Paths toward Algal Genomics

Arthur R. Grossman*

The Carnegie Institution, Department of Plant Biology, Stanford, California 94305

The last decade has led to an explosion of genomic information that is being used to help researchers understand the gene content of organisms, how gene content and expression patterns may explain the ecological niche in which the organism lives, the ways in which gene content have been arranged and modified by evolution, the movement of genes and gene clusters among different organisms, and environmental and developmental processes that modulate the expression of genes. In this introductory manuscript, I discuss select algae and how genomics is impacting our understanding of these organisms. Four algae for which near-full genome information has become or will shortly become available are the red alga *Cyanidioshyzon merolae*, the green alga *Chlamydomonas reinhardtii*, the diatom *Thalassiosira pseudonana*, and the marine picoeukraryote *Ostreococcus tauri*. There is also the full sequence of the vestigial red algal genome associated with the nucleomorph of the Cyptomonad *Guillardia theta*. A number of other algal genomes, such as that of *Phaeodactylum tricornutum*, are currently being sequenced. Furthermore, there has been a substantial body of cDNA sequence information generated from various algae. Algae are important contributors to global productivity and biogeochemical cycling, but genomics of these organisms is still in its infancy, and the resources to support large scale projects concerning algal genomes and global gene expression are limited. However, it is useful to discuss the algae that are currently being examined using genomic technologies, some of the information that has been generated from genomic analyses, criteria that may be used for choosing specific organisms for future genome studies (and viable candidates for such studies), and how the information gained might help us better understand structural, functional, developmental, and evolutionary aspects of photosynthetic organisms.

Genomics is often viewed as the generation and analyses of nucleotide sequences of the full or near-full genome as well as cDNAs collections. From sequence information, researchers identify individual genes and repeat elements, analyze the organization and arrangement of genes, and make comparisons among genomes with respect to gene arrangement and sequence identity/similarity; sometimes descriptions of genomics extend to the use of methods for examining global gene expression using microarray technology.

A number of different bacterial and mammalian systems (including humans) that serve as models for genomic studies have been developed because the information gained from such studies can be of immediate importance with respect to human health. However, other systems, including the algae, are gradually benefiting from rapid, widespread use of genomic techniques. Although many would consider the development of algal genomic systems as less urgent than those associated with humans, mice, and pathogenic bacteria, the algae are critical components of many habitats on the planet and are major producers of fixed carbon, especially in marine ecosystems.

The algae are a highly diverse group of photosynthetic organisms that are ubiquitous on the Earth and are critical for maintaining terrestrial and atmospheric conditions. These organisms come in a variety of forms ranging from the tiny picoplankton that inhabit open oceans (Díez et al., 2001; Biegala et al., 2003; see also http://www.sb-roscoff.fr/Phyto/PICODIV/PICODIV_publications.html) to the macrophytic organisms that form turf meadows and forests in coastal waters (Graham and Wilcox, 2000). The diversity among the algae is enormous, not only with respect to size and shape of the organisms, but also with respect to the production of various chemical compounds through novel biosynthetic pathways. For example, the different pigments that comprise the light-harvesting antennae in algae are visually striking and biochemically diverse. In the green algae, the light-harvesting antennae contain mostly chlorophylls *a* and *b*, with a significant level of carotenoids, while the antennae pigments of the red algae and cyanobacteria are predominantly the phycobiliproteins, in which bilin chromophores (phycoerythrobilin and phycocyanobilin) are covalently bonded to apophycobiliproteins. In contrast, diatoms and dinofagellates use oxygenated carotenoids as their major light-harvesting pigments. The composition of polysaccharides and cell walls also shows enormous diversity among the algae. For example, some algae have microfibrillar walls of cellulose or other polysaccharides and others have proteinaceous or silicacious walls or scales.

Algae are also economically important since they serve as a source of food, and in many parts of the world they can be used in salads, soups, and as garnish. Most well known among algal foods is the wrap for sushi, or nori, which is derived from the

* E-mail arthurg@stanford.edu; fax 650–325–6857.

www.plantphysiol.org/cgi/doi/10.1104/pp.104.053447.

dried fronds of the red alga Porphyra. Algae are also used as a vitamin source by the health food industry (http://www.1001beautysecrets.com/nutrition/algae/), especially cyanobacteria or blue green algae (http://www.crystalpurewater.com/health.htm) since they can be rich in the vitamin A precursor β-carotene. But there is a wide range of uses for algae and algal products. They are used as feed additives for aquaculture, as coloring agents to enhance the appeal of food, and as fluorescent tags to identify, quantify, or localize surface antigens for specific medical assays. Algae also synthesize a number of different polysaccharides and lipids that, in addition to serving as carbon storage compounds, perform biological functions and have commercial value. Some of the polysaccharides are anionic and bind metal ions, chelate heavy metals, and help maintain a hydration shell around the alga. The commercially valuable polysaccharides are agar, carrageenans, alginates, and fucoids (Berteau and Mulloy, 2003; Feizi and Mulloy, 2003; Drury et al., 2004; Matsubara, 2004). Certain of these polysaccharides have anticoagulant characteristics (Matsubara, 2004), while others are used for making solid medium for growing bacteria in the laboratory, gels for the delivery of medicines, thickeners in food products such as ice cream, and numerous products including cosmetics, cleaners, ceramics, and toothpaste (http://www.nmnh.si.edu/botany/projects/algae/Alg-Prod.htm). Furthermore, both diatoms and dinoflagellates synthesize long chain polyunsaturated fatty acids (fish oils) that appear to be beneficial for mammalian brain development (Chamberlain, 1996; Salem et al., 2001); these fatty acids are sold as health food products but are also being incorporated into baby formula in many countries throughout the world.

While most algae thrive as free-living organisms, some are more prevalent in symbiotic associations, and still others have evolved into parasites (Goff and Coleman, 1995). Many of the symbiotic associations established by algae are critical for survival of the heterotrophic host organism in environments with low levels of organic carbon compounds. For example, the dinoflagellate *Symbiodinium* sp. populates and transfers fixed carbon to the tissue of corals, allowing for the establishment and maintenance of the coral reefs that physically stabilize the coastal environment (Murdoch, 1996). Rising temperatures are causing bleaching of the reefs, which could have a pronounced impact on the environment (Coles and Brown, 2003). The growth of specific algae in oceans, estuaries, and lakes can be of concern since they can attain very high densities or blooms that stimulate the proliferation of consumers and the generation of anoxic conditions that can suffocate aquatic animals. A number of the algae and cyanobacteria that form such blooms also produce neurotoxins and are a threat to global water supplies and fisheries (especially with respect to the shell fish industry). Furthermore, the composition of phytoplankton communities has implications with respect to carbon fluxes and the trophic transfer of carbon in food chains.

One difficulty facing algal biologists is the challenge to move from morphological, chemical, and geophysical descriptors of algal/bacterial communities to more molecular descriptors that include both gene content and expression levels. Indeed, our understanding of biological, biophysical, and geochemical processes will all be informed by the wealth of data that can be acquired using a spectrum of biotechnological methods that have been developed over the last 20 years. Much of this information will have its origins in acquiring the full-gene content of an organism, combined with tools to determine the level of expression of specific genes under different environmental conditions, at different developmental stages, and in different tissue types. Naturally, genomic studies are expensive and the resources to support such studies are limited. It is critical that societies and scientific communities with knowledge of the scientific and economic importance of particular groups of organisms, such as the algae, make informed choices as to which organisms would be of most benefit for genomic examination, whether involving whole genome or cDNA projects. It would be most efficient to solicit the aid of large, well-equipped centers that have an expert staff to complete the required sequencing tasks efficiently. However, the first important step for the scientific community with a working knowledge of the field is to define the organisms for which full-genome and cDNA sequences should be obtained, to develop collaborations to facilitate the generation and analysis of genomic information, to petition various agencies for the funds required to obtain the sequence information, and to help train the community, either through courses or workshops and tutorials over the internet, in ways in which the genomic information can be used and extended.

SEQUENCED GENOMES

Sequence information for the genomes of organelles, and especially chloroplasts, is available for a number of the algae including those of the green algae *Chlamydomonas reinhardtii* (http://bti.cornell.edu/bti2/chlamyweb/default.html), *Nephroselmis olivacea* (Turmel et al., 1999), *Chaetosphaeridium globosum* (Turmel et al., 2002), *Chlorella vulgaris* (Wakasugi et al., 1997), and *Mesostigma viride* (Lemieux et al., 2000); the cryptophyte *Guillardia theta* (Douglas and Penny, 1999); the stramenopile (diatom) *Odontella sinensis* (Tada et al., 1999; Chu et al., 2004); the stramenopile *Heterosigma akashiwo* (Velluppillai et al., 2003); the glaucophyte *Cyanophora paradoxa* (Stirewalt et al., 1995); the red alga *Cyanidium caldarium* (Glockner et al., 2000); and the euglenophyte *Euglena gracilis* (Hallick et al., 1993). Interestingly (and surprisingly), the plastid genes of dinoflagellates are unique in that each gene appears to be on its own minicircle (Zhang et al., 1999, 2002).

Sequences of chloroplast genomes of both algae and plants can be accessed at http://megasun.bch.umontreal.ca/ogmp/projects/other/cp_list.html.

Currently, there are few algae for which the nuclear genome has been sequenced. Recently, complete or nearly completed sequences of the genomes of the red alga *Cyanidioschyzon merolae* (http://merolae.biol.s.u-tokyo.ac.jp/; Matsuzaki et al., 2004), the diatom *Thalassiosira pseudonana* (http://genome.jgi-psf.org/thaps1/thaps1.home.html; Armbrust et al., 2004), and the green alga *C. reinhardtii* (http://genome.jgi-psf.org/chlre2/chlre2.home.html) have been made publicly available. Other genomes either sequenced and not released or in the process of being sequenced include *Ostreococcus tauri* (http://www.iscb.org/ismb2004/posters/stromATpsb.ugent.be_844.html; Derelle et al., 2002), *Volvox carteri*, and *Phaeodactylum tricornutum* (see http://trace.ensembl.org/perl/traceview?attr=tt_ce_sp&tt_1=1). In addition, the complete sequences of the three chromosomes that constitute the nucleomorph genome of *G. theta*, which represents a vestigial red algal genome, have been reported (Douglas et al., 2001). But this is just the beginning of an era that is triggering an explosion of information on gene content, gene organization, and the sequences that control gene expression from numerous organisms within the different kingdoms of life. Below, I discuss the algae for which there is significant genomic sequence information (discussed in various articles in this issue of *Plant Physiology*, especially for *C. reinhardtii*), but I also try to raise issues concerning the direction of algal genomics and ways to decide on organisms for which full genome sequences will be most immediately useful.

Nucleomorph Genome of *G. theta*

Of the chlorophyll *c*-containing chromophytic algae, the Cryptomonads are the only organisms to retain the enslaved red algal nucleus that resulted from a secondary endosymbiotic event (Cavalier-Smith, 2000; Maier et al., 2000). This reduced nucleus or nucleomorph has an envelop membrane with nuclear pores, but the genetic content of the nucleomorph is highly reduced relative to a red algal genome. The DNA of the nucleomorph of the Cryptomonad *G. theta* has now been sequenced.

The nucleomorph of *G. theta* contains 3 mini-chromosomes that together constitute 551 kb. This genome is predicted to have 464 genes encoding polypeptides, of which nearly one-half encode proteins of unknown function. The genes are highly compacted in the genome (which has almost no noncoding DNA), and only 17 of the protein coding genes contain introns that can be removed by a spliceosome. Most of the introns are near the 5′ ends of the transcripts, and 11 of these 17 intron-containing genes encode ribosomal proteins.

There are a number of interesting aspects with respect to the protein coding sequences of the nucleomorph genome. Most proteins encoded on the nucleomorph genome are needed for the replication of the chromosomes, gene expression, and perpetuation of periplastid ribosomes, with few required for other cellular functions. For example, a number of the nucleomorph-encoded proteins participate in the processing of mRNA, the removal of tRNA introns, and the maturation of rRNA. However, the genome does contain 30 chloroplast targeted proteins, 3 transporters, and a few enzymes (one anabolic and some regulatory). Since the plastid genome houses a small percentage of the genes required for the biogenesis of functional chloroplasts, and the nucleomorph only encodes an additional 30 chloroplast localized proteins, most of the proteins that function in the chloroplast must be synthesized in the cytoplasm of the cell and traverse the rough endoplasmic reticulum (ER), the periplastid membrane, pass through the periplastid space, and then cross the double envelop membrane of the plastid to reach their site of function within the organelle. The arrangement of these membranes and the location of the nucleomorph within the periplastid space are clearly diagrammed by Douglas et al. (2001).

Of the plastid-localized polypeptides encoded on the nucleomorph genome, only a few function in photosynthesis (rubredoxin and HLIP; the latter is a small protein in the light-harvesting complex (LHC) protein family important for survival during high light stress in cyanobacteria [He et al., 2001]), plastid division and gene expression, nucleic acid metabolism, and protein translocation into the plastid and thylakoids. The nucleomorph encoded plastid proteins have amino terminal extensions that, in the case of rubredoxin, have been shown to function as a transit peptide that enables the protein to traverse the plastid envelop membrane (Wastl et al., 2000). The nucleomorph genome also encodes RNA polymerase subunits, regulatory proteins that may influence starch accumulation, protein synthesis, and nucleomorph DNA replication and division; three core histones plus a histone acetylase and deacetylase; and proteins of the ubiquitin-proteasome degradation pathway. There are also proteins essential for nucleomorph functions that are not encoded by the nucleomorph genome; these proteins, which include the subunits of DNA polymerase, would have to be routed from the cytoplasm of the cell to the nucleomorph.

Elucidating steps involved in the biosynthesis of the plastid, the nucleomorph, and periplastid compartment, and developing an understanding of coordinate expression of genes encoded on the nuclear, plastid, and nucleomorph genomes will increase our understanding of the roles of the various compartments in cellular processes, the communications between the different genetic compartments of a cell, and the ways in which proteins and metabolites are exchanged among these compartments. Ultimately, defining the genetic content of all of the different genomes in the Cryptomonads will help elucidate the loss of genetic information in the genome of the endosymbiont following the secondary endosymbiotic event and the exchange of genetic information among the genomes.

C. merolae

The Cyanidiales is a group of unicellular, asexual red algae that grow at high temperatures and under acidic conditions. This group includes the genera Cyanidium, Cyanidioschyzon, and Galdieria, although recent work suggests an unexpectedly high level of genetic diversity among the Cyanidiales (Ciniglia et al., 2004). The first algal nuclear genome to be sequenced was that of a member of the Cyanidiales, *C. merolae*, whose genome is among the smallest that occurs in photosynthetic eukaryotes. *C. merolae* is an organism that grows in the hotspring (45°C) at a pH of 1.5 and is considered one of the most primitive algal species (Kuroiwa et al., 1998; Matsuzaki et al., 2004; Nozaki et al., 2004). Its subcellular structure is relatively simple with a single Golgi apparatus and ER and a relatively small number of internal membrane structures. The plastid genome of this organism, which is about 150 kb and contains 243 genes, has been sequenced (Ohta et al., 2003). Interestingly, there is an overlap between the protein coding sequences for many of these genes (40%), which has resulted in a highly compacted plastid genome.

C. merolae has also been the subject of a number of interesting studies concerning mechanisms by which mitochondria and plastids divide (Kuroiwa et al., 1998; Miyagishima et al., 1999; Kuroiwa, 2000; Miyagishima et al., 2001a, 2001b, 2001c, 2003; Nishida et al., 2004). Furthermore, it may be possible to introduce exogenous DNA into these organisms by electroporation; the introduced DNA appears to integrate into the nuclear genome by homologous recombination (Minoda et al., 2004).

The recently sequenced nuclear genome of *C. merolae* (which still contains some gaps) is approximately 16.5 Mb, with 5,331 genes packed into 20 chromosomes. Within the genome there are only three rDNA units that are not tandemly arranged but define separate loci (Maruyama et al., 2004). The nucleolus is small and not associated with chromatin, which might make it a relatively simple model for defining the composition and biochemical features of a minimal nucleolus. Of the predicted genes contained in the nuclear genome, only 26 have introns and all but 1 of these have single introns. This organism has a very minimal set of motor proteins that includes a set of tubulin subunits, two actins, and both intermediate filament and kinesin family proteins. Furthermore, there are only 2 dynamin encoding genes (most organisms have a family of dynamin genes containing at least 10 members), which function in mitochondrion and chloroplast division, and no genes encoding myosin or dynein motors. These findings suggest that a highly reduced set of motor proteins accomplish cytokinesis and cell motility in this organism.

The analysis of the *C. merolae* genomic sequence also has implications with respect to the endosymbiont origins of the plastid. The enzymes of the Calvin cycle originated from a combination of genes derived from a cyanobacterial endosymbiont and its eukaryotic host. This mosaic gene composition is similar in *C. merolae* and Arabidopsis (*Arabidopsis thaliana*), suggesting that they originated from a common ancestral organism and that this composition remained stable even after the separation of the two lineages. There are many other interesting observations/deductions developing from the sequence of the *C. merolae* genome, including the finding that the tRNAs contain ectopic introns, that there are no genes encoding two of the major classes of photoreceptors associated with plants (the phototropins, which are blue UV-A light photoreceptors and the phytochromes, which are red light photoreceptors), and that there is only a single His kinase and no response regulators other than those encoded on the plastid genome. A seemingly limited repertoire of signaling elements encoded on the nuclear genome of this alga may reflect the specialized environmental niche in which this organism grows. It would also be interesting to learn more about the transport proteins associated with the cytoplasmic membrane of this and related organisms and the mechanisms by which it deals with the low external pH of the environment (the pumps and exclusion mechanisms that may be associated with maintaining the pH of the cytoplasm of the cell).

T. pseudonana

Diatoms are a diverse group of organisms present in marine, freshwater, and terrestrial environments. They are estimated to be represented by tens-of-thousands of species on the Earth (Round et al., 1990) and may be responsible for as much as 20% of global primary productivity. These organisms can have different gross morphologies (pennate, centric, coccoid, triangular) with precisely patterned and beautifully ornamented silicified cell walls or frustules. Recent work on diatoms has employed sophisticated molecular techniques, and many different diatom species can now be transformed using biolistic procedures (Dunahey et al., 1995; Apt et al., 1996; Zaslavskaia et al., 2000). Reporter genes have also been successfully introduced into diatoms to study gene expression; these reporters include the *Escherichia coli uidA* gene encoding β-glucuronidase, the Tn9-derived *cat* gene encoding chloramphenicol acetyl transferase, the firefly *luc* gene encoding luciferase (Falciatore et al., 1999), a variant of the green fluorescent protein gene (*egfp*), and the aequorin gene from the jellyfish *Aequorea victoria* (Falciatore et al., 2000). Genes encoding proteins fused to GFP have been introduced into the diatoms and the fusion proteins targeted to various subcellular compartments, including the lumen of the ER (Apt et al., 2002), the chloroplast (Apt et al., 2002), and the cytoplasmic membranes (Zaslavskaia et al., 2001). A chimeric gene encoding the human Glc transporter fused to GFP was introduced into *P. tricornutum*. The expressed protein integrated into the cytoplasmic membranes and converted this diatom

from an obligate photoautotroph to a heterotroph (growth in the dark on exogenous Glc; Zaslavskaia et al., 2001). One significant problem in working with the diatoms stems from the fact that they are diploid and researchers have not been able to consistently achieve sexual crosses, making it difficult to obtain mutants in which both alleles for a specific gene have been modified. Hopefully, continued analyses of the life cycle of the diatoms will help reveal factors that elicit and control sexuality in these organisms (Vaulot et al., 1986, 1987; Armbrust and Chisholm, 1990; Mann, 1993; Armbrust, 1999; Mann et al., 1999; Armbrust and Galindo, 2001).

The choice of the diatom species used in the development of genomic studies was based on several criteria including ecological importance, the capacity of the organism for biomineralization, the ease with which the organism can be manipulated at genetic and molecular levels, and the estimated size of the genome; there is an obvious bias toward sequencing small genomes. There is little information on the sizes of diatom genomes, with most of it coming from the studies of Veldhuis et al. (1997), which estimate the genome sizes of seven diatom species by staining the DNA in the cells with PicoGreen or SYTOX Green and monitoring fluorescence of the individual cells using flow cytometry. The sizes of the genomes varied from 34 to approximately 700 Mb. Ultimately, the centric diatom *T. pseudonana* and the pennate diatom *P. tricornutum* were considered to be the most appropriate for generating genomic information. *T. pseudonana*, a silicified diatom, represents a species with a small genome (estimated by Veldhuis et al. to be 34 Mb) in which other members of the group are ubiquitous and ecologically important; *Thalassiosira weissflogii* appears to be much more ecologically relevant than *T. pseudonana*, but the former was found to have a genome that is approximately 20 times larger than that of the latter. The *T. pseudonana* strain that was sequenced, CCMP 1335, was collected from Moriches Bay (Long Island, NY) in 1958 and is available from the Center for Culture of Marine Phytoplankton (http://ccmp.bigelow.org/). The physiological knowledgebase for *T. pseudonana* is not well developed, and most molecular tools (e.g. transformation) have not been tested with this organism. The sequence of the *T. pseudonana* nuclear genome has been completed (Armbrust et al., 2004) by the Joint Genome Institutes (JGI; http://genome.jgi-psf.org/thaps1/thaps1.home.html), with many cDNA sequences to help identify coding regions of genes. The genome size, based on sequence analyses, was found to be very close to the fluorescence-based size estimate (approximately 34 Mb), and, from an optical map (Jing et al., 1998), the genome was determined to consist of 24 chromosomes ranging in size from 0.66 to 3.32 Mb. The nucleotide sequencing of the genome predicts at least 11,242 protein coding genes and that the organism contains a number of metabolic pathways associated with heterotrophic growth. The genome, among the smallest diatom genomes known (the genome of *P. tricornutum* is smaller), has few repeat elements, and much of the interspersed repeats represent remnants of transposable elements.

There are numerous areas of biology for which genetic and genomic analyses of diatoms would be extremely valuable. One of the major areas of interest over the last decade concerns cell wall or frustule formation. Frustules are silicified cell walls of the diatoms in which the deposition of the silica creates a precise, nano-scale pattern; these structures have the potential for exploitation as substrates for nanotechnology development. Furthermore, researchers are just beginning to gain an understanding of the transport of silicic acid into diatom cells (Hildebrand et al., 1997, 1998; Hildebrand and Wetherbee, 2003); there is little understanding of the intracellular movement of silica and the processes involved in the assembly of this compound into a precisely patterned frustule. Analyses of cell wall biogenesis and the ability to manipulate cell wall structure may provoke the development of new strategies for silicon-based fabrication technology. From a biological perspective, understanding the synthesis of wall components and how they are put together will enhance our knowledge of factors that modulate the assembly of an extracellular matrix, the ways in which this matrix is patterned, the role of patterning in biological function, and the means for modifying biological patterns. It has been known for quite a while that silica polymerization in diatoms occurs in the silica deposition vesicle, a specialized compartment within the cell delimited by a membrane called the silicalemma (Reimann et al., 1966; Crawford and Schmid, 1986). Cytoskeletal components such as microtubules and actin function in silicification; the former is involved in positioning the site at which silicification is initiated and may also influence valve morphology (Pickett-Heaps and Kowalski, 1981; Pickett-Heaps, 1983). The recently characterized polyanionic phosphoproteins of the cell wall, the silaffins (Kröger et al., 1999, 2000, 2002; Poulsen et al., 2003; Poulsen and Kröger, 2004), are associated with silica deposition and cell wall patterning processes; there are five silaffin encoding genes on the *T. pseudonana* genome. Other components of the cell wall that appear to function in silica polymerization are linear, long-chain polyamines (Kröger et al., 2000). A number of copies of genes thought to be involved in the synthesis of spermine and spermidine, which are likely intermediates in the biosynthesis of long chain polyamines, have also been identified on the *T. pseudonana* genome. Another family of genes associated with cell wall structure encodes the frustulins, wall glycoproteins that may be important for wall biogenesis but not specifically for the assembly of the silica building blocks (Vrieling et al., 1999). Interestingly, the diatoms appear to have a complete urea cycle, which probably occurs in mitochondria, and they can use urea as a sole nitrogen source. Ornithine, an intermediate in this cycle, is a precursor of the metabolites

spermine and spermidine (Morgan, 1999; Igarashi and Kashiwagi, 2000). The urea cycle may also serve in the generation of creatine phosphate, a high energy molecule that can drive certain cellular processes.

There are a number of other areas that will be interesting to explore with respect to sequence analyses of the diatom genome. These include the ways in which diatoms position themselves in the water column, the function and evolution of light-harvesting components (Buchel, 2003; Oeltjen et al., 2004), the mechanisms associated with nonphotochemical quenching of excess absorbed light energy (Lohr and Wilhelm, 1999; Lavaud et al., 2002, 2003), carbon metabolism and the potential role of the C4 pathway in CO_2 fixation (Reinfelder et al., 2004), the biosynthesis of long chain polyunsaturated fatty acids (Lebeau and Robert, 2003; Wen and Chen, 2003; Tonon et al., 2004), the role of Ca^{2+} in signaling in cellular processes (Falciatore et al., 2000), the identification and functional analyses of photoreceptors, the development of different cell morphotypes, and the control of morphogenesis.

Some diatoms, including those in the Thalassiosira genera, can control their position in the water column, which can influence light and nutrient availability, via extrusion of chitin fibers through frustule pores (Round et al., 1990). There are numerous genes encoding enzymes involved in the biosynthesis and degradation of chitin that may help regulate the dynamics of chitin extrusion and help the organism modulate its position in the environment.

The PSBS protein, a member of the extended LHC protein family, is critical for xanthophyll cycle-mediated energy dissipation in plants (Li et al., 2000; Peterson and Havir, 2001; Aspinall-O'Dea et al., 2002). The diatoms also have a xanthophyll cycle that is thought to be involved in the dissipation of excess absorbed light energy (Lohr and Wilhelm, 1999; Lavaud et al., 2002). Interestingly, no gene encoding PSBS has been identified on the genome of *T. pseudonana*, although such a gene has recently been discovered in the genome sequence database of *C. reinhardtii* (Gutman and Niyogi, 2004). It will be important to determine if there is a protein that is functionally analogous to PSBS and which diatom proteins are important for xanthophyll-dependent energy dissipation. Furthermore, *T. pseudanana* has no identified genes encoding LHC-like, stress-associated ELIPs and SEPs, although there are two genes encoding the related HLIPs.

Several other findings concerning genes present (or absent) in the *T. pseudonana* genome are interesting to note. While many of the enzymes involved in C4 metabolism are present on the *T. pseudonana* genome, an enzyme that would decarboxylate C4 acids in the plastid to generate the CO_2 substrate for ribulose 1,5-bisphosphate carboxylase was not identified; this is intriguing since the C4 pathway appears important for the fixation of inorganic carbon in *T. weissflogii* (Reinfelder et al., 2004). Whether the gene encoding the decarboxylating enzyme was just missed in the analyses of the genome or whether a novel (or highly diverged) enzyme functions in this capacity remains to be established. Also, a high proportion of the fatty acids synthesized by *T. pseudonana* are the commercially valuable long chain polyunsaturated fatty acids eicosapentaenoic and docosahexaenoic acids. The genes involved in their biosynthesis have been identified on the genome. With respect to photoreceptors, genes encoding members of both the phytochrome and cryptochrome families have been identified on the *T. pseudonana* genome, although there do not appear to be genes encoding the phototropin or rhodopsin photoreceptors. Currently, there needs to be much more extensive analyses of the *T. pseudonana* genome and efforts to link the genomic information with physiological/ecological processes.

It has recently been announced that JGI will sequence the full genome of *P. tricornutum*. Completion of this sequence will allow a comparison between centric and pennate species and may also help clarify the genetic basis of morphotype differentiation. Like *T. pseudonana*, *P. tricornutum* is also not considered to be very ecologically important (it is considered an atypical diatom), but a number of molecular tools including reporter genes, selectable markers, and a transformation system have been well developed for this organism. Furthermore, its genome is very small (approximately 20 Mb), and there is an abundant literature on the morphology, physiology, and ecology of this organism. There is also a relatively large-scale expressed sequence tag project and a queryable database (http://avesthagen.sznbowler.com/chris/bowler/WEB/FRAMESET/frameset.php) that is helping in the analysis of the genomic sequences.

C. reinhardtii

Genetic, molecular, physiological, and genomic features have made *C. reinhardtii*, a unicellular green alga, ideal for the elucidation of biological processes critical to both plants and animals. This organism has been used for numerous studies relating to photosynthetic processes as well as the biogenesis and function of the flagella. There are many recently developed tools and applications that are facilitating these biological studies. Plastid and nuclear genomes of *C. reinhardtii* are readily transformed (Debuchy et al., 1989; Kindle et al., 1989; Diener et al., 1990; Mayfield and Kindle, 1990; Shimogawara et al., 1998) using any of a number of different selectable markers (Debuchy et al., 1989; Fernandez et al., 1989; Kindle, 1990; Goldschmidt-Clermont, 1991; Nelson et al., 1994; Stevens et al., 1996; Lumbreras et al., 1998; Auchincloss et al., 1999; Kovar et al., 2002). Plasmid, cosmid, and bacterial artificial chromosome libraries (Purton and Rochaix, 1994; Zhang et al., 1994; Lefebvre and Silflow, 1999) are available for identification of genes that rescue specific *C. reinhardtii* (Funke et al., 1997; Randolph-Anderson

et al., 1998; Wykoff et al., 1998) or *E. coli* (Yildiz et al., 1996; Palombella and Dutcher, 1998) mutant strains. Methods have been developed for generating tagged mutant alleles (Tam and Lefebvre, 1993; Davies et al., 1994, 1996; Smith and Lefebvre, 1996; Koutoulis et al., 1997; Smith and Lefebvre, 1997; Zhang and Lefebvre, 1997; Asleson and Lefebvre, 1998; Davies et al., 1999; Wykoff et al., 1999), and alleles not tagged can be isolated by map-based cloning (Vysotskaia et al., 2001; Kathir et al., 2003). Gene function can be evaluated using antisense or RNAi suppression of gene activity (Schroda et al., 1999; Jeong et al., 2002; Sineshchekov et al., 2002; Wilson and Lefebvre, 2002), and reporter genes are available to elucidate sequences involved in controlling gene expression (Davies et al., 1992; Fuhrmann et al., 1999; Minko et al., 1999; Mayfield et al., 2003) and identifying specific regulatory factors (Davies et al., 1994; Quinn and Merchant, 1995; Jacobshagen et al., 1996; Ohresser et al., 1997; Villand et al., 1997; Fuhrmann et al., 2002; Komine et al., 2002). The chloroplast transformation system (Boynton et al., 1988; Newman et al., 1990) has made possible the inactivation of specific plastid genes and site-directed mutagenesis for evaluation of gene function (Whitelegge et al., 1992; Hong and Spreitzer, 1994; Takahashi et al., 1994; Hallahan et al., 1995; Webber et al., 1996; Zhu and Spreitzer, 1996; Fischer et al., 1997; Lardans et al., 1997; Larson et al., 1997; Melkozernov et al., 1997; Xiong et al., 1997; Finazzi et al., 1999; Higgs et al., 1999).

The use of *C. reinhardtii* to dissect photosynthesis and the functions of pigment-protein complexes is aided by the finding that this haploid alga can grow heterotrophically in the dark using acetate as the sole source of fixed carbon and that dark-grown cells maintain normal chloroplast structure and resume photosynthetic CO_2 fixation upon illumination. These features of *C. reinhardtii* have enabled researchers to isolate a broad range of mutants that adversely affect photosynthetic function (Harris, 1989, 2001). Indeed, using the genetic manipulations first elegantly demonstrated by Sager (1960), Levine and his colleagues began to delineate the pathway of photosynthetic electron transport and the regulation of the photosynthetic activity (Gorman and Levine, 1966; Bennoun and Levine, 1967; Givan and Levine, 1967; Lavorel and Levine, 1968; Levine, 1969; Levine and Goodenough, 1970; Moll and Levine, 1970; Sato et al., 1971). The identification of motility mutants and the biochemical characterization of flagella in such mutants have also made this organism ideal for dissecting flagella function. A number of polypeptides associated with flagella assembly or function are similar to proteins altered in diseased mammalian cells (Pazour et al., 2000; Pennarun et al., 2002; Li et al., 2004; Snell et al., 2004). Therefore, *C. reinhardtii* is serving as an important model system for elucidating the biology of both photosynthetic and nonphotosynthetic eukaryotes.

Over the last decade, global gene expression has been examined in a number of organisms, both mutant and wild-type strain, under a number of different environmental conditions using high density DNA microarrays. With the generation of both cDNA and genomic information (Dutcher, 2000; Grossman, 2000; Dent et al., 2001; Lilly et al., 2002; Simpson and Stern, 2002; Grossman et al., 2003; Shrager et al., 2003; http://genome.jgi-psf.org/chlre2/chlre2.home.html), DNA microarrays and macroarrays have been used to study biological processes in *C. reinhardtii* (Im et al., 2003; Miura et al., 2004; Yoshioka et al., 2004; Zhang et al., 2004). Furthermore, genome-wide and proteomic approaches are currently being used to understand the dynamics of the photosynthetic apparatus in response to nutrient conditions (Im et al., 2003; Zhang et al., 2004; Y. Wang, Z. Sun, M.H. Horken, C.S. Im, Y. Xiang, A.R. Grossman, and D.P. Weeks, unpublished data), light and circadian programs (Im et al., 2003; Wagner et al., 2004), composition of pigment protein complexes (Stauber et al., 2003; Elrad and Grossman, 2004), identification of components involved in iron assimilation (La Fontaine et al., 2002), and the polypeptide components of the flagella and basal body (Li et al., 2004).

There is a wealth of information contained within the genomic sequence of *C. reinhardtii*. The genome is approximately 110 Mb, with nearly 95 Mb of the sequence completed; but the sequence information is still dispersed over approximately 3,000 individual scaffolds. These scaffolds contain over 19,000 gene models (although some of the small scaffolds may ultimately be incorporated into the larger ones and some of the gene models will be lost), many of which are supported by expressed sequence tag data. Currently, intense sequence efforts by JGI are being focused on joining many of the scaffolds. Recently, analyses of the genomic information suggests that the genome contains a low level of nuclear plastid DNA segments, relative to the Arabidopsis or *Oryza sativa* genomes (Richly and Leister, 2004). Both the cDNA and genomic sequence information has helped elucidate the *LHC* gene family with respect to genes encoding both LHCB and LHCA polypeptides (for PSII and PSI, respectively), suggesting which of the family members are highly expressed, which of the encoded polypeptides are associated with the trimeric or monomeric light-harvesting complexes, and which may be posttranslationally modified (Elrad and Grossman, 2004). Also identified are genes for ELIPs and other LHC polypeptides that might be involved in the management of absorbed excitation energy (e.g. LI818), as well as PSBS (Gutman and Niyogi, 2004), which is involved in xanthophyll cycle-dependent quenching. Numerous genes involved in chromatin structure, nutrient (nitrogen, sulfur, phosphorus, and iron) acquisition and assimilation, and carbon metabolism have also been identified. For example, there are at least six genes encoding Na^+/Pi symporters and another four genes encoding H^+/Pi cotransporters; these transporters are likely involved in the delivery of phosphate to various compartments of the

cell (J. Moseley, C.W. Chang, and A.R. Grossman, unpublished data). The genome/cDNA data has also led to the identification of genes associated with the copper-dependent iron uptake pathway that was first defined in *Saccharomyces cerevisiae*, which includes FOX1 (a multicopper ferroxidase), FTR1 (an iron permease), ATX1 (a copper chaperone), and a copper-transporting ATPase. All of these genes have coordinated induction when the cells are experiencing iron deprivation. The FOX1 mRNA was also regulated by copper availability at the posttranscriptional level (La Fontaine et al., 2002). The results clearly demonstrate a role for copper in the assimilation of iron in a photosynthetic organism, although copper deficient *C. reinhardtii* does not show signs of iron deficiency (probably because the organism also has a copper-independent system). Interestingly, the FOXA component of the iron assimilation system is most similar to mammalian hephaestin and ceruloplasmin proteins. The genes encoding the major transition metal transporters of *C. reinhardtii* have also been identified and characterized (Rosakis and Koster, 2004).

Many genes of the *C. reinhardtii* genome encode proteins that are similar to those of animal cells. A comparison of the *C. reinhardtii* gene models generated by JGI with proteins encoded by the human genome generated 4,348 matches (based on a match cutoff E value of 10^{-10}; Li et al., 2004). Dutcher and colleagues have been interested in genes required for the function and biogenesis of the basal body and flagella; the basal body can be converted to centrioles and is essential for cilia assembly in animals and for flagella biogenesis in *C. reinhardtii*. Of the 4,348 matched proteins encoded on the *C. reinhardtii* and human genomes, there was a subset of 688 that did not match any predicted proteins encoded on the genome of Arabidopsis. Since Arabidopsis does not have either basal bodies or flagella/cilia, it was hypothesized that many of the 150 and 250 proteins required for basal body and flagella formation/function (Dutcher, 1995a, 1995b), respectively, would be present in this subset. Indeed, this pool of genes did encode a number of known flagellar and basal body polypeptides and also contained genes associated with human diseases resulting from impairment of cilia or basal body function. For example, there were six genes (*BBS*1, 2, 4, 5, 7, and 8) associated with Bardet-Biedl syndrome, a human disease characterized by retinal dystrophy, obesity, polydactyly, renal and genital malformation, and learning disabilities. Suppression of synthesis of the BBS5 protein in *C. reinhardtii* using RNAi technology resulted in strains completely or partially lacking flagella and that exhibited a weak cleavage furrow defect, supporting a role for the *BBS5* gene product in flagellar and basal body function and assembly. Hence, *C. reinhardtii* can be exploited as a relatively simple genetic/molecular system that can help researchers gain significant insights into mechanistic aspects of human diseases centriole cilia associated with defects in centriole and cilia function and assembly.

C. reinhardtii genomic information is also being coupled to technologies for gene expression analyses and proteomic studies. The generation and use of a partial genome microarray (close to 3,000 distinct array elements; a second generation array containing approximately 10,000 array elements is currently under construction) has demonstrated that nutrient stress leads to the up-regulation of many of the genes encoding enzymes involved in nutrient assimilation, but also leads to increased levels of transcripts for genes involved in stress responses (Zhang et al., 2004; J. Moseley, C.W. Chang, and A.R. Grossman, unpublished data), and that excess light causes elevated expression of genes encoding proteins with antioxidant activities (Ledford et al., 2004). Both microarrays and macroarrays have also been used to define specific sets of genes that are controlled during nutrient stress by specific regulatory elements including CCM1 or CIA5 (Miura et al., 2004; Y. Wang, Z. Sun, M.H. Horken, C.S. Im, Y. Xiang, A.R. Grossman, and Weeks DP, unpublished data), SAC1 (Zhang et al., 2004), and PSR1 (J. Moseley, C.W. Chang, and A.R. Grossman, unpublished data). Nearly all of the genes that are induced under low CO_2 conditions and encode components of the carbon concentrating mechanism are controlled by CCM1/CIA5 (Miura et al., 2004), while SAC1 appears to control both sulfur assimilation genes as well as a subset of genes for proteins associated with oxidative stress and restructuring the photosynthetic apparatus (Zhang et al., 2004). An inability of the cells to acclimate to sulfur deprivation (in the *sac*1 mutant) leads to very high levels of certain stress-associated transcripts, including two small chaperones that may be located in the chloroplast. Furthermore, a mutant defective in the generation of photoprotective carotenoids (the *npq1 lor1* mutant) exhibits a complex response in which some genes associated with oxidative stress responses that are not activated in the wild-type strains become active in the mutant (Ledford et al., 2004). A number of researchers are also beginning to use the genomic information as a foundation for proteomic approaches. For example, Hippler and colleagues (Stauber et al., 2003) correlated polypeptides of LHC resolved by two-dimensional gel electrophoresis with specific *LHC* genes and also demonstrated specific N-terminal processing of the LHCBM3 and LHCBM6 polypeptides. A proteomic approach coupled with the use of genomic information allowed for the identification of specific proteins under circadian control, including proteins that might be part of a complex that binds to RNA (Wagner et al., 2004; Zhao et al., 2004).

OTHER ALGAE

Full-genome sequences are being generated for the multicellular green alga *V. carteri* (evolutionarily close to *C. reinhardtii*), the diatom *P. tricornutum*, and the picoeukaryote *O. tauri*. The *O. tauri* genome sequence is

nearly finished and recent biochemical and molecular work on this organism has been initiated (Fouilland et al., 2004; Guillou et al., 2004; Khadaroo et al., 2004; Meyer et al., 2004; Ral et al., 2004). This picoeukaryotic, photosynthetic organism has 18 chromosomes with a genome size of approximately 11.5 Mb. A 7-fold sequence coverage of the genome has been completed, and 4,000 open reading frames on the genome were annotated; this information is not currently available to the public (http://www.blackwellpublishing.com/febsabstracts2004/abstract.asp?id=17225). A number of algae are also being used for the generation of cDNA sequence information. Recently, a cDNA library was constructed for the dinoflagellate *Alexandrium tamarense* and 3,628 unique cDNAs identified (see http://genome.uiowa.edu/projects/dinoflagellate/). cDNA libraries have also been made for the dinoflagellates *Lingulodinium polyedrum* and *Amphidinium carterae* (Bachvaroff et al., 2004). Interestingly, many genes normally found on the plastid DNA in photosynthetic organisms have moved into the dinoflagellate nuclear genome (Bachvaroff et al., 2004; Hackett et al., 2004). Analyses of cDNA libraries of *Porphyra yezeonsis* (Nikaido et al., 2000; http://www.kazusa.or.jp/en/plant/porphyra/EST/), *P. tricornutum* (Scala et al., 2002), and *Laminaria digitata* (Crepineau et al., 2000) have also begun.

Molecular and genomic analyses of other algae are in the planning stages or are just beginning, and groups of researchers are developing ecological, evolutionary, physiological, genetic, and economic criteria to identify those systems that should be given priority for sequence analysis. It is especially important to develop a diverse set of systems, representing algae in different phylogenetic groupings that exhibit both unique and important biological characteristics. Criteria being used to decide upon those algae that should be targeted for genomic studies, and some of the top algal candidates, are summarized below.

Criteria to Consider for Selection of Organisms

Genomics has moved in many directions over the past several years and has advanced from the sequencing of individual genomes to the generation of metagenomic information in which DNA isolated from environmental samples is randomly sequenced. While this new direction is valuable with respect to gene discovery and has already begun to reveal biological processes potentially important in specific environments (Beja et al., 2001; de la Torre et al., 2003; Venter et al., 2004), it is still full-genome sequence information for a particular organism that will offer scientists a more complete vision of the genetic potential of an organism and foster the development of informed experimentation. A number of diverse criteria are being used to select algae for genomic studies. In a general sense, the issues will center on how important the organisms are from a biological and economic perspective, how easy it is to grow the organism in laboratory cultures, the potential for exploiting acquired genomic information based on previous ecological, physiological, biochemical, and molecular knowledge, and the extent to which sophisticated analytical tools have been developed for each of these areas. From a practical perspective, the size and repeat content of the genome needs to be considered since a larger genome with extensive repeat structures will be difficult to sequence and assemble. Of course, no algae will satisfy all of the issues raised, there will be disagreement as to the relative importance of some of the features when deciding on subject organisms, and dominant personalities with strong biases are likely to influence the direction of the field. Recently, Waaland et al. (Waaland et al., 2004) have reviewed some of these issues, especially with respect to macrophytic marine algae. My perspectives on these issues are briefly summarized below.

Growth of Organism as Axenic or Unialgal Culture on Defined Medium

Some algae are not readily cultured and may require environmental conditions and temperatures that are difficult to maintain in the laboratory; this is especially true of some of the large macrophytic algae. It is important to be able to grow the organism on defined medium to study various aspects of metabolism and acclimation, which can be strongly influenced by the composition of the medium.

Defined Sexual Life Cycle That Can Be Controlled

Many marine algae have complex life histories and sometimes the sexual cycle, even if known, is difficult to control (this is the case for many diatoms). The life histories themselves are intriguing from a developmental perspective, with many macrophytic organisms alternating between morphologically distinct gametophyte and sporophyte phases (some of the organisms have triphasic life cycles). Furthermore, the occurrence of individuals of separate sexes allows for the engineering of specific crosses that can unveil mutant phenotypes, generate strains with multiple mutations, and ultimately allow for the map-based cloning of mutant alleles.

Generate Mutants

The generation and analysis of mutants can lead to a broad understanding of biological processes and help identify associated protein factors. The mutant phenotypes can be most successfully used in organisms amenable to controlled genetic crosses. While genetic tools have been instrumental in the development of *C. reinhardtii* as a model organism, the sophisticated use of genetics will be more of a challenge with many of the macrophytic algae and even with the

unicellular diploid organisms for which no or poorly developed genetic systems have been established.

Uninucleate Cells

Many algae are multinucleate, including a number of the green algae that lie within the evolutionary lineage that evolved into land plants. It is likely to be easier to transform and segregate lesions in uninucleate forms.

Prior Knowledge

A strong knowledge base with respect to biological and molecular aspects of an organism will have a major impact on the exploitation of genomic information. Prior knowledge provides information about protein and gene sequences, physiological processes, and the conditions under which the organism thrives. As with any biological problem, a body of prior knowledge greatly enhances the value of later exploration; therefore, it would be advantageous to secure genomic information for candidate taxa supported by an extensive history of research.

Evolutionary Interest and Fossil Record

Working with genera that have a large number of different species would propel the genomic work into comparative analyses and yield insights into the evolution of a genus and the factors that led to its diversification. It is also critical to focus on distinct species positioned at important evolutionary branchpoints. Finally, having a fossil record of the genus will help calibrate the evolution of specific features (showing their earliest occurrence) that characterize that genus.

Ecological Importance

Vital information for managing the environment will depend on orienting genomic studies toward algae that are dominant components of important aquatic and terrestrial communities and that occupy critical ecological niches. Genomic information may help establish an understanding of the reasons that these organisms thrive in their respective communities and provide insights about the ways in which they interact with their biotic and abiotic environments.

Economic Importance

Several algae serve as a source of food or are used for the production of compounds that have economic value. A number of macroalgae synthesize commercially important polysaccharides while several microalgae synthesize high levels of long chain polyunsaturated fatty acids or pigment molecules; the uses for these compounds were discussed in the introductory section. Enzymes that are key components of the biosynthetic pathways that function in the synthesis of some of the unique compounds produced by the algae could have bioengineering applications.

Genome Size and Repeat Structure

Most sequencing facilities are trying to find genomes that aren't too large and that don't have a high repeat content. Many of the dinoflagellate genomes are very large, and there is currently no effort that I know of to generate a complete dinoflagellate genome sequence. For ecologically important organisms with large genomes, such as the dinoflagellates, it might be best to first generate cDNA information or to use technologies that enrich for expressed regions of the genome (Mayer and Mewes, 2002; Whitelaw et al., 2003).

Establishment of a Well-Organized Community

This is a more practical issue. It takes a lot of work and infrastructure to develop a strong genomic project. This means that the community, which often doesn't have working experience with either the technologies used for sequencing the genome and examining global gene expression or the informatic tools that are used for assembly and the analysis of the sequence information, has to establish links with sequencing centers and recruit experts that would help organize and mine the data. Such a project is time consuming and requires a concerted effort from a group of committed individuals and is often aided by strong interactions with program managers at the granting agencies.

A Brief View of Specific Organisms That Might Be Considered for Algal Genomics

There are many algae that can be included on a wish list of genomes to be sequenced. In my opinion, there are a number of unicellular organisms for which genomic information would unveil mechanistic aspects of many processes including the establishment of the chloroplast and chloroplast genomes through endosymbiotic associations, nutrient cycling and the deposition of carbonates in marine environments, and the partnering of photosynthetic and heterotrophic organisms in symbiotic associations and how that reflects specific ecological conditions. One organism to consider for genomic studies is *C. paradoxa* (a representative glaucophyte), for which extensive genetic/genomic information might help elucidate events leading to the establishment and evolution of plastids (Delwiche and Palmer, 1997). It is important to generate sequence information from a haptophyte such as *Emiliania huxleyi*, and indeed JGI has initiated a genome project with this organism, which has an estimated genome size of 220 Mb (http://www.jgi.doe.gov/sequencing/seqplans.html). This extremely abundant coccolithophore synthesizes plates of calcium carbonate (that can be shed from the surface of

the organism) and forms blooms that reduce the capture of light/heat by the oceans by reflecting it back into the atmosphere and can also impact CO_2 levels in the atmosphere. It will also be important to generate more genetic/genomic information for dinoflagellates of the *Symbiodinium* species. These organisms serve as common endosymbionts that populate various heterotrophic hosts, including corals and sea anemones, providing them with fixed carbon in environments that may be severely limited for that resource.

It will also be important to develop genomic information for additional green algal species. Green algae form the eukaryotic base of the evolutionary tree for vascular plants. Like plants, these organisms perform oxygenic photosynthesis using chlorophyll *a* and *b* as the pigments of their major light-harvesting complexes. They exist as different mating types, show cell polarity, have central vacuoles that confer turgor to the cells, exhibit phototropic responses and circadian rhythms, and even produce some hormones that are synthesized by plants. While *C. reinhardtii* has been the primary green algal system developed (as discussed above) and genome sequence of the primitive green alga *O. tauri* has been completed, other systems being explored are Volvox, Acetabularia, Caulerpa, and Chara/Nitella. Molecular and genomic examination of the close relationship between *C. reinhardtii* and *V. carteri* may provide insights into the evolution of multicellular photosynthetic organisms. In contrast, the unicellular, algae Caulerpa and Acetabularia have siphonous body plans that have a superficial, morphological similarity to that of vascular plants. Acetablularia has been used for grafting experiments and experiments that exploit the ease with which the cell can be enucleated (it has a single giant nucleus until reproduction). It is an organism that can readily be used to study those transcripts generated in the nucleus (which is located in the cell rhizoid) and transmitted/accumulated in more distal locations of the cell where they control biological processes. Developmental studies concerning Acetabularia have recently been discussed (Mandoli, 1998), and a comparison of cDNAs from the juvenile and adult stage has been initiated (Henry et al., 2004). An equally interesting organism is Caulerpa. The root and leaf-like structures of this organism are supported by a stem-like structure, and even though it is a single cell, it may extend for well over a meter (Graham and Wilcox, 2000). This alga does not produce specialized reproductive cells (unlike Acetabularia, which produces a reproductive cap-like structure; see Mandoli, 1998) but directs all of its vegetative resources into the generation of progeny. Furthermore, there are *Caulerpa* species that respond to gravity like flowering plants and that synthesize and respond to plant hormones. These single-celled organisms also exhibit tip growth, even though they lack a meristem, have distinct juvenile and adult developmental phases, make an elaborate thallus structure, are amenable to grafting experiments, and synthesize secondary metabolites (especially the sesquiterpenoids) that can act as neurotoxins (Brunelli et al., 2000; Mozzachiodi et al., 2001). There are many developmental mutants in these organisms, and there is a well-preserved fossil record of related organisms dating back 570 million years (http://ifaa.port5.com/index.html).

The charophycean algae such as Chara (clade Charales) are evolutionarily close to vascular plants based on morphological, developmental, and molecular features. Like plants, they have apical cell division, generate branching filaments with nodal and internodal structures, exhibit asymmetric cell division in which the plane of division is controlled, synthesize a phragmoplast during cell division, make a cellulosic cell wall, and develop both plasmodesmata and specific reproductive organs in which sexual cells are encapsulated and protected by vegetative cells (Graham and Wilcox, 2000). Coleochaete (clade Coleochaetales) has many characteristics similar to those of Chara and features that are also important for growth in terrestrial environments (Kranz et al., 1995; Petersen et al., 2003). They are frequently present in shallow waters where they may be exposed to desiccation conditions. They have evolved a number of features that allow them to conserve water, including the accumulation of ridge-shaped mucilage depositions with ultrastructural similarities to a cuticle (which help plants retain water). The thylakoid membranes of *Coleochaete* species are arranged in grana, similar to the grana observed in land plants. Furthermore, some *Coleochaete* species form egg cells on the thallus, and these egg cells are encased in vegetative tissue that provides the embryo with nutrients. Finally, *Mesostigma viride* is a primitive, scaly alga that occupies the earliest branch in the green algal lineage that led to the evolution of the charophyceans and land plants (Karol et al., 2001). The basal position of this organism with respect to the evolution of the charophycean algae is supported by sequence data from both mitochondrial and chloroplast genomes. Obtaining nuclear genomic information from *Mesotigma, Coleochaete*, and *Chara* or *Nitella* species will capture important chapters in the evolution of land plants.

Another green alga that would be appropriate for genomic analyses is Ulva, which now includes the genus Enteromorpha (Hayden et al., 2003). The ulvophycean algae form a blade that can reach one meter in length, but that consists of only two cell layers, or tubes with one layer of cells. These algae are widespread, used as a source of food, can form ''green tides,'' grow rapidly in culture and are represented by axenic lines, have a well-characterized life history, and have been used for both genetic and mutant studies (Fjeld and Lovle, 1976). The ulvophyceaens represent an ancient algal lineage (Caulerpa and Acetablularia are also in this class) with fossil records for some calcified members of the group, although Ulva is not among the calcified taxa, that date from 0.7 to 0.8 billion years ago (Wray, 1977; Butterfield et al., 1988).

The brown algae and red algae are important algal lineages to be considered for genomic analyses. Many of these algae form large forests or kelp beds that populate coastal regions, while others carpet the rocky coasts. The Laminariales, commonly referred to as kelps, are physically the largest seaweeds and represent an economically and ecologically important group of organisms that are found in temperate waters throughout the world. The life cycles of many of the kelps are well characterized and can be controlled by environmental factors, and some have been used for significant molecular analyses (Billot et al., 1998; Crepineau et al., 2000; De Martino et al., 2000; Yoon et al., 2001). However, it is costly to culture these organisms in the laboratory since the sporophytes range from 1 to more than 50 m.

The Fucales or rockweeds are another brown algal group that is ecologically important. These algae have a well-characterized life history and have been used for studies concerning cell and tissue polarity (Quatrano et al., 1991; Shaw and Quatrano, 1996; Brownlee and Bouget, 1998), but it is difficult to grow fronds to maturity in the laboratory. Another potential brown algal candidate (perhaps the most reasonable to consider) for genomic studies is Ectocarpus, which is relatively easy to maintain in the laboratory and has been used for numerous physiological studies, including the characterization of pheromones (Muller and Schmid, 1988). A lysogenic virus of the family Phycodnaviridae associated with Ectocarpus has been identified, and recently the DNA of the viral genome was sequenced (Delaroque et al., 1999, 2001; Van Etten et al., 2002). It might be possible to engineer such a virus for the facilitation of gene transfer into this alga.

The red algae represent an economically, ecologically, and evolutionarily important and diverse group of organisms. They are widely distributed in the marine environment and occupy intertidal habitats where they may experience desiccation and exposure to excess excitation energy and deep ocean habitats where they may receive almost no excitation energy (Littler et al., 1985). The evolutionary origin of red algae is not clear (Ragan and Gutell, 1995), but recent evidence suggests that they represent a sister group to the green plants (Moreira and Philippe, 2001) or are very closely related to a group of multicellular eukaryotes with complex patterns of ontogenetic and tissue-specific development (Stiller and Hall, 2002). The genus Porphyra serves as an important food product, while other members of this group are cultured for phycolloids and carageenan. Furthermore, carbonate skeletons from some red algal species that grow in coralline reefs of the tropical seas are mined as marl. These coralline red algae are important members of the reef communities and are represented by extensive fossil records.

A high priority for sequence analyses of a macrophytic alga, and one that was placed as the top priority by Waaland et al. (2004), is *P. yezoensis*. The farming of Porphyra is the basis of the multibillion dollar nori industry, and its economic importance has elicited numerous experimental studies. It has a well-defined life cycle with a gametophytic stage consisting of a blade that is two cell layers thick and a sporophytic phase represented by the tiny Conchocelis filament. Neither gametophytes nor sporophytes of Porphyra are very large, and both grow relatively rapidly in the laboratory, allowing for the generation of large amounts of biological material. In many instances, the gametophytes can reproduce asexually through the generation of single-celled spores, providing a genetically homogeneous source of biological material. Furthermore, protoplasts from *Porphyra* species can be readily obtained and regenerated into whole plants (Waaland et al., 1990), and these protoplasts can undergo fusion (Chen, 1992; Chen et al., 1995; Mizukami et al., 1995; Cheney, 1999) and possibly Agrobacterium-mediated transformation (Cheney et al., 2001). Genetic studies on Porphyra mutants have demonstrated that the blades are genetic chimeras (Mitman and van der Meer, 1994; Yan et al., 2000).

A significant body of molecular work has been generated with Porphyra. The complete nucleotide sequence of the mitochondrial and chloroplast genomes of *Porphyra purpurea* have been reported (Reith and Munholland, 1995; Burger et al., 1999), and extensive cDNA analysis for *P. yezoensis* has been initiated (Nikaido et al., 2000; http://www.kazusa.or.jp/en/plant/porphyra/EST/). The size of the Porphyra genome has been estimated to be 300 Mb (Kapraun et al., 1991), which is only twice the size of the genome of Arabidopsis. The genus has fossil records that extend back over 500 million years (Campbell, 1980; Xiao et al., 1998), while Bangia, a sibling genus of Porphyra, is represented by fossils that are 1.2 billion years old (Butterfield, 2000). There are certainly many arguments that place *P. yezoensis* at or near the top of the list of macroalgae to be analyzed at the genomic level.

CONCLUDING REMARKS

Morphological, physiological, and molecular aspects of many of the algal groups are fascinating and often markedly different from those of land plants. The diversity of algal form and function is important when considering specific members of this group of organisms for future genomic studies. There are only a small number of algae for which genomic studies have been initiated, and those that are being examined with genomic tools represent neither the ecologically nor economically most important algal species. Furthermore, while this article provides some information and opinions concerning algae that would be most useful for future genomic studies, there are numerous other algae with fascinating physiological, ecological, and evolutionary characteristics that could be included as candidate organisms for genomic studies.

Some of the greatest deficiencies in our knowledge of the algae concern the marine organisms, organisms that hold critical information with respect to numerous physiological and evolutionary processes, the development and control of symbiosis, and the cycling of nutrients. And while the genomic sequences of many algae will not have immediate benefits with respect to the health and nutrition of humans (major considerations that guide much of the genomic work), they will in many cases be extremely relevant to the evolving physical and biological aspects of the Earth and give us a better appreciation of the role of the marine environment in the cycling of carbon and other nutrients and the influence of these cycles on the climate and the health of both aqueous and terrestrial ecosystems.

ACKNOWLEDGMENTS

I thank NSF for supporting genomic research using *Chlamydomonas reinhardtii* and awarding us grant MCB 0235878. I also thank Dan Rokhsar and Diego Martinez at the Joint Genome Institutes and members of the Consortium involved in developing the tools and infrastructure for securing and examining *C. reinhardtii* cDNA and genomic information and for providing stimulating discussions and valuable insights, and Bob Waaland for critically reading this manuscript. This is Carnegie Institution Publication number 1683.

Received September 14, 2004; returned for revision October 30, 2004; accepted November 12, 2004.

LITERATURE CITED

Apt KE, Kroth-Pancic PG, Grossman AR (1996) Stable nuclear transformation of the diatom *Phaeodactylum tricornutum.* Mol Gen Genet **252:** 572–579

Apt KE, Zaslavskaia LA, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR, Kroth PG (2002) In vivo characterization of diatom multipartite plastid targeting signals. J Cell Sci **115:** 4061–4069

Armbrust EV (1999) Identification of a new gene family expressed during the onset of sexual reproduction in the centric diatom *Thalassiosira weissflogii.* Appl Environ Microbiol **65:** 3121–3128

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, et al (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science **306:** 79–86

Armbrust EV, Chisholm SW (1990) Role of light and cell cycle on the induction of spermatogenesis in a centric diatom. J Phycol **26:** 470–478

Armbrust EV, Galindo HM (2001) Rapid evolution of a sexual reproduction gene in centric diatoms of the genus *Thalassiosira.* Appl Environ Microbiol **67:** 3501–3513

Asleson CM, Lefebvre PA (1998) Genetic analysis of flagellar length control in *Chlamydomonas reinhardtii*: a new long-flagella locus and extragenic suppressor mutations. Genetics **148:** 693–702

Aspinall-O'Dea M, Wentworth M, Pascal A, Robert B, Ruban A, Horton P (2002) In vitro reconstitution of the activated zeaxanthin state associated with energy dissipation in plants. Proc Natl Acad Sci USA **99:** 16331–16335

Auchincloss AH, Loroch AI, Rochaix JD (1999) The arginosuccinate lyase gene of *Chlamydomonas reinhardtii*: cloning of the cDNA and its characterization as a selectable shuttle marker. Mol Gen Genet **261:** 21–30

Bachvaroff TR, Concepcion GT, Rogers CR, Herman EM, Delwiche CF (2004) Dinoflagellate expressed sequence tag data indicate massive transfer of chloroplast genes to the nuclear genome. Protist **155:** 65–78

Beja O, Spudich EN, Spudich JL, Leclerc M, DeLong EF (2001) Proteorhodopsin phototrophy in the ocean. Nature **411:** 786–789

Bennoun P, Levine RP (1967) Detecting mutants that have impaired photosynthesis by their increased level of fluorescence. Plant Physiol **42:** 1284–1287

Berteau O, Mulloy B (2003) Sulfated fucans, fresh perspectives: structures, functions, and biological properties of sulfated fucans and an overview of enzymes active toward this class of polysaccharide. Glycobiology **13:** 29R–40R

Biegala IC, Not F, Vaulot D, Simon N (2003) Quantitative assessment of picoeukaryotes in the natural environment by using taxon-specific oligonucleotide probes in association with tyramide signal amplification-fluorescence in situ hybridization and flow cytometry. Appl Environ Microbiol **69:** 5519–5529

Billot C, Rousvoal S, Estoup A, Epplen JT, Saumitou-Laprade P, Valero M, Kloareg B (1998) Isolation and characterization of microsatellite markers in the nuclear genome of the brown alga *Laminaria digitata* (Phaeophyceae). Mol Ecol **7:** 1778–1780

Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, Randolph-Anderson BL, Robertson D, Klein TM, Shark KB, et al (1988) Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. Science **240:** 1534–1538

Brownlee C, Bouget FY (1998) Polarity determination in *Fucus*: from zygote to multicellular embryo. Semin Cell Dev Biol **9:** 179–185

Brunelli M, Garcia-Gil M, Mozzachiodi R, Roberto M, Scuri R, Traina G, Zaccardi ML (2000) Neurotoxic effects of caulerpenyne. Prog Neuropsychopharmacol Biol Psychiatry **24:** 939–954

Buchel C (2003) Fucoxanthin-chlorophyll proteins in diatoms: 18 and 19 kDa subunits assemble into different oligomeric states. Biochemistry **42:** 13027–13034

Burger G, Saint-Louis D, Gray MW, Lang BF (1999) Complete sequence of the mitochondrial DNA of the red alga *Porphyra purpurea*: cyanobacterial introns and shared ancestry of red and green algae. Plant Cell **11:** 1675–1694

Butterfield NJ (2000) *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology **26:** 386–404

Butterfield NJ, Knoll AH, Swett K (1988) Exceptional preservation of fossils in the Upper Proterozoic. Nature **334:** 424–427

Campbell SE (1980) *Paleoconchocelis starmachii*, a carbonate boring microfossil from the upper Silurian of Poland (425 million years old): implications for the evolution of the Bangiaceae (Rhodophyta). Phycologia **19:** 25–36

Cavalier-Smith T (2000) Membrane heredity and early chloroplast evolution. Trends Plant Sci **5:** 174–182

Chamberlain JG (1996) The possible role of long-chain, omega-3 fatty acids in human brain phylogeny. Perspect Biol Med **39:** 436–445

Chen C (1992) Electrofusion of protoplasts from *Porphyra haitanensis* and *P. yezoensis* thalli (Rhodophyta). Chin J Biotechnol **8:** 33–39

Chen LCM, McCracken IR, Xie ZK (1995) Electrofusion of protoplasts of two species of *Porphyra* (Rhodophyta). Bot Mar **38:** 335–338

Cheney D, Metz B, Stiller J (2001) Agrobacterium-mediated genetic transformation in the macroscopic marine red alga *Porphyra yezoensis.* J Phycol (Suppl) **37:** 11

Cheney DP (1999) Strain improvement of seaweeds through genetic manipulation: current status. World Aquaculture **30:** 55–56

Chu KH, Qi J, Yu ZG, Anh V (2004) Origin and phylogeny of chloroplasts revealed by a simple correlation analysis of complete genomes. Mol Biol Evol **21:** 200–206

Ciniglia C, Yoon HS, Pollio A, Pinto G, Bhattacharya D (2004) Hidden biodiversity of the extremophilic Cyanidiales red algae. Mol Ecol **13:** 1827–1838

Coles SL, Brown BE (2003) Coral bleaching: capacity for acclimatization and adaptation. Adv Mar Biol **46:** 183–223

Crawford RM, Schmid A-MM (1986) Ultrastructure of silica deposition in diatoms. *In* BS Leadbeater, R Riding, eds, Biomineralization in Lower Plants and Animals, Vol 30. The Systematics Association, Oxford University Press, New York

Crepineau F, Roscoe T, Kaas R, Kloareg B, Boyen C (2000) Characterisation of complementary DNAs from the expressed sequence tag analysis of life cycle stages of *Laminaria digitata* (Phaeophyceae). Plant Mol Biol **43:** 503–513

Davies J, Yildiz F, Grossman AR (1996) Sac1, a putative regulator that is critical for survival of *Chlamydomonas reinhardtii* during sulfur deprivation. EMBO J **15:** 2150–2159

Davies JP, Weeks DP, Grossman AR (1992) Expression of the arylsulfatase

gene from the β_2-tubulin promoter in *Chlamydomonas reinhardtii*. Nucleic Acids Res **20:** 2959–2965

Davies JP, Yildiz F, Grossman AR (1994) Mutants of *Chlamydomonas reinhardtii* with aberrant responses to sulfur deprivation. Plant Cell **6:** 53–63

Davies JP, Yildiz FH, Grossman AR (1999) Sac3, an Snf1-like serine/threonine kinase that positively and negatively regulates the responses of *Chlamydomonas* to sulfur limitation. Plant Cell **11:** 1179–1190

de la Torre JR, Christianson LM, Beja O, Suzuki MT, Karl DM, Heidelberg J, DeLong EF (2003) Proteorhodopsin genes are distributed among divergent marine bacterial taxa. Proc Natl Acad Sci USA **100:** 12830–12835

De Martino A, Douady D, Quinet-Szely M, Rousseau B, Crepineau F, Apt K, Caron L (2000) The light-harvesting antenna of brown algae: highly homologous proteins encoded by a multigene family. Eur J Biochem **267:** 5540–5549

Debuchy R, Purton S, Rochaix JD (1989) The argininosuccinate lyase gene of *Chlamydomonas reinhardtii*: an important tool for nuclear transformation and for correlating the genetic and molecular maps of the *ARG7* locus. EMBO J **8:** 2803–2809

Delaroque N, Maier I, Knippers R, Muller DG (1999) Persistent virus integration into the genome of its algal host, *Ectocarpus siliculosus* (Phaeophyceae). J Gen Virol **80:** 1367–1370

Delaroque N, Muller DG, Bothe G, Pohl T, Knippers R, Boland W (2001) The complete DNA sequence of the *Ectocarpus siliculosus* virus EsV-1 genome. Virology **287:** 112–132

Delwiche CF, Palmer JD (1997) The origin of plastids and their spread via secondary symbiosis. Plant Syst Evol (Suppl) **11:** 53–86

Dent RM, Han M, Niyogi KK (2001) Functional genomics of plant photosynthesis in the fast lane using *Chlamydomonas reinhardtii*. Trends Plant Sci **6:** 364–371

Derelle E, Ferraz C, Lagoda P, Eychenié S, Regad F, Sabau X, Courties C, Demaille J, Picard A, Moreau H (2002) DNA libraries for sequencing the genome of *Ostreococcus tauri* (Chlorophytae, Prasinophyceae): the smallest free-living eukaryotic cell. J Phycol **38:** 1150–1156

Diener DR, Curry AM, Johnson KA, Williams BD, Lefebvre PA, Kindle KL, Rosenbaum JL (1990) Rescue of a paralyzed flagella mutant of *Chlamydomonas* by transformation. Proc Natl Acad Sci USA **87:** 5739–5743

Díez B, Pedros-Alio C, Massana R (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. Appl Environ Microbiol **67:** 2932–2941

Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-Smith T, Maier UG (2001) The highly reduced genome of an enslaved algal nucleus. Nature **410:** 1091–1096

Douglas SE, Penny SL (1999) The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol **48:** 236–244

Drury JL, Dennis RG, Mooney DJ (2004) The tensile properties of alginate hydrogels. Biomaterials **25:** 3187–3199

Dunahey TG, Jarvis EE, Roessler PG (1995) Genetic transformation of the diatoms *Cyclotella cryptica* and *Navicula saprophila*. J Phycol **31:** 1004–1012

Dutcher SK (1995a) Flagellar assembly in two hundred and fifty easy-to-follow steps. Trends Genet **11:** 398–404

Dutcher SK (1995b) Purification of basal bodies and basal body complexes from *Chlamydomonas reinhardtii*. Methods Cell Biol **47:** 323–334

Dutcher SK (2000) *Chlamydomonas reinhardtii*: biological rationale for genomics. J Eukaryot Microbiol **47:** 340–349

Elrad D, Grossman AR (2004) A genome's-eye view of the light-harvesting polypeptides of *Chlamydomonas reinhardtii*. Curr Genet **45:** 61–75

Falciatore A, Casotti R, Leblanc C, Abrescia C, Bowler C (1999) Transformation of nonselectable reporter genes in marine diatoms. Mar Biotechnol **1:** 239–251

Falciatore A, d'Alcala MR, Croot P, Bowler C (2000) Perception of environmental signals by a marine diatom. Science **288:** 2363–2366

Feizi T, Mulloy B (2003) Carbohydrates and glycoconjugates. Glycomics: the new era of carbohydrate biology. Curr Opin Struct Biol **13:** 602–604

Fernandez E, Schnell R, Ranum LPW, Hussey SC, Silflow CD, Lefebvre PA (1989) Isolation and characterization of the nitrate reductase structural gene of *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **86:** 6449–6453

Finazzi G, Furia A, Barbagallo RP, Forti G (1999) State transitions, cyclic and linear electron transport and photophosphorylation in *Chlamydomonas reinhardtii*. Biochim Biophys Acta **1413:** 117–129

Fischer N, Setif P, Rochaix J-D (1997) Targeted mutations in the *psaC* gene of *Chlamydomonas reinhardtii*: Preferential reduction of FB at low temperature is not accompanied by altered electron flow from Photosystem I to ferredoxin. Biochemistry **36:** 93–102

Fjeld A, Lovle A (1976) Genetics of multicellular algae. *In* RA Lewin, ed, The Genetics of Algae. Blackwell Scientific, Oxford, pp 219–235

Fouilland E, Descolas-Gros C, Courties C, Collos Y, Vaquer A, Gasc A (2004) Productivity and growth of a natural population of the smallest free-living eukaryote under nitrogen deficiency and sufficiency. Microb Ecol **48:** 103–110

Fuhrmann M, Ferbitz L, Eichler-Stahlberg A, Hausherr A, Hegemann P (2002) Promoter activity monitored by heterologous expression of *Renilla reniformis* luciferase in *Chlamydomonas reinhardtii*. *In* Tenth International *Chlamydomonas* Conference, June 2002, Vancouver

Fuhrmann M, Oertel W, Hegemann P (1999) A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. Plant J **19:** 353–361

Funke RP, Kovar JL, Weeks DP (1997) Intracellular carbonic anhydrase is essential to photosynthesis in *Chlamydomonas reinhardtii* at atmospheric levels of CO_2. Demonstration via genomic complementation of the high-CO_2-requiring mutant *ca-1*. Plant Physiol **114:** 237–244

Givan AL, Levine RP (1967) The photosynthetic electron transport chain of *Chlamydomonas reinhardtii*. VII. Photosynthetic phosphorylation by a mutant strain of *Chlamydomonas reinhardtii* deficient in active P700. Plant Physiol **42:** 1264–1268

Glockner G, Rosenthal A, Valentin K (2000) The structure and gene repertoire of an ancient red algal plastid genome. J Mol Evol **51:** 382–390

Goff LJ, Coleman AW (1995) Fate of parasite and host organelle DNA during cellular transformation of red algae by their parasites. Plant Cell **7:** 1899–1911

Goldschmidt-Clermont M (1991) Transgenic expression of aminoglycoside adenine transferase in the chloroplast: a selectable marker for site-directed transformation of *Chlamydomonas*. Nucleic Acids Res **19:** 4083–4089

Gorman DS, Levine RP (1966) Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **54:** 1665–1669

Graham LE, Wilcox LW (2000) Algae. Prentice Hall, Upper Saddle River, NJ

Grossman A (2000) *Chlamydomonas reinhardtii* and photosynthesis: genetics to genomics. Curr Opin Plant Biol **3:** 132–137

Grossman AR, Harris EE, Hauser C, Lefebvre PA, Martinez D, Rokhsar D, Shrager J, Silflow CD, Stern D, Vallon O, et al (2003) *Chlamydomonas reinhardtii* at the crossroads of genomics. Eukaryot Cell **2:** 1137–1150

Guillou L, Eikrem W, Chretiennot-Dinet MJ, Le Gall F, Massana R, Romari K, Pedros-Alio C, Vaulot D (2004) Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. Protist **155:** 193–214

Gutman BL, Niyogi KK (2004) Chlamydomonas and Arabidopsis. A dynamic duo. Plant Physiol **135:** 607–610

Hackett JD, Yoon HS, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Nosenko T, Bhattacharya D (2004) Migration of the plastid genome to the nucleus in a peridinin dinoflagellate. Curr Biol **14:** 213–218

Hallahan BJ, Purton S, Ivison A, Wright D, Evans MCW (1995) Analysis of the proposed Fe-Sx binding region in *Chlamydomonas reinhardtii*. Photosynth Res **46:** 257–264

Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, Spielmann A, Stutz E (1993) Complete sequence of *Euglena gracilis* chloroplast DNA. Nucleic Acids Res **21:** 3537–3544

Harris EH (1989) The *Chlamydomonas* Sourcebook. A Comprehensive Guide to Biology and Laboratory Use. Academic Press, San Diego

Harris EH (2001) *Chlamydomonas* as a model organism. Annu Rev Plant Physiol Plant Mol Biol **52:** 363–406

Hayden HS, Blomster J, Maggs CA, Silva PC, Stanhope MJ, Waaland JR (2003) Linnaeus was right all along: *Ulva* and *Enteromorpha* are not distinct genera. Eur J Phycol **38:** 277–294

He Q, Dolganov N, Bjorkman O, Grossman AR (2001) The high light-inducible polypeptides in *Synechocystis* PCC6803. Expression and function in high light. J Biol Chem **276:** 306–314

Henry IM, Wilkinson MD, Hernandez JM, Schwarz-Sommer Z,

Grotewold E, Mandoli DF (2004) Comparison of ESTs from juvenile and adult phases of the giant unicellular green alga *Acetabularia acetabulum*. BMC Plant Biol **4:** 3

Higgs DC, Shapiro RS, Kindle KL, Stern DB (1999) Small cis acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. Mol Cell Biol **19:** 8479–8491

Hildebrand M, Dahlin K, Volcani BE (1998) Characterization of a silicon transporter gene family in *Cylindrotheca fusiformis*: sequences, expression analysis, and identification of homologs in other diatoms. Mol Gen Genet **260:** 480–486

Hildebrand M, Volcani BE, Gassmann W, Schroeder JI (1997) A gene family of silicon transporters. Nature **385:** 688–689

Hildebrand M, Wetherbee R (2003) Components and control of silicification in diatoms. Prog Mol Subcell Biol **33:** 11–57

Hong S, Spreitzer RJ (1994) Nuclear mutation inhibits expression of the chloroplast gene that encodes the large subunit of ribulose-1,5-bisphosphate carboxylase-oxygenase. Plant Physiol **106:** 673–678

Igarashi K, Kashiwagi K (2000) Polyamines: mysterious modulators of cellular functions. Biochem Biophys Res Commun **271:** 559–564

Im CS, Zhang Z, Shrager J, Chang CW, Grossman AR (2003) Analysis of light and CO_2 regulation in *Chlamydomonas reinhardtii* using genome-wide approaches. Photosynth Res **75:** 111–125

Jacobshagen S, Kindle KL, Johnson CH (1996) Transcription of CABII is regulated by the biological clock in *Chlamydomonas reinhardtii*. Plant Mol Biol **31:** 1173–1184

Jeong BR, Wu-Scharf D, Zhang C, Cerutti H (2002) Suppressors of transcriptional transgenic silencing in *Chlamydomonas* are sensitive to DNA-damaging agents and reactivate transposable elements. Proc Natl Acad Sci USA **99:** 1076–1081

Jing J, Reed J, Huang J, Hu X, Clarke V, Edington J, Housman D, Anantharaman TS, Huff EJ, Mishra B, et al (1998) Automated high resolution optical mapping using arrayed, fluid-fixed DNA molecules. Proc Natl Acad Sci USA **95:** 8046–8051

Kapraun DF, Hinson TK, Lemus AJ (1991) Karyology and cytophotometric estimation of inter- and intraspecific nuclear DNA variation in four species of *Porphyra* (Rhodophyta). Phycologia **30:** 458–466

Karol KG, McCourt RM, Cimino MT, Delwiche CF (2001) The closest living relatives of land plants. Science **294:** 2351–2353

Kathir P, LaVoie M, Brazelton WJ, Haas NA, Lefebvre PA, Silflow CD (2003) Molecular map of the *Chlamydomonas reinhardtii* nuclear genome. Eukaryot Cell **2:** 362–379

Khadaroo B, Robbens S, Ferraz C, Derelle E, Eychenie S, Cooke R, Peaucellier G, Delseny M, Demaille J, Van de Peer Y, et al (2004) The first green lineage cdc25 dual-specificity phosphatase. Cell Cycle **3:** 513–518

Kindle KL (1990) High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **87:** 1228–1232

Kindle KL, Schnell RA, Fernández E, Lefebvre PA (1989) Stable nuclear transformation of *Chlamydomonas* using the *Chlamydomonas* gene for nitrate reductase. J Cell Biol **109:** 2589–2601

Komine Y, Kikis E, Schuster G, Stern D (2002) Evidence for in vivo modulation of chloroplast RNA stability by 3′-UTR homopolymeric tails in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **99:** 4085–4090

Koutoulis A, Pazour GJ, Wilkerson CG, Inaba K, Sheng H, Takada S, Witman GB (1997) The *Chlamydomonas reinhardtii ODA3* gene encodes a protein of the outer dynein arm docking complex. J Cell Biol **137:** 1069–1080

Kovar JL, Zhang J, Funke RP, Weeks DP (2002) Molecular analysis of the acetolactate synthase gene of *Chlamydomonas reinhardtii* and development of a genetically engineered gene as a dominant selectable marker for genetic transformation. Plant J **29:** 109–117

Kranz HD, Miks D, Siegler ML, Capesius I, Sensen CW, Huss VA (1995) The origin of land plants: phylogenetic relationships among charophytes, bryophytes, and vascular plants inferred from complete small-subunit ribosomal RNA gene sequences. J Mol Evol **41:** 74–84

Kröger N, Deutzmann R, Bergsdorf C, Sumper M (2000) Species-specific polyamines from diatoms control silica morphology. Proc Natl Acad Sci USA **97:** 14133–14138

Kröger N, Deutzmann R, Sumper M (1999) Polycationic peptides from diatom biosilica that direct silica nanosphere formation. Science **286:** 1129–1132

Kröger N, Lorenz S, Brunner E, Sumper M (2002) Self-assembly of highly phosphorylated silaffins and their function in biosilica morphogenesis. Science **298:** 584–586

Kuroiwa T (2000) The discovery of the division apparatus of plastids and mitochondria. J Electron Microsc (Tokyo) **49:** 123–134

Kuroiwa T, Kuroiwa H, Sakai A, Takahashi H, Toda K, Itoh R (1998) The division apparatus of plastids and mitochondria. Int Rev Cytol **181:** 1–41

La Fontaine S, Quinn JM, Nakamoto SS, Page MD, Gohre V, Moseley JL, Kropat J, Merchant S (2002) Copper-dependent iron assimilation pathway in the model photosynthetic eukaryote *Chlamydomonas reinhardtii*. Eukaryot Cell **1:** 736–757

Lardans A, Gillham NW, Boynton JE (1997) Site-directed mutations at residue 251 of the photosystem II D1 protein of *Chlamydomonas* that result in a nonphotosynthetic phenotype and impair D1 synthesis and accumulation. J Biol Chem **272:** 210–216

Larson EM, O'Brien CM, Zhu G, Spreitzer RJ, Portis AR Jr (1997) Specificity for activase is changed by a Pro-89 to Arg substitution in the large subunit of ribulose-1,5-bisphosphate carboxylase-oxgenase. J Biol Chem **272:** 17033–17037

Lavaud J, Rousseau B, Etienne AL (2003) Enrichment of the light-harvesting complex in diadinoxanthin and implications for the nonphotochemical fluorescence quenching in diatoms. Biochemistry **42:** 5802–5808

Lavaud J, Rousseau B, van Gorkom HJ, Etienne AL (2002) Influence of the diadinoxanthin pool size on photoprotection in the marine planktonic diatom *Phaeodactylum tricornutum*. Plant Physiol **129:** 1398–1406

Lavorel J, Levine RP (1968) Fluorescence properties of wild-type *Chlamydomonas reinhardtii* and three mutant strains having impaired photosynthesis. Plant Physiol **43:** 1049–1055

Lebeau T, Robert JM (2003) Diatom cultivation and biotechnologically relevant products. Part II: current and putative products. Appl Microbiol Biotechnol **60:** 624–632

Ledford HK, Baroli I, Shin JW, Fischer BB, Eggen RIL, Niyogi KK (2004) Comparative profiling of lipid soluble antioxidants and transcripts reveals two phases of photo-oxidative stress in a xanthophyll-deficient mutant of *Chlamydomonas reinhardtii*. Mol Genet Genomics **272:** 470–479

Lefebvre PA, Silflow CD (1999) *Chlamydomonas*: the cell and its genomes. Genetics **151:** 9–14

Lemieux C, Otis C, Turmel M (2000) Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. Nature **403:** 649–652

Levine RP (1969) The analysis of photosynthesis using mutant strains of algae and higher plants. Annu Rev Plant Physiol **20:** 523–540

Levine RP, Goodenough UW (1970) The genetics of photosynthesis and of the chloroplast in *Chlamydomonas reinhardii*. Annu Rev Genet **4:** 397–408

Li JB, Gerdes JM, Haycraft CJ, Fan Y, Teslovich TM, May-Simera H, Li H, Blacque OE, Li L, Leitch CC, et al (2004) Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. Cell **117:** 541–552

Li X, Bjorkman O, Shih C, Grossman A, Rosenquist M, Jansson C, Niyogi KK (2000) A pigment-binding protein essential for regulation of photosynthetic light harvesting. Nature **403:** 391–395

Lilly JW, Maul JE, Stern DB (2002) The *Chlamydomonas reinhardtii* organellar genomes respond transcriptionally and post-transcriptionally to abiotic stimuli. Plant Cell **14:** 2681–2706

Littler MM, Littler DS, Blair SM, Norris NJ (1985) Deepest known plant life discovered on an uncharted seamount. Science **227:** 57–59

Lohr M, Wilhelm C (1999) Algae displaying the diadinoxanthin cycle also possess the violaxanthin cycle. Proc Natl Acad Sci USA **96:** 8784–8789

Lumbreras V, Stevens DR, Purton S (1998) Efficient foreign gene expression in *Chlamydomonas reinhardtii* mediated by an endogenous intron. Plant J **14:** 441–447

Maier UG, Douglas SE, Cavalier-Smith T (2000) The nucleomorph genomes of cryptophytes and chlorarachniophytes. Protist **151:** 103–109

Mandoli DF (1998) Elaboration of body plan and phase change during development of *Acetabularia*: How is the complex architecture of a giant unicell built? Annu Rev Plant Physiol Plant Mol Biol **49:** 173–198

Mann DG (1993) Patterns of sexual reproduction in diatoms. Hydrobiologia **269/270:** 11–20

Mann DG, Chepurnov VA, Droop SJM (1999) Sexuality, incompatibility, size variation, and preferential polyandry in natural populations and clones of *Sellaphora pupula* (*Bacilliarophyceae*). J Phycol **35:** 152–170

Maruyama S, Misumi O, Ishii Y, Asakawa S, Shimizu A, Sasaki T,

Matsuzaki M, Shin-i T, Nozaki H, Kohara Y, et al (2004) The minimal eukaryotic ribosomal DNA units in the primitive red alga *Cyanidioschyzon merolae.* DNA Res **11:** 83–91

Matsubara K (2004) Recent advances in marine algal anticoagulants. Curr Med Chem Cardiovasc Hematol Agents **2:** 13–19

Matsuzaki M, Misumi O, Shin IT, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Yoshida Y, et al (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature **428:** 653–657

Mayer K, Mewes HW (2002) How can we deliver the large plant genomes? Strategies and perspectives. Curr Opin Plant Biol **5:** 173–177

Mayfield SP, Franklin SE, Lerner RA (2003) Expression and assembly of a fully active antibody in algae. Proc Natl Acad Sci USA **100:** 438–442

Mayfield SP, Kindle KL (1990) Stable nuclear transformation of *Chlamydomonas reinhardtii* by using a *C. reinhardtii* gene as the selectable marker. Proc Natl Acad Sci USA **87:** 2087–2091

Melkozernov AN, Su H, Lin S, Bingham S, Webber AN, Blankenship RE (1997) Specific mutations near the primary donor in Photosystem I from *Chlamydomonas reinhardtii* alters the trapping time and spectroscopic properties of P700. Biochemistry **36:** 2898–2907

Meyer A, Kirsch H, Domergue F, Abbadi A, Sperling P, Bauer J, Cirpus P, Zank TK, Moreau H, Roscoe TJ, et al (2004) Novel fatty acid elongases and their use for the reconstitution of docosahexaenoic acid biosynthesis. J Lipid Res **45:** 1899–1909

Minko I, Holloway SP, Nikaido S, Carter M, Odom OW, Johnson CH, Herrin DL (1999) *Renilla* luciferase as a vital reporter for chloroplast gene expression in *Chlamydomonas.* Mol Gen Genet **262:** 421–425

Minoda A, Sakagami R, Yagisawa F, Kuroiwa T, Tanaka K (2004) Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, *Cyanidioschyzon merolae* 10D. Plant Cell Physiol **45:** 667–671

Mitman GG, van der Meer JP (1994) Meiosis, blade development, and sex determination in *Porphyra purpurea* (Rhodophyta). J Phycol **30:** 147–159

Miura K, Yamano T, Yoshioka S, Kohinata T, Inoue Y, Taniguchi F, Asamizu E, Nakamura Y, Tabata S, Yamato KT, et al (2004) Expression profiling-based identification of CO_2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii.* Plant Physiol **135:** 1595–1607

Miyagishima S, Itoh R, Aita S, Kuroiwa H, Kuroiwa T (1999) Isolation of dividing chloroplasts with intact plastid-dividing rings from a synchronous culture of the unicellular red alga *Cyanidioschyzon merolae.* Planta **209:** 371–375

Miyagishima S, Kuroiwa H, Kuroiwa T (2001a) The timing and manner of disassembly of the apparatuses for chloroplast and mitochondrial division in the red alga *Cyanidioschyzon merolae.* Planta **212:** 517–528

Miyagishima S, Takahara M, Kuroiwa T (2001b) Novel filaments 5 nm in diameter constitute the cytosolic ring of the plastid division apparatus. Plant Cell **13:** 707–721

Miyagishima S, Takahara M, Mori T, Kuroiwa H, Higashiyama T, Kuroiwa T (2001c) Plastid division is driven by a complex mechanism that involves differential transition of the bacterial and eukaryotic division rings. Plant Cell **13:** 2257–2268

Miyagishima SY, Nishida K, Mori T, Matsuzaki M, Higashiyama T, Kuroiwa H, Kuroiwa T (2003) A plant-specific dynamin-related protein forms a ring at the chloroplast division site. Plant Cell **15:** 655–665

Mizukami Y, Okauchi M, Kito H, Ishimoto SI, Ishida T, Fuseya M (1995) Culture and development of electrically fused protoplasts from red marine algae, *Porphyra yezoensis* and *P. suborbiculata.* Aquaculture **132:** 361–367

Moll B, Levine RP (1970) Characterization of a photosynthetic mutant strain of *Chlamydomonas reinhardi* deficient in phosphoribulokinase activity. Plant Physiol **46:** 576–580

Moreira D, Philippe H (2001) Sure facts and open questions about the origin and evolution of photosynthetic plastids. Res Microbiol **152:** 771–780

Morgan DM (1999) Polyamines. An overview. Mol Biotechnol **11:** 229–250

Mozzachiodi R, Scuri R, Roberto M, Brunelli M (2001) Caulerpenyne, a toxin from the seaweed *Caulerpa taxifolia*, depresses afterhyperpolarization in invertebrate neurons. Neuroscience **107:** 519–526

Muller DG, Schmid CE (1988) Qualitative and quantitative determination of pheromone secretion in female gametes of *Ectocarpus siliculosus* (Phaeophyceae). Biol Chem Hoppe Seyler **369:** 647–653

Murdoch L (1996) Discovering the Great Barrier Reef. Harper Collins, Sydney

Nelson JAE, Savereide PB, Lefebvre PA (1994) The *CRY1* gene in *Chlamydomonas reinhardtii:* structure and use as a dominant selectable marker for nuclear transformation. Mol Cell Biol **14:** 4011–4019

Newman SM, Boynton JE, Gillham NW, Randolph-Anderson BL, Johnson AM, Harris EH (1990) Transformation of chloroplast ribosomal RNA in *Chlamydomonas*: molecular and genetic characterization of integration events. Genetics **126:** 875–888

Nikaido I, Asamizu E, Nakajima M, Nakamura Y, Saga N, Tabata S (2000) Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, *Porphyra yezoensis.* DNA Res **7:** 223–227

Nishida K, Misumi O, Yagisawa F, Kuroiwa H, Nagata T, Kuroiwa T (2004) Triple immunofluorescent labeling of FtsZ, dynamin, and EF-Tu reveals a loose association between the inner and outer membrane mitochondrial division machinery in the red alga *Cyanidioschyzon merolae.* J Histochem Cytochem **52:** 843–849

Nozaki H, Matsuzaki M, Misumi O, Kuroiwa H, Hasegawa M, Higashiyama T, Shin IT, Kohara Y, Ogasawara N, Kuroiwa T (2004) Cyanobacterial genes transmitted to the nucleus before divergence of red algae in the Chromista. J Mol Evol **59:** 103–113

Oeltjen A, Marquardt J, Rhiel E (2004) Differential circadian expression of genes *fcp2* and *fcp6* in *Cyclotella cryptica.* Int Microbiol **7:** 127–131

Ohresser M, Matagne RF, Loppes R (1997) Expression of the arylsulphatase reporter gene under the control of the *NIT1* promoter of *Chlamydomonas reinhardtii.* Curr Genet **31:** 264–271

Ohta N, Matsuzaki M, Misumi O, Miyagishima SY, Nozaki H, Tanaka K, Shin IT, Kohara Y, Kuroiwa T (2003) Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae.* DNA Res **10:** 67–77

Palombella AL, Dutcher SK (1998) Identification of the gene encoding the tryptophan synthase beta-subunit from *Chlamydomonas reinhardtii.* Plant Physiol **117:** 455–464

Pazour GJ, Dickert BL, Vucica Y, Seeley ES, Rosenbaum JL, Witman GB, Cole DG (2000) *Chlamydomonas* IFT88 and its mouse homologue, polycystic kidney disease gene tg737, are required for assembly of cilia and flagella. J Cell Biol **151:** 709–718

Pennarun G, Bridoux AM, Escudier E, Dastot-Le Moal F, Cacheux V, Amselem S, Duriez B (2002) Isolation and expression of the human hPF20 gene orthologous to *Chlamydomonas* PF20: evaluation as a candidate for axonemal defects of respiratory cilia and sperm flagella. Am J Respir Cell Mol Biol **26:** 362–370

Petersen J, Brinkmann H, Cerff R (2003) Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. J Mol Evol **57:** 16–26

Peterson RB, Havir EA (2001) Photosynthetic properties of an *Arabidopsis thaliana* mutant possessing a defective *PsbS* gene. Planta **214:** 142–152

Pickett-Heaps JD (1983) Valve morphogenesis and the microtubule center in three species of the diatom *Nitzschia.* J Phycol **19:** 269–281

Pickett-Heaps JD, Kowalski SE (1981) Valve morphogenesis and the microtubule center of the diatom *Hantzschia amphioxysis.* Eur J Cell Biol **25:** 150–170

Poulsen N, Kröger N (2004) Silica morphogenesis by alternative processing of silaffins in the diatom *Thalassiosira pseudonana.* J Biol Chem **279:** 42993–42999

Poulsen N, Sumper M, Kröger N (2003) Biosilica formation in diatoms: characterization of native silaffin-2 and its role in silica morphogenesis. Proc Natl Acad Sci USA **100:** 12075–12080

Purton S, Rochaix J-D (1994) Complementation of a *Chlamydomonas reinhardtii* mutant using a genomic cosmid library. Plant Mol Biol **24:** 533–537

Quatrano RS, Brian L, Aldridge J, Schultz T (1991) Polar axis fixation in *Fucus* zygotes: components of the cytoskeleton and extracellular matrix. Dev Suppl **1:** 11–16

Quinn JM, Merchant S (1995) Two Copper-responsive elements associated with the *Chlamydomonas Cyc6* gene function as targets for transcriptional activators. Plant Cell **7:** 623–638

Ragan M, Gutell R (1995) Are red algae plants? Bot J Linn Soc **118:** 81–105

Ral JP, Derelle E, Ferraz C, Wattebled F, Farinas B, Corellou F, Buleon A, Slomianny MC, Delvalle D, d'Hulst C, et al (2004) Starch division and partitioning. A mechanism for granule propagation and maintenance in

the picophytoplanktonic green alga *Ostreococcus tauri*. Plant Physiol **136:** 3333–3340

Randolph-Anderson BL, Sato R, Johnson AM, Harris EH, Hauser CR, Oeda K, Ishige F, Nishio S, Gillham NW, Boynton JE (1998) Isolation and characterization of a mutant protoporphyrinogen oxidase gene from *Chlamydomonas reinhardtii* conferring resistance to porphyric herbicides. Plant Mol Biol **38:** 839–859

Reimann BEF, Lewin JC, Volcani BE (1966) Studies on the biochemistry and fine structure of silica shell formation in diatoms. II. The structure of the cell wall of *Navicula pelliculosa* (Breb.) Hilse. J Phycol **2:** 74–84

Reinfelder JR, Milligan AJ, Morel FM (2004) The role of the C4 pathway in carbon accumulation and fixation in a marine diatom. Plant Physiol **135:** 2106–2111

Reith ME, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol Biol Report **13:** 333–335

Richly E, Leister D (2004) NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs. Mol Biol Evol **21:** 1972–1980

Rosakis A, Koster W (2004) Transition metal transport in the green microalga *Chlamydomonas reinhardtii*: genomic sequence analysis. Res Microbiol **155:** 201–210

Round FE, Crawford RM, Mann DG (1990) The Diatoms. Cambridge University Press, Cambridge, UK

Sager R (1960) Genetic systems in *Chlamydomonas*. Science **132:** 1459–1465

Salem N Jr, Moriguchi T, Greiner RS, McBride K, Ahmad A, Catalan JN, Slotnick B (2001) Alterations in brain function after loss of docosahexaenoate due to dietary restriction of n-3 fatty acids. J Mol Neurosci **16:** 299–307

Sato V, Levine RP, Neumann J (1971) Photosynthetic phosphorylation in *Chlamydomonas reinhardti*. Effects of a mutation altering an ATP-synthesizing enzyme. Biochim Biophys Acta **253:** 437–448

Scala S, Carels N, Falciatore A, Chiusano ML, Bowler C (2002) Genome properties of the diatom *Phaeodactylum tricornutum*. Plant Physiol **129:** 993–1002

Schroda M, Vallon O, Wollman FA, Beck CF (1999) A chloroplast-targeted heat shock protein 70 (HSP70) contributes to the photoprotection and repair of photosystem II during and after photoinhibition. Plant Cell **11:** 1165–1178

Shaw SL, Quatrano RS (1996) The role of targeted secretion in the establishment of cell polarity and the orientation of the division plane in *Fucus* zygotes. Development **122:** 2623–2630

Shimogawara K, Fujiwara S, Grossman AR, Usuda H (1998) High efficiency transformation of *Chlamydomonas reinhardtii* by electroporation. Genetics **148:** 1821–1828

Shrager J, Hauser C, Chang CW, Harris EH, Davies J, McDermott J, Tamse R, Zhang Z, Grossman AR (2003) *Chlamydomonas reinhardtii* genome project. A guide to the generation and use of the cDNA information. Plant Physiol **131:** 401–408

Simpson CL, Stern DB (2002) Mining the treasure trove of algal chloroplast genomes: Surprises in architecture and gene content, and their functional implications. Plant Physiol **129:** 957–966

Sineshchekov OA, Jung KH, Spudich JL (2002) The rhodopsins mediate phototaxis to low- and high-intensity light in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **99:** 225–230

Smith EF, Lefebvre PA (1996) PF16 encodes a protein with armadillo repeats and localizes to a single microtubule of the central apparatus in *Chlamydomonas* flagella. J Cell Biol **132:** 359–370

Smith EF, Lefebvre PA (1997) PF20 gene product contains WD repeats and localizes to the intermicrotubule bridges in *Chlamydomonas* flagella. Mol Biol Cell **8:** 455–467

Snell WJ, Pan J, Wang Q (2004) Cilia and flagella revealed: from flagellar assembly in *Chlamydomonas* to human obesity disorders. Cell **117:** 693–697

Stauber EJ, Fink A, Markert C, Kruse O, Johanningmeier U, Hippler M (2003) Proteomics of *Chlamydomonas reinhardtii* light-harvesting proteins. Eukaryot Cell **2:** 978–994

Stevens DR, Rochaix J-D, Purton S (1996) The bacterial phleomycin resistance gene *ble* as a dominant selectable marker in *Chlamydomonas*. Mol Gen Genet **251:** 23–30

Stiller JW, Hall BD (2002) Evolution of the RNA polymerase II C-terminal domain. Proc Natl Acad Sci USA **99:** 6091–6096

Stirewalt VL, Michalowski CB, Loffelhardt W, Bohnert HJ, Bryant D (1995) Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*. Plant Mol Biol **13:** 327–332

Tada N, Shibata S, Otsuka S, Namba K, Oyaizu H (1999) Comparison of gene arrangements of chloroplasts between two centric diatoms, *Skeletonema costatum* and *Odontella sinensis*. DNA Seq **10:** 343–347

Takahashi Y, Matsumoto H, Goldschmidt-Clermont M, Rochaix J-D (1994) Directed disruption of the *Chlamydomonas* chloroplast *psbK* gene destabilizes the photosystem II reaction center complex. Plant Mol Biol **24:** 779–788

Tam L-W, Lefebvre PA (1993) Cloning of flagellar genes in *Chlamydomonas reinhardtii* by DNA insertional mutagenesis. Genetics **135:** 375–384

Tonon T, Harvey D, Qing R, Li Y, Larson TR, Graham IA (2004) Identification of a fatty acid Delta11-desaturase from the microalga *Thalassiosira pseudonana*. FEBS Lett **563:** 28–34

Turmel M, Otis C, Lemieux C (1999) The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. Proc Natl Acad Sci USA **96:** 10248–10253

Turmel M, Otis C, Lemieux C (2002) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. Proc Natl Acad Sci USA **99:** 11275–11280

Van Etten JL, Graves MV, Muller DG, Boland W, Delaroque N (2002) Phycodnaviridae: large DNA algal viruses. Arch Virol **147:** 1479–1516

Vaulot D, Olson RJ, Chisholm SW (1986) Light and dark control of the cell cycle in two marine phytoplankton species. Exp Cell Res **167:** 38–52

Vaulot D, Olson RJ, Merkel S, Chisholm SW (1987) Cell cycle response to nutrient starvation in two phytoplankton species, *Thalassiosira weissflogii* and *Hymenomonas carterae*. Mar Biol **95:** 625–630

Veldhuis MJW, Cucci TL, Sieracki ME (1997) Cellular DNA content of marine phytoplankton using two new fluorophores: taxonomic and ecological implications. J Phycol **33:** 527–541

Velluppillai JM, Jacobs MA, Duplessis MR, Choi L, Cattolico RA (2003) The chloroplast genome of the toxic stramenopile *Heterosigma akashiwa* (Raphidophyceae). J Phycol (Suppl) **39:** 57

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, et al (2004) Environmental genome shotgun sequencing of the Sargasso Sea. Science **304:** 66–74

Villand P, Ericksson M, Samuelsson G (1997) Regulation of genes by the environmental CO_2 level. Plant Physiol **114:** 258–259

Vrieling EG, Gieskes WWC, Beelen TPM (1999) Silicon deposition in diatoms: control by the pH inside the silicon deposition vesicle. J Phycol **35:** 548–559

Vysotskaia VS, Curtis DE, Voinov AV, Kathir P, Silflow CD, Lefebvre PA (2001) Development and characterization of genome-wide single nucleotide pholymorphism markers in the green alga *Chlamydomonas reinhardtii*. Plant Physiol **127:** 386–389

Waaland JR, Dickson LG, Watson BA (1990) Protoplast isolation and regeneration in the marine red alga *Porphyra nereocystis*. Planta **181:** 522–528

Waaland JR, Stiller JW, Cheney DP (2004) Macroalgal candidates for genomics. J Phycol **40:** 26–33

Wagner V, Fiedler M, Markert C, Hippler M, Mittag M (2004) Functional proteomics of circadian expressed proteins from *Chlamydomonas reinhardtii*. FEBS Lett **559:** 129–135

Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, Tsudzuki J, Nakashima K, Tsudzuki T, Suzuki Y, et al (1997) Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. Proc Natl Acad Sci USA **94:** 5967–5972

Wastl J, Duin EC, Iuzzolino L, Dorner W, Link T, Hoffmann S, Sticht H, Dau H, Lingelbach K, Maier UG (2000) Eukaryotically-encoded and chloroplast-located rubredoxin is associated with photosystem II. J Biol Chem **275:** 30058–30063

Webber AN, Su H, Binghma SE, Kass H, Krabben L, Kuhn M, Schlodder E, Lubitz W (1996) Site-directed mutations affecting the spectroscopic characteristics and mid-point potential of the primary donor in photosystem I. Biochemistry **39:** 12857–12863

Wen ZY, Chen F (2003) Heterotrophic production of eicosapentaenoic acid by microalgae. Biotechnol Adv **21:** 273–294

Whitelaw CA, Barbazuk WB, Pertea G, Chan AP, Cheung F, Lee Y, Zheng L, van Heeringen S, Karamycheva S, Bennetzen JL, et al (2003)

Enrichment of gene-coding sequences in maize by genome filtration. Science **302:** 2118–2120

Whitelegge JP, Koo D, Erickson J (1992) Site-directed mutagenesis of the chloropolast *psbA* gene encoding the D1 polypeptide of photosystem II in *Chlamydomonas reinhardtii* changes at aspartate 170 affect the assembly of a functional water-splitting manganese cluster. *In* N Murata, ed, Research in Photosynthesis, Vol II. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 151–154

Wilson NF, Lefebvre PA (2002) Characterization of GSK3, a flagellar kinase with a putative role in the regulation of flagella length. *In* Tenth International *Chlamydomonas* Conference, June 2002, Vancouver

Wray J (1977) Calcareous Algae. Elsevier, New York

Wykoff D, Davies J, Grossman A (1998) The regulation of photosynthetic electron transport during nutrient deprivation in *Chlamydomonas reinhardtii.* Plant Physiol **117:** 129–139

Wykoff D, Grossman A, Weeks DP, Usuda H, Shimogawara K (1999) Psr1, a nuclear localized protein that regulates phosphorus metabolism in *Chlamydomonas.* Proc Natl Acad Sci USA **96:** 15336–15341

Xiao S, Zhang Y, Knoll A (1998) Three-dimensional preservation of algae and animal embryosin, a neoproterozoic phosphorite. Nature **391:** 553–558

Xiong J, Hutchinson RS, Sayre RT, Govindjee (1997) Modification of the photosystem II acceptor side function in a D1 mutant (arginine-269-glycine) of *Chlamydomonas reinhardtii*. Biochim Biophys Acta **1322:** 60–76

Yan X, Fujita Y, Aruga Y (2000) Induction and characterization of pigmentation mutants in *Porphyra yezoensis* (Bangiales, Rhodophyta). J Appl Phycol **12:** 69–81

Yildiz FH, Davies JP, Grossman AR (1996) Sulfur availability and the *SAC1* gene control adenosine triphosphate sulfurylase gene expression in *Chlamydomonas reinhardtii.* Plant Physiol **112:** 669–675

Yoon HS, Lee JY, Boo SM, Bhattacharya D (2001) Phylogeny of Alariaceae, Laminariaceae, and Lessoniaceae (Phaeophyceae) based on plastid-encoded RuBisCo spacer and nuclear-encoded ITS sequence comparisons. Mol Phylogenet Evol **21:** 231–243

Yoshioka S, Taniguchi F, Miura K, Inoue T, Yamano T, Fukuzawa H (2004) The novel Myb transcription factor LCR1 regulates the CO_2-responsive gene *Cah1*, encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii.* Plant Cell **16:** 1466–1477

Zaslavskaia LA, Lippmeier JC, Kroth PG, Grossman AR, Apt KE (2000) Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. J Phycol **36:** 379–386

Zaslavskaia LA, Lippmeier JC, Shih C, Ehrhardt D, Grossman AR, Apt KE (2001) Trophic conversion of an obligate photoautotrophic organism through metabolic engineering. Science **292:** 2073–2075

Zhang D, Lefebvre PA (1997) FAR1, a negative regulatory locus required for the repression if the nitrate reductase gene in *Chlamydomonas reinhardtii.* Genetics **146:** 121–133

Zhang H, Herman PL, Weeks DP (1994) Gene isolation through genomic complementation using an indexed library of *Chlamydomonas reinhardtii* DNA. Plant Mol Biol **24:** 663–672

Zhang Z, Cavalier-Smith T, Green BR (2002) Evolution of dinoflagellate unigenic minicircles and the partially concerted divergence of their putative replicon origins. Mol Biol Evol **19:** 489–500

Zhang Z, Green BR, Cavalier-Smith T (1999) Single gene circles in dinoflagellate chloroplast genomes. Nature **400:** 155–159

Zhang Z, Shrager J, Chang C-W, Vallon O, Grossman AR (2004) Genome based analysis of sulfur deprivation of wild-type cells and the *sac1* mutant of *Chlamydomonas.* Eukaryot Cell **3:** 1331–1348

Zhao B, Schneid C, Iliev D, Schmidt EM, Wagner V, Wollnik F, Mittag M (2004) The circadian RNA-binding protein CHLAMY 1 represents a novel type heteromer of RNA recognition motif and lysine homology domain-containing subunits. Eukaryot Cell **3:** 815–825

Zhu G, Spreitzer RJ (1996) Directed mutagenesis of chloroplast ribulose-1,5-bisphosphate carboxylase-oxygenase. Loop 6 substitutions complement for structual stability but decrease catalytic efficiency. J Biol Chem **271:** 18494–18498