, Beanland T, Larkum A, Lockhart P: **Plastid origins.** *Trends Ecol Evol* 1992, **7**:378-383.
3. Martin W, Sommerville C, Gor SL-D: **Molecular phylogenies of plastid origins and algal evolution.** *J Mol Evol* 1992, **35**:385-404.
4. •• Palmer JD, Delwiche CF: **Second-hand chloroplasts and the case of the disappearing nucleus.** *Proc Natl Acad Sci USA* 1996, **93**:7432-7435.

This poignant review slots the results of nucleomorph research into the broader context of eukaryotic cell evolution, namely in the methods employed by eukaryotes to acquire the power of photosynthesis by the endosymbiosis of cyanobacteria and eukaryotic algae. The review also describes recent advances in phylogenetic techniques used to resolve the evolution of rapidly evolving eukaryotic genomes.

5. Bhattacharya D, Medlin L: **The phylogeny of plastids: a review based on comparison of small-subunit ribosomal RNA coding regions.** *J Phycol* 1995, **31**:489-498.
6. Reith M: **Molecular biology of rhodophyte and chromophyte plastids.** *Annu Rev Plant Physiol Plant Mol Biol* 1995, **46**:459-575.
7. McFadden GI, Gilson PR: **Something borrowed, something green: lateral transfer of chloroplasts by secondary endosymbiosis.** *Trends Ecol Evol* 1995, **10**:12-17.
8. • Liepe D: **Biodiversity, genomes and DNA sequence databases.** *Curr Opin Genet Dev* 1996, **6**:686-691.

This is a timely review which serves to remind us that the species sampled within the molecular sequence databases are highly biased towards relatively few species. Most molecular data recorded – for example, from humans and other mammals – only cover a narrow window of the evolutionary spectrum. Recently, however, the number of database entries derived from ancient eukaryotic groups, bacteria, and archaea have been growing rapidly – no doubt aided by the initiation of several new genome projects of these organisms.

9. • McFadden GI, Gilson PR, Douglas SE, Cavalier-Smith T, Hofmann CJB, Maier U-G: **Bonsai genomics: sequencing the smallest eukaryotic genomes.** *Trends Genet* 1997, **13**:46-49.

The establishment of projects to sequence the nucleomorph genomes of two eukaryotic algae is described in this review along with some recent molecular data generated from these projects.

10. Douglas SE, Murphy CA, Spencer DF, Gray MW: **Cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes.** *Nature* 1991, **350**:148-151.
11. McFadden GI, Gilson PR, Hill DRA: ***Goniomonas*: rRNA sequences indicate that this phagotrophic flagellate is a close relative of the host component of cryptomonads.** *Eur J Phycol* 1994, **29**:29-32.
12. • Van de Peer Y, Rensing SA, Maier U-G, deWachter R: **Substitution rate calibration of small subunit rRNA identifies chlorarachniophyte endosymbionts as remnants of green algae.** *Proc Natl Acad Sci USA* 1996, **93**:7732-7736.

The rapidly evolving highly divergent gene sequences of some eukaryotes have long confounded standard phylogenetic methods. The authors of this paper describe two recent methodological advances in phylogenetic analysis and successfully apply them to resolve the origin of chlorarachniophyte endosymbionts.

13. Cavalier-Smith T, Couch JA, Thorsteinsen KE, Gilson PR, Deane JA, Hill DRA, McFadden GI: **Cryptomonad nuclear and nucleomorph 18S rRNA phylogeny.** *Eur J Phycol* 1996, **31**:315-328.

For phycological buffs only; the authors of this paper take all known cryptomonad ribosomal RNA gene sequences and try to reconstruct some of the major evolutionary events that have influenced the creation of modern cryptomonads.

14. Bhattacharya D, Helmchen T, Melkonian M: **Molecular evolutionary analysis of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphidae and the Chlorarachniophyta.** *J Eukaryotic Microbiol* 1995, **42**:65-69.

15. Hansmann P, Eschbach S: **Isolation and preliminary characterization of the nucleus and the nucleomorph of a cryptomonad, *Pyrenomonas salina*.** *Eur J Cell Biol* 1990, **52**:373-378.

16. Maier U-G, Hofmann C, Eschbach S, Wolters J, Igloi G: **Demonstration of nucleomorph-encoded eukaryotic small subunit ribosomal RNA in cryptomonads.** *Mol Gen Genet* 1991, **230**:155-160.

17. McFadden GI, Gilson PR, Douglas SE: **The photosynthetic endosymbiont in cryptomonad cells produces both chloroplast and cytoplasmic-type ribosomes.** *J Cell Sci* 1994, **107**:649-657.

18. McFadden GI, Gilson PR, Hofmann CJ, Adcock GJ, Maier U-G: **Evidence that an amoeba acquired a chloroplast by retaining part of an engulfed eukaryotic alga.** *Proc Natl Acad Sci USA* 1994, **91**:3690-3694.

19. Rensing SA, Goddemeier M, Hofmann CJB, Maier U-G: **The presence of a nucleomorph *hsp70* is a common feature of Cryptophyta and Chlorarachniophyta.** *Curr Genet* 1994, **26**:451-455.

20. • Gilson PR, McFadden GI: **Good things in small packages: the tiny genomes of chlorarachniophyte endosymbionts.** *Bioessays* 1997, **19**:167-173.

This is a highly colourful review, with lots of pretty pictures, that summarizes the history of chlorarachniophyte endosymbiont research and provides a molecular update of the nucleomorph genome sequencing project.

21. •• Gilson PR, McFadden GI: **The miniaturized nuclear genome of a eukaryotic endosymbiont contains genes that overlap, genes that are contranscribed, and the smallest known spliceosomal introns.** *Proc Natl Acad Sci USA* 1996, **93**:7737-7742.

This is the first detailed paper that lists some of the molecular data emerging from the efforts to sequence the nucleomorph genome of chlorarachniophyte algae. In addition to simple gene data, we describe the highly parsimonious DNA usage employed by these miniaturised genomes such as overlapping and co-transcribed genes and the reduction of spliceosomal introns to the smallest size found within any eukaryote.

22. Murray A, Szostak J: **Chromosome segregation in mitosis and meiosis.** *Annu Rev Cell Biol* 1985, 1:289-315.

23. Moore G: **Cereal genome evolution: pastoral pursuits with 'Lego' genomes.** *Curr Opin Genet Dev* 1995, **5**:717-724.

24. Crosland M, Crozier R: ***Myrmecia pilosula*, an ant with only one pair of chromosomes.** *Science* 1986, **231**:1278.

25. Hofmann CJ, Rensing SA, Hauber MM, Martin WF, Muller SB, Couch J, McFadden GI, Igloi GL, Maier UG: **The smallest known eukaryotic genomes encode a protein gene: towards an understanding of nucleomorph functions.** *Mol Gen Genet* 1994, **243**:600-604.

26. • Liaud M-F, Brandt U, Scherzinger M, Cerff R: **Evolutionary origin of cryptomonad microalgae: two novel GAPDH genes as potential markers of ancestral endosymbiont and host cell components.** *J Mol Evol* 1997, **44(Suppl)**:S28-S37.

A description of nuclear genes encoding GAPDH proteins targeted into the cryptomonad plastid. Amino-terminal leaders comprising a signal peptide followed by a putative transit peptide similar to that found in red alga are identified. The signal peptide may direct co-insertional translation of the host cell synthesised protein to the ER. The protein migrates through the ER and, by an as yet undetermined mechanism, is transported into the endosymbiont's cytoplasm. The transit peptide component of the amino-terminal leader possibly directs the import of the protein into the plastid.

27. McFadden GI, Gilson PR: **What's eating Eu? The role of eukaryote/eukaryote endosymbioses in plastid origins.** *Endocyt Cell Res* 1997, in press.

28. Gilson PR, McFadden GI: **The chlorarachniophyte: a cell with two different nuclei and two different telomeres.** *Chromosoma* 1995, **103**:635-641.

29. Wu K-S, Tanksley SD: **Genetic and physical mapping of telomeres and macrosatellites of rice.** *Plant Mol Biol* 1993, **22**:861-872.

30. Petracek M, Lefebvre P, Silflow C, Berman J: **Chlamydomonas telomere sequences are A+T-rich but contain three consecutive GC base pairs.** *Proc Natl Acad Sci USA* 1990, **87**:8222-8226.

31. Feagin J, Lanzer M: **The three genomes of *Plasmodium*.** In *Molecular Biology of Parasitic Protozoa*. Edited by Smith D, Parsons M. Oxford: IRL Press; 1996:35-54.

32. Donelson J: **Genome research and evolution in trypanosomes.** *Curr Opin Genet Dev* 1996, **6**:699-703.

33. Blumenthal T, Spieth J: **Gene structure and organization in *Caenorhabditis elegans*.** *Curr Opin Genet Dev* 1996, **6**:692-698.

34. Blumenthal T: ***Trans*-splicing and polycistronic transcription in *Caenorhabditis elegans*.** *Trends Genet* 1995, **11**:132-136.

35. Vanhamme L, Pays E: **Control of gene expression in trypanosomes.** *Microbiol Rev* 1995, **59**:223-240.

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

37. Howe C: **RNA polymerases and plastid evolution.** *Trends Plant Sci* 1996, 1:323-324.

38. • Moran NA: **Accelerated evolution and Muller's rachet in endosymbiotic bacteria.** *Proc Natl Acad Sci USA* 1996, **93**:2873-2878.

The genomic parallels between nucleomorphs and the organellar genomes of mitochondria and chloroplasts are remarkable, with compact gene arrangements, biased nucleotide compositions and accelerated evolutionary rates featuring prominently. We can now add *Buchnera spp.* to this list for this paper shows that the AT-rich genomes of these bacterial endosymbionts of aphids are evolving far more rapidly than their free-living relatives.

39. Liaud M-F, Brandt U, Cerff R: **The marine red alaga *Condrus crispus* has a highly divergent β-tubulin gene with a characteristic 5' intron: functional and evolutionary implications.** *Plant Mol Biol* 1995, **28**:313-325.

40. •• Palmer J: **Organelle genomes: going, going, gone!** *Science* 1997, **275**:790-791.

Commentary on the pervasive trend for transfer of genes from organelle genomes such as mitochondria, plastids, and nucleomorphs into the host nucleus. The author evaluates the emerging evidence suggesting that the trichomonad hydrogenosome is a modified mitochondrion which lacks a genome.

41. • Koonin E, Mushegian A: **Complete genome sequences of cellular life forms: glimpses of theoretical evolutionary genomics.** *Curr Opin Genet Dev* 1996, **6**:757-762.

As the list of completed genomes grows, comparative genomic analysis will provide a powerful means of understanding genome evolution and the diversification of lifeforms. This review provides a stimulating comparison between the two recently completed bacterial genomes and *S. cerevisiae*, the only fully sequenced eukaryote to date.

42. Itaya M: **An estimation of minimal genome size required for life.** *FEBS Lett* 1995, **362**:257-260.

43. • Palmer J: **The mitochondrion that time forgot.** *Nature* 1997, **387**:454-455.

A "News and views" piece comparing the residues in different mitochondrial genomes. This is a powerful advocacy of comparative genomics for defining principles and trends in intracellular gene transfer.

44. • Lang B, Berger G, O'Kelly C, Cedergren R, Golding G, Lemieux C, Sankoff D, Turmel M, Gray M: **An ancestral mitochondrial DNA resembling a eubacterial genome in miniature.** *Nature* 1997, **387**:493-497.

The complete mitochondrial DNA sequence from a flagellate protist; this unusual mitochondrion harbours more genes than any characterised to date. Relics from its bacterial ancestry include eubacterial-type RNA polymerase, eubacterial-type translation factors, conservation of ancient operon structure, and Shine-Dalgarno sequences on the mRNAs.

45. Shpakovski GV, Acker J, Wintzerith M, Lacroix J-F, Thuriaux P, Vigneron M: **Four subunits that are shared by the three classes of RNA polymerase are functionally interchangeable between *Homo sapiens* and *Saccharomyces cerevisiae*.** *Mol Cell Biol* 1995, **15**:4702-4710.