Phylogeny and Molecular Evolution of the Green Algae

Frederik Leliaert,1 David R. Smith,2 Hervé Moreau,3 Matthew D. Herron,4 Heroen Verbruggen,1 Charles F. Delwiche,5 and Olivier De Clerck1

1Phycology Research Group, Biology Department, Ghent University 9000, Ghent, Belgium
2Canadian Institute for Advanced Research, Evolutionary Biology Program, Department of Botany, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada
3Observatoire Océanologique, CNRS–Université Pierre et Marie Curie 66651, Banyuls sur Mer, France
4Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada
5Department of Cell Biology and Molecular Genetics and the Maryland Agricultural Experiment Station, University of Maryland, College Park, MD 20742, USA

Table of Contents

I. THE NATURE AND ORIGINS OF GREEN ALGAE AND LAND PLANTS ................................................................. 2

II. GREEN LINEAGE RELATIONSHIPS ......................................................................................................................... 5
   A. Morphology, Ultrastructure and Molecules ............................................................................................................. 5
   B. Phylogeny of the Green Lineage .............................................................................................................................. 6
      1. Two Main Lineages: Chlorophyta and Streptophyta ............................................................................................ 6
      2. Early Diverging Chlorophyta: The Prasinophytes .............................................................................................. 6
      3. The Core Chlorophyta: Ecological and Morphological Diversification .............................................................. 9
      4. Streptophyta: Charophyte Green Algae and the Origin of Land Plants ............................................................... 15

III. SPREAD OF GREEN GENES IN OTHER EUKARYOTES ......................................................................................... 17

IV. GREEN ALGAL EVOLUTION: INSIGHTS FROM GENES AND GENOMES ............................................................ 19
   A. Organelle Genome Evolution .................................................................................................................................... 20
   B. Ecology and Molecular Evolution of Oceanic Picoplanktonic Prasinophytes ......................................................... 25
   C. Genomic Insights into the Evolution of Complexity in Volvocine Green Algae ......................................................... 26
   D. Genetic Codes and the Translational Apparatus in Green Seaweeds .................................................................... 29
   E. Molecular Evolution in the Streptophyta and the Origin of Land Plants ............................................................... 30

V. CONCLUSIONS AND PERSPECTIVES .................................................................................................................. 31

ACKNOWLEDGMENTS ............................................................................................................................................... 32

REFERENCES ............................................................................................................................................................... 32
The green lineage (Viridiplantae) comprises the green algae and their descendants the land plants, and is one of the major groups of oxygenic photosynthetic eukaryotes. Current hypotheses posit the early divergence of two discrete clades from an ancestral green flagellate. One clade, the Chlorophyta, comprises the early diverging prasinophytes, which gave rise to the core chlorophytes. The other clade, the Streptophyta, includes the charophyte green algae from which the land plants evolved. Multi-marker and genome scale phylogenetic studies have greatly improved our understanding of broad-scale relationships of the green lineage, yet many questions persist, including the branching orders of the prasinophyte lineages, the relationships among core chlorophyte clades (Chlorodendraophyceae, Ulvophyceae, Trebouxiophyceae and Chlorophyceae), and the relationships among the streptophytes. Current phylogenetic hypotheses provide an evolutionary framework for molecular evolutionary studies and comparative genomics. This review summarizes our current understanding of organelle genome evolution in the green algae, genomic insights into the ecology of oceanic picoplanktonic prasinophytes, molecular mechanisms underlying the evolution of complexity in volvocine green algae, and the evolution of genetic codes and the translational apparatus in green seaweeds. Finally, we discuss molecular evolution in the streptophyte lineage, emphasizing the genetic facilitation of land plant origins.

Keywords Chlorophyta, Charophyta, endosymbiosis, molecular evolution, origin of embryophytes, Prasinophyceae, phylogeny, Streptophyta

I. THE NATURE AND ORIGINS OF GREEN ALGAE AND LAND PLANTS

The green lineage or Viridiplantae includes the green algae and land plants, and is one of the major groups of oxygenic photosynthetic eukaryotes. Green algae are diverse and ubiquitous in aquatic and some terrestrial habitats, and they have played a crucial role in the global ecosystem for hundreds of millions of years (Falkowski et al., 2004; O’Kelly, 2007; Leliaert et al., 2011). The evolution of land plants from a green algal ancestor was a key event in the history of life and has led to dramatic changes in the earth’s environment, initiating the development of the entire terrestrial ecosystem (Kenrick and Crane, 1997).

The green lineage originated following an endosymbiotic event in which a heterotrophic eukaryotic host cell captured a cyanobacterium that became stably integrated and ultimately turned into a plastid (Archibald, 2009; Keeling, 2010). This primary endosymbiosis, which likely happened between 1 and 1.5 billion years ago (Hedges et al., 2004; Yoon et al., 2004), marked the origin of the earliest oxygenic photosynthetic eukaryotes. The subsequent diversification of this primary plastid-containing eukaryote gave rise to the green lineage, as well as the red algae and the glaucophytes. From this starting point, photosynthesis spread widely among diverse eukaryotic protists via secondary and tertiary endosymbioses, which involved captures of either green or red algae by non-photosynthetic protists (Keeling, 2010). Secondary endosymbioses involving green algae as the autotrophic partner have given rise to three groups of algae: the chlorarachniophytes, the photosynthetic euglenids and the “green” dinoflagellates (see section III. Spread of green genes in other eukaryotes). The other eukaryotic algal groups, the cryptophytes, haptophytes, photosynthetic stramenopiles (e.g., diatoms, chrysophytes and brown seaweeds) and dinoflagellates, have acquired plastids from a red algal ancestor, either by a single or multiple endosymbiotic events (Archibald, 2009; Bodily et al., 2009; Baurain et al., 2010).

The green lineage is ancient, and dating its origin has been a difficult task because of the sparse fossil record of the group. The earliest fossils attributed to green algae date from the Precambrian (ca. 1200 mya) (Tappan, 1980; Knoll, 2003). The nature of these early fossils, however, remains controversial (e.g., Cavalier-Smith, 2006). The resistant outer walls of prasinophyte cysts (phycocysts) are well preserved in fossil deposits and especially abundant and diverse in the Paleozoic era (ca. 250–540 mya) (Parke et al., 1978; Tappan, 1980; Colbath, 1983). A filamentous fossil (Proterocladus) from the middle Neoproterozoic deposits (ca. 750 mya) has been attributed to siphonocladous green algae (Cladophorales) (Butterfield et al., 1994; Butterfield, 2009), while the oldest reliable records of the siphonous seaweeds (Bryopsidales, Dasycladales) and stoneworts (Charophyceae) are from the Paleozoic (Hall and Delwiche, 2007; Verbruggen et al., 2009a). The earliest land plant fossils are Mid-Ordovician in age (ca. 460 mya) (Kenrick and Crane, 1997; Steemans et al., 2009). Molecular clock analyses have estimated the origin of the green lineage between 700 and 1500 mya (Douzery et al., 2004; Hedges et al., 2004; Berney and Pavlowski, 2006; Roger and Hug, 2006; Herron et al., 2009). These estimates are sensitive to differences in methodology and interpretation of fossils and tend to yield older dates than are well supported by the fossil record. This could be attributable to miscalibration of the molecular clock estimates or to taphonomic bias and the difficulty of interpreting fossils with no modern exemplars. Molecular phylogenetic evidence has provided a substantially improved understanding of the relationships among major lineages. Reconstruction of ancestral character states could assist in the reinterpretation of known specimens of uncertain affinity, and this, combined with continued paleontological investigation, holds out hope for reconciliation of molecular and fossil evidence.

Green algae are characterized by a number of distinct features, many of which are also shared with the land plants (van den Hoek et al., 1995; Graham et al., 2009). The chloroplasts are enclosed by a double membrane with thylakoids grouped in lamellae, and contain chlorophyll a and b along with a...
set of accessory pigments such as carotenes and xanthophylls. Pyrenoids, when present, are embedded within the chloroplast and are surrounded by starch, which is the main reserve polysaccharide. Most green algae have firm cell walls with a fiber matrix generally composed of cellulose. The flagellate cells are isokont, which means that the two flagella are similar in structure, although they may differ in length. The flagellar transition zone (i.e., the region between the flagellum and its basal body) is typically characterized by a stellate structure, which is a nine-pointed star, visible in cross section using an electron microscope, linking nine pairs of microtubules (Melkonian, 1984).

Several green algae serve as model systems or are of economic importance. Melvin Calvin used cultures of *Chlorella* to elucidate the light-independent reactions of photosynthesis. Now known as the Calvin cycle (Calvin and Benson, 1948). Transplant experiments with the giant-celled *Acetabularia*, conducted by Joachim Hämmerling, demonstrated that the nucleus of a cell contains the genetic information that directs cellular development, and postulated the existence of messenger RNA before its structure was determined (Hämmerling, 1953). *Acetabularia*, along with other giant-celled green algae (*Valonia, Chara* and *Nitella*), has also served as an experimental organism for electro-physiological research and studies of cell morphogenesis (Menzel, 1994; Mandoli, 1998; Shepherd *et al.*, 2004; Bisson *et al.*, 2006; Mine *et al.*, 2008). The charophyte *Mougeotia* played a key role in outlining the role of phytochrome in plant development (Winands and Wagner, 1996). The biochemistry and physiology of the unicellular, halophilic *Dunaliella salina* have been studied in great detail. This alga is among the most industrially important microalgae because it can produce massive amounts of β-carotene that can be collected for commercial purposes, and because of its potential as a feedstock for biofuels production (Oren, 2005; Gouveia and Oliveira, 2009; Tafresh and Shariati, 2009). The unicellular flagellate *Chlamydomonas reinhardtii* has long been used as a model system for studying photosynthesis, chloroplast biogenesis, flagellar assembly and function, cell-cell recognition, circadian rhythm and cell cycle control because of its well-defined genetics, and the development of efficient methods for nuclear and chloroplast transformation (Rochaix, 1995; Harris, 2001; Grossman *et al.*, 2003; Breton and Kay, 2006). The colonial green alga *Volvox* has served as a model for the evolution of multicellularity, cell differentiation, and colony motility (Kirk, 1998; Kirk, 2003; Herron and Michod, 2008; Herron *et al.*, 2009).

Analysis of the complete nuclear genome sequence of *C. reinhardtii* greatly advanced our understanding of ancient eukaryotic features such as the function and biogenesis of chloroplasts, flagella and eyespots, and regulation of photosynthesis (Merchant *et al.*, 2007; Kreimer, 2009; Peers *et al.*, 2009). Genome data are rapidly accumulating and to date seven complete green algal genomes have been sequenced: the prasinophytes *Ostreococcus tauri* (Derelle *et al.*, 2006), *O. lucimaricus* (Palenik *et al.*, 2007) and two isolates of *Micromonas pusilla* (Worden *et al.*, 2009), the chlorophytes *C. reinhardtii* (Merchant *et al.*, 2007) and *Volvox carteri* (Prochnik *et al.*, 2010), and the trebouxiophyte *Chlorella variabilis* (Blanc *et al.*, 2010). Several other genome projects are ongoing, including the complete sequencing of *Coccomyxa, Dunaliella, Bathycoccus, Botryococcus* and additional *Ostreococcus* and *Micromonas* strains (Tirichine and Bowler, 2011). These data provide a great resource for in-depth analysis of genome organization and the processes of eukaryotic genome evolution. In addition, green algal genomes are important sources of information for the evolutionary origins of plant traits because of their evolutionary relationship to land plants.
FIG. 1. Taxonomical, morphological and ecological diversity among green algae. A: *Pterosperma* (Pyramimonadales), a marine prasinophyte characterized by quadriflagellate unicells (photo by Bob Andersen, reproduced under license from microscope.mbl.edu). B: *Nephroselmis* (Nephroselmidophyceae), a prasinophyte with bean-shaped cells and two unequal flagella occurring in marine or freshwater environments. C: *Palmyphylum* (Palmyphyllales) forming lobed crusts composed of coccoid cells embedded in a gelatinous matrix, growing in deep-water marine habitats (photo by Jordi Regas). D: *Tetraselmis* (Chlorodendrophyceae), quadriflagellate unicells from marine or freshwater habitats. E: *Chlorella* (Trebouxiophyceae, Chlorellales), coccoid cells, endosymbiontic inside the single-celled protozoan *Paramacium* (photo by Antonio Guillén). F: *Oocystis* (Trebouxiophyceae, Oocystaceae), small colonies of coccoid cells within a thin mucilaginous envelope from freshwater habitats (image copyright Microbial Culture Collection, NIES). G: *Haematococcus* (Chlorophyceae, Chlamydomonadales), a freshwater biflagellate unicell (photo by William Bourland, reproduced under license from microscope.mbl.edu). H: *Pediastrum* (Chlorophyceae, Sphaeropleales), a coenobial colony of non-motile cells arranged in a circular plate, occurring in freshwater habitats. I: *Bulbochaete* (Chlorophyceae, Oedogoniales), branched filaments with terminal hair-cells. J: *Chaetophora* (Chlorophyceae, Chaetophorales), highly branched filaments, occurring in freshwater habitats (H–J: photos by Jason Oyadomari). K: *Ulothrix* (Ulvophyceae, Ulotrichales), unbranched filaments from marine or brackish areas (photo by Giuseppe Vago). L: *Ulva* (Ulvophyceae, Ulvales), sheet-like plants, mainly from marine habitats (photo by Tom Schils). M: *Cladophora* (Ulvophyceae, Cladophorales), branched filament with cells containing numerous chloroplasts and nuclei (photo by Antonio Guillén). N: *Boergesenia* (Ulvophyceae, Cladophorales), plants composed of giant, multinucleate cells, from tropical marine habitats (photo by HV). O: *Acetabularia* (Ulvophyceae, Dasyycladales), siphonous plants (i.e. single-celled) differentiated into a stalk and a flattened cap, with a single giant nucleus situated at the base of the stalk; typically found in subtropical marine habitats (photo by Antoni López-Arenas). P: *Caulerpa* (Ulvophyceae, Bryopsidophyceae), siphonous plant differentiated into creeping stolons anchored by rhizoids and erect photosynthetic fronds, containing millions of nuclei; occurring in (sub)tropical marine waters (photo by FL). Q: *Klebsormidium* (Klebsormidophyceae), unbranched filamentous charophyte, mainly from moist terrestrial habitats (photo by Jason Oyadomari). R: *Spirotaenia*, a unicellular charophyte with typical spiral chloroplast; generally growing in acidic freshwater habitats (photo by Antonio Guillén). S: *Nitella* (Charophyceae), morphologically complex charophyte from freshwater habitats, consisting of a central stalk and whorls of branches radiating from nodes that bear oogonia and antheridia (photo by Nadia Abdelahad). T: *Microasterias* (Zygnematophyceae, Desmidiales), characterized by non-motile unicells constricted in two parts with ornamented cell wall; generally found in acidic, oligotrophic freshwater habitats (photo by Antonio Guillén). U: *Coleochaete* (Coleochaetophyceae), branched filaments adherent to form a disc-like, parenchymatous thallus; found in freshwater habitats, often as epiphytes on submerged vascular plants (photo by CFD). (Color figure available online.)
Reconstruction of the phylogenetic relationships among green plants is essential to identifying the innovations underlying the diversity of green algae and land plants. Molecular phylogenetics has dramatically reshaped our views of green algal relationships and evolution. This review summarizes the current understanding of green algal phylogeny, focusing primarily on relationships among the major lineages of green algae, which are usually classified as divisions, classes and orders. Current phylogenetic hypotheses have provided an evolutionary framework for molecular evolutionary studies and comparative genomics. In this review, we highlight a number of topics, including the evolution of organellar genomes, ecology and molecular evolution of marine picoplanktonic prasinophytes, genomic insights into the evolution of complexity in volvocine green algae, molecular evolution in the green seaweeds, and molecular evolution in the streptophyte green algae and the origin of land plants.

II. GREEN LINEAGE RELATIONSHIPS

A. Morphology, Ultrastructure and Molecules

Early hypotheses of green algal phylogeny were based on the concept that evolution follows trends in levels of morphological complexity (Fritsch, 1935; Fott, 1971). Unicellular flagellates were believed to have initially evolved into non-motile unicells (coccoid) and loose packets of cells (sarcinoid), followed by various multicellular forms and siphonous algae. This hierarchy reflected the view that the morphologies that are organized in two- and three-dimensional space require more elaborate developmental controls, and hence would be expected to appear later in an evolutionary sequence. In this view the land plants were derived from more complex, filamentous green algae (Blackman, 1900; Pascher, 1914).

A large amount of new information was gathered in the 1970s and 1980s, mainly from investigations of the fine structures of green algal cells and life cycles (Round, 1984). These data led to a thorough reevaluation of evolutionary relationships and a revised classification of green algae, primarily based on flagellar ultrastructure and processes of mitosis and cell division (Picket-Heaps and Marchant, 1972; Melkonian, 1982; Mattox and Stewart, 1984; Melkonian, 1984; O’Kelly and Floyd, 1984a; O’Kelly and Floyd, 1984b; van den Hoek et al., 1988). These features, which apply to most (but not all) green algae, were believed to accurately reflect phylogenetic relationships because of their involvement in fundamental processes of cell replication and cell motility, and thus to be less liable to convergent evolution than gross morphological traits. Ultrastructure-based phylogenetic hypotheses posited an early diversification of flagellate unicells, resulting in a multitude of ancient lineages of flagellates, some of which then evolved into more complex green algae. Although ultrastructural data have laid the foundations for a natural classification of the green algae, analyses of these data have not resolved the phylogenetic relationships among the main green algal lineages.

The introduction of molecular phylogenetic methods provided a new framework for reconstructing the evolutionary history of the green lineage. Analyses of DNA sequence data took a start in the mid 1980s with the first phylogenies of green plants inferred from 5.8S nrDNA sequences (Hori et al., 1985; Hori and Osawa, 1987), soon followed by 18S and 28S nrDNA sequence analyses (Gunderson et al., 1987; Perasso et al., 1989; Buchheim et al., 1990; Zechman et al., 1990; Chapman et al., 1991; Mishler et al., 1992; Chapman et al., 1998). Ribosomal sequences were chosen at the time because enough RNA could be obtained for direct sequencing and (later) because regions of the gene were conserved enough to make universal primers. Ribosomal RNA is also found in all known living cells, and although it is typically present in many copies in the genome, concerted evolution homogenizes sequence diversity, and consequently is assumed to reduce the risk of sequencing paralogs. Early molecular studies were pre-PCR and hence involved laboratory cloning and generally few taxa. Since 1990 rapid advancements in techniques in molecular biology (e.g., the utilization of PCR) and bioinformatics made it possible to generate and analyze larger datasets. Nuclear-encoded 18S rDNA sequences have been, until recently, the primary source of data for inferring phylogenetic relationships among green algae (Pröschold and Leliaert, 2007), supplemented by 28S rDNA (e.g., Buchheim et al., 2001; Shoup and Lewis, 2003), actin (An et al., 1999) and the chloroplast genes rbcL, tufA and atpB (e.g., Daugbjerg et al., 1994; Daugbjerg et al., 1995; McCourt et al., 2000; Hayden and Waaland, 2002; Nozaki et al., 2003; Zechman, 2003; Rindi et al., 2007).

These initial molecular phylogenetic investigations have generally corroborated the ultrastructure-based higher-level classification of the green algae, but have also revised the circumscriptions of several lineages (McCourt, 1995). However, analyses of individual genes have only partly resolved the relationships among the main green algal lineages. It is now clear that a large number of genes from many species must be analysed to arrive at a reliable phylogenetic resolution for an ancient group such as the green algae (Philippe and Telford, 2006). These datasets have mainly involved concatenated sequences of protein-coding genes that are shared among green algal chloroplast genomes. To date, 26 complete green algal plastid genomes have been sequenced and assembled (Wakasugi et al., 1997; Turmel et al., 1999b; Lemieux et al., 2000; Maul et al., 2002; Turmel et al., 2002b; Pombert et al., 2005; Turmel et al., 2005; Bélanger et al., 2006; de Cambiaire et al., 2006; Pombert et al., 2006; Turmel et al., 2006; de Cambiaire et al., 2007; Lemieux et al., 2007; Robbens et al., 2007a; Brouard et al., 2008; Turmel et al., 2008; Turmel et al., 2009a; Turmel et al., 2009b; Zuccarello et al., 2009; Brouard et al., 2010; Brouard et al., 2011), in addition to more than 30 angiosperm plastid genomes (Soltis et al., 2009). Chloroplast genomes are particularly useful for phylogenetic reconstruction because of their relatively high and condensed gene content, in comparison to nuclear genomes. Furthermore, in contrast to many nuclear genes that are multi-copy...
in nature, which can confound phylogenetic reconstruction, organelar genes are typically single-copy and do not pose these problems. Only recently have multi-gene analyses of nuclear genes been carried out (e.g., Rodriguez-Espeleta et al., 2007; Cocquyt et al., 2010b; Finet et al., 2010).

A persistent problem in many of the multi-gene phylogenetic investigations thus far is sparse and incomplete taxon sampling, which may result in systematic errors in phylogenetic reconstruction. For this reason, phylogenetic hypotheses derived from such studies are preferably supported by independent data, such as structural genomic features like gene content, gene order, gene structure or intron distribution (Lemieux et al., 2007). But ideally, these systematic errors would be overcome by increasing taxon sampling. Recent advances in high-throughput DNA sequencing, including Roche-454 and Illumina-Solexa (Shendure and Ji, 2008; Metzker, 2010) facilitate rapid sequencing of organelar genomes, transcriptomes and entire nuclear genomes. This fast accumulation of genomic data in conjunction with new developments in phylogenetic inference techniques is creating unprecedented research opportunities. The reconstruction of large-scale multi-gene phylogenies and studies of the molecular mechanisms underlying this diversity is now a short-term feasible prospect.

B. Phylogeny of the Green Lineage

1. Two Main Lineages: Chlorophyta and Streptophyta

Current hypotheses on green algal evolution posit the early divergence of two discrete lineages: the Chlorophyta and Streptophyta (Figure 2) (Picket-Heaps and Marchant, 1972; Bremer, 1985; Lemieux et al., 2007). The Chlorophyta includes the majority of described species of green algae. The Streptophyta are comprised of the charophytes, a paraphyletic assemblage of freshwater algae, and the land plants.

Charophytes range in morphology from unicellular to complex multicellular organisms (Figure 1Q–U) and primarily occur in freshwater and, to a lesser extent, terrestrial habitats. They display a number of ultrastructural and biochemical traits that are shared with land plants but not with the Chlorophyta: motile cells (when present) with two subapically inserted flagella and an asymmetrical flagellar apparatus that contains a distinctive multilayered structure (MLS) and parallel basal bodies, open mitosis and persistent mitotic spindle, a phragmoplast-type cytokinesis (in many species), and the presence of several unique enzyme systems such as Cu/Zn superoxide dismutase, class I aldolases and glycolate oxydase (Frederick et al., 1973; Mattox and Stewart, 1984; de Jesus et al., 1989; Jacobshagen and Schnarrenberger, 1990; Graham and Kaneko, 1991). The Streptophyta are more broadly characterized by discrete molecular features, such as a GapA/B gene duplication in the common ancestor of the clade (Petersen et al., 2006), while other molecular characteristics, such as polyadenylation signals, show a more complex phylogenetic distribution (Wodniok et al., 2007).

Members of the Chlorophyta are common inhabitants of marine, freshwater and terrestrial environments and exhibit a remarkable morphological diversity (Figure 1A–P). The flagellar apparatus is typically characterized by a symmetrical cruciate root system wherein rootlets of variable (X) numbers of microtubules alternate with rootlets composed of two microtubules to form a “X-2-X-2” arrangement (Moestrup, 1978). The orientation of the flagellar root system has served as an important character for defining the main groups of Chlorophyta (Mattox and Stewart, 1984; O’Kelly and Floyd, 1984b). When viewed from the anterior (flagellar) end of the cell, the flagellar basal bodies and rootlets can have a perfect cruciate pattern (i.e., with basal bodies directly opposed, DO) or they are offset in a counter-clockwise (CCW) or clockwise (CW) position (for a detailed overview, see van den Hoek et al., 1995; Graham et al., 2009). Traditionally, four classes are recognized within the Chlorophyta: the freshwater or terrestrial Trebouxiophyceae and Chlorophyceae, the coastal Ulvophyceae, and the unicellular, predominantly marine planktonic Prasinophyceae.

Resolution of the earliest green algal divergences often remains elusive because of the antiquity of the green lineage and the rapidity of the early evolutionary radiations (McCourt, 1995). Figures 2 and 3 illustrate our current understanding of green algal relationships. The topology presented is a somewhat conservative interpretation of available data with polytomies representing uncertainties or conflicts between different studies. The evidence supporting these relationships is discussed below on a clade-by-clade basis.

2. Early Diverging Chlorophyta: The Prasinophytes

The Prasinophyceae originally comprised motile unicells with organic body scales (Mattox and Stewart, 1984). The ultrastructure of the body scales and the flagellar apparatus, along with cell division and mitotic features, have long served as the main characters to distinguish among the major groups of Prasinophyceae (Melkonian, 1990b; Sym and Pienaar, 1993; Becker et al., 1994; Marin and Melkonian, 1994). Molecular data have provided clear evidence that the class forms a paraphyletic assemblage of early diverging lineages (Kantz et al., 1990; Steinkötter et al., 1994; Nakayama et al., 1998; Leliaert et al., 2011), which is reflected by the wide diversity of cell shapes, body scale morphologies, flagellar behaviour, mitotic and cytokinetic processes, and biochemical features (Sym and Pienaar, 1993; Latasa et al., 2004).

Based on the early diverging nature of the Prasinophyceae, the ancestor of the green lineage is believed to be a flagellate unicell. However, the nature of this ancestral green flagellate (AGF) remains tentative. Simple flagellates have been proposed to most closely resemble the AGF (Moestrup, 1991). Alternatively, the food uptake apparatus of some complex mixotrophic prasinophytes (e.g., Cymbomonas) has been interpreted as a feature inherited from a phagotrophic ancestor of the green lineage that was subsequently lost in most green algae (Moestrup et al., 2003; O’Kelly, 2007; Turmel et al., 2009a).

The circumscription of the prasinophytes has changed considerably with the discovery of several new taxa and
MOLECULAR EVOLUTION OF GREEN ALGAE

FIG. 2. Overview phylogeny of the green lineage (top) and spread of green genes in other eukaryotes (bottom). (Color figure available online.)

environmental sequencing (Hasegawa et al., 1996; Nakayama et al., 1998; Fawley et al., 2000; Zingone et al., 2002; Guillou et al., 2004; Vaulot et al., 2008; Viprey et al., 2008; Lepère et al., 2009; Shi et al., 2009). The prasinophytes, as presently conceived, include a heterogeneous assemblage of scaled (e.g., Nephroselmis, Pyramimonas and Mamiella) and naked (e.g., Ostreococcus, Micromonas and Scourfieldia) unicells with diverse morphologies with one to eight flagella. In addition, several coccolid forms have been found, which are distributed among at least four clades (Fawley et al., 2000; Guillou et al., 2004; Jouenne et al., 2011). Prasinophytes are predominantly marine planktonic but also include several freshwater representatives (Marin and Melkonian, 2010). Although sexual reproduction has rarely been documented in prasinophytes (but see Suda et al., 2004), some members (e.g., Pyramimonadales) produce resistant cysts (phycomata) containing two chloroplasts, indicative of sexual reproduction (Moestrup et al., 2003). Sexual reproduction has also been implied in Micromonas and Ostreococcus based on the presence of meiosis-specific and sex-related genes in the four sequenced genomes (see below) (Derelle et al., 2006; Worden et al., 2009; Grimsley et al., 2010).

The paraphyletic nature of prasinophytes has important consequences for understanding early green algal evolution, and a well-resolved phylogeny has the potential to shed light on the nature of the common ancestor of the Chlorophyta and on the origin of the core chlorophytes (Turmel et al., 2009a). Multi-gene phylogenetic analyses are just beginning to shed light on these early divergences. To date five chloroplast genomes from prasinophyte representatives of four clades (Mamiellaphyceae, Nephroselmidiphyceae, Pycnococcaceae, and Pyra-
FIG. 3. A consensus reconstruction of green algal relationships, based on molecular data. Details and support for various clades are discussed in the text. Poorly resolved or conflicting relationships are shown as polytomies. Non-monophyletic lineages are indicated by dotted lines. Branch lengths roughly represent genetic distance deduced from different phylogenetic studies. (Color figure available online.)
mimonadales) have been sequenced, in addition to Mesostigma, a former member of the prasinophytes that is now known to be an early-branching streptophyte (Turmel et al., 1999b; Lemieux et al., 2000; Robbens et al., 2007a; Turmel et al., 2009a).

The Mamiellophyceae is the largest lineage of prasinophytes and unites the morphologically and ecologically diverse Mamiellales, the Monomastigales (Monomastix) and Dolichomastidales (Dolichomastix and Crustomastix) (Marin and Melkonian, 2010). The Mamiellales includes marine and freshwater flagellates with one or two laterally inserted flagella, as well as cocccoid forms, with or without body scales. They include some of the smallest eukaryotes known (e.g., Ostreococcus, Micromonas), which are prominent in the oceanic picoplankton (Courties et al., 1994; O’Kelly et al., 2003; Guillou et al., 2004; Not et al., 2004). The clade is further characterized by prasinoxanthin, a pigment that is also found in some other clades, such as the Pycnococcaceae (Zingone et al., 2002; Latasa et al., 2004). The phylogenetic affinity of the freshwater, unflagellate Monomastix has long been uncertain because it lacks scales and has atypical surface structures (Manton, 1967; Melkonian, 1990b; Sym and Pienaar, 1993). Analyses of 18S sequence and chloroplast genome data clearly show that Monomastix is sister to the Mamiellales (Turmel et al., 2009a). The biflagellate Dolichomastidales are also morphologically distinct: Dolichomastix has atypical scales and Crustomastix has a specialized cell covering rather than the more typical body scales (Nakayama et al., 2000; Zingone et al., 2002). Their phylogenetic placement in Mamiellophyceae has been demonstrated by 18S data (Zingone et al., 2002). The relationships among the Mamiellales, Monomastigales and Dolichomastidales, however, have not been resolved.

Chloroplast phylogenomic analyses resolved a sister relationship between the Mamiellophyceae and the Pyramimonadales (Turmel et al., 2009a), a clade of relatively large quadriflagellates with diverse and complex body scale coverings that occur in marine and freshwater habitats (Melkonian, 1990b). As mentioned above, some Pyramimonadales (e.g., Cymbomonas) are unique among green algae in possessing a food uptake apparatus (Moestrup et al., 2003).

The Pycnococcaceae is a small clade containing the marine biflagellate Pseudosocourfieldia and the naked cocoidal Pycnococcus. Some 18S studies have related this clade with the Nephrozelmidophyceae (although never with strong support; Fawley et al., 2000; Guillou et al., 2004), which is corroborated by some morphological similarities (e.g., Nephrozelmis and Pseudosocourfieldia both have two unequal flagella and two body and flagellar scale layers) (Nakayama et al., 2007). However, chloroplast multi-gene analyses do not support a relationship between the two clades and have identified the Nephrozelmidophyceae as an early diverging prasinophyte lineage, while the phylogenetic position of the Pycnococcaceae remains equivocal (Turmel et al., 2009a).

Environmental sequencing of photosynthetic picocyanobacteria communities in the Mediterranean Sea and SE Pacific Ocean has identified two additional prasinophyte clades (clades VIII and IX) (Viprey et al., 2008; Lepère et al., 2009; Shi et al., 2009). However, the nature and phylogenetic affinities of these clades remain elusive.

Zecharia et al. (2010) provided evidence for another early diverging lineage of green algae, the Palmophyllales (Leliaert et al., 2011). This lineage includes Palmophyllum and Verdigellas, green seaweeds that thrive in deepwater and other dimly lit, benthic marine habitats (Womersley, 1984; Nelson and Ryan, 1986; Ballantine and Norris, 1994). Palmophyllales feature a unique type of multicellularity, forming firm, well-defined macroscopic thalli composed of isolated spherical cells in a gelatinous matrix (Pueschel et al., 1997). Analysis of two plastid genes (atpB and rbcL) placed the Palmophyllales as the sister clade to all other Chlorophyta while 18S phylogenetic analysis allied the Palmophyllales with the Prasinococcales clade. This clade includes the cocoidal prasinophytes Prasinococcus and Prasinoderma, and it has also been shown to form an early diverging prasinophyte clade based on 18S data (Guillou et al., 2004; Turmel et al., 2009a). The possible relationship between the Palmophyllales and Prasinococcales is supported by a number of shared cytogenetical features and similarities in cell division (O’Kelly, 1988; Hasegawa et al., 1996; Zecharia et al., 2010; Jouenne et al., 2011).

Sequencing of environmental samples and cultures has identified a clade of cocoidal prasinophytes (CCMP1205 clade) that, together with the saline lake dwelling cocoidal Picocystis, emerges as a sister lineage to the core chlorophytes, although strong support for this relationship is still lacking (Guillou et al., 2004; Marin and Melkonian, 2010).

3. The Core Chlorophyta: Ecological and Morphological Diversification

The prasinophytes have given rise to the morphologically and ecologically diverse core chlorophytes. This group includes the early diverging Chlorodendrophyceae, and three major classes: Ulvophyceae, Trebouxiophyceae, and Chlorophyceae (UTC). The core chlorophytes are characterized by a new mode of cell division that is mediated by a phycoplast (i.e., a system of microtubules that develops parallel to the plane of nuclear division), which was subsequently lost in the Ulvophyceae.

The Chlorodendrophyceae is a small clade unifying the marine or freshwater scaly quadriflagellates Tetraselmis and Scherffelia (Guillou et al., 2004). These unicells were traditionally regarded as members of the Prasinophyceae but they share several features with the UTC clades, including closed mitosis and a phycoplast (Mattox and Stewart, 1984). The close relationship with the UTC classes has been confirmed by 18S and multi-gene phylogenetic data (Massjuk, 2006; Cocquyt et al., 2010b).

3.1. Radiation of the Ulvophyceae, Trebouxiophyceae and Chlorophyceae. The UTC classes are species-rich and morphologically and ecologically diverse. Ecophysiological adaptations have likely led to the success of the Chlorophyceae and
Trebouxiophyceae in freshwater and terrestrial environments, while the Ulvophyceae mainly diversified in coastal ecosystems. Marine versus freshwater lifestyles also coincide with differentiations in life histories. Whereas the marine Ulvophyceae mainly have life cycles involving an alternation between a free-living haploid, gametophytic and a free-living diploid, sporophytic multicellular generation, most freshwater green algae have a haploid vegetative phase and a single-celled, often dormant zygote as the diploid stage. In terrestrial members of the core chlorophytes, sexual reproduction has rarely been documented (Rindi, 2011).

Resolving the phylogenetic relationships among and within the UTC classes can provide important insights into the evolution of these ecophysiological, life history and morphological traits. Molecular phylogenetic studies based on 18S sequences (Zechman et al., 1990; López-Bautista and Chapman, 2003; Watanabe and Nakayama, 2007), chloroplast and mitochondrial multigene data (Pombert et al., 2004; Pombert et al., 2005; Turmel et al., 2009b) have yielded ambivalent results. All possible relationships have been hypothesized, depending on interpretation of ultrastructural characters, gene and taxon sampling, and phylogenetic methods used. Furthermore, the monophyly of the Trebouxiophyceae and Ulvophyceae has not been unequivocally demonstrated (Pröschold and Leliaert, 2007; Zuccarello et al., 2009). The unstable relationships exhibited among these three classes likely result from their antiquity and the short time span over which they diverged from one another (O’Kelly, 2007; Cocquyt et al., 2010b). The fossil record indicates the presence of the classes in the mid-Neoproterozoic and molecular clock estimates situate the UTC divergence in the early Neoproterozoic (Butterfield et al., 1994; Douzery et al., 2004; Herron et al., 2009).

Some early 18S phylogenies showed a sister relationship between Chlorophyceae and Trebouxiophyceae (e.g., Krienitz et al., 2001), while more recent studies with increased taxon sampling revealed a sister relationship between the Chlorophyceae and Ulvophyceae (e.g., Friedl and O’Kelly, 2002; Watanabe and Nakayama, 2007; De Wever et al., 2009). Chloroplast multi-gene phylogenetic analyses have generally supported a sister relationship between Ulvophyceae and Trebouxiophyceae (Pombert et al., 2005; Turmel et al., 2009b), although some analyses based on nucleotide sequences suggest a sister relationship between Ulvophyceae and Chlorophyceae (Turmel et al., 2008; Turmel et al., 2009a). The latter topology is also supported by a mitochondrial multi-gene analysis (Pombert et al., 2004), a phylogenetic analysis of eight nuclear and two plastid genes (Cocquyt et al., 2010b), and structural chloroplast genome data (Pombert et al., 2005; Turmel et al., 2009b). Ultrastructural data have been interpreted as either providing supports for a sister relationship between Chlorophyceae and Trebouxiophyceae or between Trebouxiophyceae and Ulvophyceae. A relationship between Chlorophyceae and Trebouxiophyceae has been suggested based on the shared presence of a non-persistent mitotic spindle (Mattox and Stewart, 1984). A relationship between Trebouxiophyceae and Ulvophyceae was proposed based on a counter-clockwise orientation of the flagellar apparatus (Sluiman, 1989). In contrast, molecular data generally support an early diverging Trebouxiophyceae. This would imply the ancestral status of a CCW orientation of the flagellar basal bodies, which evolved to a DO and CW orientation in the Chlorophyceae. This interpretation is congruent with a CCW flagellar root system in the Chlororaphidophyceae.

The genus-level systematics of the core chlorophytes has been profoundly mislead by convergent evolution toward reduced morphologies such as unicells or simple filaments, and many genera defined on these features have been shown to be polyphyletic (Lewis and McCourt, 2004). For example, strains conforming to the morphological circumscription of the genera Neochloris, Characium and Planophila have been found in all three classes (Watanabe and Floyd, 1989; Lewis et al., 1992; Watanabe et al., 2000; Friedl and O’Kelly, 2002). Similarly, species of the coccosoid genus Chlorococcum and the filamentous genus Uronema were present in the Chlorophyceae as well as the Ulvophyceae (Watanabe et al., 2001; Krienitz et al., 2003; Leliaert et al., 2009a). Species of Chlorella were shown to be distributed between the Chlorophyceae and Trebouxiophyceae (Huss and Sogin, 1990; Huss et al., 1999; Luo et al., 2010).

3.2 Trebouxiophyceae. The Trebouxiophyceae was originally defined (as Pleurostrophophyceae) based on ultrastructural features (CCW flagellar apparatus orientation, non-persistent metacentric spindle, and phycoplast-mediated cytokinesis) (Mattox and Stewart, 1984) and its circumscription was later refined by 18S sequence data (Kantz et al., 1990; Friedl, 1995; Wolf et al., 2003) (see Lewis and McCourt, 2004, for a taxonomic and nomenclatural history of the class). As presently conceived, the class encompasses motile and non-motile unicells, colonies and multicellular filaments or blades from freshwater or terrestrial habitats, with some species penetrating in brackish or marine waters. Several species engage in symbiotic relationships with fungi to form lichens (Friedl and Büdel, 1996; Friedl and Bhattacharya, 2002). Others are photosynthetic endosymbionts in various freshwater and marine protists, invertebrates and plants, including ciliates, mussels, hydra, sea anemones and Ginkgo (“zoochlorellae”) (Karakashian and Karakashian, 1965; Tremouillaux-Guiller et al., 2002; Lewis and Muller-Parker, 2004; Hoshina and Imamura, 2008; Summerer et al., 2008; Letsch et al., 2009). Molecular phylogenetic data have shown that many of these endosymbiotic relationships evolved multiple times independently (Hoshina and Imamura, 2008; Pröschold et al., 2011). A number of trebouxiophytes have lost photosynthetic capacity and have evolved heterotrophic free-living or highly adapted parasitic lifestyles (e.g., Prototheca and Helicosporidium) (de Koning and Keeling, 2006; Pombert and Keeling, 2010).

Several distinct lineages within the Trebouxiophyceae have been recovered by molecular data, including the Chlorellales, Trebouxiales, Microthamniales, Prasiola-clade and several
clades that have not yet received a formal name (Katana et al., 2001; Friedl and O’Kelly, 2002; Krienitz et al., 2003; Zhang et al., 2008; Neustupa et al., 2011). Resolving the branching patterns among these lineages can provide important insights into morphological and ecological evolution, and provide clues about the origins and adaptations of symbiotic and parasitic lifestyles, ultimately leading to obligate heterotrophic lifestyles. Knowledge of trebouxiophycean interrelationships, however, is limited. Single gene (mostly 18S) analyses were not able to resolve the relationships among the main lineages, and some analyses even cast doubt on the monophyly of the group (Krienitz et al., 2003).

Molecular data have revealed extreme polyphyly of morphologically simple genera indicating convergent evolution toward reduced morphology. In addition, polyphyly has been demonstrated in genera with more complex colonial forms (e.g., Dictyosphaerium, Microactinium and Diacanthos) (Krienitz et al., 2010; Pröschold et al., 2010). On the other hand, molecular phylogenies have shown close relationships between highly dissimilar morphologies, demonstrating that evolution of vegetative morphology can be rapid (Lewis and McCourt, 2004).

The Chlorellales mainly includes freshwater and terrestrial coccolid forms and a few marine members (Aslam et al., 2007). The non-photosynthetic, parasitic Prototheca and Helicosporidium are also included in the Chlorellales based on nuclear and chloroplast gene data (Huss and Sogin, 1990; Tartar et al., 2003; Ueno et al., 2003; de Koning et al., 2005; Ueno et al., 2005). Although plastids are not apparent in Helicosporidium cells, molecular evidence indicates that it maintains a functional plastid genome (Tartar et al., 2003; de Koning and Keeling, 2004; Tartar and Boucias, 2004; de Koning and Keeling, 2006). Diversity and phylogenetic relationships within the Chlorellales have been well studied (Huss et al., 2002; Krienitz et al., 2003; Henley et al., 2004; Zhang et al., 2008; Luo et al., 2010; Pazoutova et al., 2010; Pröschold et al., 2010). 18S and chloroplast genome data provided evidence that the Oocystaceae, including the semi-colonial Oocystis and the spiny Lagerheimia, form an early diverging clade within the Chlorella (Heppner et al., 2000; Krienitz et al., 2003; Turmel et al., 2009b).

The enigmatic Pedinomonas is likely related to the Chlorellales. This tiny naked uniflagellate has traditionally been placed in a separate order Pedinomonadales or class Pedinophyceae based on some unusual ultrastructural features (Melkonian, 1990a; Moestrup, 1991). However, chloroplast multi-gene analyses and gene linkage data place Pedinomonas firmly within the Chlorellales (Turmel et al., 2009b). This placement is supported by the CCW orientation of the flagellar apparatus but the persistent telophase spindle in Pedinomonas is atypical for the Trebouxiophyceae. It should also be noted that phylogenetic analyses of seven mitochondrial genes have placed Pedinomonas sister to the Chlorophyceae on a long branch, but these analyses may suffer from low taxon sampling and systematic errors in phylogeny reconstruction (Turmel et al., 1999a; Pombert et al., 2004; Pombert and Keeling, 2010).

The Prasiola-clade is the morphologically and ecologically most diverse trebouxiophyte lineage, including unicellular (e.g., Pseudochlorella, Pseudomarvania, Stichococcus), filamentous (e.g., rosespheniella, Raphidonema) and blade-like (e.g., Prasiola) forms growing in freshwater, marine and terrestrial habitats (Karsten et al., 2005; Rindi et al., 2007; Elias and Neustupa, 2009).

The Trebouxiales is best known as the “lichen algae” group and includes the common phycobiont genera Trebouxia and Asterochloris (Friedl and Bhattacharya, 2002; Blaha et al., 2006; Skaloud and Peksa, 2010). The order also includes several free-living soil algae (e.g., Myrmecia) (Friedl, 1995). The Microthamniales, including the branched filamentous Microthamnion and the cocolid Fusochloris, is thought to be allied with the Trebouxiales (Lewis and McCourt, 2004), but strong support for this relationship is lacking (Sluiman et al., 2008; Neustupa et al., 2009).

Several other distinct clades of trebouxiophytes, composed of taxa with diverse morphologies and ecologies, have been characterized but their phylogenetic affinities to the better-studied lineages remain uncertain. These clades include the Choricystis lineage, comprising Choricystis, Botryoccus and Coccomyxa (Karsten et al., 2005), the Watanabea clade uniting Dictychloropsis and Watanabea (Karsten et al., 2005), and various other species that form distinct lineages, Lobosphaera, Parietochloris, Coenocystis, Picoclorum and Xylochloris (Lewis et al., 1992; Hanagata and Chihara, 1999; Henley et al., 2004; Karsten et al., 2005; Neustupa et al., 2011).

The phylogenetic placement of the terrestrial, filamentous genus Leptosira is uncertain (Lokhorst and Rongen, 1994). 18S analyses placed this genus within the Trebouxiophyceae on a long branch with uncertain position (Friedl, 1996; De Wever et al., 2009; Neustupa et al., 2011), while chloroplast multi-gene analyses placed Leptosira on a long branch, sister to the Chlorophyceae with moderate support (Turmel et al., 2009a; Turmel et al., 2009b; Zuccarello et al., 2009). Given the fact that Leptosira was the only trebouxiophycean representative outside the Chlorellales in these analyses, this relationship may be the result of systematic error in phylogenetic reconstruction (Turmel et al., 2009b).

3.3. Chlorophyceae. The Chlorophyceae are a large and morphologically diverse group, including non-motile and motile unicells, colonies, branched and unbranched filaments, and blade-like thalli. Reproduction is equally diverse, including various asexual and sexual modes (van den Hoek et al., 1995). Chlorophycean algae are especially abundant in freshwater but also occur in terrestrial habitats. The class is characterized by closed mitosis during cell division, phycoplast-mediated cytokinesis, and diverse configurations of the flagellar apparatus of motile cells.

Molecular phylogenetic analyses have drastically reshaped the classification of the class. Traditional orders and genera were found to be polyphyletic while others were moved to other
classes (e.g., Chlorellales) (see Lewis and McCourt, 2004 for a taxonomic history). Molecular and ultrastructural data have identified five major clades (Booton et al., 1998a; Buchheim et al., 2001; Wolf et al., 2002; Wolf et al., 2003; Turmel et al., 2008). The Chlamydomonadales is characterized by a CW flagellar apparatus orientation in biflagellate members, but quadriflagellate representatives may display various other orientations (Nakayama et al., 1996a; Nozaki et al., 2003; Watanabe et al., 2006b). The Sphaeropleales include nonmotile unicells and colonies that produce biflagellate zoospores with a DO basal body configuration (Deason et al., 1991). The Chaetophorales and Chaetopeltidales are both characterized by quadriflagellate motile cells. The orientation of the flagellar apparatus of the Chaetophorales is variable (DO or CW) (Manton, 1964; Melkonian, 1975; Floyd et al., 1980; Watanabe and Floyd, 1989), whereas the Chaetopeltidales have a DO orientation (O’Kelly and Floyd, 1984a; O’Kelly et al., 1994). The Oedogoniales are characterized by zoospores with an anterior ring of numerous flagella (stephanokont) (Pickett-Heaps, 1975). The atypical flagellar apparatus in the Chaetopeltidales, Oedogoniales and Sphaeropleales has hampered homology assessment with flagellar characters occurring in the other two clades (Turmel et al., 2008).

The phylogenetic relationships among the five main clades of Chlorophyceae have proven difficult to resolve. Phylogenetic analyses of nuclear rDNA data have suggested a sister relationship between the Chlamydomonadales and Sphaeropleales, with the Chaetophorales, Oedogoniales, and Chaetopeltidales forming early diverging clades with uncertain interrelationships (Booton et al., 1998a; Buchheim et al., 2001; Shoup and Lewis, 2003; Müller et al., 2004; Alberghina et al., 2006). Chloroplast multi-gene analyses identified two main lineages within the Chlorophyceae: a clade uniting the Chlamydomonadales and Sphaeropleales (CS clade), and a clade uniting the Oedogoniales, Chaetophorales, and Chaetopeltidales (OCC clade) (Turmel et al., 2008; Brouard et al., 2010). This dichotomy was independently supported by molecular signatures in chloroplast genes, such as presence of indels and transspliced introns. Within the OCC clade, the sister relationship observed for the Chaetophorales and Chaetopeltidales has also been supported by rDNA data (Buchheim et al., 2001; Müller et al., 2004; Caisovsky et al., 2011), similarities in zoospore structure (O’Kelly et al., 1994) and the occurrence of two trans-spliced group II introns at identical positions (Turmel et al., 2008). These results have strengthened the notion that the ancestor of the Chlorophyceae was likely a quadriflagellate with a DO+DO configuration of the flagellar apparatus, and that the CW configuration evolved independently in the Chlamydomonadales and the Chaetophorales (Buchheim et al., 2001; Turmel et al., 2008) (but see Nozaki et al., 2003, discussed below), and further suggest that the stephanokont zoospores of the Oedogoniales arose from the DO condition (Turmel et al., 2008).

The Chlamydomonadales forms the largest group of Chlorophyceae and has a complex taxonomic history. Molecular phylogenetic analyses of nuclear ribosomal DNA and chloroplast sequences have profoundly changed the concept of the class, which now includes a diverse range of taxa formally placed in the Dunaliellales, Chlorococcales, Tetrarogales, Chlororosarcinales, Volvocales, and Chaetophorales (Lewis et al., 1992; Wilcox et al., 1992; Buchheim et al., 1994; Nakayama et al., 1996a; Nakayama et al., 1996b; Booton et al., 1998b; Krienitz et al., 2003; Lewis and McCourt, 2004; Watanabe et al., 2006a; Watanabe et al., 2006b; Nakada et al., 2008a; Nakada et al., 2008b). As presently conceived, the class includes non-motile or motile unicells with two or four flagella, biflagellate colonies, filaments and cells that are imbedded in a mucilage envelope (Nakazawa et al., 2004; Nozaki et al., 2006b; Yamada et al., 2008; Nakada and Nozaki, 2009; Matsuza et al., 2010; Novis et al., 2010; Nozaki et al., 2010). The Chlamydomonadales mainly comprise freshwater and terrestrial green algae, including several psychrophilic and/or halotolerant species (e.g., Dunaliella and Chlamydomonas nivalis) (Eddie et al., 2008; Buchheim et al., 2010; Muramoto et al., 2010; Remias et al., 2010). A few species have evolved an obligate heterotrophic life style (e.g., Polyionema, Polytonella) and have plastids with highly modified functions and genomes (Vernon et al., 2001). 18S phylogenetic analyses revealed 21 main clades with a poorly resolved backbone (Nakada et al., 2008a). Several genera are polyphyletic, with the most extreme example being the large genus Chlamydomonas, which is distributed in at least five distinct lineages within the Chlamydomonadales (Buchheim et al., 1990; Buchheim et al., 1996; Pröschold et al., 2001; Nakada et al., 2008a). Phylogenetic analyses based on chloroplast genes and 18S sequences showed that quadriflagellate members with CCW or CW basal body orientation (Golenkinia, Carteria, Pseudocarteria, Hafniomonas, Trochisica, and Cylindrocapsa) form early diverging lineages, suggesting that the CW orientation of most chlamydomonadanean genera may have evolved from the CCW orientation in ancestral quadriflagellate members (Nozaki et al., 2003; Nakada et al., 2008a).

Because the Chlamydomonadales include unicellular flagellates (e.g., Chlamydomonas) and colonial forms (e.g., Volvox), its members have been extensively studied as models of multicellular evolution (Kirk, 2003; Kirk, 2005; Herron and Michod, 2008; Herron et al., 2009; Prochnik et al., 2010; Ueki et al., 2010). Interestingly, several phylogenetic studies have demonstrated non-monophyly of colonial flagellates and the absence of a phylogenetic trend in the level of colony complexity, complicating evolutionary interpretations (Buchheim and Chapman, 1991; Buchheim et al., 1994; Herron and Michod, 2008; Nakada et al., 2008a; Nakada et al., 2010). We will return to the subject of evolution of multicellularity in the section “Genomic insights into the evolution of complexity in volvocine green algae.”

The Sphaeropleales form another large group of Chlorophyceae, including some of the most common freshwater phytoplankters such as Scenedesmus, Desmodesmus and Pediasstrum, as well as picoplanktonic members (Wolf et al., 2002; Krienitz et al., 2011). Representatives are non-motile unicells, colonies
or filaments, which produce biflagellate zoospores (zoosporic) or non-motile spores (autosporic). The filamentous \textit{Microspora} has been allied with the coccoid genus \textit{Bracteacoccus} based on ultrastructure of the flagellar apparatus and 18S data (Lokhorst and Star, 1999; Buchheim and Buchheim, 2001). Monophyly of the Sphaeropleales has been supported by unique rDNA ITS2 secondary structure characteristics (Keller et al., 2008) and phylogenetic analysis of nuclear rDNA atpB and rbcL sequences (Vergheze, 2007). Diversity and phylogenetic relationships within the Sphaeropleales have been studied by Lewis (1997), Krienitz et al. (2001, 2003, 2011), Buchheim et al. (2001, 2005), Shoup and Lewis (2003), Wolf et al. (2003), Manus and Lewis (2005, 2011) and Johnson et al. (2007).

In contrast to the species-rich CS clade, the OCC orders are much less diverse. The Chaetophorales of about 10 genera include unbranched (e.g., \textit{Uronema}) or branched (e.g., \textit{Chaetophora}, \textit{Stigeoclonium}) filaments that produce quadriflagellate zoospores. Phylogenetic relationships within the order have been studied based on 18S sequence data (Caisová et al., 2011). The Chaetopeltidales, originally erected based on ultrastructural data (O’Kelly et al., 1994), includes vegetative disk-like thalli (e.g., \textit{Chaetopeltis}, \textit{Pseudulvella}) (Sanchez-Puerta et al., 2006), or three-dimensional packets of cells (e.g., \textit{Floydella}) that reproduce asexually by quadriflagellate spores. 18S phylogenetic relationships demonstrated that \textit{Planophila}, which was originally included in the order, is polyphyletic and distributed among the three core ulvophycean classes (Friedel and O’Kelly, 2002). The Oedogoniales includes the filamentous genera \textit{Oedogonium}, \textit{Oedocladium} and \textit{Bulbochaete}. They share a unique form of cytokinesis and a specialized form of oogamous sexual reproduction involving the production of stephanokont motile cells (Pickett-Heaps, 1975). Phylogenetic relationships within the order have been studied by Alberghina et al. (2006) and Mei et al. (2007).

3.4. Ulvophyceae. The Ulvophyceae was originally defined based on a suite of ultrastructural characteristics, including a CCW orientation of the flagellar root system, cytokinesis by furrowing, closed mitosis with a persistent telophase spindle and the absence of a phycoplast (Mattox and Stewart, 1984; O’Kelly and Floy, 1984a; O’Kelly and Floy, 1984b; Sluiman, 1989). The class is best known for its morphocarpean marine representatives (the green seaweeds), but several members also occur in freshwater or damp subaerial habitats such as soil, rocks, tree bark and leaves (Chihara et al., 1986; López-Bautista and Chapman, 2003; Rindi et al., 2006; Watanabe and Nakayama, 2007).

In terms of diversity in thallus complexity and cellular sophistication, the Ulvophyceae far exceed the other chlorophycean classes. Their morphologies range from microscopic unicells to macroscopic multicellular plants, and giant-celled organisms with unique cellular and physiological characteristics (Mine et al., 2008). Four main cytological types can be distinguished (Copquyt et al., 2010b). The first type comprises non-motile uninucleate unicells, and is present in some Ulotrichales (Chihara et al., 1986; Nakayama et al., 1996a; Friedl and O’Kelly, 2002; Watanabe and Nakayama, 2007) and a number of genera of uncertain affinity (e.g., \textit{Oltmannsielopsis} and \textit{Ignatiuss}). The second type consists of multicellular filaments or blades composed of uninucleate cells; it characterizes most Ulvales, Ulotrichales and Trentephyliales. The third type has multicellular bodies composed of multinucleate cells with nuclei organized in regularly spaced cytoplasmic domains (McNaughton and Goff, 1990; Motomura, 1996). This is known as the siphonoclados type and characterizes the Cladophorales and \textit{Blastophysea} and some members of the Ulotrichales (e.g., \textit{Urospora} and \textit{Acrosiphonia}). The fourth type is better known as the siphonous type and is characterized by plants consisting of a single giant tubular cell. It is present in the orders Bryopsidales and Dasyycladales. Siphonous cells generally contain thousands to millions of nuclei. Many Dasyycladales, however, remain uninucleate throughout much of their life cycle with a giant diploid nucleus that only divides at the onset of reproduction (Berger and Kaever, 1992). In contrast to the siphonoclados type, the cytoplasm of siphonous algae exhibits vigorous streaming, enabling transportation of transcripts across the plant (Menzel, 1987; Menzel, 1994; Mine et al., 2001). Although some siphonous algae are tiny microscopic siphons, many form large and complex seaweeds that exhibit morphological differentiation into structures that resemble the roots, stems, and leaves of land plants and have similar functions (Chisholm et al., 1996). The evolution of siphonoclados and siphonous architectures coincided with several cytological and cytoskeletal specializations such as unique mechanisms of wounding response (Menzel, 1988; La Claire, 1992; Kim et al., 2001; Mine et al., 2008). In addition to facilitating transport of transcripts as mentioned above, the evolution of cytoplasmic streaming in siphonous algae also allowed transport of nutrients and organelles throughout the siphonous algal body. In combination with morphological changes, this allows nutrient uptake from marine sediments (Chisholm et al., 1996) and chloroplast migration to optimize photosynthesis and avoid herbivory by micrograzers. Such innovations have most likely had selective advantages and contribute to the ecological dominance of siphonous algae in tropical and warm-temperate coastal ecosystems (Vroom and Smith, 2001).

Understanding the phylogenetic relationships among the ulvophycean orders can provide important insights into the diversification of cytological types and evolution of morphological complexity in the class. Analyses of 18S datasets have supported the circumscription of traditional orders based on cytological characteristics but could not resolve the relationships among them (Chappell et al., 1991; López-Bautista and Chapman, 2003; Watanabe and Nakayama, 2007; Lelievaert et al., 2009a). Moreover, the monophyly of the class has been questioned because it lacks unique ultrastructural synapomorphies (Mattox and Stewart, 1984; O’Kelly and Floy, 1984a). Several molecular phylogenetic studies provided weak or no support for monophyly of the Ulvophyceae and recovered two distinct
lineages: the Oltmannsiellopsidales–Ulvalles–Ulotrichales lineage and a clade consisting of Trentepohliales and the siphonoclados and siphonous seaweed orders (Zechman et al., 1990; Watanabe et al., 2001; López-Bautista and Chapman, 2003; Watanabe and Nakayama, 2007; Cocquyt et al., 2009; Zuccarello et al., 2009). More recently, a 10-gene phylogenetic analysis recovered the class as a well supported monophyletic group, and confirmed the divergence of two main ulvophycean clades (Cocquyt et al., 2010b). These two discrete clades are also supported by independent molecular data, including the distribution of elongation factor genes and a deviant genetic code (Cocquyt et al., 2009; Gile et al., 2009; Cocquyt et al., 2010a). The phylogeny of Cocquyt et al. (2010b) recovered the Trentepohliales as the sister lineage of the Bryopsidales, Cladophorales, and Dasycladales. The Bryopsidales and Dasycladales were recovered as sister lineages, and the genus Blastophypha was sister to the Cladophorales. The relationship of the enigmatic endophytic green alga Blastophypha with the Cladophorales is corroborated by morphological, ultrastructural, cytological, and biochemical features (O’Kelly and Floyd, 1984a; Chappell et al., 1991). On the other hand, the grouping of the Dasycladales and Bryopsidales was surprising given that ultrastructural features of the flagellar apparatus suggested that Dasycladales may be more closely related to Cladophorales (O’Kelly and Floyd, 1984a).

The phylogeny of Cocquyt et al. (2010b) has provided the groundwork for new insights into the cytomorphic dimension of the Ulvophyceae from an ancestral unicell to a large genus Cladophora (including Oltmannsiellopsis, Boergesenia, and Pseudoneochloris) may provide important insights into the evolution of multicellularity in the Ulvophyceae, but as yet, their phylogenetic positions are poorly resolved (Watanabe et al., 2003; Leliaert et al., 2004a; O’Kelly et al., 2004a; O’Kelly et al., 2004b; O’Kelly et al., 2004c; Lindstrom and Hanic, 2005; Lindstrom et al., 2006; Loughnane et al., 2008; Heesch et al., 2009; Kraft et al., 2010; O’Kelly et al., 2010). The relationships of unicellular or sarcinoid members (e.g., Desmochloris, Halochlorococcum, Pseudoneochloris) may provide important insights into the evolution of multicellularity in the clade, but as yet, their phylogenetic positions are poorly resolved (Watanabe et al., 2003; O’Kelly et al., 2004a; O’Kelly et al., 2004c).

The Oltmannsiellopsidales is generally inferred to have diverged near the base of the Ulvales-Ulotrichales clade (Watanabe and Nakayama, 2007; Cocquyt et al., 2009; Leliaert et al., 2009a). The order consists of a small number of marine and freshwater species. Oltmannsiellopsidales includes quadriflagellate unicells or small colonies (Nakayama et al., 1996a). Dangemania microcystis and Halochlorococcum marimum are non-motile unicells, disks or packets of cells that produce quadriflagellate spores (Friedl and O’Kelly, 2002; Pröschold et al., 2002).

The Cladophorales (including Siphonocladales) evolved in four main lineages, likely from a filamentous marine ancestor (Bakker et al., 1994; Hanyuda et al., 2002; Leliaert et al., 2003; Leliaert et al., 2009a). The marine unbranched, microfilamentous Okellya forms the earliest diverging lineage, followed by the Aegagropila clade, which is often found in highly specialized freshwater habitats (e.g., lake-balls or epizoophytes on turtles and snails). The species-rich Cladophora clade, including branched (Cladophora) or unbranched (Chaetomorpha and Rhizoclonium) filaments, is predominantly marine, but at least one clade has invaded freshwater habitats. The marine tropical Siphonocladales clade is morphologically most diverse, comprising highly specialized forms and giant-cells with unique cytomorphic traits and modes of cell division (e.g., Valonia, Boergesenia, Dictyosphaerium) (Okuda et al., 1997a; Okuda et al., 1997b; Mine et al., 2008). Molecular phylogenies have shown that these specialized features have evolved independently multiple times in the clade (Leliaert et al., 2007). The large genus Cladophora, and several other genera have been found to be polyphyletic (Leliaert et al., 2003; Leliaert et al., 2009b).

The Bryopsidales (also referred to as the Caulerpaiales, Codi-ales, or Siphonales) range in morphology from simple, branched

habitats have occurred independently (Shimada et al., 2008; Ichihara et al., 2009; Suutari et al., 2010). The orders Ulvales and Ulvophyceae have traditionally been distinguished based on life history features: the Ulvales have a life cycle involving isomorphic alternations of multicellular stages, while in the Ulvophyceae the diploid, spore-producing stage is a small, thick-walled unicell that is attached to the substrate by a stalk (Codium stage) (O’Kelly et al., 2004b). Molecular phylogenetic studies provide poor or no support for the separation of the orders (e.g., O’Kelly et al., 2004a; O’Kelly et al., 2004c). Diversity and phylogenetic relationships within the clade have been studied based on plastid genes (mainly rbcL) and nuclear 18S sequence data (Hayden and Waaland, 2002; Hayden and Waaland, 2004; O’Kelly et al., 2004a; O’Kelly et al., 2004b; O’Kelly et al., 2004c; Lindstrom and Hanic, 2005; Lindstrom et al., 2006; Loughnane et al., 2008; Heesch et al., 2009; Kraft et al., 2010; O’Kelly et al., 2010). The relationships of unicellular or sarcinoid members (e.g., Desmochloris, Halochlorococcum, Pseudoneochloris) may provide important insights into the evolution of multicellularity in the clade, but as yet, their phylogenetic positions are poorly resolved (Watanabe et al., 2003; O’Kelly et al., 2004a; O’Kelly et al., 2004c).

The Oltmannsiellopsidales is generally inferred to have diverged near the base of the Ulvales-Ulotrichales clade (Watanabe and Nakayama, 2007; Cocquyt et al., 2009; Leliaert et al., 2009a). The order consists of a small number of marine and freshwater species. Oltmannsiellopsidales includes quadriflagellate unicells or small colonies (Nakayama et al., 1996a). Dangemania microcystis and Halochlorococcum marimum are non-motile unicells, disks or packets of cells that produce quadriflagellate spores (Friedl and O’Kelly, 2002; Pröschold et al., 2002).

The Cladophorales (including Siphonocladales) evolved in four main lineages, likely from a filamentous marine ancestor (Bakker et al., 1994; Hanyuda et al., 2002; Leliaert et al., 2003; Leliaert et al., 2009a). The marine unbranched, microfilamentous Okellya forms the earliest diverging lineage, followed by the Aegagropila clade, which is often found in highly specialized freshwater habitats (e.g., lake-balls or epizoophytes on turtles and snails). The species-rich Cladophora clade, including branched (Cladophora) or unbranched (Chaetomorpha and Rhizoclonium) filaments, is predominantly marine, but at least one clade has invaded freshwater habitats. The marine tropical Siphonocladales clade is morphologically most diverse, comprising highly specialized forms and giant-cells with unique cytomorphic traits and modes of cell division (e.g., Valonia, Boergesenia, Dictyosphaerium) (Okuda et al., 1997a; Okuda et al., 1997b; Mine et al., 2008). Molecular phylogenies have shown that these specialized features have evolved independently multiple times in the clade (Leliaert et al., 2007). The large genus Cladophora, and several other genera have been found to be polyphyletic (Leliaert et al., 2003; Leliaert et al., 2009b).
siphons (e.g., Bryopsis, Chlorodesmis) to more complex, differentiated thalli (e.g., Codium, Halimeda, Udotea, Caulerpa). Several species form key components of tropical marine coastal ecosystems, where they are among the major primary producers on coral reefs, in lagoons and seagrass beds. The thallus surface of several species is calcified, and some are important contributors to coral reef structure. A single genus, Dichotomosiphon, occurs in freshwater habitats. A phylogeny based on rbcL sequence data (Lam and Zechman, 2006) revealed two main clades, corresponding to the suborders Bryopsidinae and Halimedinae, which are characterized based on differences in thallus morphology, reproduction and plastid types (Hillis-Colinvaux, 1984). Increased gene and taxon sampling improved phylogenetic resolution in these clades and revealed the limestone-boring genus Ostreobium as an early diverging lineage (Verbruggen et al., 2009a). Furthermore, non-monophyly has been demonstrated for several genera and species (Verbruggen et al., 2009b).

The Dasyycladales are marine tropical algae characterized by radially symmetrical thalli encrusted with calcium carbonate. Two families have traditionally been recognized, Dasyycladales and Polypophysaceae, based on differences in the reproductive structures. Early phylogenetic studies based on 18S or rbcL sequence data indicated that at least the Dasyycladales was not monophyletic (Olsen et al., 1994; Berger et al., 2003; Zechman, 2003). A 5-gene phylogeny with extended taxon sampling demonstrated that the Dasyycladales form a paraphyletic assemblage that gave rise to a monophyletic Polypophysaceae (Verbruggen et al., 2009a). The Dasyycladales have a rich fossil record and were much more diverse historically than they are today.

The Trentepohliales is an entirely terrestrial order and was originally omitted from the Ulvophyceae based on atypical characteristics such as a multilayered structure (MLS) in the flagellar root system, phragmoplast-like cytokinesis, the presence of plasmodesmata between vegetative cells, and a unique type of sporangial reproduction implying a relationship with streptophyte green algae (Chapman, 1984; O’Kelly and Floyd, 1984a; Chapman et al., 2001). However, molecular data firmly established an alliance with the ulvophycean orders Cladophorales, Bryopsidales and Dasyycladales (Zechman et al., 1990; Lopez-Bautista and Chapman, 2003; Cocquyt et al., 2010b), either pointing towards parallel evolution of the streptophyte-like ultrastructural features or indicating that some of these characteristics (e.g., a MLS) may represent an ancestral condition in the green lineage (Lewis and McCourt, 2004); either way it is highly interesting that a phragmoplast-like cytokinesis has evolved in Trentepohliales and land plants, both found in terrestrial environments. The phylogenetic position of the Trentepohliales within a clade of marine orders suggests a sea-to-land transition, which would be unique among algae (Lewis and McCourt, 2004). Diversity and phylogenetic relationships in the order have been studied based on rbcL and 18S rDNA sequence data (Lopez-Bautista et al., 2006; Rindi et al., 2009).

4. Streptophyta: Charophyte Green Algae and the Origin of Land Plants

As one of the deepest branches of the green lineage, the Streptophyta are certainly old, quite likely a billion years or more, and both molecular sequence divergence and the few available fossils are consistent with that view. However, despite this great age, most of the subclades of Streptophyta are not particularly diverse, with the striking exception of the embryophytes (land plants), a lineage with several hundred thousand species.

Ultrastructural, biochemical and molecular data revealed six distinct groups of charophytes: Charophyceae, Coleochaetophyceae, Zygnematothyrceae, Klebsormidiothyrceae, Chlorokybophyceae and probably the Mesostigmatophyceae (Mattox and Stewart, 1984; Qiu and Palmer, 1999; McCourt et al., 2004). Considerable progress has been made during the past decade in clarifying the relationships among these lineages, and elucidating the closest living relative of the land plants (Karol et al., 2001; Turmel et al., 2003; Turmel et al., 2006; Qiu et al., 2006; Lemieux et al., 2007; Rodriguez-Ezpeleta et al., 2007; Finet et al., 2010; Wodniok et al., 2011).

4.1. The early diverging charophytes. Mesostigma and Chlorokybus likely form the earliest diverging lineage of extant streptophytes (Bhattacharya et al., 1998; Simon et al., 2006; Turmel et al., 2006; Lemieux et al., 2007; Rodriguez-Ezpeleta et al., 2007), although some early chloroplast and mitochondrial phylogenomic analyses with more limited taxon sampling suggested that Mesostigma emerged before the divergence of the Streptophyta and Chlorophyta (Lemieux et al., 2000; Turmel et al., 2002c). Additional evidence for the placement of Mesostigma in the Streptophyta comes from the presence of a MLS in the flagellar root system (Rogers et al., 1981; Melkonian, 1989), a shared multigene family (BIP) between Mesostigma and the rest of the Streptophyta (Nedelcu et al., 2006), and a shared duplication of the GapA/GapB gene (Petersen et al., 2006). Plastid phylogenomic analyses implied a sister relationship between Mesostigma and Chlorokybus (Lemieux et al., 2007; Rodriguez-Ezpeleta et al., 2007), while two nuclear multi-gene phylogenies suggested that Chlorokybus diverged after Mesostigma (Cocquyt et al., 2010b; Finet et al., 2010). The non-sister relationship between Mesostigma and Chlorokybus is further supported by the fact that Mesostigma is the only streptophyte with flagella in its vegetative stage (a presumed ancestral feature of the green plant lineage). This suggests that loss of flagella occurred after the divergence of Mesostigma (Figure 4).

Morphologically, Mesostigma and Chlorokybus are among the simplest streptophyte algae. Mesostigma (the only genus in the Mesostigmatophyceae) is a freshwater, scaly, asymmetrical unicell with two flagella (Marin and Melkonian, 1999) and a unique suite of photosynthetic pigments (Yoshii et al., 2003; Hall and Delwiche, 2007). Chlorokybus (the sole member of the Chlorokybophyceae) occurs in moist terrestrial habitats where it forms packets of a few cells that reproduce asexually.
by asymmetrical, motile spores with ultrastructural streptophyte characteristics (Rogers et al., 1980; Lokhorst et al., 1988). Sexual reproduction is unknown in the two lineages. Spirotaenia, which is normally placed in the Zygnematophyceae based on the lack of flagella and sexual reproduction by conjugation, was unexpectedly inferred as a sister lineage to Chlorokybus based on 18S rDNA and chloroplast rbcL sequence data (Gontcharov and Melkonian, 2004).

Multi-gene phylogenies show that the Klebsormidiophyceae diverged after the Mesostigmatophyceae and Chlorokybophyceae, and is sister to the clade comprising the Charophyceae, Zygnematophyceae, Coleochaetophyceae and land plants (Karol et al., 2001; Turmel et al., 2002a; Cocquyt et al., 2010b; Finet et al., 2010; Wodniok et al., 2011). This phylogenetic position is further supported by some chloroplast genomic features, such as the presence of introns in two transfer RNAs which is shared between Klebsormidium and the later-diverging streptophyte lineages, and the presence of functional plastid tufA, which is shared with the early-diverging lineages (Baldauf et al., 1990; Baldauf and Palmer, 1990; Manhart and Palmer, 1990; Turmel et al., 2005; Turmel et al., 2007b) (Figure 4). The Klebsormidiophyceae includes the freshwater or terrestrial genera Klebsormidium, Interfilum and Entranzia, which form unbranched filaments or sarcinoid packets that produce asexual motile spores with streptophyte characteristics (Cook, 2004b; Sluiman et al., 2008). Diversity and phylogenetic relationships within the Klebsormidiophyceae have been studied by Rindi et al. (2008, 2011) and Mikhailyuk et al. (2008).

In contrast to the three early-diverging streptophyte lineages (Mesostigmatophyceae, Chlorokybophyceae and Klebsormidiophyceae), which undergo cytokinesis by furrowing, the remaining lineages produce a phragmoplast and plasmodesmata (absent in the Zygnematophyceae), facilitating intercellular communication and differentiation (Pickett-Heaps, 1975; Raven, 1997). Additional traits that are uniquely shared by the remaining streptophytes include apical growth (not in the Zygnematophyceae), sexual reproduction, some biochemical features, and similar cellulose-synthesizing rosettes (Tsekos, 1999).

4.2. The closest living relative of the land plants. Numerous studies have focused on identifying the closest living relative to the land plants, and several lineages have been proposed based on morphological, ultrastructural and molecular data (Graham, 1984; Graham, 1993; Kranz et al., 1995; Bhattacharya and Medlin, 1998; An et al., 1999; Graham et al., 2000; McCourt et al., 2004). Multi-marker phylogenies have been sensitive to taxon and gene sampling and provided equivocal results, suggesting the Charophyceae (Karol et al., 2001; Turmel et al., 2007a; Cocquyt et al., 2010b), Zygnematophyceae (Turmel et al., 2006; Lemieux et al., 2007; Rodríguez-Ezpeleta et al., 2007; Wodniok et al., 2011), or Coleochaetophyceae (Turmel et al., 2009a; Turmel et al., 2009b; Finet et al., 2010) as the sister lineage of the embryophytes. Interestingly, mitochondrial and plastid data have yielded inconsistent positions for the Charales (Turmel et al., 2006; Turmel et al., 2007a), which might result from lateral transfer of large genomic segments or even whole organellar genomes (Turmel et al., 2008).

Two recent multi-gene phylogenetic analyses based on nuclear-encoded genes either placed the Coleochaetophyceae or Zygnematophyceae as sister to the land plants (Finet et al., 2010; Wodniok et al., 2011).

A first analysis based on 77 nuclear ribosomal protein genes suggested that Coleochaete represents the closest relative of
land plants (Finet et al., 2010), which agrees with earlier morphology-based hypotheses (Graham, 1984; Graham, 1993; Graham et al., 2000). *Coleochaete* ranges in morphology from filamentous to relatively complex discoid parenchymatous thalli (Graham, 1982; Graham, 1984; Dupuy et al., 2010). In some species the zygote is retained on the maternal plant and corticated after fertilization by a layer of sterile cells, and receives nourishment via placental transfer cells with wall ingrowths. The walls of the zygote possess sporopollenin, a highly resistant substance found in the outer wall of pollen (Delwiche et al., 1989). Cytokinesis and phragmoplast formation in *Coleochaete* are typical of land plants (Marchant and Pickett-Heaps, 1973; Graham et al., 2000; Cook, 2004a). However, there is also reason to be cautious. First, an unexpected result in the phylogeny of Finet et al. (2010) was the grouping of *Chaetosphaeridium* with the Zygnematophyceae. *Chaetosphaeridium* has typically been placed in the Coleochaetophyceae, a position that has been supported by molecular phylogenetic studies (Karol et al., 2001; Turmel et al., 2002a), and morpho-cytological characteristics, such as the presence of typical sheathed hairs and similar chloroplast structure (Delwiche et al., 2002; Hall and Delwiche, 2007). It is certainly possible that the two genera, both of which live on aquatic plants or stones in freshwater environments, do not form a monophyletic group, but they share enough distinctive properties to cast doubt on a phylogeny that separates them. Second, only limited taxon sampling is currently available for multi-gene analyses, and the deep divergence times involved make dense taxon sampling extremely important to avert systematic error in phylogenetic analyses.

A second analysis, based on 129 nuclear genes and slightly fewer taxa, suggested the Zygnematophyceae or a clade consisting of the Zygnematophyceae and Coleochaetophyceae as the sister lineage of the land plants (Wodniok et al., 2011). The sister relation between Zygnematophyceae and land plants is supported by some shared features, including some components of the auxin signalling machinery (De Smet et al., 2011) and chloroplast movement (Wada et al., 2003). It is more difficult to find morphological synapomorphies for the relationship between Zygnematophyceae and embryophytes, possibly as a result of secondary simplification and specialization in the Zygnematophyceae (Wodniok et al., 2011).

The Zygnematophyceae, also known as conjugating green algae, is the most species-rich and morphologically diverse lineage of charophytes. Vegetative bodies include non-motile unicells, filaments and small colonial forms. Sexual reproduction occurs by a unique process of conjugation, involving fusion of non-motile gametes. Absence of flagellate reproductive stages and basal bodies also sets this class apart from other charophytes. Traditionally, the class was divided into the Zygnematales and the Desmidiales based primarily on differences in cell wall structure, but molecular phylogenies have shown that the Zygnematales (characterized by smooth non-ornamented cell walls) is a paraphyletic assemblage that gave rise to the monophyletic Desmidiales (McCourt et al., 2000; Gontcharov et al., 2003; Gontcharov et al., 2004; Gontcharov, 2008). The latter order consists of unicells, filaments and colonies with conspicuously ornamented cell walls and cell constrictions in two half cells (semicells). The diversity and phylogenetic relationships within the lineage have been studied with nuclear rDNA and chloroplast rbcL sequence data (Denboh et al., 2001; Drummond et al., 2005; Gontcharov and Melkonian, 2005; Kim et al., 2006; Gontcharov and Melkonian, 2008; Hall et al., 2008a; Hall et al., 2008b; Gontcharov and Melkonian, 2010; Gontcharov and Melkonian, 2011).

The Charophyceae (also known as stoneworts) are large and morphologically complex macroscopic algae with a plant-like appearance, consisting of a central stalk of large, elongate and multinucleate cells and whorls of branches radiating from uninnucleate cells at the nodes. Growth is by a single apical meristematic cell. Sexual reproduction is oogamous with motile sperm produced in complex antheridia. Oogonia and antheridia are surrounded by sterile cells, and zygotes develop a thick covering of sporopollenin. Because the sporopollenin surrounding the zygotes is highly resistant to degradation and because the shallow freshwater habitats in which they live often create favourable depositional conditions, the stoneworts are well represented in the fossil record (fossilized zygote walls are called gyrogonites). Unambiguous fossils of stoneworts date back to at least 380 mya (Feist and Feist, 1997), with abundant and diverse gyrogonites in more recent sediments. Phylogenetic relationships within the class have been studied by McCourt et al. (1996, 1999) Sakayama et al. (2004, 2005, 2008, 2009).

As noted above, Karol et al. (2001) presented a four-gene phylogeny with relatively dense taxon sampling that placed the Charophyceae as the sister taxon to embryophytes. Thus, plausible recent analyses have shown almost every possible arrangement among the four lineages of “higher” streptophytes or phragmoplastophyta (embryophytes, Coleochaetophyceae, Zygnematophyceae, and Charophyceae). Given the very deep divergence times involved, it is not surprising that representation by six or eight species is not sufficient to provide a well-resolved phylogeny even when multiple genes are available for analysis. What is far less clear is whether or not enough diversity has survived extinction to permit taxon sampling alone to resolve this phylogenetic puzzle, or whether new analytical approaches will be needed to distinguish among the alternative topologies.

We will return to the subject of land plant origins in the section “Molecular evolution in the Streptophyta and the origin of land plants,” in which we discuss molecular evolution in the streptophyte lineage, emphasizing the genetic facilitation of land plant origins.

III. SPREAD OF GREEN GENES IN OTHER EUKARYOTES

A growing body of data is providing evidence for widespread movement of genetic information between distantly related eukaryotic genomes (Keeling and Palmer, 2008; Schaack et al.,...
Horizontal gene transfer (HGT, also known as lateral gene transfer) is a mechanism to rapidly spread evolutionary innovations across distinct lineages and is now regarded as an important force in the evolution of eukaryotes and their genomes (Nedelcu et al., 2008; Sun et al., 2010). Green algal genes have been incorporated into the genomes of various unrelated eukaryotes via different mechanisms (Figure 2).

Two important modes of eukaryote-to-eukaryote HGT are the gene ratchet mechanism and endosymbiosis (Dagan and Martin, 2009a; Sun et al., 2010). The gene ratchet mechanism, also known as “you are what you eat,” posits that by engulfing and digesting other cells, phagotrophic protists may occasionally incorporate genetic material of their prey into their genomes (Doolittle, 1998). In a few instances during eukaryote evolution, captured cells have been retained as intracellular symbionts (endosymbionts) and have become stably integrated inside the host cell. The two most prominent endosymbioses in eukaryote evolution have involved an alpha-proteobacterial and a cyanobacterial endosymbiont, and have given rise to mitochondria and plastids, respectively (Keeling and Palmer, 2008). Both endosymbionts have undergone massive gene loss and gene transfer from the endosymbiont genome to the host genome (endosymbiotic gene transfer, EGT), reducing them to metabolic slaves (organelles) inside their host cells (Gould et al., 2008; Tirichine and Bowler, 2011). The plastids of the green lineage, the red algae and glaucophytes are derived from a single primary endosymbiotic event, i.e. the original uptake and retention of a cyanobacterium by a nonphotosynthetic eukaryotic host cell (Rodríguez-Ezpeleta et al., 2005). The great majority of other photosynthetic eukaryotes acquired their plastids by secondary or tertiary endosymbioses, in which eukaryotes with green or red plastids were preyed upon by a heterotrophic eukaryote, followed by EGT, this time from the endosymbiont nucleus to that of the secondary or tertiary hosts (Delwiche, 1999; Keeling, 2004; Rodríguez-Ezpeleta et al., 2005; Keeling, 2010). The number of secondary (and tertiary) endosymbiotic events is heavily debated (Archibald, 2009).

Three eukaryotic lineages have fully integrated secondary plastids of green algal origin: photosynthetic euglenids, chlorarachniophytes and the “green” dinoflagellates (Delwiche, 1999; Keeling, 2004; Archibald, 2009; Keeling, 2010). While it has been proposed that the chloroplasts of chlorarachniophytes and euglenids share a common origin (the Cabozoa hypothesis, Cavalier-Smith, 1999), chloroplast multi-gene phylogenies suggest that these plastids originated independently (Rogers et al., 2007; Takahashi et al., 2007; Turmel et al., 2009a). The Pyramimonadales have been identified as the source of the chloroplasts of the euglenid *Euglena gracilis* (Turmel et al., 2009a; Matsumoto et al., 2011) (Figure 2). The chloroplast of the chlorarachniophyte *Bigelowiella natans* was found to be sister to the ulvophytes *Pseudendoclonium* and *Oltmannsiellopsis* (Turmel et al., 2009a), although this relationship was not supported in an analysis based on more taxa but fewer genes (Matsumoto et al., 2011).

Euglenids are a diverse group of mainly freshwater flagellate unicells, including photosynthetic members with plastids bounded by three membranes, as well as nonphotosynthetic members (Keeling, 2004; Triemer and Farmer, 2007). Together with the kinetoplastids and diplomemids they make up the Euglenozoa, which belong to the Excavata, a largely nonphotosynthetic eukaryotic supergroup (Baldauf, 2008). Chloroplast genomes have been characterized for two euglenids: the photosynthetic *E. gracilis* and the closely related, nonphotosynthetic *E. longa* (Hallick et al., 1993; Gockel and Hachtel, 2000). The chloroplast genome of *E. gracilis* is 143 kb in size and shows high similarities in gene order with *Pyramimonas* (Turmel et al., 2009a). The plastid genome of *Euglena longa* is highly reduced (73 kb) and has lost most photosynthesis-related genes but retained most of the other genes present in *E. gracilis*.

Chlorarachniophytes are a small group of marine amoeboflagellates bearing plastids bounded by four membranes (Keeling, 2004; Ishida et al., 2007). The discrepancy of chloroplasts with three membranes in euglenids versus four membranes in chlorarachniophytes has been suggested to have resulted from a myzocytosic aquisition of the plastids in euglenids versus a phagocytotic capture in chlorarachniophytes (Keeling, 2010). Chlorarachniophytes belong to the Rhizaria, which, like the Excavata, are primarily nonphotosynthetic (Baldauf, 2008). Chlorarachniophytes are one of only two groups of photosynthetic eukaryotes (the other one being the red plastid-containing cryptophytes) in which remnant nuclei of the eukaryotic endosymbionts, called nucleomorphs, have been retained (Cavalier-Smith, 2002). Chlorarachniophyte nucleomorphs contain highly reduced genomes (330–610 kb) and most essential genes have been transferred to the host’s nuclear genome. The nucleomorphs contain three small linear chromosomes and a gene density similar to that seen in prokaryotes (Moore and Archibald, 2009). Remarkably, the nucleomorph genomes of chlorarachniophytes and cryptophytes have evolved similar genomic features, including a three-chromosome architecture with subtelomeric rRNA operons. The 69 kb chloroplast genome of *Bigelowiella natans* is smaller than plastids of most other green algae and is closer in size to genomes of several nonphotosynthetic plastids (Rogers et al., 2007). Unlike nonphotosynthetic plastids, however, the *B. natans* chloroplast genome is highly compacted and encodes most of the genes found in other photosynthetic green algal plastids.

Dinoflagellates are a diverse group of photosynthetic or heterotrophic flagellates. While most photosynthetic dinoflagellates have secondary or tertiary plastids of red algal origin, members of the “green” dinoflagellate genus *Lepidodinium* have replaced their original red plastids with a new secondary endosymbiont from the green lineage (termed serial secondary endosymbiosis) (Watanabe et al., 1990; Keeling, 2010). Although prasinophytes had been hypothesized as the chloroplast source of *Lepidodinium* based on pigment composition (Watanabe et al., 1990), recent chloroplast multi-gene phylogenetic analyses suggest that these plastids were derived from an early
representative of the core chlorophytes, but the exact donor lineage remains equivocal (Takisita et al., 2008; Matsumoto et al., 2011).

In addition to these stable endosymbioses, there is also evidence for early stages of plastid acquisition via secondary endosymbioses (Gould et al., 2008). The katablepharid flagellate Hatena arenicola harbours a prasinophyte green algal endosymbiont of the genus Nephroselmis (Okamoto and Inouye, 2005; Okamoto and Inouye, 2006). The single membrane-bound endosymbiont exhibits extensive structural changes when within the Hatena cell and is ultrastructurally tightly associated with its host. However, the integration of the symbionts is not entirely stable as the division of the symbiont is not coordinated with that of the host: cell division results in only one symbiont-bearing daughter cell while the other daughter cell re-establishes a phagotrophic lifestyle by capturing a new Nephroselmis symbiont from the environment. As yet, it is unknown if this endosymbiotic association has resulted in EGT.

A remarkable case of endosymbiosis is the plastid theft performed by sacoglossan molluscs (Elysia and relatives) (Gould et al., 2008). These sea slugs feed upon siphonous and siphonous cladous green algae (Ulvophyceae) and siphonous xanthophytes (e.g., Vaucheria). Some species are able to sequester chloroplasts from these algae into their gut cells (Händeler et al., 2009; Händeler et al., 2010). Despite the absence of algal nuclei, these so-called kleptoplasts remain transcriptionally and translationally active and allow the slugs to rely on photosynthesize for weeks to months (Mujer et al., 1996). The kleptoplasts are not permanently acquired; they are unable to divide inside the host and are not passed on from one slug generation to another. The molecular basis of plastid retention is not well understood but has been suggested to involve gene transfer from the algal food source to the slug’s genome. Indeed, recent studies have provided evidence of gene transfer in a species retaining xanthophyte plastids (Rumpho et al., 2008; Schwartz et al., 2010). Conversely, analysis of transcriptome data from two other species maintaining green algal plastids did not reveal any transfer of algal genes specific to photosynthetic function, suggesting that these slugs do not express genes acquired from algal nuclei to maintain plastid function (Wägele et al., 2011).

There are several other examples of green algae that persist as endosymbions in various eukaryotic hosts. The ciliate Paramecium bursaria engages in close physiological association with trebouxiophyte green algal endosymbionts (Chlorella and relatives) that reside in vacuoles close to the cell surface and are vertically transmitted to daughter cells upon host cell division (Hoshina and Imamura, 2008; Nowack and Melkonian, 2010). This symbiosis has been found to be facultative in lab conditions since both the Paramecium and the algae can be cultivated separately. Other eukaryotes that harbour green algal symbionts (including parasites) include fungi, dinoflagellates, foraminifers, radiolarians, sponges, marine flatworms, cnidarians, molluscs (nudibranchs and giant clams) and vertebrates (Parke and Manton, 1967; Sweeney, 1976; Cachon and Caram, 1979; Williamson, 1979; Friedl and Bhattacharya, 2002; Lewis and Muller-Parker, 2004; Rodriguez et al., 2008; Kovacevic et al., 2010; Nowack and Melkonian, 2010; Kerney et al., 2011). As yet, it remains unclear whether these associations have resulted in endosymbiont-to-host gene transfer.

Remarkably, green algal-derived genes have also been detected in the genomes of a diverse array of other eukaryotes without green algal plastids, including algae with red-type plastids and eukaryote lineages without plastids. Phylogenomic analyses of two diatom genomes revealed that a considerable proportion of nuclear genes are of green algal origin, which is surprising given that diatom plastids are of red algal origin (Moustafa et al., 2009). These findings have been interpreted as evidence for the presence of a green algal endosymbiont early in the evolution of the chromalveolate lineage, which was later replaced by an endosymbiont with a red-type plastid (Moustafa et al., 2009). An ancient green algal endosymbiont may also explain the presence of numerous “green” genes in oomyceses (nonphotosynthetic chromalveolates) and apicomplexan parasites (chromalveolates containing nonphotosynthetic plastids of red-algal origin) (Huang et al., 2004; Tyler et al., 2006; Janouskovec et al., 2010). Similarly, the presence of genes related to green algal sequences in trypanosomatid parasites (kinetoplastids) has been explained by EGT of an ancient green algal endosymbiont (Hannaeit et al., 2003), but these data are open to interpretation and scenarios of ancient cryptic secondary endosymbioses in the chromalveolates and other eukaryotes have been questioned (Dagan and Martin, 2009b; Elias and Archibald, 2009; Stiller et al., 2009; Sun et al., 2010). In Monosiga, a member of the Choanozoa, which forms the sister group of the Metazoa, numerous genes of putative algal origin are present, including several genes with green algal affinities (Nedelcu et al., 2008; Sun et al., 2010). Based on the taxonomic distribution pattern of the algal genes and the phagotrophic nature of choanoflagellates, the presence of algal genes in Monosiga has been explained by the gene ratchet mechanism rather than ancient cryptic endosymbiosis (Sun et al., 2010).

IV. GREEN ALGAL EVOLUTION: INSIGHTS FROM GENES AND GENOMES

The large diversity of green plants offers a unique opportunity to study the molecular mechanisms underlying its evolution. The rapid accumulation of genomic data, along with current phylogenetic hypotheses, is greatly advancing our understanding of molecular evolution in the green lineage. Apart from vertical descent with modification, various other evolutionary events may affect gene histories, including gene duplication, gene loss, horizontal gene transfer and gene rearrangements (Koonin, 2005). Gene duplications and horizontal gene transfer have been suggested to contribute to biological novelty in the green lineage. Large-scale gene and whole genome duplication events have been well-characterized in embryophyte lineages (Flagel and Wendel, 2009) but have also played an important
role in the evolution of phenotypic novelty within green algal lineages (Petersen et al., 2006; Robbens et al., 2007b; Becker and Hoef-Emden, 2009). Several studies have indicated horizontal transfers of genetic information from green algae to other eukaryotes or vice versa (Bhattacharya et al., 1996; Friedl et al., 2000; Cocquyt et al., 2009; Ghoshroy et al., 2010).

Below we focus on a number of specific topics dealing with molecular evolution in the green lineage. A first section summarizes our current understanding of chloroplast and mitochondrondal genome evolution in green algae. A second section reviews genomic insights into the ecology of oceanic picoplanktonic prasinophytes. Thirdly, molecular mechanisms underlying the evolution of multicellularity and cellular differentiation in volvocine green algae are discussed. A fourth section discusses the evolution of genetic codes and the translational apparatus in green seaweeds. Finally, we discuss recent progress in our understanding of molecular evolution in the streptophyte lineage, focusing on the genetic bases underlying the origin of land plants.

A. Organelle Genome Evolution

The exploration of organelle genomes—the DNA molecules found within the mitochondria and plastids of eukaryotic cells—has revealed an astonishing array of genomic architectures (Palmer, 1985; Lang et al., 1999). Some of the most diverse and unusual mitochondrial and plastid DNAs (mtDNAs and ptDNAs) from all eukaryotes come from green algae (Gray, 1999). Indeed, among green algal groups there is an impressive range of organelle genome sizes (13 to > 525 kb), conformations (circular or linear), chromosome numbers (monomorphic, bipartite, or highly fragmented), compactnesses (20–95% coding DNA), gene repertoires (10 to > 130 genes), and nucleotide compositions (20 to > 60% guanine and cytosine)—see Table 1 for examples and references. Investigations of green algal organelle DNAs have helped unravel the origins and interrelationships of green plants and eukaryotes as a whole (Lemieux et al., 2000; Turnel et al., 2006; Archibald, 2009), provided insights into the forces driving genome evolution (Popescu and Lee, 2007; Smith and Lee, 2010), and contributed to changing the misconception that organelle genomes are relatively homogeneous in structure and content (Palmer, 1985; Gray et al., 2004).

In a broad sense, the organelle genomic architectures of green algae differ from those of land plants (Table 2). Green algae tend to have relatively compact, intron-poor mitochondrial genomes (with some major exceptions), whereas land plant mtDNAs are capacious and intron dense. And green algal plastid genomes are generally larger and more bloated than their mitochondrial counterparts, but for land plants the opposite is true. But as we will show below, it is hard to talk about major trends when dealing with green algal organelle DNAs, as the mitochondrial and plastid genome diversity within green algal groups is often greater than that between groups.

Although organelle genomes tend to be rich in adenine and thymine (AT-rich), data from certain green algae have shown that this is not always the case. The mtDNA of the non-photosynthetic, unicellular chlamydomonadalean Polyoomella capuana is 57% guanine and cytosine (GC), making it one of only a few GC-biased organelle genomes sequenced to date (Smith and Lee, 2008). Other green algae also harbour GC-rich organelle DNA: Analyses of the mtDNA-encoded cox1 and/or cob from Oogamochoygya gigantea, Lobochlamys segnis, and Lobochlamys culleus (all from the Oogamochoygylynea-clade, sensu Nakada et al. (2008a)) revealed average GC compositions of 50, 54, and 63%, respectively, suggesting that the overall mitochondrial GC content of at least some of these species is higher than that of P. capuana (Borza et al., 2009); and the unpublished mitochondrial and plastid genome sequences of the trebouxiophyte Coccovymya sp. C-169 (GenBank accession numbers HQ874522 and HQ693844) are 53% and 51% GC, respectively—only the second example (the other being the lycophyte Selaginella moellendorffii [Smith, 2009]) of a species with GC-biased DNA in both the mitochondrial and plastid compartments. In land plants, the levels of G and C in the mitochondrial and plastid genomes are positively correlated with the number of C-to-U RNA editing sites (Malek et al., 1996; Jobson and Qiu, 2008; Smith, 2009), but this is not the case for green algae as there is no evidence of RNA editing in either their mitochondrial or plastid genomes (Lenz et al., 2010, and references therein). Most green algal organelle DNAs have a more “typical” nucleotide composition, showing a propensity for A and T, sometimes strongly so, as in the cases of the Chara vulgaris, Scenedesmus obliquus, Helicosporidium sp., and Leptosira terrrestris plastid genomes, which are among the most AT-biased ptDNAs sequenced to date (de Cambiaire et al., 2006; de Koning and Keeling, 2006; Turnel et al., 2006). Whether the nucleotide bias is towards A and T or G and C, it is usually expressed most strongly within green algal organelle DNAs at what are considered to be among the more neutrally evolving sites in a genome (i.e., intergenic and intronic positions and the synonymous sites of protein-coding DNA), implying that the forces driving green algal organelle nucleotide landscape are nonadaptive (Smith and Lee, 2008; Borza et al., 2009).

It was once assumed that all mitochondrial genomes are circular (or at least circular-mapping) molecules. However, it is now well established that linear and linear fragmented mtDNAs have evolved numerous times in diverse eukaryotic lineages (Nosek and Tomáška, 2003), including the Reinhardtinia clade (sensu Nakada et al., 2008a) of the Chlamydomonadales. Every Reinhardtinia-clade mtDNA examined thus far appears to have either a linear monomeric or a linear fragmented conformation (Laflamme and Lee, 2003; Mallet and Lee, 2006), with the exception of the Volvox carteri mtDNA, which assembles as a circular molecule (Smith and Lee, 2010). Conversely, all of the characterized green algal mitochondrial genomes from outside this clade map in genome assemblies and/or gel electrophoresis studies as unit-sized circular chromosomes save for...
### TABLE 1
Examples of the organelle genome architectural diversity among green algae

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Lineage</th>
<th>Genome size (kb)</th>
<th>% GC</th>
<th>% Coding&lt;sup&gt;b&lt;/sup&gt;</th>
<th># of Genes&lt;sup&gt;a&lt;/sup&gt;</th>
<th># of Introns</th>
<th>Notable features</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MITOCHONDRIAL DNAs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ostreococcus tauri</td>
<td>P</td>
<td>44.2</td>
<td>38.2</td>
<td>90</td>
<td>65</td>
<td>0</td>
<td>Compact. Gene dense. Intron poor. Contains inverted repeats.</td>
<td>(Robbens et al., 2007a)</td>
</tr>
<tr>
<td>Pycnococcus provasolli</td>
<td>P</td>
<td>24.3</td>
<td>37.8</td>
<td>87</td>
<td>36</td>
<td>0</td>
<td>Small genome. Reduced gene content. Deviation from standard genetic code.</td>
<td>(Turmel et al., 2010)</td>
</tr>
<tr>
<td>Scenedesmus obliquus</td>
<td>C</td>
<td>42.9</td>
<td>36.3</td>
<td>52</td>
<td>42</td>
<td>4</td>
<td>Deviation from standard genetic code. Fragmented rrnS and rrnL genes.</td>
<td>(Nedelcu et al., 2000)</td>
</tr>
<tr>
<td>Prototheca wickerhamii</td>
<td>T</td>
<td>55.3</td>
<td>25.8</td>
<td>66</td>
<td>61</td>
<td>5</td>
<td>Non-photosynthetic parasitic green alga. Moderate size and coding content.</td>
<td>(Wolff et al., 1994)</td>
</tr>
<tr>
<td>Pedinomonas minor</td>
<td>T</td>
<td>25.1</td>
<td>22.2</td>
<td>58</td>
<td>22</td>
<td>1</td>
<td>Small genome. Reduced gene content. Fragmented rrnS gene.</td>
<td>(Turmel et al., 1999a)</td>
</tr>
<tr>
<td>Oltmannsiellopsis viridis</td>
<td>U</td>
<td>56.7</td>
<td>33.4</td>
<td>49</td>
<td>54</td>
<td>3</td>
<td>Repeat rich. Harbours putatively horizontally acquired genes.</td>
<td>(Pombert et al., 2006)</td>
</tr>
<tr>
<td>Pseudendoclonium akinetum</td>
<td>U</td>
<td>95.8</td>
<td>39.3</td>
<td>47</td>
<td>57</td>
<td>7</td>
<td>Expanded and repeat rich. Similar repeats in plastid compartment.</td>
<td>(Pombert et al., 2004)</td>
</tr>
<tr>
<td>Chara vulgaris</td>
<td>R</td>
<td>67.7</td>
<td>40.9</td>
<td>52</td>
<td>68</td>
<td>20</td>
<td>Intron dense.</td>
<td>(Turmel et al., 2003)</td>
</tr>
<tr>
<td>Chlorokybus atmophyticus</td>
<td>R</td>
<td>201.7</td>
<td>39.8</td>
<td>21</td>
<td>70</td>
<td>27</td>
<td>Among the largest and most intron-dense mtDNAs from green algae.</td>
<td>(Turmel et al., 2007a)</td>
</tr>
<tr>
<td><strong>PLASTID DNAs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nephroselmis olivacea</td>
<td>P</td>
<td>200.7</td>
<td>42.2</td>
<td>68</td>
<td>128</td>
<td>0</td>
<td>Large genome. Intron and repeat poor.</td>
<td>(Turmel et al., 1999b)</td>
</tr>
<tr>
<td>Pyramimonas parkeae</td>
<td>P</td>
<td>101.6</td>
<td>34.7</td>
<td>80</td>
<td>110</td>
<td>1</td>
<td>Small genome. Contains a DNA primase gene putatively acquired from a virus.</td>
<td>(Turmel et al., 2009a)</td>
</tr>
<tr>
<td>Floydiella terrestris</td>
<td>C</td>
<td>521.1</td>
<td>34.5</td>
<td>60</td>
<td>97</td>
<td>26</td>
<td>Expanded and intron dense. Lacks an inverted repeat.</td>
<td>(Brouard et al., 2010)</td>
</tr>
<tr>
<td>Volvox carteri</td>
<td>C</td>
<td>&gt;525</td>
<td>~43</td>
<td>&lt;20</td>
<td>96</td>
<td>8</td>
<td>Largest green algal ptDNA sequenced to date. Repeat rich.</td>
<td>(Smith and Lee, 2010)</td>
</tr>
</tbody>
</table>

(Continued on next page)
### TABLE 1
Examples of the organelle genome architectural diversity among green algae (Continued)

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Lineage</th>
<th>Genome size (kb)</th>
<th>% GC</th>
<th>% Coding</th>
<th># of Genes&lt;sup&gt;a&lt;/sup&gt;</th>
<th># of Introns</th>
<th>Notable features</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pedinomonas minor</em></td>
<td>T</td>
<td>98.3</td>
<td>34.8</td>
<td>71</td>
<td>105</td>
<td>0</td>
<td>Small, intron-lacking genome.</td>
<td>(Turmel et al., 2009b)</td>
</tr>
<tr>
<td><em>Bryopsis hypnoides</em></td>
<td>U</td>
<td>153.4</td>
<td>33.1</td>
<td>37</td>
<td>111</td>
<td>11</td>
<td>Evidence for multimeric ptDNA molecules.</td>
<td>(Lü et al., 2011)</td>
</tr>
<tr>
<td><em>Pseudendoclonium akiinetum</em></td>
<td>U</td>
<td>195.8</td>
<td>31.5</td>
<td>54</td>
<td>105</td>
<td>27</td>
<td>Large genome. Intron and repeat rich. Similar repeats in mitochondrial compartment.</td>
<td>(Pombert et al., 2005)</td>
</tr>
<tr>
<td><em>Mesostigma viride</em></td>
<td>R</td>
<td>118.3</td>
<td>30.2</td>
<td>74</td>
<td>137</td>
<td>0</td>
<td>Intron poor. Gene dense.</td>
<td>(Lemieux et al., 2000)</td>
</tr>
<tr>
<td><em>Zygnema circumcarinatum</em></td>
<td>R</td>
<td>165.3</td>
<td>31.1</td>
<td>55</td>
<td>125</td>
<td>13</td>
<td>Lacks inverted repeat. Large intergenic spacers.</td>
<td>(Turmel et al., 2005)</td>
</tr>
</tbody>
</table>

Note: C = Chlorophyceae; P = prasinophytes; R = charophyte green algae; T = Trebouxiophyceae; U = Ulvophyceae.

<sup>a</sup>Genes present in inverted repeats were counted only once.

<sup>b</sup>Does not include introns or unclassified ORFs.

### TABLE 2
Average architectural features of green algal organelle genomes

<table>
<thead>
<tr>
<th>Phylogenetic Group</th>
<th>Genotype Size (kb)</th>
<th>% Noncoding</th>
<th># of introns</th>
<th># of genes</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MITOCHONDRIAL DNAs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorophyta</td>
<td>35.6</td>
<td>35.3</td>
<td>3.7</td>
<td>33</td>
<td>19</td>
</tr>
<tr>
<td>prasinophytes</td>
<td>40.3</td>
<td>17.6</td>
<td>1</td>
<td>57</td>
<td>4</td>
</tr>
<tr>
<td>Chlorophyceae</td>
<td>23.3</td>
<td>39.6</td>
<td>4.6</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Trebouxiophyceae</td>
<td>43.2</td>
<td>37.7</td>
<td>3.3</td>
<td>47.7</td>
<td>3</td>
</tr>
<tr>
<td>Ulvophyceae</td>
<td>76.3</td>
<td>45.9</td>
<td>5</td>
<td>55.5</td>
<td>2</td>
</tr>
<tr>
<td>Charophytes</td>
<td>92.1</td>
<td>47.2</td>
<td>15.8</td>
<td>67.5</td>
<td>4</td>
</tr>
<tr>
<td>Land plants</td>
<td>431.2</td>
<td>84.3</td>
<td>24.4</td>
<td>58.0</td>
<td>22</td>
</tr>
<tr>
<td><strong>PLASTID DNAs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorophyta</td>
<td>179.3</td>
<td>41.3</td>
<td>8.4</td>
<td>98.3</td>
<td>22</td>
</tr>
<tr>
<td>prasinophytes</td>
<td>106.9</td>
<td>26.2</td>
<td>1.5</td>
<td>100.2</td>
<td>6</td>
</tr>
<tr>
<td>Chlorophyceae</td>
<td>300.1</td>
<td>60.3</td>
<td>17.7</td>
<td>96.9</td>
<td>7</td>
</tr>
<tr>
<td>Trebouxiophyceae</td>
<td>116.9</td>
<td>31.6</td>
<td>1.8</td>
<td>96.0</td>
<td>6</td>
</tr>
<tr>
<td>Ulvophyceae</td>
<td>167.1</td>
<td>49.4</td>
<td>16.0</td>
<td>104.5</td>
<td>3</td>
</tr>
<tr>
<td>Charophytes</td>
<td>151.5</td>
<td>41.2</td>
<td>9.7</td>
<td>128.8</td>
<td>6</td>
</tr>
<tr>
<td>Land plants&lt;sup&gt;a&lt;/sup&gt;</td>
<td>147.8</td>
<td>41.9</td>
<td>23.8</td>
<td>106.9</td>
<td>107</td>
</tr>
</tbody>
</table>

Note: N, sample size (i.e., number of genomes). For the ranges within groups please refer to GenBank Organelle Genome Resources [www.ncbi.nlm.nih.gov/genome].

<sup>a</sup>For land plant plastid genomes, many of the GenBank entries were incompletely or incorrectly annotated; thus, intron and gene contents were based on the 24 entries for which we were most confident.
the mtDNAs of some Lobochlamys taxa, which may be linear fragmented (Borza et al., 2009). Studies on linear green algal mtDNAs have revealed a range of interesting telomeric sequences and structures, including inverted repeats, 3' overhangs, and closed single-stranded loops (Vahrenholz et al., 1993; Smith and Lee, 2008). Since mitochondria lack telomerase, it is presumed that the elaborate termini of linear mtDNAs help the genome overcome the end replication problem, as defined by Olovnikov (1971) and Watson (1972). All of the available ptDNA sequences for green algae assemble as genome-sized circular molecules, but, like the mitochondrial and plastid DNAs from land plants, they probably exist in vivo as multi-genome-sized branched linear forms, which can (as a byproduct of recombination-dependent replication) recombine to generate unit-sized circular molecules (Simpson and Stern, 2002; Bendich, 2004, 2007). Recent data from the ptDNA of the ulvophyte Bryopsis hypnoides support this hypothesis (Lâ et al., 2011).

Green algal organelle DNAs vary in size and compactness (Tables 1 and 2). The smallest and most reduced plastid genome observed from a green alga belongs to the non-photosynthetic parasite Helicosporidium sp. (Trebouxiophyceae)—it is 37.4 kb and 95% coding DNA (de Koning and Keeling, 2006). In contrast, the plastid genome of the free-living photosynthetic trebouxiophyte Leptosira terrestis is 195 kb, half of which is noncoding DNA (de Cambiaire et al., 2007). Compact plastid genomes are also found in some prasinophytes, including Ostreococcus tauri, Micromonas spp., and Pyrococcus provosolii whose ptDNAs are around 75 kb and 80% coding (Robbens et al., 2007a; Turmel et al., 2009a). But this is not a trend of the entire genus: the Nephrosemis olivacea ptDNA is 201 kb, a third of which represents noncoding nucleotides (Turmel et al., 1999b). The plastid genomes of the chlorophyceans Floya diella terrestris and Volvox carteri are ~525 kb and about 80% noncoding DNA (Brouard et al., 2010; Smith and Lee, 2010), making them almost 300 kb larger than any other available ptDNA sequence. Large plastid genomes are a common theme among chlorophycean algae: the sequenced ptDNAs from C. reinhardtii, Stigeoclonium helveticum, and Dunaliella salina are 204, 223, and 269 kb, respectively (Maul et al., 2002; Bélangier et al., 2006; Smith et al., 2010b), and gel electrophoresis results place the plastid genomes of both Chlamydomonas gelatinosa and Chlamydomonas moewusii at ~290 kb (Boudreau et al., 1994; Boudreau and Turmel, 1996). But the largest green algal plastid genomes probably come from certain Acetabularia species (Ulvophyceae), which are believed to have ptDNAs in excess of 2 Mb (reviewed by Palmer, 1985).

The mtDNAs from green algae can also be large, although they are generally smaller than their plastid counterparts and much smaller than land plant mtDNAs, which can achieve sizes of 3 Mb (Ward et al., 1981). The charophyte Chlorokybus amphophyctis has the largest recorded green algal mitochondrial genome (201 kb and 77% noncoding)—more than twice the size of any other sequenced green algal mtDNA (Turmel et al., 2007a). It closest rivals in this respect are the mitochondrial genomes of Chara vulgaris (67.7 kb, 47.5% noncoding) and the ulvophyte Pseudendoclonium akinetum (96 kb, 50% noncoding) (Turmel et al., 2003; Pombert et al., 2004). The most diminutive mtDNAs from the Viridiplantae are currently found in the Chlorophyceae. For example, the mtDNAs of Polytonella species range from 13–16 kb (Smith et al., 2010a) and that of C. reinhardtii varies from 16–19 kb (depending on the presence of optional introns) (Gray and Boer, 1988; Michaelis et al., 1990; Vahrenholz et al., 1993). Though small, chlorophycean mtDNAs can be quite bloated; both the D. salina and V. carteri mtDNAs are ~60% noncoding (Smith et al., 2010b). Trebouxiophytes and charophytes can also have small mtDNAs: those of Pedinomonas minor and M. viride are 25 and 42 kb, respectively (Turmel et al., 1999a; Turmel et al., 2002c). Prasinophytes tend to have the most close-packed mitochondrial genomes within the Viridiplantae, as exemplified by O. tauri, Micromonas spp., and P. provosolii, whose mtDNAs are between 82–92% coding (Robbens et al., 2007a; Turmel et al., 2008); this compact architecture is paralleled in their nuclear and plastid compartments (Derelle et al., 2006; Worden et al., 2009).

As alluded to above, differences in the amount of intronic and intergenic DNA are responsible for much of the variability in mitochondrial and plastid genome size observed among green algae. The number of introns (group I and group II) in green algal organelle DNAs varies for mtDNA from 0 (Polytonella and most prasinophyte species) to 27 (C. vulgaris) and for ptDNA from 0 (M. viride, Micromonas spp., N. olivacea, and P. minor) to maximum of 26, 27, and 36 (F. terestris, P. akinetum, and D. salina, respectively). Intergenic sequences can account for as little as 5% (e.g., the ptDNA of Helicosporidium sp.) to more than 50% (e.g., the ptDNAs of V. carteri, L. terestris, and C. reinhardtii, and the mtDNA of C. amphophyctis) of green algal organelle genomes. In many cases the intergenic regions (and sometimes the intronic DNA) of green algal mitochondrial and plastid genomes are overrun with repetitive elements (Maul et al., 2002; Pombert et al., 2005; Smith and Lee, 2009). Green algal organelle DNA repeats come in a range of sizes, orientations (e.g., tandem, palindromic, inverted), nucleotide compositions, and complexities, and can often be folded into secondary structures such as hairpin loops. Some are believed to be transposable elements, and there are examples of similar types of repeats being found in both the mitochondrial and plastid compartments of the same species (Pombert et al., 2005; Smith and Lee, 2009). It has also been argued that in some green algal lineages organelle repeats have played a role in fashioning the mitochondrial and plastid genomes by acting as catalysts for genome fragmentation and reorganization as well as gene splitting and scrambling (Boer and Gray, 1991; Nedelcu, 1998; Nedelcu and Lee, 1998; Maul et al., 2002). Moreover, certain repeats, especially those that can be folded into secondary structures, are thought to be involved in gene regulation and processing (Boer and Gray, 1986; Jiao et al., 2004; Smith and Lee, 2008).
Most plastid genomes have a “quadruplicate structure” where a single set of large inverted repeats, which typically contain the rRNA-coding genes and various other genes, divide the genome into two single-copy regions. The quadruplicate structure is believed to have been present in the plastid genome of the ancestor that gave rise to green plants, and departure from it is relatively rare among plastid-harboursing eukaryotes. Thus, it is significant that there are examples from each of the major green algal groups of species that have lost (or almost lost) their ptDNA inverted repeats. Examples include the charophyte green alga Staurastrum punctatum and Zygnema circumcarinatum (Turmel et al., 2005), the chlorophyceans F. teretris and S. helvetica (Bélanger et al., 2006; Brouard et al., 2010), the prasinophytes P. provasolii and Monomastix (Turmel et al., 2009a), the ulvophyte B. hypnoides (Lü et al., 2011), and the trebouxiophytes L. teretris and Helicosporidium sp. (de Koning and Keeling, 2006; de Cambiaire et al., 2007). These observations suggest that loss of the inverted repeat has occurred multiple times throughout the evolution of green algae—and sometimes multiple times within the same group, as in the case of trebouxiophytes (de Cambiaire et al., 2007). Interestingly, the O. tauri mtDNA has an inverted repeat region within its mitochondrial genome, which, like that of plastid genomes, harbors the RNA-coding genes as well as other types of genes (Robbens et al., 2007a).

Some green algal organelle genomes abound with genes whereas others are gene deaparate. The available green algal ptDNAs contain anywhere from around 55–135 genes, typically representing 3 rRNAs (rrnS, rrnL, and rrnF), 25–35 tRNAs, and approximately 25–100 proteins, which are involved in processes such as photosynthesis, metabolism, transcription, and translation. The sequenced mtDNAs from green algae contain around 10–70 genes, which encode 2–3 rRNAs (rrnS, rrnL, and, sometimes, rrnF), approximately 1–28 tRNAs, and around 7–37 proteins, most of which function in transcription, translation, or oxidative phosphorylation. Charophyte green algae tend to have gene-abundant mitochondrial and plastid genomes, as demonstrated by C. atmophyticus and C. vulgaris, which have 70 and 68 mitochondrial-encoded and 138 and 127 plastid-encoded genes, respectively (Turmel et al., 2006; Lemieux et al., 2007). The organelle genomes from chlorophyceans, particularly the mtDNAs of chlamydomonadalean algae, are gene poor. For instance, the C. reinhardtii mitochondrial and plastid genomes contain 13 and 94 genes, respectively (Gray and Boer, 1988; Maul et al., 2002), and the mtDNAs from Polytomella species (which are the most reduced mtDNAs observed from green plants) encode only seven proteins, two rRNAs, and one tRNA (Fan and Lee, 2002; Smith et al., 2010a). Other examples of reduced organelle gene contents include the Helicosporidium sp. ptDNA, which contains no genes for photosynthetic proteins, and the P. minor mtDNA, which has 22 genes, none of which code for ribosomal proteins.

Studies on the organization and architecture of genes within green algal organelle genomes—including their arrangement and conservation among species, intron content and insertion sites, genomic localities (i.e., inside or outside of repeated regions), and sequence divergence—have helped infer the phylogenetic relationships among green plants and provided insights into the modes and temps of green algal genome evolution. Some interesting findings to have come from these investigations include observed changes from the standard genetic code. In the S. obliquus mitochondrial genome TCA (normally a serine codon) is a stop codon and TAG (normally a stop codon) codes for leucine (Nedelcu et al., 2000), and in the P. provasolii mtDNA TGA (normally a stop codon) codes for tryptophan and TTA and TTG (normally leucine codons) are stop codons (Turmel et al., 2010). Other discoveries include the presence of fragmented and scrambled small- and/or large-subunit rRNA-coding genes in the mtDNAs from chlorophycean species and P. minor (Gray and Schnare, 1996; Nedelcu, 1997; Turmel et al., 1999a). Fragmented protein-coding genes, which are brought together at the RNA level via intron trans-splicing, have been identified in both the mitochondrial and plastid genomes of green algae (Glanz and Kuck, 2009; Pombert and Keeling, 2010). Genes acquired through horizontal gene transfer have been reported. The int and dpoB genes, which are located in the ptDNA inverted repeats of O. cardiaicum, and putatively code for a tyrosine recombinase and a DNA-directed DNA polymerase, are believed to have been obtained laterally, possibly from the mitochondrial plasmid of a fungus (Brouard et al., 2008). Also, the Olmannsiellopsis viridis mtDNA has what appears to be a recently captured integrase gene originating from a bacterium and a group II intron coming from a cryptophyte (Pombert et al., 2006), the P. parkeae ptDNA harbors a DNA primase gene putatively acquired from a virus (Turmel et al., 2009a), and the mitochondrial genome of C. reinhardtii contains a reverse-transcriptase-like gene (rtl) whose evolutionary origins and function remain unknown (Boer and Gray, 1988).

Much effort has been spent in trying to understand the evolutionary forces responsible for the architectural variation among green algal organelle DNAs. One hypothesis that has been tested on multiple fronts is the mutational-hazard hypothesis, which suggests that the primary forces governing organelle genome size and structure are mutation and random genetic drift (Lynch and Conery, 2003; Lynch et al., 2006). The hypothesis argues that genomic embellishments, such as introns and intergenic DNA, are a mutational liability because they represent targets for potentially deleterious mutations, where the higher the mutation rate the greater the burden of the embellishment. It is further argued that species with large effective population sizes (i.e., large Ne) are more efficient at perceiving and eliminating burdensome DNA than those with small effective population sizes. Two types of silent-site DNA sequence divergence data—that within species (\( \pi_{silent} \)) and that between species (ds)—have provided clues into the roles that branch-site and conservation among species, intron content and insertion sites, genomic localities (i.e., inside or outside of repeated regions), and sequence divergence—have helped infer the phylogenetic relationships among green plants and provided insights into the modes and temps of green algal genome evolution. Some interesting findings to have come from these investigations include observed changes from the standard genetic code. In the S. obliquus mitochondrial genome TCA (normally a serine codon) is a stop codon and TAG (normally a stop codon) codes for leucine (Nedelcu et al., 2000), and in the P. provasolii mtDNA TGA (normally a stop codon) codes for tryptophan and TTA and TTG (normally leucine codons) are stop codons (Turmel et al., 2010). Other discoveries include the presence of fragmented and scrambled small- and/or large-subunit rRNA-coding genes in the mtDNAs from chlorophycean species and P. minor (Gray and Schnare, 1996; Nedelcu, 1997; Turmel et al., 1999a). Fragmented protein-coding genes, which are brought together at the RNA level via intron trans-splicing, have been identified in both the mitochondrial and plastid genomes of green algae (Glanz and Kuck, 2009; Pombert and Keeling, 2010). Genes acquired through horizontal gene transfer have been reported. The int and dpoB genes, which are located in the ptDNA inverted repeats of O. cardiaicum, and putatively code for a tyrosine recombinase and a DNA-directed DNA polymerase, are believed to have been obtained laterally, possibly from the mitochondrial plasmid of a fungus (Brouard et al., 2008). Also, the Olmannsiellopsis viridis mtDNA has what appears to be a recently captured integrase gene originating from a bacterium and a group II intron coming from a cryptophyte (Pombert et al., 2006), the P. parkeae ptDNA harbors a DNA primase gene putatively acquired from a virus (Turmel et al., 2009a), and the mitochondrial genome of C. reinhardtii contains a reverse-transcriptase-like gene (rtl) whose evolutionary origins and function remain unknown (Boer and Gray, 1988).
for \( N_e(\mu) \) from \( V. \) carteri are on average very low (around 0.0004, 0.0006, and 0.005 for the mtDNA, ptDNA, and nucDNA, respectively), especially relative to other protists (Smith and Lee, 2009; Smith and Lee, 2010). This suggests that \( V. \) carteri has a small \( N_e(\mu) \) and therefore a reduced ability to detect and eradicate excess DNA, which could help explain why its organelle and nuclear genomes are among the most inflated from green algae. Studies on the relative dS values (which can be used to estimate relative mutation rates) from \( Chlamydomonas \) and \( Mesostigma \) indicate that in each of these algae the mutation rates of the mitochondrial, plastid, and nuclear compartments are similar (Popescu and Lee, 2007; Hua et al., 2012). This is in sharp contrast to the situation for most land plants where the silent-site substitution rate of the mtDNA is estimated to be three-fold lower than that in the ptDNA and 12-times lower than that in the nuclear DNA (Wolfe et al., 1987; Wolfe et al., 1989). This may help explain why, broadly speaking, there is less disparity in genomic complexity between the mitochondrial and plastid genomes of green algae than there is for land plant mitochondrial and plastid DNAs. The average fraction of noncoding DNA in the available mtDNA and ptDNA sequences from chlorophytes is 0.35 and 0.41, respectively, and for chlorophyte green algae the corresponding values are 0.47 (mtDNA) and 0.41 (ptDNA). Conversely, the proportion of noncoding nucleotides in the mtDNA of land plants averages 0.84 but is 0.42 for the ptDNA (Table 2). Thus, the comparable architectures of green algal plastid and mitochondrial DNAs may be a consequence of these genomes having similar mutation rates, as reflected in their similar rates of silent-site substitution. As more green algal organelle DNA data become available—especially data from different members of the same population and from closely related species—we will get a better idea of now \( N_e \) and \( \mu \) are shaping green algal genome architecture.

**B. Ecology and Molecular Evolution of Oceanic Picoplanktonic Prasinophytes**

The marine planktonic compartment of picoeukaryotes has been previously overlooked because of the extremely small size of its constituents (1–3\( \mu \)m). However, its photosynthetic representatives, although present in lower concentration than cyanobacteria, are metabolically very active and can represent up to 80% of the biomass primary production in coastal ecosystems (Li, 1994; Worden et al., 2004). Picoeukaryote composition is highly diverse, encompassing several main branches of the eukaryotic crown (Lopez-Garcia et al., 2001; Moon-van der Staay et al., 2001; Vaulot et al., 2008), and, among this diversity, the prasinophytes are constantly present and emerged as a major component, especially in coastal areas (Diez et al., 2001; Massana et al., 2004; Not et al., 2004; Piganeau and Moreau, 2007; Viprey et al., 2008). The prasinophytes form a paraphyletic assemblage of at least nine clades, with the Mamiellephyceae (prasinophyte clade II) being common in mesotrophic coastal areas (Zhu et al., 2005; Countway and Caron, 2006; Marie et al., 2006) and clades VIII and IX (known exclusively from environmental sequencing) in oligotrophic oceanic waters (Lepère et al., 2009; Shi et al., 2009). Sequences belonging to the Pico-cystis clade (prasinophyte clade VII) correspond to strains originally isolated from open oceanic areas but are also encountered in clone libraries originating from coastal regions (Shi et al., 2009).

*Micromonas pusilla* (Butcher, 1952) was the first species to be discovered among the Mamiellephyceae. Since this date, at least seven new genera of this class have been discovered over the last decade by electron microscopic and phylogenetic analyses (Marin and Melkonian, 2010), currently giving a total of 19 species. Most of the clone library studies showed that among Mamiellephyceae, the genera *Bathyccoccus*, *Micromonas* and *Ostreococcus* are dominant in many different areas (Guillou et al., 2004; Not et al., 2004; Marie et al., 2006; Viprey et al., 2008). These three genera grow well in culture and are also well-represented in culture collections. Phylogenies based on the 18S rDNA showed that *Micromonas* and *Ostreococcus* are divided in several clades (Guillou et al., 2004) corresponding to different ecotypes and probably different species (see below) (Rodriguez et al., 2005; Slapeta et al., 2006). This has been shown on the basis of physiological properties (Rodriguez et al., 2005; Cardol et al., 2008), geographical origin (Foulon et al., 2008), and niche partitioning (Demir-Hilton et al., 2011) of the strains/ecotypes. For example, “low light” and “high light” ecotypes have been described in *Ostreococcus*, corresponding to different clades (Rodriguez et al., 2005). A recent study showed that co-occurrence of both ecotypes at the same geographical location is rare, although parameters explaining clade distribution are more complex than irradiance alone (Demir-Hilton et al., 2011). It is proposed that these two “low light” and “high light” ecotypes could rather correspond to oceanic and coastal clades/ecotypes, respectively. In contrast, all *Bathyccoccus* sequences are grouped in a homogeneous clade (Guillou et al., 2004), even with the addition of new strains isolated recently from different areas (Moreau, unpublished). The ecological importance of *Bathyccoccus* is probably overlooked (Johnson and Sieburth, 1982; Eikrem and Throndsen, 1990), but it is sporadically found in clone library studies where it can even be dominant (Marie et al., 2006).

The wide distribution of Mamielliphyceae and especially the three genera mentioned above, their ecological importance, ease of culturing and small genomes have made these eukaryotic microorganisms emerging models for ecological and biological studies of marine phytoplankton. Two *Micromonas* (Worden et al., 2009) and two *Ostreococcus* (Derelle et al., 2006; Palenik et al., 2007) genome sequences have been published, and the genome sequence of a low light strain of *Ostreococcus* is also available on the JGI web site (genome.jgi-psf.org). These genome sequences allowed a quantification of the genetic distances between the different clades within a genus. For example, in two sequenced *Ostreococcus* strains belonging to different clades, the 18S rDNA divergences was less than 0.5% whereas around 25% divergence in amino-acid identity over
their orthologous protein coding genes was observed. An even higher divergence was observed between the two sequenced Micromonas strains (Worden et al., 2009). This high genome sequence divergence and a different number of chromosomes, are not compatible with genetic exchange via meiosis, and have led to the “genomic species” definition of O. lucimarinus (Palenik et al., 2007). Sexual reproduction has never been observed between any Mamielllophyceae strains, although a complete genome set necessary to control meiosis is present in both Micromonas and Ostreococcus (Derelle et al., 2006; Worden et al., 2009). However, evidences of recombination events in several Ostreococcus tauri strains strongly suggest genetic exchanges among strains belonging to the same clade (Grimsley et al., 2010).

The five available Mamielllophyceae genomes are very gene-dense, with protein-coding sequences representing almost 70% of the genome for most strains. Their small cellular and genome sizes have led to the hypothesis that they may represent the “bare limits” of life as a free-living photosynthetic eukaryote, presumably having disposed of redundancies and presenting a simple organization. Indeed, for most of the annotated metabolic processes, a reduction in the number of gene copies was observed, usually down to a single copy. Complete genome sequences have provided insights into potential Mamielllophyceae metabolisms and adaptations, although most of them still await experimental confirmation. For example, biochemical studies have shown that the same light-harvesting chlorophyll-binding proteins (LHCs) were associated with both PSI and PSII. However, the availability of Micromonas and Ostreococcus genomes confirmed the existence of an unusual Mamielllophyceae-specific LHC, probably associated with both photosystems (Cardol et al., 2008; Six et al., 2009), but also revealed the presence of specific PSI and PSII subunits, which are expressed at low levels and which were not detected biochemically (Six et al., 2005). Because Mamielllophyceae are considered as early-diverging green algae, these findings had important evolutionary implications and supported the hypothesis of an ancient origin of LHCI genes from which LHCII and the major Mamielllophyceae LHC genes probably evolved, consolidating data into a coherent evolutionary scenario. Likewise, genes encoding all of the enzymes required for C4-photosynthesis are present in both the Micromonas and Ostreococcus genomes indicating that, like diatoms, these algae may use this pathway as a CO2 concentration mechanism (Derelle et al., 2006; Worden et al., 2009). This may, therefore, be a strategy general to planktonic microalgae, raising the possibility that, despite its high energetic cost, C4-photosynthesis might give a critical ecological advantage in CO2-limiting conditions, such as in phytoplankton blooms.

Other examples of potential metabolic adaptations can be found in the available genomes. For example, in unicellular eukaryotes a general system for iron (Fe) uptake usually involves a ferric reductase, multicopper oxidase, and a ferric permease. Ostreococcus, in contrast, appears to lack all of these iron transport components, with the possible exception of a multicopper oxidase found only in O. tauri (Palenik et al., 2007). This implies that this genus has a different system of Fe acquisition compared to those of major competitors. Furthermore, Ostreococcus shows several adaptations that reduce Fe requirements, such as the absence of Cytochrome c6 in Ostreococcus and its replacement by plastocyanin as the only electron carrier between the Cyt b6/f complex and photosystem I (Palenik et al., 2007). In silico analysis of nutrient acquisition has also been investigated by comparative analysis of genes involved in nitrogen metabolisms. Eight of the genes involved in nitrate uptake and assimilation, as well as four genes for urea assimilation are found next to each other in two clusters located on two chromosomes in Ostreococcus (Derelle et al., 2006; Palenik et al., 2007). A comparable clustering of nitrate assimilation genes is also observed in Micromonas, whereas in C. reinhardtii fewer of these genes are clustered. This organization indicates a possible selective pressure for optimization of nitrate and urea uptake and assimilation in microalgae. A recent study in Micromonas showed a high degree of “mixed lineage gene affiliations,” for nitrogen transport and assimilation genes (McDonald et al., 2010), suggesting ancient origin, complex evolution pattern of gene duplication and gene loss, or horizontal gene transfer for some of these genes.

These in silico analyses are still in their infancy and many metabolisms, potential adaptations and/or evolutive signatures still remain to be elucidated. Beside portals dedicated to the annotation of each of the different genomes, web sites focused on comparative analysis between the different organisms belonging to the green lineage, as PLAZA (bioinformatics.psb.ugent.be/plaza/) (Proost et al., 2009) and its derivative Pico-PLAZA (bioinformatics.psb.ugent.be/pico-plaza) are now available. Such tools for comparative approaches, coupled to the development of new sequencing possibilities for both new genomes and metagenomes (Cheung et al., 2010; Marie et al., 2010) from various environments open the door to a better understanding of the adaptations of the organisms to their environment.

C. Genomic Insights into the Evolution of Complexity in Volvocine Green Algae

The volvocine algae (Chlorophyceae: Chlamydomonadales) are an important model system for understanding the evolution of multicellularity and cellular differentiation. This informal grouping, imprecisely defined as “Volvox and its close relatives,” is usually taken to include the families Volvocaceae, Goniaceae and Tetrabenaecae (Nozaki et al., 2000; Nozaki et al., 2003) along with a few unicells in the genera Chlamydomonas and Vireochlamys. Their appeal as a model system stems from the relative simplicity of their development and from the wide range of sizes and degrees of complexity found in extant species. Cell numbers range from 1 to ~50,000, including nearly every power of 2 (4, 8, 16, etc.) in this range, and multicellular species may have partial, complete, or no differentiation between reproductive and somatic cells (Figure 5).
FIG. 5. Volvocine algae: evolution of complexity. A: Chlamydomonas reinhardtii, a single-celled relative of Volvox with two flagella at the anterior. C. reinhardtii diverged from the multicellular volvocine algae around 250 million years ago. B: Gonium pectorale, a flat or slightly curved plate of 8 to 32 cells (8 in this example), all oriented in the same direction (photos A and B by Deborah Shelton). C: Eudorina elegans, a spheroid with up to 32 undifferentiated cells (32 in this example) (photo by Antonio Guillén). D: Pleodorina starrii, a partially differentiated spheroid with up to 64 cells (32 in this example). The small cells near the anterior pole (top) are terminally differentiated somatic cells specialized for motility; the larger cells perform both reproductive and motility functions. E: Volvox carteri, a spheroid with ca. 2000 small somatic cells arranged at the periphery and a handful of much larger reproductive (germ) cells. F: Volvox barberi, a spheroid with ca. 30,000 small somatic cells arranged at the periphery and a handful of much larger reproductive (germ) cells. The germ cells in this colony have begun to develop into daughter colonies, and some are in the process of inversion. Volvox barberi diverged from the other Volvox species shown here around 200 million years ago and represents a remarkable example of evolutionary convergence (photos D-F by MH). (Color figure available online.)

Two species in particular, the unicellular Chlamydomonas reinhardtii and the differentiated multicellular Volvox carteri forma nagariensis (V. carteri hereafter), have been the focus of molecular-genetic investigations, and their genomes have recently been sequenced (Merchant et al., 2007; Prochnik et al., 2010). C. reinhardtii is a biflagellate, facultatively heterotrophic soil alga that serves as an important model organism for understanding eukaryotic photosynthesis, flagellar motility, and basal body function. V. carteri is a spheroid typically made up of ∼1000–2000 small, biflagellate somatic cells, ∼8–16 much larger reproductive (germ) cells, and a gelatinous extracellular matrix (ECM) that makes up ∼99% of the total volume. Several of the genes underlying multicellular development and cellular differentiation in V. carteri have been identified and cloned (Kirk, 1998; Kirk, 2005), greatly improving our understanding of the molecular and developmental bases of these traits.

In spite of the dramatic differences in size and complexity between C. reinhardtii and V. carteri, their genomes are similar in size and gene content. The V. carteri nuclear genome is ∼17% larger than that of C. reinhardtii (138 Mb vs. 118 Mb), but this difference is largely due to non-coding sequences, as V. carteri has greater repeat content and, on average, longer introns (Prochnik et al., 2010). The total number of genes is similar, with ∼14,500 genes in each species’ genome (Merchant et al., 2007; Prochnik et al., 2010). Most gene families have similar numbers of genes in each species, although several interesting exceptions occur (discussed below).

Since its divergence from unicellular ancestors ∼230 million years ago (Herrick et al., 2009), the lineage leading to V. carteri has undergone a series of developmental changes resulting in a multicellular organism with a distinct anterior-posterior polarity, a complex developmental program, and a clear distinction between reproductive and somatic cells. The evolutionary path of V. carteri from a single-celled ancestor was mapped by Kirk (2005), who divided the evolution of multicellularity and cellular differentiation into a series of twelve developmental changes. For several of these changes, information about the underlying genetics is available, and there are several genomic differences that suggest possible genes whose roles have yet to be identified.

In C. reinhardtii, and presumably in the unicellular ancestor of V. carteri, cell division occurs through palintomy rather than by binary fission; that is, cells grow to many times their original size without dividing and then undergo several rounds of rapid cell division with little or no growth between successive divisions (Sleigh, 1989). In C. reinhardtii, the resulting offspring emerge from the mother cell wall and begin their lives as separate unicellular organisms. Most multicellular volvocine algae have retained this mode of cell division, producing large reproductive cells, each of which divides n times to produce 2^n daughter cells (n ranges from 2 to 14). In the multicellular volvocine algae, the daughter cells resulting from palintomic division of a reproductive cell remain connected to form a single, multicellular offspring.

The first step in the evolutionary change from unicellularity in the volvocine algae was probably the origin of some means of keeping daughter cells connected to each other. The mechanisms by which this attachment occurs differ among volvocine species. In the Tetrabaenaceae, the smallest multicellular volvocine algae and the earliest to diverge from the lineage leading to V. carteri, palintomic division produces four daughter cells that remain embedded in a common ECM (Iyengar and Desikachary, 1981). In all other species, cells are held together by specialized cell wall attachments or by a boundary layer that surrounds the entire organism (Nozaki, 1990; Nozaki and Kuroiwa, 1992). Both the ECM and the boundary layer are homologs of portions of the C. reinhardtii cell wall (Kirk et al., 1986).

The ECM in V. carteri is a complex, multifunctional structure that makes up over 99% of the volume of mature spheroids (Hallmann, 2003). Two gene families involved in ECM synthesis and function are substantially larger in the V. carteri genome than in that of C. reinhardtii: the pherophorin family has 49 members in V. carteri and only 22 in C. reinhardtii, while the Volvox metalloproteinase family has 42 members in V. carteri and only 8 in C. reinhardtii (Prochnik et al., 2010). Expansion of these gene families may thus have been an important factor in the increase in size and complexity of ECM in
multicellular volvocine algae. The volume of ECM has undergone several expansions and contractions throughout volvocine evolution (Harr and Michod, 2008), and it would be interesting to know if the same genes were involved in lineages other than that leading to \textit{V. carteri}.

Several of the important differences between single-celled and multicellular volvocine algae can be attributed to differences in cell cycle regulation. Thanks to the palintomic cell division program, and specifically to the lack of growth between cell divisions, the reproductive cells of large multicellular species must postpone the start of cell division until they are large enough to produce all of the cells in the offspring. In addition, early in development the cells of multicellular species are connected by cytoplasmic bridges resulting from incomplete cytokinesis (Kirk, 2005). In most species these connections break down before maturity (Stein, 1965; Marchant, 1977; Fulton, 1978). Most of the genes involved in cell cycle regulation have one ortholog at each in \textit{V. carteri} and \textit{C. reinhardtii}, but one family of cell cycle-related proteins, the D cyclins, is larger in \textit{V. carteri} (Prochnik et al., 2010). Prochnick et al. (2010) suggest that the additional members of this gene family may play a role in the differences in timing and extent of cell division between single-celled and multicellular volvocine algae.

One of the important functions of the cytoplasmic bridges that result from incomplete cytokinesis is in the process of inversion, a developmental process that occurs in all members of the Volvocaceae (Green et al., 1981). In this family, at the end of cell division, embryos are cup- or bowl-shaped, but with their flagella on the inner (concave) surface (Gerisch, 1959; Marchant, 1977; Fulton, 1978; Kirk, 2005). During inversion, cells move relative to the cytoplasmic bridges in order to turn the embryo inside-out, so that the flagella end up on the outside surface (Gerisch, 1959; Marchant, 1977; Fulton, 1978; Green et al., 1981). This process seems to have evolved by co-option of genes already present in the unicellular ancestor of \textit{V. carteri}: three genes known to be involved in \textit{V. carteri} inversion (\textit{invA}, \textit{invB}, \textit{invC}) have highly conserved orthologs in \textit{C. reinhardtii} (Nishii et al., 2003; Ueki and Nishii, 2009; Nishii and Miller, 2010), and an \textit{invA} ortholog (\textit{IAR1}) in \textit{C. reinhardtii} is capable of rescuing \textit{V. carteri} \textit{invA} mutants (Nishii et al., 2003).

Changes in cell cycle regulation are also involved in cellular differentiation. The distinction between somatic and reproductive cells in \textit{V. carteri} is established early in development, when different numbers and patterns of divisions – including asymmetric divisions – produce cells of two different sizes (Starr, 1969). The ∼30-fold difference in cell size triggers two distinct patterns of gene expression that establish the fate of the large and small cells as gonidia (asexual reproductive cells) and somatic cells, respectively (Kirk and Kirk, 1985; Tam and Kirk, 1991; Kirk et al., 1993). The genes involved in asymmetric division and the differentiation of somatic cells have been subjected to intensive study, and the overall picture seems to be one of co-option of genes already present in the unicellular ancestor. For example, the \textit{glsA} (gonidialessA) gene, which is required for asymmetric division in \textit{V. carteri}, has a homolog in \textit{C. reinhardtii} that is capable of rescuing \textit{V. carteri} \textit{glsA} mutants (Cheng et al., 2003). The \textit{regA} (somatic regenerator) gene, a member of the \textit{VARL} (Volvox algal \textit{regA}-like) family, suppresses growth and reproduction in \textit{V. carteri} somatic cells (Kirk et al., 1999; Meissner et al., 1999; Duncan et al., 2007). Although \textit{C. reinhardtii} lacks an exact \textit{regA} ortholog, phylogenetic analysis of the \textit{VARL} gene family suggests that this is due to gene loss in the \textit{C. reinhardtii} lineage rather than a gene duplication in the \textit{V. carteri} lineage (Duncan et al., 2007). The most closely related \textit{C. reinhardtii \textit{regA}} homolog is induced under nutrient and light-deprivation, leading to the suggestion that the somatic differentiation function of \textit{regA} was co-opted from an environmental acclimation function in the unicellular ancestor of \textit{V. carteri} (Nedelcu and Michod, 2006; Nedelcu, 2009).

Another form of cellular differentiation is between the dissimilar gametes of anisogamous (or oogamous) mating types. Both \textit{C. reinhardtii} and \textit{V. carteri} are facultatively sexual and heterothallic (sex or mating type is genetically determined), but their sexual phase differs in several respects. In \textit{C. reinhardtii}, entry into the sexual phase of the life cycle is induced by nitrogen starvation, which causes haploid vegetative cells to differentiate into isogamous gametes (Sager and Granick, 1954). Two gametes of opposite mating type fuse to form a diploid zygote, which undergoes meiosis to produce four haploid vegetative cells. In \textit{V. carteri}, sex is induced by a sex-inducing pheromone that can be either produced by the somatic cells of haploid asexual spheroids in response to heat stress or released from the sperm packets of spontaneously occurring sexual males (Kirk and Kirk, 1986). Asexual spheroids respond to the sexual inducer by producing male or female offspring, which in turn produce either eggs or packets of 64 or 128 sperm (Kirk, 1998). Sperm and eggs fuse to form a dormant diploid zygote, which undergoes meiosis to form a single asexual spheroid and three polar bodies (Kirk, 1998).

In \textit{C. reinhardtii}, mating type is determined within the ∼200 kb \textit{MT} region, in which meiotic recombination is suppressed by differences in gene order (Ferris and Goodenough, 1994; Ferris et al., 2002). Within this region, the \textit{MID} (Minus Dominant) gene, found only in the minus mating type, acts as a dominant Mendelian trait determining mating type (Ferris and Goodenough, 1997). Discovery of a \textit{MID} ortholog in males of \textit{Pleodorina starrii}, an oogamous volvocine alga, showed that males evolved from the minus mating type; this is the first case in which homology between an oogamous sex and an isogamous mating type has been demonstrated (Nozaki et al., 2006a). As in the \textit{C. reinhardtii \textit{MT} region}, the \textit{V. carteri \textit{MT} region} includes genes that are shared between both sexes as well as several that are limited to either males or females (Ferris et al., 2010). However, the \textit{V. carteri \textit{MT} region} is around five times as large, includes substantially more genes, and has, on average, much greater sequence divergence between alleles of genes found in both sexes (Ferris et al., 2010). Most of the sex-limited genes are differentially expressed during sexual differentiation, and many
of the shared genes are differentially expressed both during sexual differentiation and between males and females (Ferris et al., 2010). Differential expression, along with the increased gene content of the *V. carteri* MT region relative to that of *C. reinhardtii*, suggests a number of candidate genes that could have been important for the evolution of oogamy.

Although *V. carteri*, with just two cell types, is often cited as an example of a simple multicellular organism, its structure and development have changed a great deal relative to its unicellular ancestors. Aside from a 2000-fold increase in cell number and a million-fold increase in total volume, the lineage leading to *V. carteri* has undergone several important changes in cell cycle regulation resulting in the establishment of cytoplasmic connections among cells and of a reproductive-somatic division of labor. Changes in flagellar arrangement and a coordinated developmental process of inversion produced a multicellular organism with a distinct anterior-posterior polarity and the ability to coordinate the movements of 4000 flagella in order to seek out optimal light levels (Hoops, 1993). Changes in the sexual phase of the life cycle have led to the establishment of males and females from isogamous ancestors, and entry into the sexual phase has become dependent on signals from other individuals rather than on direct sensing of environmental conditions.

Comparison of the *V. carteri* and *C. reinhardtii* genomes has already revealed much, including some surprises, about the evolution of complexity in the volvocine algae. In a few cases, differences in the sizes of gene families suggest that gene duplication has played a role in the evolution of multicellular development and cellular and sexual differentiation. In most cases, though, the traits that comprise a differentiated multicellular organism seem to be controlled by genes co-opted from a unicellular ancestor. Even in those cases in which genomic comparisons suggest candidate genes, their specific functions have yet to be identified, nor, for most, has their involvement in particular traits even been confirmed. Furthermore, with few exceptions, we know nothing about the genetic basis of most traits in any of the ~50 volvocine species other than *V. carteri* and *C. reinhardtii*. Publication of the *V. carteri* genome has opened up the comparative genomics phase of volvocine research, suggesting answers to some questions and posing new ones that could not previously have been framed. Thus volvocine genomics has already revealed much, including some surprises, about the genetic basis of most traits in some eukaryotes, and they have also been shown in some green algal groups. Early work on cDNA of the nuclear-encoded small subunit of the Rubisco gene showed that the stop codons UAA and UAG were reassigned to code for glutamine in the siphonous dasycladalean green seaweeds *Acetabularia* and *Batophora* (Schneider et al., 1989; Schneider and Degroot, 1991). More recently, the evolution of these codon reassignments among the Ulvophyceae was studied in more detail using multi-gene datasets (Gile et al., 2009; Cocquyt et al., 2010a). These studies have shown a complex distribution of the alternative genetic code, the canonical genetic code being used in early-branching Ulvophyceae and the Bryopsidales, and the alternative code being used in Dasycladales, Cladophorales and Trentepohliiales (Figure 6). Several explanations for the observation that the Bryopsidales, which have the canonical code, are nested in a clade of which all other representatives have the alternative code are compared by Cocquyt et al. (2010a). These authors suggest that a stepwise acquisition model involving ambiguous intermediates (Figure 6) (Schultz and Yarus, 1994; Santos et al., 2004) could have caused this pattern, but this hypothesis awaits confirmation from studies of tRNA populations and eukaryotic release factor sequences.

The elongation factor 1α gene (EF-1α), which has a function in delivering aminoacyl-tRNAs to the eukaryotic ribosome,

**D. Genetic Codes and the Translational Apparatus in Green Seaweeds**

Translation is the first step of protein biosynthesis in which the ribosome decodes messenger RNA (mRNA) to form polypeptides that will later fold into the protein’s active structure. The eukaryotic ribosome consists of two subunits, each of which is composed of a folded RNA and several ribosomal proteins. A crucial aspect in translation is the recognition of the codons in the mRNA by anticodons on the transfer RNAs (tRNA), which carry the corresponding amino acids. This recognition is generally well-conserved across eukaryotes, leading to a nearly universal genetic code linking codons to their corresponding amino acids.

Deviations from the canonical genetic code are known in some eukaryotes, and they have also been shown in some green algal groups. Early work on cDNA of the nuclear-encoded small subunit of the Rubisco gene showed that the stop codons UAA and UAG were reassigned to code for glutamine in the siphonous dasycladalean green seaweeds *Acetabularia* and *Batophora* (Schneider et al., 1989; Schneider and Degroot, 1991). More recently, the evolution of these codon reassignments among the Ulvophyceae was studied in more detail using multi-gene datasets (Gile et al., 2009; Cocquyt et al., 2010a). These studies have shown a complex distribution of the alternative genetic code, the canonical genetic code being used in early-branching Ulvophyceae and the Bryopsidales, and the alternative code being used in Dasycladales, Cladophorales and Trentepohliiales (Figure 6). Several explanations for the observation that the Bryopsidales, which have the canonical code, are nested in a clade of which all other representatives have the alternative code are compared by Cocquyt et al. (2010a). These authors suggest that a stepwise acquisition model involving ambiguous intermediates (Figure 6) (Schultz and Yarus, 1994; Santos et al., 2004) could have caused this pattern, but this hypothesis awaits confirmation from studies of tRNA populations and eukaryotic release factor sequences.

The elongation factor 1α gene (EF-1α), which has a function in delivering aminoacyl-tRNAs to the eukaryotic ribosome,

**FIG. 6.** Evolution of genetic codes and elongation factor genes in the Ulvophyceae. The first column shows the distribution of canonical and alternative genetic codes. The complex distribution of genetic codes, in which all but one representative of a clade possess an identical deviant genetic code, can be explained by a stepwise acquisition model involving ambiguous intermediates whereby the deviant code has a single initiation (i) and the final steps from an ambiguous intermediate situation to a non-canonical code have been completed (c) all orders but the Bryopsidales as a result of disappearance (d) of the ambiguously decoding tRNAs from the genome in the latter. The second column shows the distribution of EF-1α and EFL. The elongation factor gene in the Trentepohliiales is not known with certainty. (Color figure available online.)
is a crucial function, not all eukaryotes have this protein, and it has been shown that an elongation factor-like protein (EF-like) can substitute for EF-1α in these taxa (Keeling and Inagaki, 2004). Interestingly, most eukaryote genomes contain either EF-1α or EF-like, although a few have both (Kamikawa et al., 2008). In green algae, these elongation factors show a mutually exclusive but scattered distribution (Noble et al., 2007; Cocquyt et al., 2009; Gile et al., 2009). Whereas the streptophytes have EF-1α, with the exception of Mesostigma, the chlorophytes encode EF-like with the exception of certain ulvophytes. Like the prasinophytes and most core chlorophytes, the early-branching ulvophytes have the EF-like gene whereas the clade comprising Ignatius, the Trentepohliales, Dasycladales, Bryopsidales and Cladophorales have EF-1α (Figure 6). Regrettably, phylogeny-explicit models of gene gain and loss were unable to determine the presence of either one or both of the genes in the genome of the ancestor of the green lineage and the Ulvophyceae (Cocquyt et al., 2009). Nevertheless, the alternative elongation factors used in different lineages of the green seaweeds can be taken to mark a second considerable alteration of the translation pathway that occurred during the evolution of the Ulvophyceae.

In addition to the changes in genetic code and elongation factors mentioned above, two other sources of evidence point to changes in the translational apparatus of the Ulvophyceae (Cocquyt, 2009). First, patterns of codon usage of eight nuclear housekeeping genes differ considerably between the prasinophytes and Ulvophyceae. Second, the rates of molecular evolution of the 18S ribosomal RNA molecule are considerably elevated in the ulvophycean orders Cladophorales, Dasycladales and Bryopsidales. All this evidence taken together points to profound changes in the translational apparatus of the Ulvophyceae, but this hypothesis awaits confirmation from more detailed studies of the translational pathway.

E. Molecular Evolution in the Streptophyta and the Origin of Land Plants

The origin of embryophytic land plants from a green algal ancestor was a major event in the history of life, which influenced the establishment of the entire terrestrial ecosystem and had far-reaching effects on atmospheric chemistry and climate (Berner, 1997; Kenrick and Crane, 1997; Steemans et al., 2009). As noted above (Section II.B.4), although the Streptophyta are not particularly diverse (outside of the land plants) in terms of number of described species, they represent the full range of structural diversity, from scaly unicellular flagellates through unbranched and branched filaments, to the complex three-dimensionally organized tissues of plants. There is little doubt, however, of the phylogenetic placement of embryophytes among the other streptophytes. The structurally simple streptophytes retain a number of ancestral features that can shed light on the factors that led to the success of embryophytes in the terrestrial environment.

The colonization of dry land involved adaptation to new and harsh environmental conditions such as desiccation, temperature fluctuations, the need for support in the absence of a buoyant medium, and UV radiation (Raven, 2000; Waters, 2003; Floyd and Bowman, 2007; Lang et al., 2008). Ecophysiological adaptations of the first land plants included enhanced osmoregulation and osmoprotection, desiccation and freezing tolerance, and heat resistance (Rensing et al., 2008). Several biochemical innovations have been identified, including synthesis and accumulation of protective “sunscreens,” plant growth hormones, isoprenoids, phenolics, heat shock proteins, and enhanced DNA repair mechanisms (Waters, 2003; Rensing et al., 2008). In addition, several morphological innovations are believed to have allowed successful adaption to life on land and radiation into new niches (Graham et al., 2000). Some of these are found in one or more algal relatives of embryophytes, including production of extracellular matrices such as sporopollenin and perhaps lignin, the development of differentiated cells and tissues, dorsoventral development (which is important in the development of structures such as leaves), the establishment of intercellular communication networks (plasmodesmata, plant hormones, receptors and their ligands), and the perception of environmental cues (light and gravity), while others, such as life cycle involving alternation of two distinct multicellular generations (a haploid sporophyte and diploid gametophyte), protected embryos, and gas-filled spaces within the plant body appear to be unique to embryophytes (Delwiche et al., 1989; Graham et al., 2000; Bowman et al., 2007; Ligrone et al., 2008). The success of land plants has further been linked to symbiotic associations with mycorrhizal fungi (Simon et al., 1993; Heckman et al., 2001; McCourt et al., 2004).

Important evolutionary transitions and adaptive radiations, such as the origin and diversification of land plants, have been associated with gene family expansions resulting from large-scale gene duplication or whole-genome duplication events (Ohno, 1970; De Bodt et al., 2005; Flagel and Wendel, 2009). The various physiological and morphological adaptations to land were likely associated with expansion of gene families involved in signalling pathways, such as those for auxin, ABA and cytokinin (Lang et al., 2008; Rensing et al., 2008; Timme and Delwiche, 2010; De Smet et al., 2011). Morphological innovations and the evolution of morphological complexity in land plants have been linked with increased gene family complexity of several genes including actin (An et al., 1999), MADS box genes (Henschel et al., 2002; Tanabe et al., 2005), homeobox genes (Mukherjee et al., 2009) and OPR genes (Li et al., 2009). Expansion of the glutaredoxins gene family (enzymes implicated in oxidative stress response) in land plants likely resulted in genes with novel functions in development and pathogenesis response (Ziemann et al., 2009). The unique sexual life cycle of land plants likely evolved through expansion of homeodomain gene networks (e.g. MADS-box genes) (Tanabe et al., 2005; Niklas and Kutschera, 2010). Ca²⁺-dependent signalling processes, which are important in the response to many developmental and environmental stimuli, have been found to be very different in green algae and land plants, with several Ca²⁺ signalling mechanisms
having apparently been lost in land plants after their divergence from charophyte algae (Wheeler and Brownlee, 2008).

Although there is still uncertainty concerning the precise relationships between land plants and their green algal relatives (see Section II.B.4) it is clear that several fundamental land plant features were inherited from their charophyte green algal progenitors, including cellulosic cell walls, multicellularity, cytokinetic phragmoplast, plasmodesmata, apical meristematic cells (although the meristematic regions in the Charophyceae, Coleochaetophyceae and Zygnematophyceae are very different from that in the embryophytes), asymmetrical cell division, cell specialization, branching, three-dimensional organization, zygote retention, and placenta (Graham et al., 2000). It has been hypothesized that charophyte green algae were physiologically pre-adapted to life on land by their primarily freshwater life style, which exposed them to periodic exposure and desiccation, and allowed a gradual shift towards moist terrestrial habitats, and ultimately the colonization of dry land (Becker and Marin, 2009). It is striking that several streptophyte lineages other than embryophytes are found in terrestrial habitats (i.e., Chlororhybus atmosphyticus, many species of Klebsormidium, and several Zygnematales), which emphasizes the close relationship between freshwater and moist terrestrial habitats.

A growing body of evidence indicates that the molecular bases of many morphological, life-history, cytological, biochemical and physiological features of embryophytes lie in an era prior to the colonization of land (Becker and Marin, 2009). This has been shown by several studies elucidating the evolution of land plant genes in the green lineage, including actin (An et al., 1999), cellulose synthase (Roberts and Roberts, 2004; Yin et al., 2009), class III homeodomain–leucine zipper genes (Floyd et al., 2006), MADS-box genes (Tanabe et al., 2005), vacuolar sorting receptors (Becker and Hoef-Emden, 2009), arabinogalactan-like proteins and hemicelluloses (e.g. endotransglucosylase, responsible for cell wall loosening and cell expansion) (Van Sandt et al., 2007; Eder et al., 2008; Fry et al., 2008). Expressed sequence tag (EST) sequencing of various charophytes and land plants has been used in comparative plant genomics studies to uncover the origins of land plant genes and their associated molecular pathways (Nedelcu et al., 2006; Simon et al., 2006; Timme and Delwiche, 2010). Overall, these studies indicate that several land plant characteristics evolved before the transition to land. For example, EST analysis of Mesostigma (Mesostigmatophyceae) suggests that important physiological changes involving regulation of photosynthesis and photorespiration took place early in the evolution of the Streptophyta (Simon et al., 2006). Similarly, putative homologs of genes involved in the development of three-dimensional tissues through asymmetrically dividing apical meristematic cells (BB1 gene family) have been identified in Mesostigma (Nedelcu et al., 2006). Several other genes that have been hypothesized to be important in the colonization of land plants (Graham et al., 2000) may have true orthologs in Coleochaete (Coleochaetozychyceae) and Spirogyra (Zygnematophyceae) but are apparently absent in the ESTs of the earlier diverging Mesostigma (Timme and Delwiche, 2010). These include genes involved in cellulose synthesis (RSW1), phragmoplast mediated cytokinetic (GEM1/MOR1), formation of plasmodesmata (CRT1) and development of a multicellular sporophyte (MERISTEM LAYER1). Two genes associated with asymmetric cell division (WUSCHEL and GNOM) were only found in Coleochaete, while the plant cell wall loosening expansin genes (EXP) were only present in Spirogyra. In addition, several ethylene pathway genes, long thought to be unique to land plants, have been identified in Coleochaete and Spirogyra.

Cell walls are believed to have played crucial roles in the colonization of land by plants (Sorensen et al., 2010). Although some cell wall components do appear to be land plant innovations, cell wall evolution after the colonization of land appears to be characterized mostly by the elaboration of a pre-existing set of cell wall polysaccharides and the enzymes that synthesize them (e.g., cellulose synthase and wall-remodelling enzymes), rather than substantial innovation (Roberts and Roberts, 2004; Domozych et al., 2007; Van Sandt et al., 2007; Eder et al., 2008; Fry et al., 2008; Domozych et al., 2009; Yin et al., 2009; Popper and Tuohy, 2010; Popper et al., 2011).

There are several other examples where the genetic potential for plant specific features has been suggested in green algae. For example, genes involved in auxin signalling (central to land plant growth and development) have been detected in various members of Chlorophyta and Streptophyta, indicating that auxin response and transport mechanisms were likely present before the evolution of land plants (De Smet et al., 2011). Similarly, the B3 DNA Binding superfamily, including genes mainly involved in hormone signaling pathways such as those for auxin, abscisic acid, brassinosteroid and gibberellins, are also present in Chlorophyta but have undergone extensive duplication events during land plant evolution (Romanel et al., 2009).

V. CONCLUSIONS AND PERSPECTIVES

20 years ago, a review article in this journal reported on the state of knowledge on green algal relationships, gathered from the few pioneering years of ribosomal DNA-based phylogenetic research (Chapman et al., 1991). The past two decades have witnessed profound changes in our understanding of the evolution of green algae. 18S phylogenies have made way for multi-gene and genome scale analyses. These studies have greatly improved our understanding of the deepest relationships of the green lineage (Figure 3); however, many questions remain.

A large body of molecular evidence has confirmed the ultrastructural-based hypothesis that the green lineage diverged into two discrete clades: the Chlorophyta, which includes the majority of described species of green algae, and the Streptophyta, which is comprised of the charophytes, a paraphyletic assemblage of freshwater algae from which the land plants have evolved. The prasinophytes take up a critical position,
diverging early from the remaining Chlorophyta, but the relationships among these lineages remain largely unresolved, mainly because multi-gene data are only available for a limited number of taxa. Similarly, phylogenetic relationships among and within the main clades of the core chlorophytes (Ulvo- phyceae, Trebouxiophyceae and Chlorophyceae) have not been fully resolved. Other outstanding issues include elucidation of the early diversification of the Streptophyta and identification of the closest living relative of the land plants.

It has become clear that an accurate phylogenetic reconstruction of an ancient group like the green plants will require a rich sampling both in terms of exemplar taxa and molecular markers, along with the application of state of the art phylogenetic techniques. The rapid increase in the amount of genomic data from a wide range of green algae has great potential to resolve large-scale green algal relationships. In addition these data form the basis for investigations of molecular evolution of genes and genomes, providing valuable insight into the evolutionary histories of the green algae.

ACKNOWLEDGMENTS

We are grateful to Pavel Škaloud, Fabio Rindi, Juan López-Bautista, Frederick Zechman, Marvin Fawley, Aurora Nedelcu, Marc Buchheim, Andrey Gontcharov, Burkhard Becker and Bautista, Frederick Zechman, Marvin Fawley, Aurora Nedelcu, Tom Schils, Antoni L´opez-Arenas, Nadia Abdelahad and Deb- orah Shelton for providing photographs. Support was provided by FWO-Flanders (FL, HV and ODC).

REFERENCES


Boer, P. H. and Gray, M. W. 1988. Genes encoding a subunit of respiratory NADH dehydrogenase (ND1) and a reverse transcriptase-like protein (RTL) are linked to ribosomal RNA gene pieces in Chlamydomonas reinhardtii mitochondrial DNA. EMBO J. 7: 3501–3508.

Boudreau, E. and Turmel, M. 1996. Extensive gene rearrangements in the chloro-
Chlamydomonas reinhardtii (Chlorophyceae) high GC content and
highly rearranged chloroplast genomes of Chlamydomonas moewusii and
Boudreau, E. and Turmel, M. 1996. Extensive gene rearrangements in the chloro-
plast DNAs of Chlamydomonas species featuring multiple dispersed repeats.
Bremner, K. L. 1985. Summary of green plant phylogeny and classification. Cladis-
tics 1: 369–385.
green algae inferred from nuclear small-subunit ribosomal DNA. J. Phycol.
34: 306–311.
DNA in the Oogamochlamys clade (Chlorophyceae): high GC content and
novel genes acquired through horizontal transfer.
flagellates: a study of 18S rRNA sequence data. [cited 11 October 2011]
Phylogeny of the filamentous green algal orders Chaetophorales and
Oedogoniales based on 18S rRNA gene sequences. [cited 11 October 2011]
1993. Complex phylogenetic distribution of a non-canonical genetic code
Chapman, R. L., Buchheim, M. A., Delwiche, C. F., Friedl, T., Huss, V. A.
Karol, K. G., Lewis, L. A., Manhart, J., McCourt, R. M., Olsen, J. L., and
J. J. Eds., Kluwer Academic, Boston, MA.
gene sequences: Analysis and significance in the phylogeny and taxonomy of
Res. 173: 201–211.
Butterfield, N. J., Knoll, A. H., and Swett, K. 1994. Paleobiology of the Neo-
proterozoic Svanbergfjellet Formation, Spitsbergen. Fossils and Strata 34:
1–84.
Cachon, M. and Caram, B. 1979. A symbiotic green alga, Pedinomonas symbiot-
ica sp. nov. (Prasinophyceae), in the radiolarian Thalassolasme margarodes.
Polyphyly of Chroptophora and Stigeoclonium within the Chaeotophorales
(Chlorophyceae), revealed by sequence comparisons of nuclear-encoded SSU
Cardol, P., Bailleul, B., Rappaport, F., Derelle, E., Beal, D., Breyton, C., Bailey,
original adaptation of photosynthesis in the marine green algae Ostreococcus.
14: 461–481.
Philos. Soc. 73: 203–266.
Cavalier-Smith, T. 1999. Principles of protein and lipid targeting in second-
sory symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid
origins and the eukaryote family tree. J. Eukaryot. Microbiol. 46: 347–
366.
Microbiol. 5: 612–619.
Chapman, R. L. 1984. An assessment of the current state of our knowl-
London.
Chapman, R. L., Borisenious, O. L., Brown, R. C., Henk, M. C., and Waters,
51: 759–765.
Chapman, R. L., Buchheim, M. A., Delwiche, C. F., Friedl, T., Huss, V. A.
Karol, K. G., Lewis, L. A., Manhart, J., McCourt, R. M., Olsen, J. L., and
J. J. Eds., Kluwer Academic, Boston, MA.
gene sequences: Analysis and significance in the phylogeny and taxonomy of
biflagellate zoospores of the enigmatic marine green-alga Blastophysa
role of GlrA in the evolution of asymmetric cell division in the green algae
2010. Composition and genetic diversity of picoeukaryotes in subtropi-
cal coastal waters as revealed by 454 pyrosequencing. ISME J. 4: 1053–
1059.
Chihara, M., Inouye, I., and Takahata, N. 1986. Oltmannsieliopsis, a new genus
of marine flagellate (Dinulillaceae, Chlorophyceae). Arch. Protistenk.
132: 313–324.
Chisholm, J. R. M., Dauga, C., Ageron, E., Grimont, P. A. D., and Jaubert, J.
Cocquyt, E. 2009. Phylogeny and Molecular Evolution of Green Algae. PhD
thesis, Ghent University.
Cocquyt, E., Gile, G., Leliaert, F., Verbruggen, H., Keeling, P. and De Clerck,
O. 2010a. Complex phylogenetic distribution of a non-canonical genetic code
Cocquyt, E., Gile, G., Leliaert, F., Verbruggen, H., Keeling, P. and De Clerck,
O. 2010b. Complex phylogenetic distribution of a non-canonical genetic code


de Koning, A. P. and Keeling, P. J. 2006. The complete plastid genome sequence of the parasitic green alga Helicosporidium sp. is highly reduced and structured. BMC Biology 4: art no. 12.


Eddie, B., Krens, C., and Neuer, S. 2008. Characterization and growth response to temperature and salinity of psychrophilic, halotolerant Chlamydo-


Eikrem, W. and Thronsen, J. 1990. The ultrastructure of Bathycoccus gen. nov. and Bathycoccus prasinos sp. nov., a nonmotile picoplanktonic alga (Chloro-


MOLECULAR EVOLUTION OF GREEN ALGAE


