European Journal of Phycology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713725516

Cryptomonad nuclear and nucleomorph 18S rRNA phylogeny

T. Cavalier-Smith [a]; J. A. Couch [ab]; K. E. Thorsteinsen [a]; P. Gilson [c]; J. A. Deane [c]; D. R. A. Hill [d]; G. I. Mcfadden [c]
[a] Evolutionary Biology Program, Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, Vancouver, Canada
[b] Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA
[c] Plant Cell Biology Research Centre, School of Botany, University of Melbourne, Parkville, Australia
[d] School of Botany, University of Melbourne, Parkville, Australia

Online Publication Date: 01 November 1996
To cite this Article: Cavalier-Smith, T., Couch, J. A., Thorsteinsen, K. E., Gilson, P., Deane, J. A., Hill, D. R. A. and Mcfadden, G. I. (1996) 'Cryptomonad nuclear and nucleomorph 18S rRNA phylogeny', European Journal of Phycology, 31:4, 315 - 328
To link to this article: DOI: 10.1080/09670269600651541
URL: http://dx.doi.org/10.1080/09670269600651541

Cryptomonad nuclear and nucleomorph 18S rRNA phylogeny

T. CAVALIER-SMITH[1], J. A. COUCH[1,*], K. E. THORSTEINSEN[1], P. GILSON[2], J. A. DEANE[2], D. R. A. HILL[3] AND G. I. MCFADDEN[2]

[1]*Evolutionary Biology Program, Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, Vancouver, B.C., Canada V6T 1Z4*
[2]*Plant Cell Biology Research Centre, School of Botany, University of Melbourne, Parkville, 3052, Australia*
[3]*School of Botany, University of Melbourne, Parkville, 3052, Australia*

(Received 15 April 1996; accepted 15 June 1996)

Nuclear and nucleomorph 18S ribosomal RNA genes from six cryptomonads were amplified by the polymerase chain reaction and sequenced. Phylogenetic trees were constructed by distance, parsimony, and maximum likelihood methods for all available cryptomonad nuclear and nucleomorph 18S rRNA sequences. Nuclear and nucleomorph trees are largely congruent and clearly disprove the idea of polyphyletic origins for cryptomonad chloroplasts. Both show the leucoplast-containing *Chilomonas* as the sister to all photosynthetic cryptomonads. Using 11 cryptomonad nucleomorph sequences gives more convincing evidence than before that cryptomonad nucleomorphs originated from a red alga and are not specifically related to *Chlorarachnion* nucleomorphs. Both trees show as a clade the genera with nucleomorphs embedded in a chloroplast-envelope invagination into the pyrenoid (*Storeatula, Rhinomonas, Rhodomonas*). This monophyly of embedded nucleomorphs supports the recent creation of the order Pyrenomonadales for such cryptomonads. Cryptomonads ancestrally having free nucleomorphs are much more diverse. *Komma* and *Chroomonas*, with the blue accessory pigment phycocyanin, form a clade, as do *Guillardia* and *Cryptomonas* Φ, both with the red pigment phycoerythrin. The nucleomorph trees strongly show the blue *Chroomonas/Komma* clade as sister to all red-pigmented genera, but nuclear sequences support this weakly, if at all, being sensitive to taxon sampling. Red and blue cryptomonads probably diverged early by differential pigment loss. Nuclear sequences provide no clear evidence for the nature of the host that engulfed the ancestral symbiont. Our nuclear trees using an extensive selection of outgroups, and recent evidence from chloroplast DNA, are consistent with but do not positively support the view that the closest relatives of Cryptista (i.e. Cryptophyceae plus Goniomonadea) are the Chromobiota (i.e. Haptophyta plus Heterokonta, the latter including heterokont algae—phylum Ochrophyta), and that Cryptista and Chromobiota are appropriately classified as subkingdoms of the kingdom Chromista. Maximum likelihood often groups *Goniomonas* with *Chilomonas*, suggesting that *Goniomonas* may have lost both nucleomorph and plastid and that the cryptist common ancestor was photosynthetic.

Key words: Chromista, endosymbiosis, *Geminigera, Komma caudata*, nucleomorphs, plastid origins, *Rhinomonas pauca, Rhodomonas abbreviata*, rRNA phylogeny, *Storeatula major*

Introduction

Cryptomonads, though ecologically very important (Klaveness, 1988), are neglected because their taxonomy is less well developed at the generic, familial and ordinal levels than that of any other major algal group (Santore, 1984; Klaveness, 1985; Santore & Leedale, 1985; Hill, 1991*a*, *b*; Novarino & Lucas, 1993*a*). The backwardness of cryptomonad systematics stems from the great morphological uniformity of the class in the light microscope, making electron microscopy essential for defining taxa (Santore, 1982*a*, 1984, 1985, 1986, 1987; Hill, 1991*a*, *b*; Hill & Wetherbee, 1988, 1989, 1990; Novarino, 1991*a*, *b*; Novarino & Lucas, 1993*a*,*b*). Though only a small number of taxa have yet been studied in the electron microscope, such studies have recently led to major changes in the circumscription of key genera (Hill, 1991*a*, *b*: Hill & Wetherbee, 1989; Novarino *et al.*, 1994) and to the establishment of several new genera (Hill & Wetherbee, 1986, 1988, 1990). Molecular methods can provide an independent test of these systematic innovations and also a means of extending them to new strains of cryptomonads more rapidly than is possible by electron microscopy alone. To provide an initial basis for such an approach we have sequenced the small subunit ribosomal RNA genes of several cryptomonad strains that have recently been subject to such reassignments based on electron microscopy and used them for a first molecular phylogenetic analysis of the cryptomonads.

Cryptomonads, like all other chromistan algae (Cavalier-Smith, 1986, 1995), acquired their chloroplasts through a secondary symbiotic event: the engulfment of a photosynthetic eukaryote by a phagotrophic eukaryote (Greenwood *et al.*, 1977; Whatley *et al.*, 1979; Cavalier-Smith, 1986, 1989; Douglas *et al.*, 1991; Sitte, 1993; McFadden, 1993; McFadden & Gilson, 1995). All cryptomonads still contain a relic of the nucleus of the photosynthetic symbiont, known as the nucleomorph

This paper is dedicated to the memory of the late Denis Greenwood, the discoverer of the cryptomonad nucleomorph.

Correspondence to: T. Cavalier-Smith.

*Present address: Department of Genetics, Washington University School of Medicine, 4444 Forest Park Drive, St Louis, MO 63108, USA.

(Greenwood, 1974; Greenwood *et al.*, 1977; Gillott & Gibbs, 1980; Santore, 1982*b*), which contains DNA (Ludwig & Gibbs, 1985; Hansmann *et al.*, 1986) arranged in three linear chromosomes (Eschbach *et al.*, 1991*a*), and undergoes a non-mitotic division (Morrall & Greenwood, 1982; McKerracher & Gibbs, 1982). The colourless zooflagellate genus *Goniomonas*, though indubitably related to cryptomonads (McFadden *et al.*, 1994*a*), has recently been removed from the class Cryptomonadea (syn. Cryptophyceae) and placed in its own class Goniomonadea within the phylum Cryptista (syn. Cryptophyta), on account of the absence of both nucleomorph and plastids, and differences in ciliary hairs and cell shape (Cavalier-Smith, 1989, 1993*a*), and so will not be referred to here as a cryptomonad. Some authors accept the placement of *Goniomonas* in a separate order Goniomonadales, but not in a separate class (Novarino & Lucas, 1993*a*). For the most recent discussions of cryptist nomenclature see Novarino & Lucas (1995) and Cavalier-Smith (1995).

Cryptomonad nucleomorphs and chloroplasts are located together within a periplastic membrane, the relic of the endosymbiont's plasma membrane (Whatley *et al.*, 1979; Cavalier-Smith, 1982, 1986), which also encloses starch grains and 80S ribosomes (McFadden, 1993). The periplastid membrane and the enclosed nucleomorph, chloroplast and 80S ribosomes are located within the lumen of the perinuclear cisterna/rough endoplasmic reticulum. The 18S rRNA of the 80S ribosomes within the periplastid membrane are encoded by the nucleomorph 18S rRNA gene whereas those of the cytosol are encoded by the nucleus (McFadden *et al.*, 1994*b*).

Cryptomonads are therefore the only chromists in which it is possible to study directly the nuclear genome of the algal part of the cellular chimaera and to use its genes, as well as those of the former host component of the cell, to study molecular phylogeny. The nucleus and nucleomorph of the same cell can provide important internal controls for each other to help guard against systematic bias in trees, which can be a real problem in molecular phylogeny (Swofford & Olsen, 1990). We have therefore sequenced both the nuclear and the nucleomorph 18S rRNA genes of six diverse cryptomonads in order to see whether or not the trees derived from the two separate genomes are congruent. Together with previous data from five other cryptomonads (Douglas *et al.*, 1991; Eschbach *et al.*, 1991*a*; Maier *et al.*, 1991; Cavalier-Smith *et al.*, 1994) this enables us to clarify several important aspects of cryptomonad phylogeny. We particularly wished to examine two key questions.

The first question is the taxonomic significance of the location of the nucleomorph in the cell (Santore, 1982*b*). In three genera [*Rhodomonas* (syn. *Pyrenomonas*), *Storeatula*, *Rhinomonas*] the nucleomorph is located within an invagination of the chloroplast envelope into a deep groove within the pyrenoid. This unusual location of the nucleomorph has facilitated the purification of the nucleomorph and pyrenoid as a single organellar complex in these genera (Hansmann & Eschbach, 1990; McFadden, 1993; McFadden & Gilson, 1995). In most cryptomonad genera, however, the nucleomorphs lie freely in the periplastid space and are not embedded in the pyrenoid, and their nucleomorphs have not been purified. Whether nucleomorphs are embedded in the pyrenoid or not is an important taxonomic character, which has been used in the more precise taxonomic separation of *Rhodomonas* Karsten 1898 *emend.* Hill & Wetherbee (1989) (syn. *Pyrenomonas* Santore 1984), which has the pyrenoid embedded (Santore, 1986; Hill & Wetherbee, 1989; Novarino, 1991*a*), from *Cryptomonas* (pyrenoid not embedded: Hill, 1991*a*; Novarino, 1991*a*). The shared presence of an embedded nucleomorph was part of the reason for the recent synonymization of *Pyrenomonas* and *Rhodomonas* (Hill & Wetherbee, 1989); whether the relatively new name *Pyrenomonas* (Santore, 1984; Novarino, 1991*a*) or a refinement of the old name *Rhodomonas* (Hill & Wetherbee, 1989) is the more appropriate name for the genus has been controversial. More recently, Novarino & Lucas (1993*a*) used this character as a basis for defining a new cryptophyte order, Pyrenomonadales. In order to test the validity of this order we analysed all three cryptomonad genera with embedded nucleomorphs and a selection of cryptomonads without an embedded nucleomorph. In view of the presence of pyrenoid-embedded nucleomorphs in *Chlorarachnion* (McFadden *et al.*, 1994*c*) and the remote possibility that the *Chlorarachnion* nucleomorph might be related to that of cryptomonads (Cavalier-Smith, 1993*a*, 1994; Sitte, 1993; Cavalier-Smith *et al.*, 1994; Gilson & McFadden, 1995) we also wished to determine whether an embedded nucleomorph was ancestral or secondarily derived within cryptomonads.

The second important question concerned the significance of the fact that some cryptomonads have the red antennal pigment phycoerythrin and others have the blue pigment phycocyanin (Hill & Rowan, 1989). Some specialists have considered this difference to be of minor taxonomic importance, whereas others have given it considerable weight in the grouping of species into higher level taxa (Santore, 1987; Hill, 1991*b*; Novarino & Lucas, 1993*a*). Margulis & Schwartz (1988) even proposed that differently pigmented cryptomonads acquired their chloroplasts in several independent symbioses. By including both blue- and red-pigmented species in our analysis we have been able definitively to disprove this polyphyletic hypothesis, and to show that all cryptomonads acquired their chloroplasts from the same, probably red algal, source. Red- and blue-pigmented cryptomonads clearly diverged by the differential loss of phycocyanin and phycoerythrin pigments, both of which would have been present in the ancestral unicellular red algal endosymbiont.

Materials and methods

The species and strains used and the GenBank accession

numbers of the resulting nuclear and nucleomorph sequences are shown in Table 1. DNA was isolated using a modification of the method of Sepsenwol (1973) for the unidentified cryptomonad, and by the method of Rochaix *et al.* (1988) for all the other species. Small subunit rRNA genes were amplified by the polymerase chain reaction (PCR), using universal eukaryotic primers (Medlin *et al.*, 1988) for the unidentified species and *Storeatula major* and for the nuclear genes of all species. For *Storeatula*, amplification was done using DNA extracted from purified isolated nuclei and nucleomorphs (McFadden, 1993), whereas for the other species total DNA was used. The *Rhinomonas pauca, Rhodomonas abbreviata, Cryptomonas* Φ, *Komma* and *Geminigera* nucleomorph genes were amplified using a nucleomorph-specific primer (5′ CAG TAG TCA TAT GCT TGT CTT AAG 3′) in conjunction with a universal primer. PCR conditions for the unidentified species were as follows: an initial denaturing step of 95°C for 8 min, annealing at 65°C for 4 min, and extending at 72°C for 3 min; 28 cycles of 93°C for 2 min, 65°C for 2 min, and ramping through the cycles at 62°C for 3–7 min. PCR conditions for the other species were as described previously (McFadden *et al.*, 1995). The unidentified cryptomonad's nucleomorph PCR product contained no internal *Pst* I or *Sal* I sites (sites contained in the termini of the primers) and was cloned, using these sites, into pGEM3Zf (Promega). All the other genes were similarly cloned using *Pst* I and *Bam* HI into either pGEM3Zf or pGEM5Zf.

All clones were sequenced with conserved internal primers (Elwood *et al.*, 1985), using an ABI 373A automated sequencer and the dideoxy terminator sequencing kit (as instructed by ABI). Sequences were edited on a Sun workstation using the trace editor Ted (Gleeson & Hillier, 1991) or Sequencher 3.0 for the Macintosh (Gene Codes Inc.). Contiguous sequences were assembled from sequences of both strands of each clone. The sequences were aligned manually with over 300 other eukaryotic 18S rRNA sequences obtained from GenBank (or our own unpublished sequences in the case of *Prymnesium* and *Pavlova*) using the Genetic Data Environment software (Smith *et al.*, 1994). For outgroups we selected 66 non-cryptophyte species representing the greatest diversity available of sequences of neokaryotes (i.e. eukaryotes branching above Euglenozoa: Cavalier-Smith, 1993*a*). Phylogenetic trees were calculated using Phylip v.3.5 (Felsenstein, 1993) for parsimony, distance methods, maximum likelihood and neighbor-joining distance methods. The fastDNAml package v.1.0.6 (Olsen *et al.*, 1994) was used for maximum likelihood on a Sunsparc 10 workstation with four 80 MHz processors. All trees were calculated with a random order of addition of taxa. All parsimony and likelihood trees were calculated with equal weighting of the 1632 most conserved alignment positions. Maximum likelihood used empirical base frequencies and the fast add option. All Kimura and maximum likelihood calculations used a transition/transversion ratio of 2. Bootstrap calculations, which give a numerical indication of the degree to which different parts of the molecule support a given clade (not, as sometimes incorrectly thought, of its 'correctness'), were based on 100, 200 or 1000 pseudoreplicates depending on the method and size of the data set.

Results

In all six species the nucleomorph sequences were several hundred nucleotides longer (range 1839–2054 nucleotides) than the nuclear sequences (range 1747–1774 nucleotides, including amplification primers), owing to lengthy (often very AT-rich) insertions in less well conserved regions. Long insertions in these regions

Table 1. Strains used for DNA extraction and sequence accession numbers

Species	Strain	GenBank nucleus	GenBank nucleomorph
Komma caudata Hill (1991*b*)	MUCC Cr#10	U53122	U53121
Geminigera cryophila (Taylor & Lee) Hill (1991*a*)	CSIRO Marine Culture Collection (CS-138)	U53124	U53123
Cryptomonas sp. Φ	CCMP 325	U53126	U53125
Rhodomonas abbreviata Hill (1991*a*) (syn. *Pyrenomonas/Cryptomonas abbreviata*)	CCAP 976/16	U53128	U53127
Storeatula major Hill (1991*a*)	CCMP320	U53130	U53129
Rhinomonas pauca Hill & Wetherbee (1989)	MUCC Cr#47	U53132	U53131
Unidentified cryptomonad	(Culture now dead)	U53191[a]	No sequence

[a] This sequence grouped deeply with the *Chroomonas* sequence but was excluded from the present analysis as we do not have a corresponding nuclear sequence and are uncertain of its identity and detailed phenotype.

Additional nomenclatural notes:

1. The *Guillardia theta* Hill & Wetherbee (1990) sequences in our trees were published by Douglas *et al.* (1991) under the name *Cryptomonas* sp. Φ; owing to an earlier mix-up of cultures the Douglas laboratory was provided originally not with *Cryptomonas* sp. Φ but with *Cryptomonas* sp. Θ (now named formally as *Guillardia theta* Hill & Wetherbee, 1990) mislabelled as Φ. Therefore all the papers published from the Douglas laboratory on *Cryptomonas* sp. Φ actually refer to *Guillardia theta* (further details will be published elsewhere).

2. The *Rhodomonas salina* sequences on our trees were originally published as *Pyrenomonas salina* (Maier *et al.* 1991; Eschbach *et al.* 1991*b*); we regard *Pyrenomonas* as an unnecessary junior synonym for *Rhodomonas*).

were mostly absent from all eukaryotic nuclear sequences and present only in chlorarachnean nucleomorph rRNA genes. But there was no evidence of homology between the cryptomonad and chlorarachnean insertions; even within cryptomonads these nucleomorph gene insertions were very variable in length, poorly conserved, and hard to align except for closely related species. They were therefore excluded from the phylogenetic analysis. The cryptophyte nuclear 18S rRNA genes are a little shorter than those of most neokaryotes, partly because of numerous very short phylum-specific deletions in moderately well conserved regions of the molecule; these regions were included in the phylogenetic analyses after preliminary tests showing that it made no difference to the position of cryptomonads on the trees whether they were included or not.

Maximum likelihood and parsimony analyses

Since maximum likelihood has been shown to be superior to either distance or parsimony methods for molecular phylogenetics over a wide range of parameters and is less vulnerable to violations of its basic assumptions (Nei, 1991; Kuhner & Felsenstein, 1994) we chose it as our primary method of analysis, despite its intensive use of computer time. Fig. 1 shows the relationships calculated for the 1632 most conserved and best-aligned nucleotide positions for the small subunit ribosomal RNA genes of our 92 sequence data set.

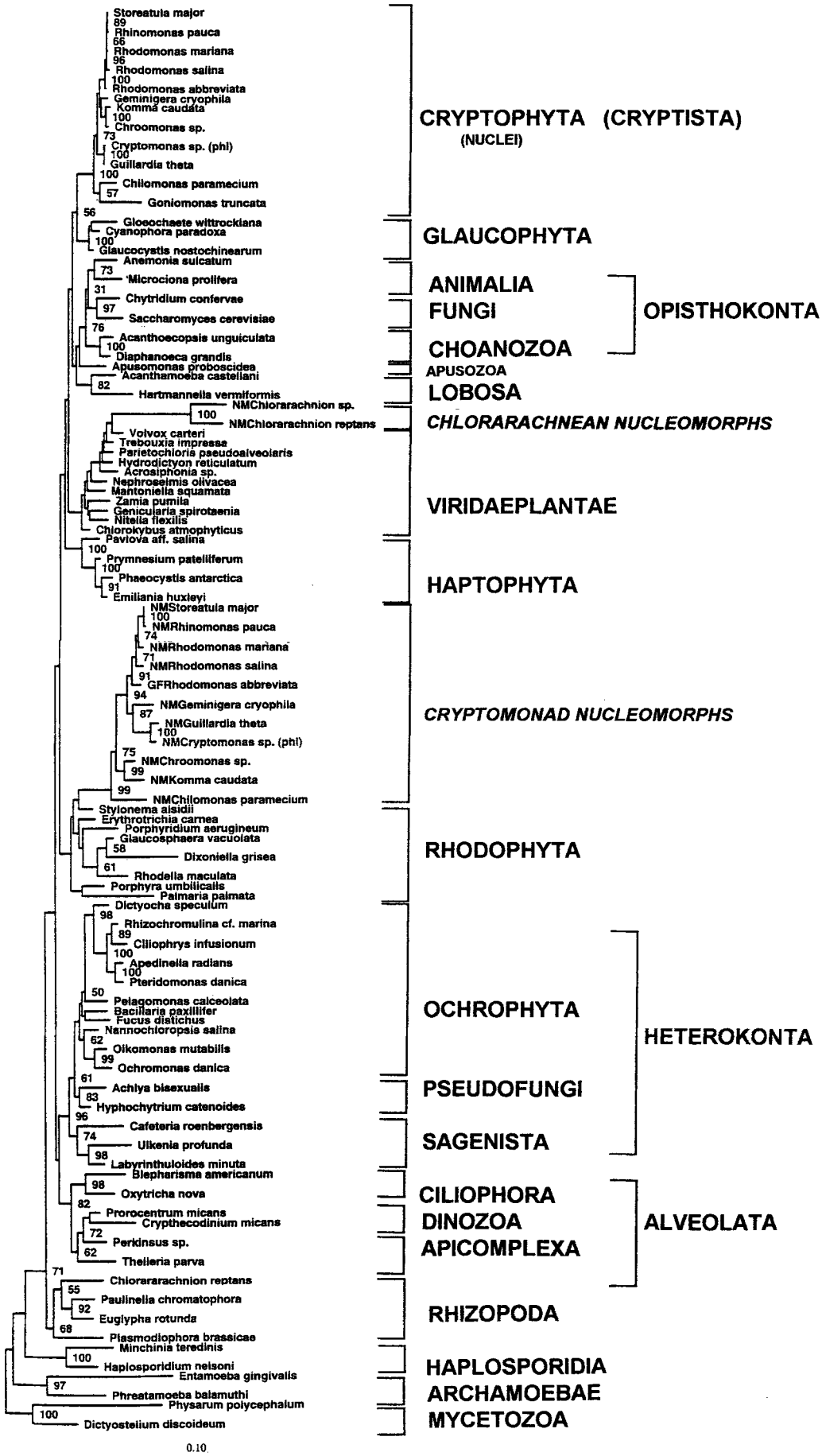
We were not able to carry out bootstrap analysis of our 92 sequence maximum likelihood trees as this would have taken 10 years of computer time. But as we find that parsimony often gives quite similar trees to maximum likelihood, except where particularly long branches are involved, the bootstrap values for a separate parsimony analysis are shown on Fig. 1. The shortest trees found by unbootstrapped parsimony analysis after several separate runs of the data were four equally parsimonious trees of length 10 823 steps. The major difference seen on these parsimony trees compared with the maximum likelihood tree was, in fact, the placement of the exceptionally long-branched *Chlorarachnion* nucleomorph clade, which appeared at the base of the tree just above the *Haplosporidia*, much as in a previous neighbor-joining study (Cavalier-Smith 1993*a*), and not in its very different position as the sister to *Volvox* (Fig. 1) which was also observed by Cavalier-Smith (1995) for a smaller data set. This deep position of the *Chlorarachnion* nucleomorph clade makes no biological sense and can be attributed to the long-branch-attracts artefact to which parsimony analysis is particularly prone (Felsenstein, 1978). In contrast, the position in Fig. 1 is concordant with other data (McFadden *et al.*, 1995; van de Peer *et al.*, 1996). The four equally parsimonious trees differed from each other only in the branching order within the green algae and the rhizopods. Branching within the cryptomonad nuclear and nucleomorph clades was identical to that on the maximum likelihood trees, except that the position of the *Geminigera* nuclear sequence was as in the marginally less likely tree mentioned in the legend to Fig. 1.

Both the maximum likelihood and the parsimony trees placed *Goniomonas* within the Cryptophyceae as sister to *Chilomonas*, suggesting that it may have arisen by loss of plastids and nucleomorph after the loss of photosynthesis in a common ancestor of *Chilomonas* and *Goniomonas*. This position, however, has very low bootstrap support; moreover an only slightly less likely maximum tree (ln likelihood −50398·33) placed *Goniomonas* as sister to cryptomonads as a whole, as previously seen with a much smaller data set (McFadden *et al.*, 1994*a*: Cavalier-Smith *et al.*, 1994).

With both methods the branching order for the cryptomonad nuclear sequence differs in certain respects from that for the nucleomorph sequences. But as the bootstrap support for the nucleomorph topology is much greater than the nuclear one, and the nucleomorph topology is more congruent with non-sequence data (see Discussion) and with the distance analyses described below, it is more likely to be correct.

With both maximum likelihood and parsimony the cryptist nuclei were grouped weakly with the Glaucophyta, but their common stem was so short, and the bootstrap support so low (about 35%), that we do not regard this congruence as significant. Unlike maximum likelihood, parsimony did not place the cryptomonad nucleomorphs within the red algae, but as sisters to them, although with low bootstrap support. The parsimony and maximum likelihood analyses differed substantially in the branching order within the green and red algae; as similar differences were noted with distance analyses according to whether the nucleomorphs were included or not within these clades (see below), we attribute them to perturbation by the exceptionally long nucleomorph sequences. In all other respects the maximum likelihood and parsimony trees were very similar, except that parsimony placed Pseudofungi (the oomycete *Achlya* and hypochytrid *Hypochytrium*) within the Ochrophyta, not as their sister group, thus supporting their origin by chloroplast loss (Cavalier-Smith, 1995). Parsimony therefore gives evidence for four independent losses of chloroplasts within the Heterokonta [i.e. *Oikomonas, Ciliophrys, Pteridomonas* (see Fig. 1 also), in addition to the pseudofungi].

Fig. 1. Maximum likelihood tree for the 92-sequence data set. The bootstrap values for a separate maximum parsimony analysis (100 pseudoreplicates) using the same 1632 alignment positions are shown at the nodes, except for those under 50%. Three separate fastDNAml trees were calculated using a different random order of taxon addition and allowing regional rearrangements across 10 branches. The tree shown had the highest log likelihood (−50397·931463). An insignificantly less likely tree (ln likelihood −50397·96) was identical except for the position of *Geminigera*, which was as in Fig. 2. Overall over 350 000 trees were examined. The scale bar indicates a sequence divergence of 10%.


Storeatula major
89
Rhinomonas pauca
66
Rhodomonas mariana
96
Rhodomonas salina
100
Rhodomonas abbreviata
Geminigera cryophila
Komma caudata
100
Chroomonas sp.
73
Cryptomonas sp. (phi)
100
Guillardia theta
100
Chilomonas paramecium
57
Goniomonas truncata
56
Gloeochaete wittrockiana
Cyanophora paradoxa
100
Glaucocystis nostochinearum
Anemonia sulcatum
73
Microciona prolifera
31
Chytridium confervae
97
Saccharomyces cerevisiae
76
Acanthoecopsis unguiculata
100
Diaphanoeca grandis
Apusomonas proboscidea
Acanthamoeba castellani
82
Hartmannella vermiformis
NMChlorarachnion sp.
100
NMChlorarachnion reptans
Volvox carteri
Trebouxia impressa
Parietochloris pseudoalveolaris
Hydrodictyon reticulatum
Acrosiphonia sp.
Nephroselmis olivacea
Mantoniella squamata
Zamia pumila
Genicularia spirotaenia
Nitella flexilis
Chlorokybus atmophyticus
Pavlova aff. salina
100
Prymnesium patelliferum
100
Phaeocystis antarctica
91
Emiliania huxleyi
NMStoreatula major
100
NMRhinomonas pauca
74
NMRhodomonas mariana
71
NMRhodomonas salina
91
GFRhodomonas abbreviata
94
NMGeminigera cryophila
87
NMGuillardia theta
100
NMCryptomonas sp. (phi)
75
NMChroomonas sp.
99
NMKomma caudata
99
NMChilomonas paramecium
Stylonema alsidii
Erythrotrichia carnea
Porphyridium aerugineum
Glaucosphaera vacuolata
58
Dixoniella grisea
61
Rhodella maculata
Porphyra umbilicalis
Palmaria palmata
Dictyocha speculum
98
Rhizochromulina cf. marina
89
Ciliophrys infusionum
100
Apedinella radians
100
Pteridomonas danica
50
Pelagomonas calceolata
Bacillaria paxillifer
Fucus distichus
Nannochloropsis salina
62
Oikomonas mutabilis
99
Ochromonas danica
61
Achlya bisexualis
83
Hyphochytrium catenoides
96
Cafeteria roenbergensis
74
Ulkenia profunda
98
Labyrinthuloides minuta
Blepharisma americanum
98
Oxytricha nova
82
Prorocentrum micans
Crypthecodinium micans
72
Perkinsus sp.
62
Theileria parva
71
Chlorarachnion reptans
55
Paulinella chromatophora
92
Euglypha rotunda
68
Plasmodiophora brassicae
Minchinia teredinis
100
Haplosporidium nelsoni
Entamoeba gingivalis
97
Phreatamoeba balamuthi
Physarum polycephalum
100
Dictyostelium discoideum
0.10
CRYPTOPHYTA (CRYPTISTA)
(NUCLEI)
GLAUCOPHYTA
ANIMALIA
FUNGI
CHOANOZOA
OPISTHOKONTA
APUSOZOA
LOBOSA
CHLORARACHNEAN NUCLEOMORPHS
VIRIDAEPLANTAE
HAPTOPHYTA
CRYPTOMONAD NUCLEOMORPHS
RHODOPHYTA
OCHROPHYTA
PSEUDOFUNGI
SAGENISTA
HETEROKONTA
CILIOPHORA
DINOZOA
APICOMPLEXA
ALVEOLATA
RHIZOPODA
HAPLOSPORIDIA
ARCHAMOEBAE
MYCETOZOA

Distance analyses by neighbor joining

One limitation of both parsimony and maximum likelihood for such a large number of taxa is that it is impossible to test all possible trees, so we cannot be certain that we have actually found the best tree by either method. It is therefore important also to carry out a distance analysis that does not have this problem. However, distance analysis as usually practised (Woese, 1987; Cavalier-Smith, 1993*a*; van de Peer *et al.*, 1993; Ragan & Guttell, 1995) shares with the most widely used forms of maximum likelihood and parsimony analyses the problem that it uses equal weighting for all nucleotide positions in the molecule, thereby assuming an equal rate of evolution throughout the molecule. This assumption has long been known to be grossly violated for ribosomal RNA (Olsen, 1987), and it is also known that neighbor joining (NJ) is more affected by this violation than maximum likelihood and is therefore more likely reproducibly to give the wrong tree. We have therefore carried out distance analysis using the Jin & Nei (1990) method to calculate the distance matrix, which assumes a variation in rates according to a γ distribution. Though still oversimplified (see Felsenstein & Churchill, 1996) this method should theoretically be distinctly superior to the usual Jukes–Cantor and Kimura corrections for calculating distance matrices, and it has recently begun to be used for real data sets (e.g. Winnepenninckx *et al.*, 1995).

Fig. 2 is a Jin & Nei NJ tree for the same data set as Fig. 1 using the value 1·6 for the shape parameter *a*. Though not totally congruent with Fig. 1, the tree is very similar and places the chlorarachnean nucleomorphs within the green algae as in Fig. 1, and not as sisters to the cryptomonad nucleomorphs as in previous distance analyses using the Jukes–Cantor correction (Cavalier-Smith *et al.*, 1994). When, however, the same data set was used to calculate a NJ tree using a Jukes–Cantor corrected distance matrix, or when a Jin & Nei matrix with a lower value of 0·72 was used for *a* as in Winnepenninckx *et al.* (1995), the chlorarachnean nucleomorph clade moved dramatically to become the sister of the cryptomonad nucleomorph as found in several earlier studies (Cavalier-Smith, 1993*a*; Cavalier-Smith *et al.*, 1994; Bhattacharya *et al.*, 1995). A position of chlorarachneans as sister to the cryptomonad nucleomorphs is also not in accord with other evidence (McFadden *et al.*, 1995; Gilson & McFadden 1995) and was previously (Cavalier-Smith *et al.*, 1994) attributed to the long-branches-attract artefact. Trial and error using *a* values in the range 0·25 to 1·55 showed that none of these was sufficient to prevent the probably artefactual association of the two nucleomorph clades. In Fig. 2, the chlorarachnean clade is a sister not to *Volvox* alone but to a clade consisting of Chlorophyceae/Trebouxiophyceae/Ulvophyceae. The position of the cryptomonad nucleomorphs was also slightly different; they were placed as sister not to *Stylonema* (Fig. 1) but to a clade comprising *Stylonema, Porphyridium* and *Erythrotrichia*. However when an *a* value of 2 was used to calculate the Jin & Nei tree the cryptomonad nucleomorphs were placed as sisters to *Stylonema* alone and the chlorarachnean nucleomorphs as sisters to *Volvox* alone. This tree also placed *Goniomonas* as sister to *Chilomonas* as in Fig. 1, not lower down as in Fig. 2. It also altered the position of the *Chilomonas* nucleomorph clade, placing it as sister to the *Komma/Chroomonas* clade. As this latter change makes the nuclear and nucleomorph part of the tree incongruent, and as the $a = 2$ tree also places the Archamoebae in the biologically unreasonable position of sisters to the piroplasm *Theileria*, we have chosen, somewhat arbitrarily, to show the $a = 1{\cdot}6$ tree.

For values of *a* between 0·25 and 1·8 the cryptomonad nuclear and nucleomorph branching order was totally congruent and identical with that of the nucleomorphs found in the maximum likelihood and parsimony analyses. NJ trees calculated from a distance matrix corrected by either the Jukes–Cantor or the Kimura methods were identical to each other in topology and virtually indistinguishable also in relative branch lengths. In the Jukes–Cantor and Kimura trees, the internal branching orders of the cryptomonad nuclei and nucleomorphs were totally congruent and identical to those of Fig. 2; but these trees placed the cryptomonad and chlorarachnean clades as a whole not within the red and green algae respectively as in Figs 1 and 2, but as separate clades just above the Haplosporidia as in an earlier Jukes–Cantor analysis (Cavalier-Smith, 1993*a*).

All the NJ trees that did not place the nucleomorphs within the red and green algae showed the same branching order for the red and green algae as did the parsimony trees. By contrast the Jin & Nei NJ trees that placed the nucleomorphs within the red and green algae showed the very different branching order for the red and green algae seen in the Fig. 1 maximum likelihood tree, which placed the nucleomorphs within these groups. From these comparisons it is evident that any tree, no matter by what method it is generated, that places the exceedingly long nucleomorph branches within the red and green algae thereby distorts the branching order of these two algal groups. We have checked this by calculating separate trees for red and green algae with much larger data sets that exclude the nucleomorph sequences. The branching order for these two groups is then as in Fig. 2 not Fig. 1. With distance trees a different kind of distortion is seen compared with those where the cryptomonad nucleomorphs are excluded from the red algae. On these trees the red algae are relatively lower in the tree, and just above the nucleomorphs, than on trees where the nucleomorphs are within the red algae. It is as though the NJ algorithm is trying to place the two groups near each other even at the expense of placing the red algae too low in the tree.

The Kimura and Jukes–Cantor NJ trees placed the cryptist nuclei as sisters to the glaucophytes as in Fig. 1, whereas Jin & Nei NJ trees with *a* values in the range 0·25–1·8 placed them as sisters to the opisthokont/*Apusomonas* clade as in Fig. 2. But the common stems are so short and the bootstrap values so low that both

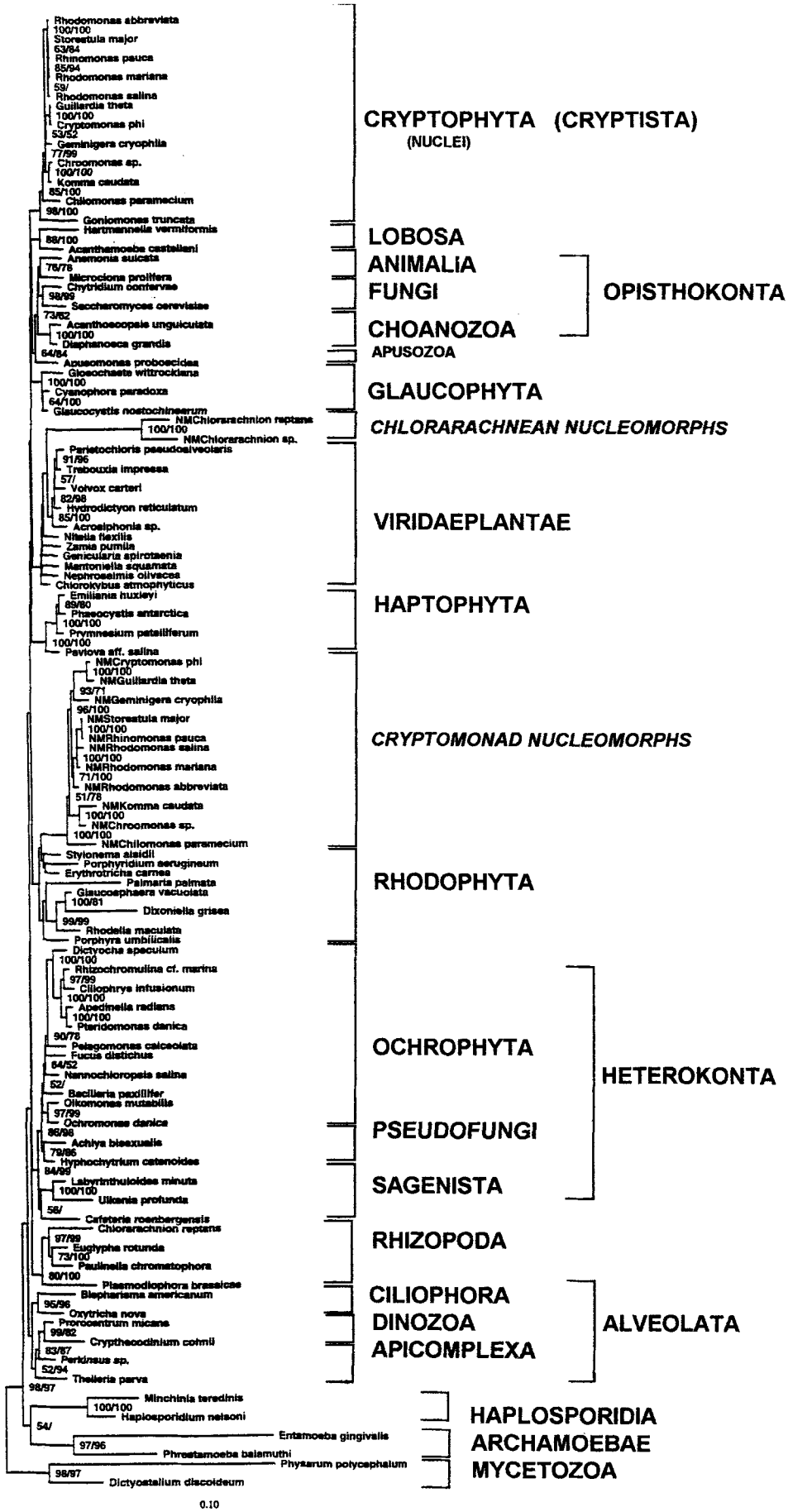


Fig. 2. Neighbor-joining (NJ) tree for the same data set as in Fig. 1 using the Jin & Nei (1990) method for calculating the distance matrix and an α shape parameter for the γ distribution of 1·6. Bootstrap percentages (200 pseudoreplicates) are shown at the nodes, except for those under 50%; the figures on the left are for the Jin & Nei analysis; those on the right are for a separate NJ analysis using the Kimura correction (with a transition/transversion ratio of 2) alone to calculate the distance matrix. The scale bar indicates a sequence divergence of 10%.

positions must be strongly suspect as accidental and biologically irrelevant.

The congruence of the nucleomorph and nuclear cryptomonad topologies in our distance analysis and the low bootstrap values for the latter are consistent with our interpretation that the partially different branching orders seen for the nuclear sequences in the maximum likelihood and parsimony analyses are artefactual. Despite these partial incongruences there were four reproducible major similarities between the nuclear and nucleomorph trees seen with all three methods of analysis. First, the Pyrenomonadales (*Rhodomonas, Rhinomonas, Storeatula*) formed a clade supported by high bootstrap values (100% with both parsimony and NJ) within which the branch order of the five taxa was invariant; *Storeatula* and *Rhinomonas* are very closely related and derived compared with the ancestral paraphyletic *Rhodomonas*. Second, the two blue-pigmented genera *Komma* and *Chroomonas* form a robust clade supported by high bootstrap values (100% for parsimony and for most but not all NJ trees). Third, the colourless *Chilomonas* appears to be an outgroup to all the pigmented photosynthetic genera, though this conclusion does not have high bootstrap support and was not so in the $a = 2$ Jin & Nei tree. The red-pigmented species *Guillardia theta* and *Cryptomonas* Φ form a robust clade with 100% bootstrap support.

Where the nuclear tree differs from, and is sometimes in apparent conflict with, the more robust nucleomorph tree is in the relative branching order of the Pyrenomonadales, the *Guillardia/Cryptomonas* ϕ clade, the *Komma/Chroomonas* clade and *Geminigera*. In all nucleomorph trees, and some nuclear trees, *Geminigera* groups deeply with the *Guillardia/Cryptomonas* ϕ clade, but sometimes in nuclear trees it groups instead (deeply) with the Pyrenomonadales. In all nucleomorph trees, and in all distance trees, the blue *Komma/Chroomonas* clade is a sister to all the eight red-pigmented taxa, whereas in the nuclear trees parsimony and likelihood trees the *Guillardia/Cryptomonas* ϕ clade branches slightly more deeply.

There is a conflict between the methods for the position of *Goniomonas*. With distance methods that do not allow for variable rates of evolution between sites or with Jin & Nei corrections with low values of a it is the sister of all the cryptomonads with nucleomorphs, as originally observed (McFadden *et al.*, 1994*a*), but with maximum likelihood or parsimony or with Jin & Nei NJ trees with high values of a it lies within the cryptomonads as the sister to *Chilomonas*.

Effects of taxon sampling

Because of these uncertainties concerning the cryptomonad nuclear part of the tree, we also studied the effect of taxon sampling since it is commonly found that this has a greater effect on tree topology than variations in calculation method. We therefore also calculated trees with more restricted outgroups.

One set of trees was calculated by excluding all the distant outgroups (the lower 16 species of Fig. 1) that were used to root the trees shown in Figs 1 and 2, as well as reducing the number of heterokonts so as to reduce the total number of species to 70 to hasten the calculations. Three independent maximum likelihood calculations were done, using global rearrangements (i.e. allowing crossing of 67 branches; each took about 2 weeks, and about 300 000 trees were examined), and all gave identical trees with exactly the same branching order within the nucleomorph and cryptomonad nuclear clades as in Fig. 1. The cryptomonad nuclear part of the tree therefore still showed exactly the same inconsistencies with the nucleomorph branching order as with the 92-species data set. However, the chlorarachnean nucleomorph clade moved to join the cryptomonad nucleomorphs, emphasizing the

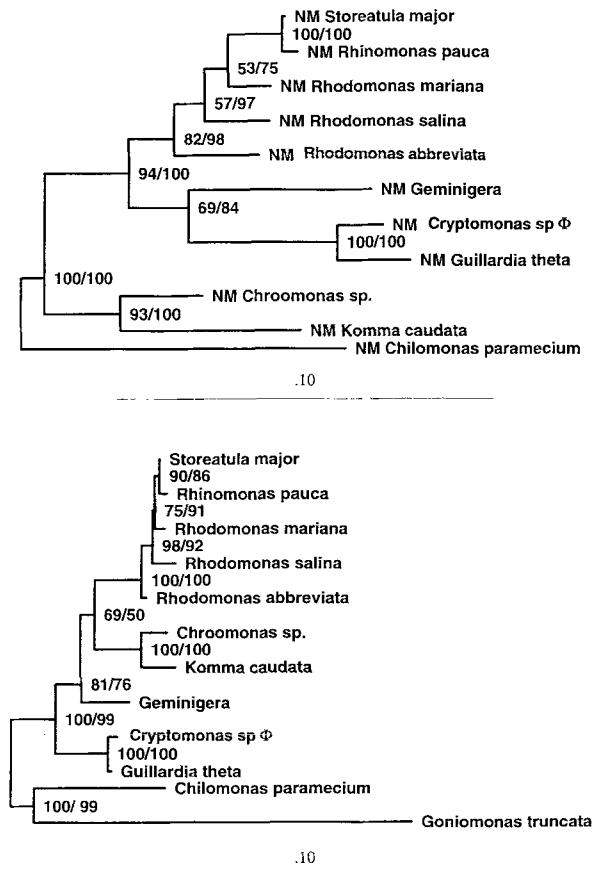


Fig. 3. Maximum likelihood tree for the cryptomonad nucleomorphs (upper figure) and cryptist nuclei (lower figure) excluding all outgroups. The bootstrap values for a separate maximum parsimony analysis (200 pseudoreplicates; left figure) and a NJ analysis (1000 pseudoreplicates; Jukes–Cantor correction; right figure) are shown at the nodes, except for those under 50%. Three separate fastDNAml trees were calculated using a different random order of taxon addition and all gave exactly the same trees (upper figure: log likelihood = −5801·36034; transition/transversion parameter 1·494376; 267 distinct data patterns; 1125 trees examined; the single most parsimonious tree in the corresponding parsimony analysis had 711 steps; lower figure: log likelihood = −4473·63203; transition/transversion parameter 1·516765; 186 distinct data patterns; single most parsimonious tree with 456 steps). The scale bars indicate a sequence divergence of 10%.

instability of its position on trees; though this joint nucleomorph clade was sister to the red algae, the branching order within the green algae as well as the red algae was distorted (as discussed above) in comparison with trees from which nucleomorphs were excluded. As with the 92-sequence data set, the 70-species distance trees showed the same branching order for cryptomonad nuclei and nucleomorphs as in Fig. 2 (including the placement of *Goniomonas* as sister to *Chilomonas*), but the parsimony trees agreed with the likelihood ones.

To eliminate any possible perturbing effects of distant outgroups we calculated separate nuclear and nucleomorph trees for the cryptist sequences alone (Fig. 3). The maximum likelihood and parsimony trees have an identical branching order to the 70-species trees with outgroups and preserve the same incongruencies. Paradoxically, the distance trees for the cryptomonad nuclear sequences are not identical to the maximum likelihood and parsimony trees, and are therefore incongruent with the nucleomorph branching order, which remains unchanged. This suggests that the nuclear sequences have some sort of systematic bias that distorts the tree in comparison with the nucleomorph tree, which can be compensated for by the distance methods but not by the other two when outgroups are present but not when they are absent. In this respect, distance methods appear to be superior to either maximum likelihood or parsimony, which highlights the importance of not relying solely on one or two of these three methods in phylogenetic analysis.

Because the *Goniomonas* branch is particularly long and present only in the nuclear tree, we calculated separate trees excluding it (not shown). As these were identical in topology by all methods to those shown here, a systematic bias must be present in the pattern of evolution of the 11 cryptomonad nuclear sequences themselves. Possibly the problem lies in the incorrect position of the long-branch *Chilomonas* sequence in one or both trees; excluding it makes the unrooted trees for the photosynthetic cryptomonads fully congruent by all three methods.

Discussion

Monophyletic origin of the cryptomonad nucleomorph from the nucleus of a red algal ancestor

Contrary to the proposal by Margulis & Schwartz (1988) of several separate symbiotic origins for the chloroplasts of cryptomonads from differently pigmented ancestors, all our trees, by all methods of analysis and with different taxon samples, robustly show cryptomonad nucleomorphs to be monophyletic. For both distance and parsimony methods their monophyly has 100% bootstrap support, and the length of the stem at their base is so great with both maximum likelihood and distance trees that there can be no doubt that all cryptomonad nucleomorphs originated from the same secondary symbiosis (Cavalier-Smith, 1982, 1986).

The nature of the eukaryotic alga from which cryptomonad nucleomorphs and chloroplasts both evolved is also clarified by our new trees. It has long been clear from the similarity of their phycobiliprotein pigments that cryptomonad chloroplasts must be related to those of red algae (Glazer & Appell, 1977). But there have been two different hypotheses as to the precise nature of this relationship. According to one, the ancestral endosymbiont was a red alga (Whatley *et al.*, 1979); a possible alternative was that it was a phyletic intermediate between the ancestors of red algae and dinoflagellates that possessed both phycobilins and chlorophyll *c* (Cavalier-Smith, 1982). The former idea that it was an actual red alga, and not a phyletic intermediate, is clearly supported by our present trees, which thus confirm the earlier tentative support for a direct red algal ancestry based on relatively few sequences (Douglas *et al.*, 1991; Maier *et al.*, 1991; Cavalier-Smith *et al.*, 1994; Ragan *et al.*, 1994; McFadden *et al.*, 1994*a*; Ragan & Guttell, 1995; Cavalier-Smith, 1995). As red algae do not have chlorophyll *c*, and unlike cryptomonads have phycobilisomes, this makes it likely that the chlorophyll *c* of cryptomonads evolved after the initial endosymbiotic event, possibly contemporaneously with the loss of phycobilisome linker proteins and allophycocyanin that must have occurred in the ancestor of extant cryptomonads (Cavalier-Smith, 1982; Ludwig & Gibbs, 1989).

It has long been obvious that if the endosymbiont was a red alga it must have been a unicellular red alga and not a multicellular member of the Bangiales or Florideophyceae, unless it was a spermatium or carpospore of a multicellular species, which seems a little improbable. Though our trees, unlike those of Ragan *et al.* (1994) and Ragan & Guttell (1995), include several unicellular red algae (*Glaucosphaera, Rhodella, Porphyridium*), none of them group specifically and reproducibly with the cryptomonad nucleomorphs on our best trees, though *Porphyridium* did so on one slightly suboptimal parsimony tree. Instead, the nucleomorphs seem to be closest to the minute, branched, filamentous epiphyte *Stylonema* (illustrated in Bird & McLachlan, 1992) on maximum likelihood trees and on the Jin & Nei tree that weighted the most conserved parts of the sequences the most strongly. Possibly, when genes from a greater diversity of unicellular red algae are sequenced, one of them will prove to be even more closely related to the cryptomonad nucleomorphs. However, the nucleomorph 18S rRNA sequences are so rapidly evolving that one would only be confident of their precise position within the red algae if it were corroborated by several different genes.

The red algal ancestry of the cryptomonad nucleomorph is also supported by chloroplast 16S rRNA gene phylogenies (McFadden *et al.*, 1995; Bhattacharya & Medlin, 1995; Leblanc *et al.*, 1995), but as these do not at present include sequences from dinoflagellates they are less decisive in discriminating between the two theories discussed above. The evidence from Rubisco is also most simply consistent with a red algal ancestry, since cryptomonads (like all chromists) have the type I purple bacterial form of

Rubisco otherwise known only in red algae (Martin *et al.*, 1992), whereas dinoflagellates have the radically different type II Rubisco (Morse *et al.*, 1995; Whitney *et al.*, 1995; Palmer, 1995, 1996; Rowan *et al.*, 1996).

Probable early divergence of blue- and red-pigmented cryptomonads

The phycobilisomes of the red alga that was ancestral to the cryptomonad chloroplast must have had both red phycoerythrin and blue phycocyanin and allophycocyanin pigments. The fact that photosynthetic cryptomonads contain either phycocyanin or phycoerythrin but never both, and never have allophycocyanin, means that the ancestor(s) of red-pigmented cryptomonads lost phycocyanin and the ancestor(s) of blue-pigmented cryptomonads lost phycoerythrin. *A priori*, the most parsimonious interpretation of these facts would be that this loss occurred once only for each type, and that red- and blue-pigmented cryptomonads diverged from a common ancestor with both pigments, and are therefore sister groups. On this simple interpretation there would never have been a changeover from red pigment to blue pigment or vice versa.

Our nucleomorph sequence trees strongly support this parsimonious theory since they show the blue *Komma/Chroomonas* clade as a sister to all the eight red-pigmented taxa. Unfortunately the nuclear sequences are more ambiguous: for distance methods the blue and red clades are sisters, but for parsimony and maximum likelihood the red *Guillardia/Cryptomonas* ϕ clade tends to branch slightly lower than the blue clade. If this latter topology were correct, it would imply that there had been a changeover from red to blue pigments, which is chemically very improbable. Both for this reason, and because the branching order of nuclear sequences is congruent with the much more robust nucleomorph sequence tree in our distance trees, we suggest that the nucleomorph branching order is almost certainly correct, and that the somewhat discordant topology of the parsimony and likelihood nuclear trees is probably an artefactual perturbation of the branching order, possibly caused by the misplacement of the long *Chilomonas* branch. This interpretation could be tested by obtaining sequences from more species, especially ones more closely related to *Chilomonas* such as *Campylomonas*.

The reason why the nucleomorph trees are much more robust than the nuclear tree is probably that, as the nucleomorph genes are evolving more rapidly, a larger number of nucleotide substitutions occurred in each of the short internal segments of the nucleomorph tree so that each clade is supported by the larger number of shared derived characters. This will make the nucleomorph trees more resistant to changes in method, as well as to variations in taxon sampling, and increase the bootstrap support for the branching order.

Our results give strong molecular support to the use of variations in the phycobilin pigments in the suprageneric classification of cryptomonads (Hill & Rowan, 1989; Novarino & Lucas, 1993*a*).

Monophyly of cryptomonads with embedded nucleomorphs (Pyrenomonadales)

Our results clearly show that *Rhodomonas*, *Rhinomonas* and *Storeatula* form a natural group of cryptomonads characterized by possessing nucleomorphs embedded in the pyrenoid. This conclusion is totally robust, regardless of variations in analytic method, taxon sampling, and intramolecular sampling (as shown by the 100% bootstrap values for the clade).

Thus, the recently established order Pyrenomonadales (Novarino & Lucas, 1993*a*) is a very well supported holophyletic taxon. The ancestral state for cryptomonads was clearly that with nucleomorphs not embedded in the pyrenoid, making Cryptomonadales a paraphyletic taxon. In future, when a larger number of sequences are available, it will probably be desirable to subdivide Cryptomonadales into several orders, not because they are paraphyletic but because they are phenotypically too diverse (with blue, red and colourless genera for a start) to be included in a single order. Our molecular phylogenetic results strongly support the recent use of well-chosen ultrastructural characters in revising the limits of cryptomonad genera and in their higher-level classification.

Origin of the chlorarachnean endosymbiont

Our present results highlight the extremely volatile position of the *Chlorarachnion* nucleomorphs on 18S rRNA trees. Our $a = 1{\cdot}6$ Jin & Nei NJ and 92-sequence maximum likelihood trees, however, support the earlier evidence reviewed by Cavalier-Smith (1995) and McFadden *et al.* (1995) for a green algal ancestry for the chlorarachnean nucleomorph and chloroplast, and are consistent with the interpretation of Cavalier-Smith *et al.* (1994) and van de Peer *et al.* (1996) that the frequent grouping together of chlorarachnean and cryptomonad nucleomorphs on 18S rRNA trees is an artefact caused by their very long branches, especially those of chlorarachneans. Our present trees suggest that the ancestor of the chlorarachnean nucleomorph and chloroplast may have been a chlorophycean green alga related to the Volvocales, rather than a prasinophyte as earlier tentatively suggested (McFadden *et al.*, 1995; Cavalier-Smith, 1995). This evidence clearly favours the original exclusion (Cavalier-Smith, 1986) of chlorarachneans from the Chromista, rather than their inclusion (Cavalier-Smith, 1993*a*, 1994) based primarily on the earlier, misleading 18S rRNA trees. Location of the nucleomorph within a pyrenoid invagination in cryptomonads of the order Pyrenomonadales and *Chlorarachnion reptans* is apparently a case of evolutionary convergence, perhaps attributable to requirements of partitioning daughter nucleomorphs during plastid fission (McFadden, 1993).

Relationship of cryptists to other chromists

Our results do little to clarify the controversial question of the relationship of cryptists themselves to other organisms. In none of our trees is the position of the cryptists supported by high bootstrap values. In some of our trees they are weakly grouped with glaucophytes as observed by Bhattacharya *et al.* (1995), whereas in others they are not, as observed by Cavalier-Smith (1995). Because of the many contradictory positions of the cryptists shown on published trees, including our own, we do not think that any of them can be relied upon. We know of no particular biological reason to think that cryptists actually are related to glaucophytes.

Reasons for classifying cryptists together with Heterokonta and Haptophyta in the kingdom Chromista have been extensively discussed elsewhere (Cavalier-Smith, 1986, 1989, 1995). As emphasized previously (Cavalier-Smith, 1993*a*, *b*; Cavalier-Smith *et al.*, 1994), 18S rRNA trees neither refute nor confirm this classification. Our present trees are no different from earlier ones in this respect, but they do offer new insight with respect to the position of *Goniomonas*. Previous analyses in which *Goniomonas* was positioned below the cryptomonads (McFadden *et al.*, 1994*a*; Cavalier-Smith *et al.*, 1994; Bhattacharya *et al.*, 1995; Cavalier-Smith, 1995) may have suffered from a long-branch artefact. The *Goniomonas* sequence is markedly different from that of all other cryptists, and it always appears as a long branch. The fact that our maximum likelihood tree (Fig. 1), parsimony trees, and Jin & Nei distance tree with an *a* value of 2·0 (slowly evolving parts of the molecule weighted more strongly than the more rapidly evolving parts) placed *Goniomonas* as sister to *Chilomonas* could suggest that the *Goniomonas* sequence was placed too low on earlier trees because of its ultra-rapid rate of evolution. In the absence of an appropriate outgroup sequence for cryptists, sequencing other *Goniomonas* species and *Chilomonas* relatives (ideally with shorter branches than the *Chilomonas paramecium* used here) will be necessary to refute or confirm this possibility. If, however, our maximum likelihood, parsimony, and high *a* Jin & Nei trees are correct, then *Goniomonas* is derived from photosynthetic ancestors, and the common ancestor of all cryptists could have been photosynthetic, as postulated by the monophyletic hypothesis of the kingdom Chromista (Cavalier-Smith, 1986, 1989, 1995).

Our suggestive, but by no means conclusive, evidence that *Goniomonas* also may have lost its chloroplasts, when added to the new extensive molecular evidence for multiple losses of chloroplasts within the Heterokonta (Cavalier-Smith *et al.*, 1995, 1996; see also Figs 1 and 2 of the present paper), can be interpreted as additional support for the thesis that all chromists are descended from a photosynthetic ancestor that originated in a single symbiogenetic event (Cavalier-Smith, 1986, 1989, 1995). The idea of a single symbiotic event involving the uptake of a red alga by the ancestral chromistan host is supported by the fact that on chloroplast 16S rRNA trees (McFadden *et al.*, 1995; Bhattacharya & Medlin, 1995; van de Peer *et al.*, 1996) haptophytes and heterokonts, like cryptomonads, branch within the red algal part of the tree, even though haptophytes and heterokonts lack the phycobilin evidence for a red algal ancestry. Moreover, as mentioned above, these two groups have a Rubisco of the red algal type, which also supports the origin of their chloroplasts from red algae. Unfortunately the branching positions of haptophytes, heterokonts and cryptists on both the 16S rRNA and Rubisco trees (McFadden *et al.*, 1995) are too closely spaced and non-robust for them to be any more reliable than is nuclear 18S rRNA in discriminating unequivocally between a monophyletic and a polyphyletic origin of the Chromista. Some 16S rRNA trees group haptophytes with cryptomonads (Nelissen *et al.*, 1995), some group them with heterokonts, and others do neither.

This very closeness of the branching of the three major chromistan taxa on all three types of molecular trees is congruent with the hypothesis that they diverged from each other very rapidly following the symbiogenetic origin of the kingdom Chromista (Cavalier-Smith, 1982, 1986, 1989, 1995). Furthermore, the fact that on our 18S rRNA trees, and many other published ones, these three taxa are close to, and usually intermingled with, the three main plant groups (red algae, green plants and glaucophytes), which are proposed to have diverged from a common ancestor almost immediately following the symbiogenetic origin of the chloroplast (Cavalier-Smith, 1995), is consistent with the thesis that the chimaeric origin of the Chromista also occurred very shortly after the origin of chloroplasts (Cavalier-Smith, 1982, 1986). The position of the cryptomonad nucleomorph clade on our trees relatively close to the base of the red algal radiation also supports this thesis. By contrast, the position of the chlorarachnean nucleomorphs relatively high up amongst the chlorophycean algae suggests that the endocytobiotic acquisition of green algal chloroplasts by chlorarachneans may have occurred very substantially after the origin of chloroplasts.

Utility of nuclear and nucleomorph rRNA sequences for cryptomonad taxonomy

Our results demonstrate the great utility of 18S rRNA phylogeny for resolving some problems in the taxonomy of cryptomonads, an algal class within which the delimitation of genera and families has proved particularly difficult (Klaveness, 1985; Hill & Wetherbee, 1989; Hill, 1991*a*, *b*). Ribosomal RNA sequencing also has the potential to give much more information than restriction site studies (Chesnick *et al.*, 1991). Sequencing the 18S rRNA of other genera will be important for showing the major features of cryptomonad phylogeny and in particular whether any photosynthetic genera branch markedly lower down than *Goniomonas*, in order to test more thoroughly the suggestion that it was derived from cryptomonads by the loss of plastids and nucleomorphs, as argued by Cavalier-Smith (1989, 1995), rather than

being primitively without these organelles as reasoned by McFadden *et al.* (1994*a*). Since, however, the delimitation of most cryptomonad genera is still so uncertain, it will be important in future to sequence several representatives of each presently recognized genus in order to test their validity.

The fact that the nucleomorph sequences have been evolving at over twice the rate of the nuclear sequences provides more phylogenetically informative changes, and therefore suggests that nucleomorph sequences will be more useful than nuclear ones, especially at lower taxonomic levels. The systematic differences in rate of evolution between two eukaryotic genomes that have been part of the same cell for several hundred million years clearly show that there is no cell-wide molecular clock. The simplest explanation for the higher rate of evolution of the cryptomonad nucleomorph genome is that its great reduction in coding capacity has reduced the strength of stabilizing selection against random changes in ribosome structure. The apparently rapid evolution of the chlorarachnean nucleomorph sequences is consonant with the fact that they have experienced a nearly two-fold greater degree of genome reduction than the cryptomonad nucleomorphs (Gilson & McFadden, 1995). The insertion of extra sequences in similar positions within the molecule in both cryptomonads and chlorarachneans is probably a convergent consequence of the independent reduction in genome size and consequently of the strength of stabilizing selection in these two types of nucleomorph.

Even though the nuclear trees do not always support the nucleomorph trees in every detail, their very substantial congruence (and total congruence by distance methods) gives us considerable confidence in the validity of this approach to cryptomonad systematics. Because of their exceptionally rapid evolution, the long insertions in the nucleomorph DNA, though not useful for the overall phylogeny of cryptomonads considered here, could be of particular systematic value at the intrageneric, and possibly even intraspecific, levels.

Acknowledgements

We thank E. E. Chao for help in preparing the manuscript. T.C.-S. thanks NSERC for a research grant, the Canadian Institute for Advanced Research for fellowship support, G. de Jager for use of the Sparc10, and E. Harley and M. Berman for hospitality in their department while writing the paper. G.I.McF. is supported as an Australian Research Council (ARC) Senior Fellow and his project by a grant from the ARC. P.G. and J.A.D. are recipients of Australian Postgraduate Awards. D.R.A.H. is supported as an ARC Research Fellow.

References

BHATTACHARYA, D. & MEDLIN, L. (1995). The phylogeny of plastids: a review based on comparisons of small-subunit ribosomal RNA coding regions. *J. Phycol.*, **31:** 489–498.

BHATTACHARYA, D., HELMCHEN, T., BIBEAU, C. & MELKONIAN, M. (1995). Comparisons of nuclear-encoded small-subunit ribosomal RNAs reveal the evolutionary position of the Glaucocystophyta. *Mol. Biol. Evol.*, **12:** 415–420.

BIRD, C.J. & MCLACHLAN, J.L. (1992). *Seaweed Flora of the Maritimes. 1. Rhodophyta: The Red Algae.* Biopress, Bristol.

CAVALIER-SMITH, T. (1982). The origins of plastids. *Biol. J. Linn. Soc.*, **17:** 289–306.

CAVALIER-SMITH, T. (1986). The kingdom Chromista: origin and systematics. In *Progress in Phycological Research* vol. 3 (Round, F. & Chapman, D.J., editors), 309–347. Biopress, Bristol.

CAVALIER-SMITH, T. (1989). The kingdom Chromista. In *The Chromophyte Algae: Problems and Perspectives* (Green, J.C., Leadbetter, B.S.C. & Diver, W.L., editors), 381–407. Clarendon Press, Oxford.

CAVALIER-SMITH, T. (1993*a*). Kingdom Protozoa and its 18 phyla. *Microbiol. Rev.*, **57:** 953–994.

CAVALIER-SMITH, T. (1993*b*). Evolution of the eukaryotic genome. In *The Eukaryotic Genome* (Broda, P., Oliver, S.G. & Sims, P., editors), 333–385. Cambridge University Press, Cambridge.

CAVALIER-SMITH, T. (1994). Origin and relationships of Haptophyta. In *The Haptophyte Algae* (Green, J.C. & Leadbeatter, B.S.C., editors), 413–435. Clarendon Press, Oxford.

CAVALIER-SMITH, T. (1995). Membrane heredity, symbiogenesis, and the multiple origins of algae. In *Biodiversity and Evolution* (Arai, R., Kato, M. & Doi, Y., editors), 75–114. National Science Museum, Tokyo.

CAVALIER-SMITH, T., ALLSOPP, M.T.E.P. & CHAO, E.E. (1994). Chimeric conundra: are nucleomorphs and chromists monophyletic or polyphyletic? *Proc. Natl. Acad. Sci., U.S.A.*, **91:** 11368–11372.

CAVALIER-SMITH, T., CHAO, E.E. & ALLSOPP, M.T.E.P. (1995). Ribosomal RNA evidence for chloroplast loss within Heterokonta: pedinellid relationships and a revised classification of ochristan algae. *Arch. Protistenk.*, **145:** 209–220.

CAVALIER-SMITH, T., CHAO, E.E., THOMPSON, C.E. & HOURIHANE, S.L. (1996). *Oikomonas*, a distinctive zooflagellate related to chrysomonads. *Arch. Protistenk.*, **146:** 273–279.

CHESNICK, J.M., KUGRENS, P. & CATTOLICO, R.A. (1991). The utility of mitochondrial DNA restriction fragment length polymorphisms in cryptomonad phylogenetic assessment. *Mol. Mar. Biol. Biotech.*, **1:** 18–26.

DOUGLAS, S.E., MURPHY, C.A., SPENCER, D.F. & GRAY, M.W. (1991). Cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes. *Nature*, **350:** 148–151.

ELWOOD, H.J., OLSEN, G.J. & SOGIN, M.L. (1985). The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. *Mol. Biol. Evol.*, **2:** 399–410.

ESCHBACH, S., HOFMANN, C.J.B., MAIER, U.-G., SITTE, P. & HANSMANN, P. (1991*a*). A eukaryotic genome of 660 kb: electrophoretic karyotype of nucleomorph and cell nucleus of the cryptomonad alga, *Pyrenomonas salina*. *Nucleic Acids Res.*, **19:** 1779–1781.

ESCHBACH, S., WOLTERS, J. & SITTE, P. (1991*b*). Primary and secondary structure of the nuclear small subunit ribosomal RNA of the cryptomonad *Pyrenomonas salina* as inferred from the gene sequence: evolutionary implications. *J. Mol. Evol.*, **32:** 247–252.

FELSENSTEIN, J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.*, **27:** 401–410.

FELSENSTEIN, J. (1993). *PHYLIP version 3.5.1.* University of Washington, Seattle.

FELSENSTEIN, J. & CHURCHILL, G.A. (1996). A hidden Markov model approach to variation among sites in rate of evolution. *Mol. Biol. Evol.*, **13:** 93–104.

GILLOTT, M.A. & GIBBS, S.P. (1980). The cryptomonad nucleomorph: its ultrastructure and evolutionary significance. *J. Phycol.*, **16:** 558–568.

GILSON, P.R. & MCFADDEN, G.I. (1995). The chlorarachniophyte: a cell with two different nuclei and two different telomeres. *Chromosoma*, **103:** 635–641.

GLAZER, A.N. & APPELL, G.S. (1977). A common evolutionary origin for the biliproteins of cyanobacteria, Rhodophyta, and Cryptophyta. *Fed. Eur. Microbiol. Soc., Microbiol. Lett.*, **1:** 113–116.

GLEESON, T. & HILLIER, L. (1991). A trace display and editing program for data from fluorescence-based sequencing machines. *Nucleic Acids Res.*, **19:** 6481–6483.

GREENWOOD, A.D. (1974). The Cryptophyta in relation to phylogeny and photosynthesis. *Abstracts of the Eight International Congress of Electron Microscopy*, Canberra, pp. 566–567.

GREENWOOD, A.D., GRIFFITHS, H.B. & SANTORE, U.J. (1977). Chloroplasts and cell compartments in Cryptophyceae. *Br. Phycol. J.*, **12:** 119.

HANSMANN, P. & ESCHBACH, S. (1990). Isolation and preliminary characterization of the nucleus and the nucleomorph of a cryptomonad, *Pyrenomonas salina*. *Eur. J. Cell Biol.*, **52:** 373–378.

HANSMANN, P., FALK, H., SCHEER, U. & SITTE, P. (1986). Ultrastructural localization of DNA in two *Cryptomonas* species by use of a monoclonal DNA antibody. *Eur. J. Cell Biol.* **52:** 373–378.

HILL, D.R.A. (1991*a*). A revised circumscription of *Cryptomonas* (Cryptophyceae) based on examination of Australian strains. *Phycologia*, **30:** 170–188.

HILL, D.R.A. (1991*b*). *Chroomonas* and other blue-green cryptomonads. *J. Phycol.*, **27:** 133–145.

HILL, D.R.A. & ROWAN, K.S. (1989). The biliproteins of the Cryptophyceae. *Phycologia*, **28:** 455–463.

HILL, D.R.A. & WETHERBEE, R. (1986). *Proteomonas sulcata* gen. et sp. nov. (Cryptophyceae), a cryptomonad with two morphologically distinct and alternating forms. *Phycologia*, **25:** 521–543.

HILL, D.R.A. & WETHERBEE, R. (1988). The structure and taxonomy of *Rhinomonas pauca* gen. et sp. nov. (Cryptophyceae). *Phycologia*, **27:** 355–365.

HILL, D.R.A. & WETHERBEE, R. (1989). A reappraisal of the genus *Rhodomonas* (Cryptophyceae). *Phycologia*, **28:** 143–158.

HILL, D.R.A. & WETHERBEE, R. (1990). *Guillardia theta* gen. et sp. nov. (Cryptophyceae). *Can. J. Bot.*, **68:** 1873–1876.

JIN, L. & NEI, M. (1990). Limitations of the evolutionary parsimony method of phylogenetic analysis. *Mol. Biol. Evol.*, **7:** 82–102.

KLAVENESS, D. (1985). Classical and modern criteria for determining species of Cryptophyceae. *Bull. Plankton Soc. Japan*, **32:** 111–128.

KLAVENESS, D. (1988). Ecology of the Cryptomonadida: a first review. In *Growth and Reproductive Strategies of Freshwater Phytoplankton* (Sandgren, C.S., editor), 105–133. Cambridge University Press, Cambridge.

KUHNER, M.K. & FELSENSTEIN, J. (1994). A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Mol. Biol. Evol.*, **11:** 459–468.

LBBLANC, C., BOYEN, C. & GOËR, S.L. (1995). Organisation of the plastid genome from the rhodophyte *Chrondrus crispus* (Gigartinales); sequence and phylogeny of the 16S rRNA gene. *Eur. J. Phycol.*, **30:** 133–140.

LUDWIG, M. & GIBBS, S.P. (1985). DNA is present in the nucleomorph of cryptomonads: further evidence that the chloroplast evolved from a eukaryotic endosymbiont. *Protoplasma*, **127:** 9–20.

LUDWIG, M. & GIBBS, S.P. (1989). Localization of phycoerythrin at the luminal surface of the thylakoid membrane in *Rhodomonas lens*. *J. Cell Biol.*, **108:** 875–884.

MAIER, U.-G., HOFMANN, C.J.B., ESCHBACH, S., WOLTERS, J. & IGLOI, G. (1991). Demonstration of nucleomorph-encoded eukaryotic small subunit ribosomal RNA in cryptomonads. *Mol. Gen. Genet.*, **230:** 155–60.

MARGULIS, L. & SCHWARTZ, K.V. (1988). *Five Kingdoms*, 2nd edn. Freeman, New York.

MARTIN, W., SOMMERVILLE, C.C. & LOISEAUX-DE GOER, S. (1992). Molecular phylogenies of plastid origins and algal evolution. *J. Mol. Evol.*, **35:** 385–404.

MCFADDEN, G.I. (1993). Second-hand chloroplasts: evolution of cryptomonad algae. *Adv. Bot. Res.*, **19:** 189–230.

MCFADDEN, G.I. & GILSON, P.R. (1995). Something borrowed, something green: lateral transfer of chloroplasts by secondary endosymbiosis. *Trends Ecol. Evol.*, **10:** 12–17.

MCFADDEN, G.I., GILSON, P.R. & HILL, D.R.A. (1994*a*). *Goniomonas*: rRNA sequences indicate that this phagotrophic flagellate is a close relative of the host component of cryptomonads. *Eur. J. Phycol.*, **29:** 29–32.

MCFADDEN, G.I., GILSON, P.R. & DOUGLAS, S.E. (1994*b*). The photosynthetic endosymbiont in cryptomonad cells produces both chloroplast and cytoplasmic-type ribosomes. *J. Cell Sci.*, **107:** 649–657.

MCFADDEN, G.I., GILSON, P.R., HOFMANN, C.J.B., ADCOCK, G.J. & MAIER, U.-G. (1994*c*). Evidence that an amoeba acquired a chloroplast by retaining part of an engulfed eukaryotic alga. *Proc. Natl. Acad. Sci., U.S.A.*, **91:** 3690–3694.

MCFADDEN, G.I., GILSON, P.R. & WALLER, R.F. (1995). Molecular phylogeny of chlorarachniophytes based on plastid rRNA and *rbc*L sequences. *Arch. Protistenk.*, **145:** 231–239.

MCKERRACHER, L. & GIBBS, S.P. (1982). Cell and nucleomorph division in the alga *Cryptomonas*. *Can. J. Bot.*, **60:** 2440–2452.

MEDLIN, L., ELWOOD, H.J., STICKEL, S. & SOGIN, M.L. (1988). The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene*, **71:** 491–498.

MORRALL, S. & GREENWOOD, A.D. (1982). Ultrastructure of nucleomorph division in species of the Cryptophyceae and its evolutionary implications. *J. Cell Sci.*, **54:** 311–318.

MORSE, D., SALOIS, P., MARKOVIC, P. & HASTINGS, J.W. (1995). A nuclear-encoded form II RuBisCO in dinoflagellates. *Science*, **268:** 1622–1624.

NEI, M. (1991). The relative efficiency of different methods of phylogenetic reconstruction. In *Phylogenetic Analyses of DNA Sequences* (Miyamoto, M.M. & Cracraft, J., editors), 90–128. Oxford University Press, Oxford.

NELISSEN, B., VAN DE PEER, Y., WILMOTTE, A. & DE WACHTER, R. (1995). An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S rRNA sequences. *Mol. Biol. Evol.*, **12:** 1166–1173.

NOVARINO, G. (1991*a*). Observations on *Rhinomonas reticulata* comb. nov. and *R. reticulata* var. *eleniana* var. nov. (Cryptophyceae), with comments on the genera *Pyrenomonas* and *Rhodomonas*. *Nord. J. Bot.*, **11:** 243–252.

NOVARINO, G. (1991*b*). Observations on some new and interesting Cryptophyceae. *Nord. J. Bot.*, **11:** 599–611.

NOVARINO, G. & LUCAS, I.A.N. (1993*a*). Some proposals for a new classification system of the Cryptophyceae. *Bot. J. Linn. Soc.*, **111:** 3–21.

NOVARINO, G. & LUCAS, I.A.N. (1993*b*). A comparison of some morphological characters in *Chroomonas ligulata* sp. nov. and *C. placoidea* sp. nov. (Cryptophyceae). *Nord. J. Bot.*, **13:** 583–591.

NOVARINO, G. & LUCAS, I.A.N. (1995). A zoological classification system of cryptomonads. *Acta Protozool.*, **34:** 173–180.

NOVARINO, G., LUCAS, I.A.N. & MORRALL, S. (1994). Observations on the genus *Plagioselmis* (Cryptophyceae). *Cryptogamie, Algol.*, **15:** 87–107.

OLSEN, G.J. (1987). Earliest phylogenetic branchings: comparing rRNA-based evolutionary trees inferred with various techniques. *Cold Spring Harbor Symp. Quant. Biol.*, **52:** 825–837.

OLSEN, G.L., MATSUDA, H., HAGSTROM, R. & OVERBEEK, R. (1994). FastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comp. Appl. Bios.*, **10:** 41–48.

PALMER, J.D. (1995). Rubisco rules fall; gene transfer triumphs. *BioEssays*, **17:** 1005–1008.

PALMER, J.D. (1996). Rubisco surprises in dinoflagellates. *Plant Cell*, **8:** 343–345.

RAGAN, M.A. & GUTELL, R.R. (1995). Are red algae plants? *Bot. J. Linn. Soc.*, **118:** 81–105.

RAGAN, M.A., BIRD, C.J., RICE, E.L., GUTELL, R.R., MURPHY, C.A. & SINGH, R.K. (1994). A molecular phylogeny of the marine red algae (Rhodophyta) based on the nuclear small-subunit rRNA gene. *Proc. Natl. Acad. Sci., U.S.A.*, **91:** 7276–7280.

ROCHAIX, J.-D., MAYFIELD, S., GOLDSCHMIDT-CLERMONT, M. & ERICKSON, J. (1988). Molecular biology of *Chlamydomonas*. In *Plant Molecular Biology* (C. H. Shaw, editor), 253–275. Oxford University Press, Oxford.

ROWAN, R., WHITNEY, S.M., FOWLER, A. & YELLOWLEES, D. (1996). Rubisco in marine symbiotic dinoflagellates: form II enzymes in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family. *Plant Cell*, **8:** 555–564.

SANTORE, U.J. (1982*a*). Comparative ultrastructure of two members of the Cryptophyceae assigned to the genus *Chroomonas*: with comments on their taxonomy. *Arch. Protistenk.*, **125:** 5–29.

SANTORE, U.J. (1982*b*). The distribution of the nucleomorph in the Cryptophyceae. *Cell Biol. Int. Rep.*, **6:** 1055–1063.

SANTORE, U.J. (1984). Some aspects of the taxonomy in the Cryptophyceae. *New Phytol.*, **98:** 627–646.

SANTORE, U.J. (1985). A cytological survey of the genus *Cryptomonas* (Cryptophyceae) with comments on its taxonomy. *Arch. Protistenk.*, **130:** 1–52.

SANTORE, U.J. (1986). The ultrastructure of *Pyrenomonas heteromorpha* comb. nov. (Cryptophyceae). *Bot. Mar.*, **29:** 75–82.

SANTORE, U.J. (1987). A cytological survey of the genus *Chroomonas*: with comments on the taxonomy of this natural group of the Cryptophyceae. *Arch. Protistenk.*, **134:** 83–114.

SANTORE, U.J. & LEEDALE, G.F. (1985). Cryptomonadida. In *An Illustrated Guide to the Protozoa* (Lee, J.J., Hutner, S.H. & Bovee, E.C., editors), 19–22. Society of Protozoologists, Lawrence, Kansas.

SEPSENWOL, S. (1973). Leucoplast of the cryptomonad *Chilomonas paramecium*: evidence for presence of a true plastid in a colorless flagellate. *Exp. Cell Res.*, **76:** 395–409.

SITTE, P. (1993). Symbiogenetic evolution of complex cells and complex plastids. *Eur. J. Protistol.*, **29:** 131–143.

SMITH, S., WOESE, C.R., GILBERT, W., GILLEVET, P.M. (1994). The genetic data environment and expandable GUI for multiple sequence analysis. *Comp. App. Biosci.*, **10:** 671–675.

SWOFFORD, D.L. & OLSEN, G.J. (1990). Phylogeny reconstruction. In *Molecular Systematics* (Hillis, D.M. & Moritz, C., editors), 411–501. Sinauer, Sunderland, Mass.

VAN DE PEER, Y., NEEFS, J.-M., DE RIJK, P., DE WACHTER, R. (1993). Evolution of eukaryotes as deduced from small ribosomal subunit RNA sequences. *Biochem. System. Ecol.*, **21:** 43–55.

VAN DE PEER, Y., RENSING, S.A., MAIER, U.-G. & DE WACHTER, R. (1996). Substitution rate calibration of small subunit rRNA identifies chlorarachniophyte endosymbionts as remnants of green algae. *Proc. Natl. Acad. Sci., U.S.A.*, **93:** 4467–4471.

WHATLEY, J.M., JOHN, P. & WHATLEY, F.R. (1979). From extracellular to intracellular: the establishment of mitochondria and chloroplasts. *Proc. R. Soc. Lond. B*, **204:** 165–187.

WHITNEY, S.M., SHAW, D.C. & YELLOWLEES, D. (1995). Evidence that some dinoflagellates contain a ribulose-1,5-bisphosphate carboxylase–oxygenase related to that of the alpha proteobacteria. *Proc. Natl. Acad. Sci., U.S.A.*, **259:** 271–275.

WINNEPENNINCKX, B., BACKELJAU, T., MACKEY, L., BROOKS, J.M., DE WACHTER, R.D., KUMAR, S. & GAREY, J.R. (1995). 18S rRNA data indicate that Aschelminthes are polyphyletic in origin and consist of at least three distinct clades. *Mol. Biol. Evol.*, **12:** 1132–1137.

WOESE, C.R. (1987). Bacterial evolution. *Microbiol. Rev.*, **51:** 221–271.