ANNUAL REVIEWS

Annual Review of Plant Biology

The Land–Sea Connection: Insights Into the Plant Lineage from a Green Algal Perspective

Charles Bachy,[1] Fabian Wittmers,[1] Jan Muschiol,[1]
Maria Hamilton,[1] Bernard Henrissat,[2,3,4]
and Alexandra Z. Worden[1,5,6]

[1] Ocean EcoSystems Biology Unit, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

[2] Architecture et Fonction des Macromolécules Biologiques, CNRS UMR 7257, Aix-Marseille Université (AMU), Marseille, France

[3] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[4] DTU Bioengineering, Technical University of Denmark, Kongens Lyngby, Denmark

[5] Marine Biological Laboratories, Woods Hole, Massachusetts, USA

[6] Max Planck Institute for Evolutionary Biology, Plön, Germany

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

green lineage, plant evolution, core genomes, carbohydrate-active enzymes, CAZymes, peptidoglycan, phytochrome, marine algae

Abstract

The colonization of land by plants generated opportunities for the rise of new heterotrophic life forms, including humankind. A unique event underpinned this massive change to earth ecosystems—the advent of eukaryotic green algae. Today, an abundant marine green algal group, the prasinophytes, alongside prasinodermophytes and nonmarine chlorophyte algae, is facilitating insights into plant developments. Genome-level data allow identification of conserved proteins and protein families with extensive modifications, losses, or gains and expansion patterns that connect to niche specialization and diversification. Here, we contextualize attributes according to Viridiplantae evolutionary relationships, starting with orthologous protein families, and then focusing on key elements with marked differentiation, resulting in patchy distributions across green algae and plants. We place attention on peptidoglycan biosynthesis, important for plastid division and walls; phytochrome photosensors that are master regulators in plants; and carbohydrate-active enzymes, essential to all manner of carbohydrate

biotransformations. Together with advances in algal model systems, these areas are ripe for discovering molecular roles and innovations within and across plant and algal lineages.

Contents

INTRODUCTION

Major advances in plant biology have come in part from our recognition of the need for food security: Heterotrophic organisms, such as humans, are inherently dependent on the productivity of the photosynthetic world. The evolutionary story of plants is rooted in the realm of marine microbes. It was the primary endosymbiosis event, wherein a cyanobacterium was engulfed by a unicellular eukaryote and incorporated as an organelle, that gave rise to the first eukaryotic alga (1, 109). One such marine green alga, perhaps long extinct, is widely accepted as having been the progenitor of all terrestrial plant life (97, 175). Phylogenetic relationships within the Viridiplantae, which collectively houses marine and nonmarine green algae as well as multicellular land plants, continue to be resolved as genome and transcriptome projects tackle sequencing the breadth of streptophyte diversity alongside prasinodermophyte, chlorophyte, and prasinophyte algal diversity (**Figure 1**). This work has illuminated retained features that span much but not all of Viridiplantae diversity, some that lead back to the algal progenitor of plants and others that support green algal growth in modern-day ocean environments.

A key difference between terrestrial and aquatic environments is that the latter are subject to the physical influence of fluid mixing throughout a three-dimensional space. This mixing can be to both the benefit and detriment of single-celled algae, and transport between regions that have very different environmental profiles (e.g., temperature, nutrient availability) certainly impacts the evolution of these organisms. In terrestrial environments, extreme environmental conditions include desiccation or varied water availability, as well as extended exposure to direct unfiltered sunlight, which causes photooxidative and ultraviolet (UV) radiation damage (79). The effects of excessive light exposure are also relevant to aquatic green algae, but some mitigation occurs via rapid attenuation by water. Light itself is clearly critical both on land and in the water—and varied, whether by cloud cover or shading. The latter can occur due to mixing to darker depths in the water column or to the proximity of other photosynthetic organisms that block light under dense bloom conditions or, for land plants, light blockage by other plants, and cloud cover is clearly relevant in both environments. Mechanisms and sensors present in green algae are now known to

Heterotrophic:
a nutritional lifestyle wherein an organism's growth relies on acquiring externally produced organic carbon sources

Primary endosymbiosis event:
the origin of photosynthesis in eukaryotes, wherein a heterotrophic protist phagotrophically engulfed a cyanobacterium and incorporated it as an organelle

Viridiplantae:
monophyletic group of organisms commonly known as green plants, the green lineage, or Plantae; includes green algae and terrestrial plants as well as seagrasses

a

Core rosids

Saxifragales

Asterids

Caryophyllales

Ranunculids

Magnoliids

Monocots

ANA grade

Eudicots

Flowering plants

Vascular plants

Seed plants

Land plants

Green plants

Outgroup

Rhodophytes

Glaucophytes   Prasinophytes

Chlorophytes

Prasinodermophytes

Streptophyte algae

Bryophytes

Lycophytes

Ferns

Gymnosperms

Sedum *Viola* *Vicia* *Fragaria* *Brassica* *Micromonas* *Bathycoccus* *Symbiochloris* *Ulva* *Prasinoderma* *Hypnum* *Dryopteris* *Athyrium* *Pinus* *Nymphaea* *Ophrys* *Spirodela* *Magnolia* *Eschscholzia* *Mammillaria* *Dianthus* *Leucanthemum* *Mentha* *Solanum* *Hamamelis*

b

	Genome size (Mb)
Prasinococcus capsulatus	ND
Prasinoderma coloniale	25
Prasinoderma singularis	ND
Pyramimonas parkeae	ND
Cymbomonas sp.	ND
Monomastix opisthostigma	ND
Dolichomastix tenuilepis	ND
Micromonas pusilla	22
Micromonas commoda	21
Bathycoccus prasinos	15
Ostreococcus tauri	15
Ostreococcus lucimarinus	13
Coccomyxa subellipsoidea	49
Auxenochlorella protothecoides	23
Chlorella variabilis	46
Ulva mutabilis	98
Chromochloris zofingiensis	60
Chlamydomonas reinhardtii	111
Volvox carteri	131
Gonium pectorale	149
Mesostigma viride	146
Chlorokybus atmophyticus	74
Klebsormidium nitens	104
Chara braunii	1,751
Arabidopsis thaliana	120

0.3

Prasinodermophytes   Single-celled   Prasinodermophyceae   Ulvophyceae
Prasinophytes   Colonial   Pyramimonadophyceae   Chlorophyceae
Chlorophytes   Sarcinoid   Mamiellophyceae   Streptophyte algae
Streptophytes   Multicellular   Trebouxiophyceae   Rosids

c

Prasinodermophyta

Pyramimonadophyceae

Mamiellophyceae

Pycnococcaceae

Picocystophyceae

Chloropicophyceae

Nephroselmidophyceae

Pedinophyceae

Chlorodendrophyceae

Chlorophyceae

Ulvophyceae

Trebouxiophyceae

Streptophyta

Core chlorophytes

○ BS ≥90%

0.1

(*Caption appears on following page*)

Figure 1 (*Figure appears on preceding page*)

The Viridiplantae houses all land plants and four types of green algae: streptophytes, chlorophytes, prasinodermophytes, and prasinophytes. (*a*) This phylogenomic reconstruction is adapted from Reference 131 (CC BY 4.0) and was developed using 410 single-copy nuclear gene family proteins based on ASTRAL analysis. There are 1,090 Viridiplantae representative species and 66 outgroup members (the latter including cryptophytes, stramenopiles, haptophytes, rhodophytes, and glaucophytes). Colored bars surrounding the tree represent recognized divisions (or phyla) of the green lineage (i.e., Viridiplantae): Streptophyta, Prasinodermophyta, and the chlorophyte and prasinophyte algae, which together are often termed Chlorophyta. Colors on branches reflect different taxonomic clades. The micrograph illustrating *Prasinoderma* specimens is adapted from Reference 100 (CC BY 4.0); the photo of *Eschscholzia* was provided by Scott Joly. All other images are original. (*b*) This phylogenetic tree was constructed using ML methods in RAxML and MrBayes based on a concatenated sequence alignment of 256 single-copy genes published in Reference 90c. BS of 100% is indicated and supports the division of Viridiplantae into four algal clades: the prasinodermophytes, the chlorophytes, the streptophytes (also including land plants), and the prasinophytes (Pyramimonadophyceae and Mamiellophyceae). Colored circles describe morphological features (*left*), genome size is listed in the middle, colored boxes on the right follow the classification into taxonomic clades based on branches in panel *a*, and colored shading behind clades reflects divisions. Note that while *Chlamydomonas reinhardtii* can be found in palmelloid formation in response to environmental stress, its unicellular lifestyle is its primary mode in nature and culture. (*c*) This ML reconstruction is based on the 18S rRNA gene, for which sequences from a greater number of marine genera and species are available than genomic data. Representatives of major Viridiplantae lineages were included with rooting by other Archaeplastida (specifically, rhodophytes and glaucophytes; not shown). The alignment comprised 130 representative species and 1,524 nt positions after masking (positions with less than 10% gaps). A rapid bootstrap analysis and search for the best-scoring ML tree was performed using RAxML (154) under the GTR+G+I evolution model. Colors of bars on the right follow bars in panel *a*. Abbreviations: ANA, Amborellales, Nymphaeales, Austrobaileyaceae; BS, bootstrap support; ML, maximum likelihood; ND, not determined; RAxML, Randomized Axelerated Maximum Likelihood; rRNA, ribosomal RNA.

Prasinophytes:
a paraphyletic group of geographically widespread unicellular green algae that are primarily marine and also part of the Viridiplantae

Primary producers:
organisms that use light, CO_2, and H_2O in the process of oxygenic photosynthesis, generating reduced carbon compounds rich in chemical energy; they form the base of the food web

have been co-opted and/or expanded as plants adapted to life on land, and new functions evolved to allow survival under differing stressors (129, 141).

As on land, food availability in the oceans is directly connected to photosynthetic productivity—in this case by eukaryotic algae and cyanobacteria. Marine primary productivity not only is important to ocean-dwelling organisms but also underpins fisheries and their harvests worldwide (17, 78, 174). The prasinophytes, for which several genomes (30, 127, 134, 175) and many transcriptomes (82) have been sequenced, are of particular interest, as they are globally distributed in marine environments. They are the dominant green algae in the oceans, remarkable not just for being important primary producers (9, 101, 104, 176) and potential sentinels for ocean change (29, 175) but also, from a cell-biological perspective, for being the tiniest plants yet discovered (21, 176). The most abundant members are as tiny as ≤2 μm in cell diameter, the size of a typical bacterium! Their diminutive size is important because nutrient availability (e.g., nitrogen, phosphorus, iron) is often low in the open ocean, hence higher cell surface area to cell volume ratios result in a competitive advantage over larger photosynthetic cells in such environments (20). The chloroplast itself generally occupies approximately half of the cell. Most prasinophytes have scales (see *Bathycoccus* image, **Figure 1a**) and most are motile, having a motile cilium, or often two motile cilia (see the sidebar titled Motility in Green Algae Relies on Cilia, a.k.a. Flagella), whose

MOTILITY IN GREEN ALGAE RELIES ON CILIA, A.K.A. FLAGELLA

Until recently, the term flagellum was used for the cilium in prasinophytes and most motile algae and protists. Recognition that this organelle is conserved across all eukaryotes but is different than the flagellum of bacteria has led to a shift to the usage of cilium for the eukaryotic organelle. The term flagellum signifies a motile cilium in contrast to the primary, a.k.a. nonmotile, cilia that have sensory functions in mammalian cells. The word flagellum is still widely used for the eukaryotic organelle in oceanographic and ecological literature, and typically refers to longer cilia that occur in numbers of one or two per cell in unicellular organisms.

structure is conserved within all eukaryotes (36, 51, 156). Beyond food security, the photosynthetic activities of algae (including cyanobacteria) in the oceans and plants on land are each responsible for about half of the annual uptake of CO_2 from the atmosphere (42). Moreover, oceans have served as a long-term sink for this CO_2: Its continuance as a reservoir is intimately connected to environmental conditions such as temperature (34, 55).

What can a tiny prasinophyte alga tell us about a giant redwood tree? Many key functions of photosynthetic cells appear to be conserved across the entire Viridiplantae. Some conserved proteins have altered in function as algal and plant lineages diverged, and some proteins have patchy distributions across the individual Viridiplantae lineages sequenced to date. A great deal can be learned by comparing protein families and their functions in land plant and algal groups. Prasinophytes are excellent model organisms since they have tiny genomes relative to plants (**Figure 1*b***), low gene family expansion, few repeat regions, and fast division times (30, 127, 134, 175, 178). Nascent approaches for genetic manipulation are also available for several (40), including *Micromonas*—a widespread, motile, scaleless genus. Furthermore, they can easily be synchronized to a diel light cycle, such that cells move through the same cell cycle phase together, sharpening transcriptional profiles and statistical analyses.

The idea that green algae can contribute fundamental insights into the evolution of plants as a whole is not new. Decades ago, the unicellular chlorophyte freshwater alga *Chlamydomonas reinhardtii* (120), with simple cultivation requirements and available methods for genetic manipulation, was already being used to facilitate the understanding of basic pathways and cell wall construction in land plants (54, 128). *Chlamydomonas* is capable of purely heterotrophic growth via the uptake of organic carbon compounds, and therefore does not require light for growth, allowing the isolation of mutants blocked in photosynthesis. This is unlike prasinophytes (120, 161) and the newly sequenced prasinodermophytes (100), at least based on knowledge to date. Comparative studies across these green algal lineages and multicellular plants hence provide a window into the characteristics of the progenitor of the Viridiplantae and illustrate expansion of certain protein families, accompanied by functional modifications and losses, that were essential to the evolution and diversification of plants (131).

As we face major changes to land and sea environments, a cell-biological understanding of different plant and algal species rooted in evolutionary knowledge is essential to more accurately predict potential future transitions in the Viridiplantae communities responsible for the majority of global photosynthesis. Here, we review the latest understanding of relationships between plant and green algal groups and the corresponding presence/absence patterns of orthologous groups of proteins. Note that because chlorophytes (e.g., *Chlamydomonas reinhardtii*) have been extensively examined (see, e.g., 8, 161) and are not among the more abundant marine green algae, we allot more attention to comparing prasinophytes and prasinodermophytes to streptophytes. We focus on advances in knowledge of three specific features shared between these algae and plants: phytochromes, peptidoglycan biosynthesis, and carbohydrate-active enzymes (CAZymes). All three are important in plant biology but with differing levels of significance between lineages, as illustrated by studies on function as well as evolutionary patterns of retention, expansion, and loss. These events and their variations between lineages underpin the long-term success of the Viridiplantae.

GETTING INTO THE NITTY-GRITTY: EVOLUTIONARY RELATIONSHIPS ACROSS THE VIRIDIPLANTAE

The Viridiplantae, informally called the green lineage, is a well-established monophyletic group of eukaryotic organisms comprising green algae and land plants (93, 97). It is one of the richest

Diel light cycle:
24-hour periodicity associated with a day/night light/dark regime

Chlorophytes:
a Viridiplantae green algal lineage that contains unicellular, colonial, and multicellular taxa; largely freshwater-associated but also in other environments

Phytochrome: a type of light-responsive photoreceptor; typically found in plants, but also in some nonphotosynthetic organisms

Peptidoglycan:
a polymer of sugars and amino acids that forms a mesh-like wall layer around bacteria, maintaining cell shape

Carbohydrate-active enzyme (CAZyme):
an enzyme acting on carbohydrates and generally categorized according to its global activity, e.g., glycoside hydrolase, which hydrolyzes glycosidic bonds between carbohydrates

Archaeplastida:
eukaryotic supergroup defined by having so-called primary plastids directly derived from the primary endosymbiosis event

Streptophyta: a broad lineage within the Viridiplantae that includes both streptophyte green algae, such as *Chlorokybus*, and land plants, such as *Brassica*

Prasinodermophyta: a Viridiplantae lineage recently delineated from the prasinophytes as a phylum-level division that contains both unicellular and multicellular algae

eukaryotic groups with an estimated half a million land plant species (24, 136) to which an estimated ~20,000 algal species can be added (58). The plastids of the Viridiplantae originated in an ancestor of extant Archaeplastida (which also includes the Rhodophyta, i.e., red algae, and Glaucophyta) by the engulfment of a cyanobacterium and was a pivotal event in the history of life—the emergence of photosynthetic eukaryotes (1, 118). Members of the green lineage are characterized by the presence of chloroplasts (plastids) with a double-membrane envelope, thylakoids grouped in lamellae and containing chlorophylls *a* and *b*, and intraplastidial starch (92). That said, the Viridiplantae are morphologically extremely diverse, and they have colonized almost all habitats on Earth's surface, both aquatic and terrestrial. Within Viridiplantae, analysis of the current genomic data indicates three early-diverging groups: the Streptophyta; the Chlorophyta (131, 170); and the recently recognized Prasinodermophyta (100), which appears, based on very limited taxonomic sampling, to have diverged before the split of the Chlorophyta and Streptophyta.

Within the Streptophyta, basal clades are freshwater algae, and their evolution occurred alongside adaptation to life on land (26, 27). Phylogenetic analyses recovered *Mesostigma*, *Chlorokybus*, and *Spirotaenia* in a clade that is sister to the remainder of streptophyte algae (131, 170). Successive divergence then gave rise to the vast diversity of land plants comprising bryophytes, lycophytes, ferns (polypodiophytes), and seed plants (which include both gymnosperms and flowering plants) (**Figure 1a**).

The Chlorophyta group includes the monophyletic core chlorophytes and the paraphyletic prasinophytes, which each harbor considerable diversity. In phylogenomic analyses, the prasinophyte marine algal classes Mamiellophyceae and Pyramimonadophyceae tend to group as a sister clade to the core clade of chlorophytes made up of Trebouxiophyceae, Ulvophyceae, and Chlorophyceae (**Figure 1b**), but this result is dependent on the taxonomic sampling (96, 100). Phylogenetic reconstructions based on 18S ribosomal RNA (rRNA) gene sequences allow for more comprehensive taxonomic sampling because this gene is recovered in polymerase chain reaction (PCR)-based studies of both lab-cultured algae and algae in their native habitats. 18S rRNA analyses with expanded taxonomic sampling typically do not conclusively resolve prasinophytes and chlorophytes as two overarching clades within Chlorophyta. Rather, the prasinophytes are often recovered as multiple polyphyletic lineages in basal positions to the core chlorophytes (**Figure 1c**). Thus, the name prasinophyte refers to early classifications describing single-celled algae covered with scales (114) and is still used to informally designate nine clades (among which two are still uncultured) (159), but it does not refer to a phylogenetically resolved lineage. Phylogenomic analyses with improved sampling based on transcriptomic or genomic data from marine algae, cultured or not, are needed to shed light on the branching of the numerous lineages in the Chlorophyta phylum. Note that in this review, we (disputably, but for simplicity) consider the Chloropicophyceae (represented by the genome of *Chloropicon primus*) and the Picocystophyceae (represented by the genome of *Picocystis* sp. ML1) as additional members of the core chlorophytes (**Figure 1c**), while the Mamiellophyceae (represented by genomes of *Bathycoccus prasinos*, *Ostreococcus tauri*, *Ostreococcus lucimarinus*, *Ostreococcus* sp. RCC809, *Micromonas pusilla*, and *Micromonas commoda*) and the Pyramimonadophyceae are considered as prasinophytes. However, multiple clades within this region of the tree have no genome-sequenced members, and node support is not retained along the backbone of the tree (**Figure 1c**); thus, relationships and groupings should be reconsidered after analyses that achieve improved taxonomic sampling. In fact, it seems likely that the Chloropicophyceae, Picocystophyceae, Pycnococcaceae, and Nephroselmidophyceae will prove to be something other than either chlorophytes or prasinophytes, and that the Chlorophyta as a whole will undergo revisions.

While plants have held a tremendous place in science, literature, culture, and art, even millennia before the conception of Persephone's trials and the sorrow of Demeter, it was significantly

after the 1933 discovery of green microalgae and their role as a food source for bivalves (48) that they were recognized as important in the global oceans. Published studies then emerged describing the marine picophytoplankton, specifically a tiny prasinophyte alga now termed *Micromonas* (15). Thus, Mamiellophyceae were among the earliest marine green algae discovered, and their high abundances were first reported in 1951 (84). Today a plethora of oceanographic studies have pursued the prasinophyte distributions (104, 126, 160), activities (3, 110, 176), and diversity, initially with environmental clone libraries (57, 173), and later with amplicon sequencing (e.g., 9, 60, 126, 159). Many of these algae are ciliated and thus have the ability to take back some control of their own destiny, at least at the microscale (keeping in mind that viscous forces dominate over inertial forces for small cells in water), responding to local gradients in the chemical and physical environments (2, 56). The tiny size of these algae, and the fact that they cannot be distinguished using epifluorescence microscopy, likely led to the early focus on larger algae (such as the secondary plastid-bearing diatoms or dinoflagellates). With improved methodology and spatiotemporal sampling of the oceans, the influence of prasinophytes under multiple environmental scenarios has been increasingly recognized (9, 29, 133) and has contributed to the delineation of species (4, 22, 28, 150, 151). By analyzing available sequence data from survey studies, we can see that just three picoplanktonic Mamiellophyceae genera comprise a significant fraction of data from all eukaryotes captured by filtering seawater, with *Micromonas* and *Bathycoccus* ranging from the tropics to poles and *Ostreococcus* from tropical to temperate settings (**Figure 2a**). Moreover, it was recently discovered that *O. lucimarinus* and *Micromonas* spp. are dominant algae in the annual phytoplankton bloom in the North Atlantic (9), which is important to many fisheries and has classically been attributed to diatoms.

Until 2021, members of a third Viridiplantae phylum, the Prasinodermophyta, were considered prasinophytes (i.e., the Prasinococcales), despite the fact that they often branched deep within the Chlorophyta based on 18S rRNA gene (35) and plastidial gene analyses (94). Prasinodermophytes are now divided into two orders, the Palmophyllales and Prasinococcales (represented herein by the genome of *Prasinoderma coloniale*) (100). Prasinococcales have been noted in coastal ecosystems by molecular methods, but at low read numbers (160), and so far are characterized as typically being small (2.2–5.5 μm) coccoid cells lacking scales (63, 123). By contrast, the Palmophyllales are thalloid algae seemingly adapted to dimly lit environments (down to 200 m depth), raising fascinating questions about the nature of the green plant ancestor (95). Based on environmental PCR-based surveys, prasinodermophytes are globally distributed but much less abundant than prasinophytes (**Figure 2**).

As noted above, all of the relationships in the early-branching regions of the Viridiplantae tree are influenced by uneven taxon sampling (81, 82), especially prasinophytes, prasinodermophytes, and lineages basal to the core chlorophytes. Each harbors groups that are still uncultured and lack sequenced genomes. From an environmental perspective, improved sampling is important because some of these uncultured groups are reported frequently in the ocean, e.g., the so-called prasinophyte clades VIII (168) and IX (148). Even some that are cultured have yet to have their genomes sequenced, such as *Prasinococcus* and *Pyramimonas* species (Prasinodermophyta and Pyramimonadophyceae, respectively; **Figure 1c**). Moreover, even within the genus *Micromonas*, which has multiple cultured isolates from some clades, other clades have long been known that remain uncultured (149, 173), and, reportedly, new uncultured clades exist as well (151). Transcriptome sequencing of existing cultures has expanded in the last several years (82, 131). Nevertheless, the incompleteness of taxonomic sampling, and possible extinctions over the millennia, still hinder resolution of the Viridiplantae tree and affect studies of molecular attributes underpinning physiological and niche differences.
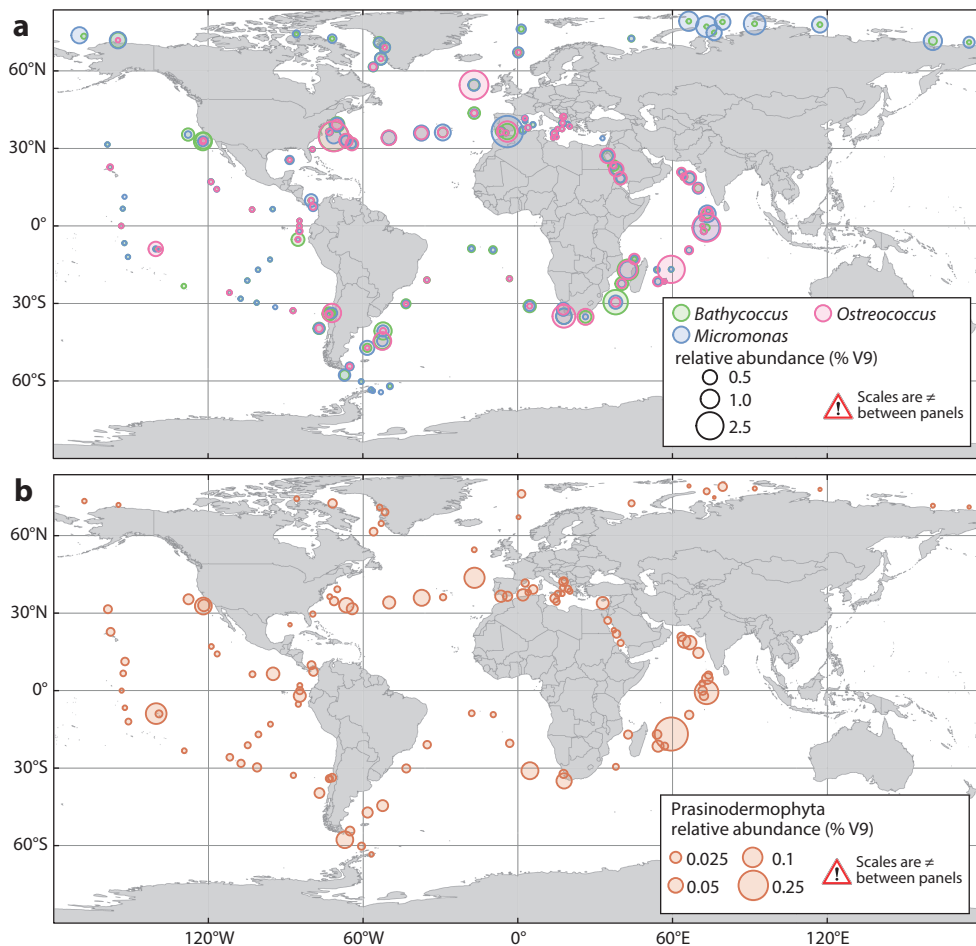
Figure 2

Distributions of prasinophytes and prasinodermophytes in surface ocean survey data. (*a*) Relative abundances
of three prasinophyte genera within the Mamiellophyceae and (*b*) relative abundance of prasinodermophytes
(summed values from both classes due to low numbers). Note the different circle size scales between panels *a*
and *b*. For both panels, V9 18S rRNA gene amplicons from *Tara Ocean*s data (69) were analyzed as ASVs,
which we classified through maximum likelihood placement using a PhyloAssigner (167) database
constructed from the phylogenetic reconstruction in **Figure 1c**. Circle size represents the relative abundance
of ASVs from the respective group, as percent of the total number of V9 amplicon sequence reads per
sample (the total includes amplicons from heterotrophic and photosynthetic eukaryotes). Abbreviations:
ASV, amplicon sequence variant; rRNA, ribosomal RNA.

FINDING THE ANCESTOR: PROTEIN DISTRIBUTIONS ACROSS LAND AND SEA GREENS

The plant–algal progenitor can be at least partially reconstructed by drawing upon existing
genome- and transcriptome-level data collections (e.g., 52, 53). These can be used to identify
conserved proteins that may have been lost (more difficult with transcriptomes) in some lineages
but retained in others. With sufficient depth or careful taxonomic sampling of specific lineages, the
appearance of gain can be differentiated from loss. Such discoveries do not ensure that functions
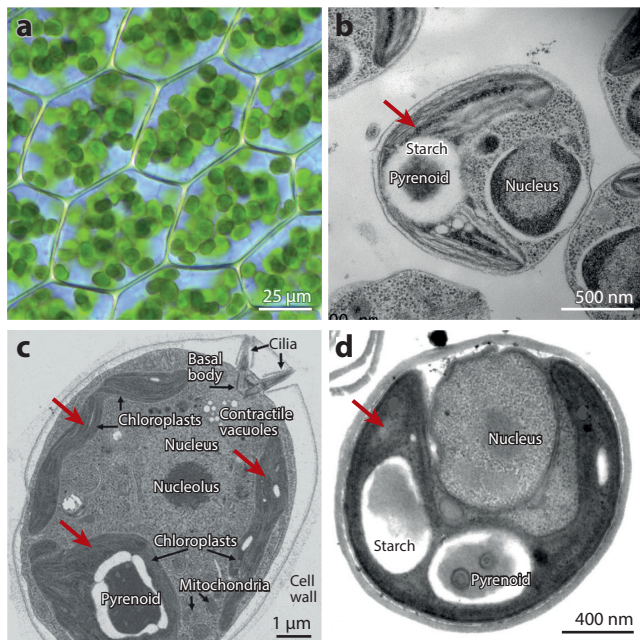
Figure 3

Up close with the Viridiplantae. (*a*) Plant cells within the moss *Plagiomnium affine* imaged using optical microscopy. The cells contain multiple chloroplasts (*green bubbles*) and are adjacent to each other with a wall around them. Image adapted from Kristian Peters/Fabelfroh (**https://commons.wikimedia.org/wiki/ File:Plagiomnium_affine_laminazellen.jpeg**) (CC BY-SA 3.0). (*b*) A *Micromonas polaris* cell imaged using high-pressure freezing and transmission electron microscopy (TEM). These free-living marine unicells have one chloroplast (*arrow*), filling much of the cell; one mitochondrion; and one nucleus. The cilium's structure (out of view) and genes encoding it are conserved with those in other eukaryotes (see, e.g., 36, 135), including the sperm of some multicellular plants, such as the bryophyte *Marchantia polymorpha* (71) and the fern *Pteris vittata* (87). (*c*) A *Chlamydomonas reinhardtii* cell TEM micrograph prepared by high-pressure freezing followed by freeze substitution and embedded in resin (micrograph adapted with permission from Reference 132), showing the cell wall, two cilia emerging from the basal bodies in the apical region, and chloroplasts (*arrows*) distributed throughout the cell. (*d*) *Prasinoderma coloniale* cell observed by TEM (micrograph adapted with permission from Reference 76) showing coccoid morphology with one chloroplast forming a large portion of the cell volume as also seen in prasinophytes (76).

have remained the same since divergence; indeed, we know that in many cases functions have also diverged.

Collectively, processes such as gain, loss, co-option, and expansion underlie diversification in the Viridiplantae. If one zooms out from the images shown in **Figure 1*a***, the overall structure of multicellular green organisms—ranging from chlorophyte algae such as *Ulva* to more familiar taxa such as ferns, *Nymphaea* (water lilies), and *Mammillaria* (cacti)—is tremendously varied. From an intracellular perspective, some similarities persist. For example, in all Viridiplantae, plastids are derived as described above, but the numbers present in each cell vary between lineages and can be influenced by multiple factors (47). Just among the cells from the four selected taxa shown in **Figure 3**, we see variations in plastid numbers. Reported numbers also vary within individual leaf mesophyll cells, such as in the *Arabidopsis thaliana* Columbia ecotype, where published numbers range from 30 to 100 (47). Differences are also seen in the number of cilia. *Chlamydomonas* has two cilia, while some prasinophytes have none (*Ostreococcus*), one (*Micromonas*), or two or more
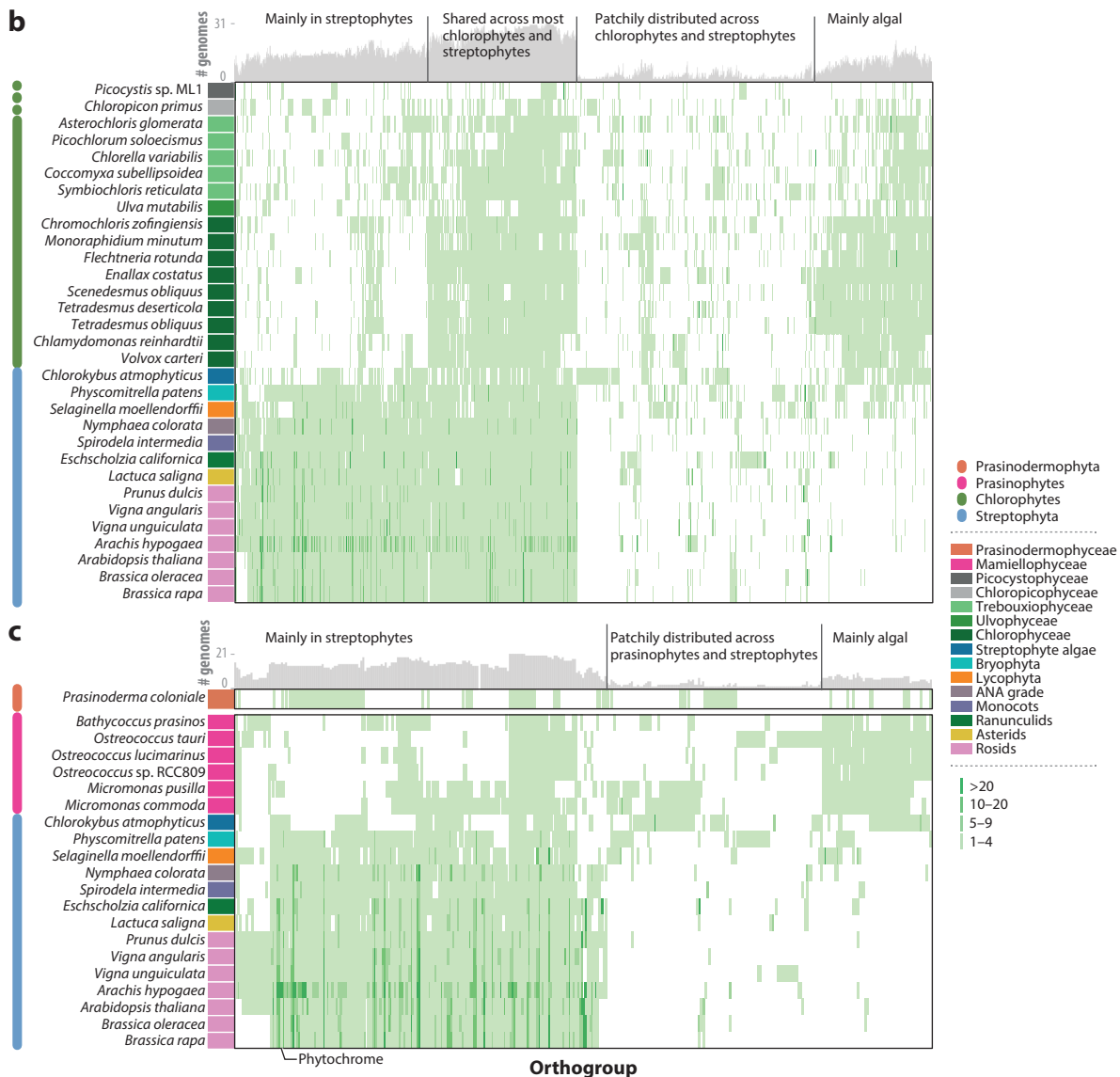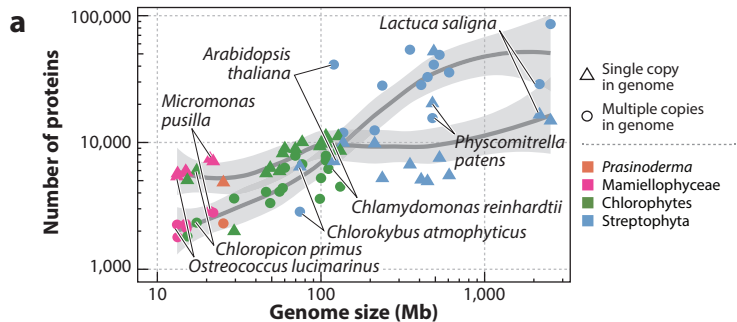
(other genera), and prasinodermophytes appear to have none. *Arabidopsis* also has no cilium, but zoospores from streptophyte algae such as *Mesostigma viride* and *Chlorokybus atmophyticus* each have two cilia, and some multicellular plants have ciliated sperm cells, such as the bryophyte *Marchantia polymorpha* (71) and fern *Pteris vittata* (87).

One way to identify proteins that are potentially important to differentiation processes, regardless of whether or not their function is known, is to compare orthologous protein groups. This is often done when comparing closely related species and is most accurate when excellent gene models are available, such that the sequence of the protein encoded is complete. Problems arise if small domains are present at the N or C terminus that have significant consequences for function but are too small to disrupt overlap criteria and orthogroup formation. For the prasinophytes, specifically the Mamiellophyceae, several studies have classified orthogroups (127, 134, 175). In the two *Micromonas* species with complete genome sequences (175), *Micromonas pusilla* and *Micromonas commoda*, 8,000 of approximately 10,000 genes in each genome appear to be contained in shared orthogroups (162). These two algae have unexpectedly different gene and genome structures (149, 162, 175). Additionally, *Micromonas pusilla* has no classically recognizable transposable elements (TEs), while *Micromonas commoda*, other Mamiellophyceae, and (to our knowledge) all other sequenced plants do have TEs. While differences in gene content identified between the two species provide a springboard for deeper investigation of differentiation and evolutionary processes that extend to plant biology (117, 162), the fact that some *Micromonas* clades lack cultured representatives signifies that there is much more to learn about diversification and growth requirements of other members of this diverse genus.

From a phylogenomic perspective, orthologous groups have been inferred in many studies with the hope to identify features that contribute more broadly to diversification or core attributes. Orthogroups have been made available from photosynthetic eukaryotes using predicted proteomes considered to be good quality, with attention to conserved proteins involved in photosynthesis and its regulation (54, 162, 163). Other studies highlight findings more relevant to understanding diversification processes in the Viridiplantae. Orthogroup analysis of the fern *Ceratopteris richardii* alongside predicted proteomes from 38 other Viridiplantae led to identification of a GRAS domain protein (an important superfamily of regulatory proteins) that is specific to *Ceratopteris richardii* (50). Analyses using transitional Viridiplantae lineages that possess both ancestral and derived plant traits, such as liverworts, mosses, and ferns, provide a step toward understanding broader plant diversification (11, 89, 98). Moreover, orthology inferences have been used to examine evolutionary processes in crops, such as a heat shock transcription factor gene family shown to be expanded in 13 cotton species by comparing cotton and *A. thaliana* (102).

To improve our understanding of Viridiplantae evolution and developmental processes as a whole, past studies on plant and algal orthologous protein families can be expanded upon by comparing shared presence/absence patterns between the different major groups of green algae and streptophytes. From a total of 28,494 orthogroups comprised of proteins from streptophytes, chlorophytes, prasinophytes, and prasinodermophytes identified for this review, 13% were found that had at least one protein from an organism in each of the four major groupings. Given the low genome availability for prasinodermophytes, we also analyzed these numbers with prasinodermophytes included with prasinophytes, leading to an increase from 13% to 19% being shared across the major groupings. Streptophytes alone (including *Chlorokybus atmophyticus*) formed 36% of orthogroups, while algae alone (excluding *Chlorokybus*) formed 38%. Although the *Chlorokybus atmophyticus* genome size (74 MB) and protein number (9,300; **Figure 4a**) are the smallest among streptophytes, its genome size is larger than those of the picoprasinophytes (**Figure 1b**), and the protein number is similar. Across orthogroups, it had fewer proteins (1.3 on average) in

---

Orthogroup: a set of protein sequences that are deduced as having derived from a single gene in a last common ancestor based on amino acid identity and protein sequence overlap

a

- Number of proteins (y-axis): 1,000 / 10,000 / 100,000
- Genome size (Mb) (x-axis): 10 / 100 / 1,000

Labeled species: *Lactuca saligna*, *Arabidopsis thaliana*, *Micromonas pusilla*, *Physcomitrella patens*, *Chlamydomonas reinhardtii*, *Chlorokybus atmophyticus*, *Chloropicon primus*, *Ostreococcus lucimarinus*

Legend:
△ Single copy in genome
○ Multiple copies in genome

- Prasinoderma
- Mamiellophyceae
- Chlorophytes
- Streptophyta

b

genomes: 31 – 0

- Mainly in streptophytes
- Shared across most chlorophytes and streptophytes
- Patchily distributed across chlorophytes and streptophytes
- Mainly algal

Species (top to bottom):
Picocystis sp. ML1
Chloropicon primus
Asterochloris glomerata
Picochlorum soloecismus
Chlorella variabilis
Coccomyxa subellipsoidea
Symbiochloris reticulata
Ulva mutabilis
Chromochloris zofingiensis
Monoraphidium minutum
Flechtneria rotunda
Enallax costatus
Scenedesmus obliquus
Tetradesmus deserticola
Tetradesmus obliquus
Chlamydomonas reinhardtii
Volvox carteri
Chlorokybus atmophyticus
Physcomitrella patens
Selaginella moellendorffii
Nymphaea colorata
Spirodela intermedia
Eschscholzia californica
Lactuca saligna
Prunus dulcis
Vigna angularis
Vigna unguiculata
Arachis hypogaea
Arabidopsis thaliana
Brassica oleracea
Brassica rapa

Legend:
- Prasinodermophyta
- Prasinophytes
- Chlorophytes
- Streptophyta

- Prasinodermophyceae
- Mamiellophyceae
- Picocystophyceae
- Chloropicophyceae
- Trebouxiophyceae
- Ulvophyceae
- Chlorophyceae
- Streptophyte algae
- Bryophyta
- Lycophyta
- ANA grade
- Monocots
- Ranunculids
- Asterids
- Rosids

- >20
- 10–20
- 5–9
- 1–4

c

genomes: 21 – 0

- Mainly in streptophytes
- Patchily distributed across prasinophytes and streptophytes
- Mainly algal

Species (top to bottom):
Prasinoderma coloniale
Bathycoccus prasinos
Ostreococcus tauri
Ostreococcus lucimarinus
Ostreococcus sp. RCC809
Micromonas pusilla
Micromonas commoda
Chlorokybus atmophyticus
Physcomitrella patens
Selaginella moellendorffii
Nymphaea colorata
Spirodela intermedia
Eschscholzia californica
Lactuca saligna
Prunus dulcis
Vigna angularis
Vigna unguiculata
Arachis hypogaea
Arabidopsis thaliana
Brassica oleracea
Brassica rapa

Phytochrome

Orthogroup

(Caption appears on following page)

Figure 4 (*Figure appears on preceding page*)

Patterns of presence or absence of orthologous protein(s) in the Viridiplantae. (*a*) Within-genome assessment of whether an orthogroup has one member (*triangle*) or multiple copies (i.e., forming an orthogroup; *circle*) with numbers representing protein counts in each of these categories. (*b*) Orthogroups shared between at least one chlorophyte and one streptophyte and not present in any prasinophyte (n = 1,671). (*c*) Orthogroups shared between at least one prasinophyte and one streptophyte and not present in chlorophytes (n = 343). The prasinodermophyte is shown in this panel due to high similarities with prasinophytes (and because of the few orthogroups uniquely present between this particular genome and streptophytes). (*a–c*) Orthogroups were computed using OrthoFinder (39) with a significance threshold of $<1 \times 10^{-03}$, and barplots along heatmap tops indicate the number of genomes in which an orthogroup was found at least once. Orthogroups were clustered based on Euclidean distance of presence/absence patterns. Note that 38 high-quality Viridiplantae genomes and predicted proteomes were analyzed and sorted based on the taxonomy in **Figure 1**. Shades of green indicate the number of copies per orthogroup and genome pair (numbers ≥20 are reflected by the same green shade). Databases presented in References 52 and 53 were used to collect genomes and protein annotations analyzed here.

each than occurred in other streptophytes (3.9 on average) and chlorophytes (1.5) but more than in prasinophytes (1.2 on average).

Protein distribution comparisons revealed that orthogroups representing expanded families where multiple orthologs were present from an individual organism increased disproportionately in multicellular streptophytes (**Figure 4a**). This did not occur for the multicellular algae *Ulva mutabilis* and colony-forming *Volvox carteri*, both of which had low numbers of proteins in expanded families, even when compared to other chlorophytes. Additionally, no orthogroups were detected that were shared only between multicellular chlorophytes and multicellular streptophytes, illustrating the multiple possible evolutionary routes to multicellularity (120, 161). Finally, some orthogroups appear to be purely algal—even some of those that are shared across the four primary organismal groupings compared herein. For example, the streptophyte alga *Chlorokybus* shares distinct orthogroups with chlorophytes, and others with prasinophytes, that are not present in sequenced multicellular streptophytes (**Figure 4b,c**). Presumably, these findings will provide insights into distinct factors required for life as an alga, present in the green algal progenitor but subsequently lost from land plants and, in some cases, differentially lost from chlorophytes or prasinophytes.

Other orthogroups shared across the organismal groupings highlight factors that have been differentially lost from one or more major groups but were presumably present in the Viridiplantae progenitor. With respect to prasinodermophytes, just 51 orthogroups were identified that were exclusively shared with streptophytes, most of which lack clear functional classifications. Chlorophytes and streptophytes share 1,331 orthogroups, based only on the genomes analyzed herein (**Figure 4b**). Those shared exclusively between prasinodermophytes, prasinophytes, and streptophytes number 343 and include the photosensor phytochrome, a master regulator in extant land plants (**Figure 4c**).

SENSING AND RESPONDING TO THE LIGHT ENVIRONMENT: PHYTOCHROMES

Plants and algae—and even other nonphotosynthetic eukaryotes—use photosensory proteins as switches that convert light information into biological signals. Phytochromes are among these and comprise a photosensory module (PCM) at the N-terminal region of the protein and a histidine kinase output module (OPM) at the C terminus (**Figure 5**). Phytochrome activities impact many light-dependent aspects of streptophyte development, including germination, seedling photomorphogenesis, flowering time, and shade avoidance (13, 19, 32, 35). In most streptophytes, phytochrome has undergone considerable gene family expansion but, again, with a generally conserved architecture where the number of OPMs can occasionally vary (19). The PCM contains
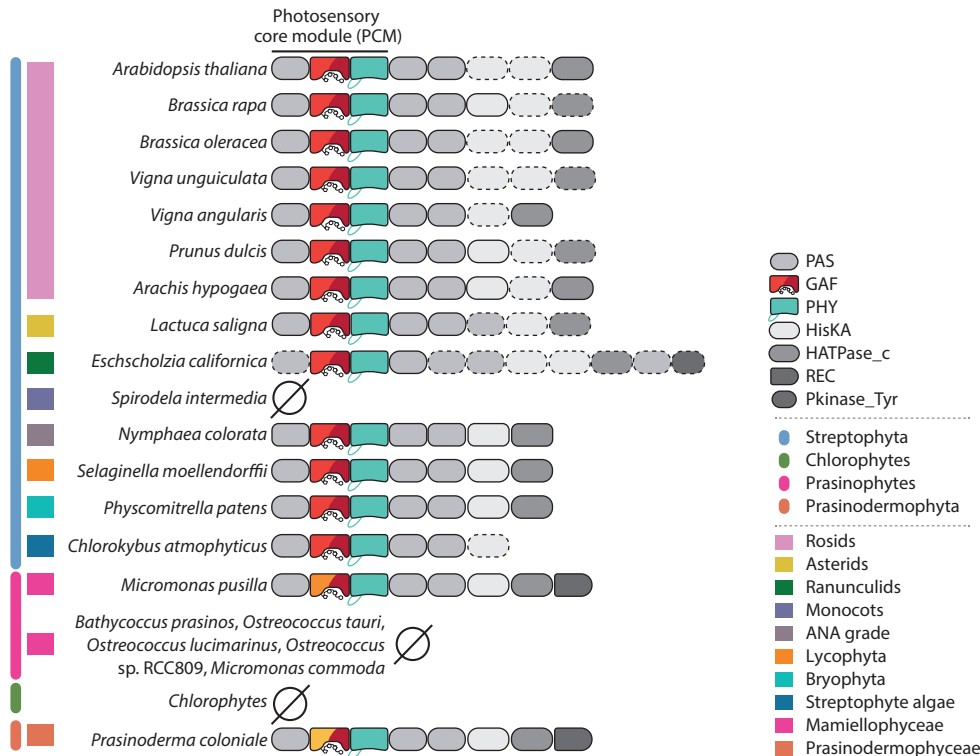
Figure 5

Domain structures and light-sensing of phytochromes. The N-terminal PCM of phytochromes is composed of PAS, GAF, and PHY domains. Colors on the chromophore-binding GAF domains correspond to those of the two reversibly photointerconverting states of each phytochrome where known; those in prasinophytes and prasinodermophytes examined thus far through biochemical studies perceive light wavelengths (orange-shifted) that penetrate deeper in seawater than red wavelengths (35, 142), as indicated by domain color coding. C-terminal output modules of phytochromes from all Viridiplantae lineages typically contain one or two PAS domains adjacent to HKMs composed of HisKA and HATPase_c domains. A lack of C-terminal REC domains in streptophyte phytochromes contrasts with their presence in prasinophyte phytochromes. Dashed outlines indicate domains that are not always observed in predicted proteins from genome projects. Null symbols indicate taxa in which phytochrome is lacking in genome data available to date. Note that data presented here are based on presence/absence detection from the 38 high-quality Viridiplantae genomes/ predicted proteomes analyzed in **Figure 4**. Abbreviations: GAF, cGMP phosphodiesterase/adenylate cyclase/FhlA; HATPase_c, ATPase domain of the histidine kinase; HisKA, dimerization/phosphorylation domain of the histidine kinase; HKM, histidine kinase module; PAS, Per/Arnt/Sim; PCM, photosensory core module; PHY, phytochrome; Pkinase_Tyr, Protein tyrosine kinase; REC, response regulator receiver.

PAS, GAF, and PHY domains, and this is where covalent interaction with a heme-derived bilin chromophore occurs (**Figure 5**). Light perception by phytochromes relies on this chromophore and in streptophytes function via red/far-red wavelengths: In the dark-adapted state, phytochromes absorb red light and, in the light, a photoisomerization reaction leads to far-red absorption. In this manner, plants are able to detect the light environment, and downstream cascades lead to the regulation of different processes and, ultimately, essential responses such as shade avoidance.

The first green algae to be sequenced, the chlorophyte *Chlamydomonas* and the prasinophyte *Ostreococcus* (30, 120), lacked phytochromes altogether (35). Thus, phytochromes might have been

PHY: the phytochrome-specific domain, one of three N-terminal domains in the photosensory core module of the phytochrome

thought to be a later-stage invention or the result of the co-option of a horizontally transferred cyanobacterial phytochrome-like protein at the time streptophytes diverged. They have tremendously important, well-studied roles in plants. However, earlier presence of phytochromes in the Viridiplantae was discovered upon analysis of the genome-sequenced prasinophyte *Micromonas pusilla* alongside transcriptome sequencing of other prasinophytes and glaucophytes (35). Our analysis of the *Prasinoderma coloniale* genome data (100) confirmed the structure of the prasinodermophyte phytochrome (**Figure 5**) deduced from earlier transcriptomic data (35). These results collectively indicate that phytochrome was already present in the green algal progenitor of land plants. The domain structure observed in major model land plants, such as *Arabidopsis* and *Physcomitrella* (moss), has since been found to be conserved in liverworts, hornworts and charophyte algae (99). Like plant phytochromes, non-streptophyte Archaeplastida algal phytochromes contain core light input and signaling output architecture (**Figure 5**). However, given architectural similarities between glaucophyte, prasinophyte, and prasinodermophyte phytochromes, the structure in Archaeplastida as a whole prior to the divergence of these lineages appears to have included one or more C-terminal response regulator receiver (REC) domains, subsequently lost from canonical streptophyte OPMs (35, 143).

Phytochromes actually span diverse eukaryotic lineages, extending beyond the Viridiplantae into other eukaryotic supergroups, including algal members of the Heterokonta and Cryptophyta, as well as fungi (35, 41, 99). With respect to loss, patterns within the Viridiplantae are striking. Most prasinophytes have phytochromes, which we know from analysis of transcriptome data (35), but some diminutive genera whose genomes have been sequenced, especially *Ostreococcus* and *Bathycoccus*, and even one *Micromonas* species, *Micromonas commoda*, have lost them (**Figure 5**). Moreover, all core chlorophytes have lost phytochromes (35), a pattern upheld in analyses of genomes that have lately become available.

The function and roles of the different phytochrome proteins in plant development have been studied since their discovery (18, 19, 80, 143). Bilin photoisomerization, as described above, initiates downstream signaling events associated with translocation into the nucleus and interactions with transcription factors, specifically PHYTOCHROME-INTERACTING FACTORS (PIFs) (91). PIFs belong to the helix–loop–helix family of transcription factors and bind to the light-activated protein after translocation to the nucleus, where they are known to regulate many target promoters to repress light responses in the dark in plants. The inhibition or degradation of PIFs by phytochromes alters the expression of many genes to induce a range of responses to light (75). Homologs of PIFs are now known in mosses, ferns, bryophytes, and streptophyte algae, where their functions and phytochrome interactions are just beginning to be elucidated (137, 177). With respect to phytochrome proteins themselves in plants, they are abundant, which has complicated functional studies. For example, in *Arabidopsis*, five different phytochrome genes have been identified (19, 143), and the functions of Phytochrome A (PHYA) and PHYC are the best understood. These functional studies extend to many other agriculturally important plants; for example, recent proteomic analysis of the PHYA homolog in tomato indicates that it has a key role in establishing the mature seed proteome (157). Roles of PHYB in mitigating thermal and light stress have now been demonstrated for a variety of streptophytes (19), alongside variations in response levels dependent on the latitude to which a particular plant was adapted (70).

Relatively little is known about the role of phytochrome in unicellular algae. Some roles observed in plants are not relevant for algae due to inherent differences between organismal morphologies as well as environments inhabited. For example, light transmission differs in water and air, in that far-red light penetrates well through air but is rapidly attenuated in the first few meters of seawater. Thus, wavelength detection in prasinophytes is generally blue-shifted relative

Glaucophytes: an archaeplastidal lineage of freshwater algae colloquially referred to as blue-green plants

Eukaryotic supergroups: major groupings of eukaryotic lineages; originally 5–8 phylogenetically derived supergroups, but continually under revision as new data improve orphan lineage placement

Heterokonta: a group of photosynthetic (e.g., diatoms), mixotrophic, and heterotrophic protists as well as some macroscopic multicellular seaweeds

Cryptophyta: eukaryotic algal group with red algal–derived plastids; the latter are considered the source of most secondary and tertiary plastids present in non-Archaeplastida supergroups

to the red/far-red sensing in streptophytes, such that in *Micromonas pusilla* it has an orange/far-red photocycle (35). Other prasinophytes and the prasinodermophyte *Prasinoderma coloniale* also have blue-shifted photocycles, with the latter responding to yellow/far-red light (113, 142). Interestingly, though, a recent study of a phytochrome protein in diatoms, a widespread algal group within the Heterokonta, shows that they have a red/far-red responsive phytochrome, which is more similar to those of streptophytes. The authors postulated that chlorophyll fluorescence and Raman scattering deeper into the water column could generate red/far-red photons below the surface (46). Importantly, in *Arabidopsis*, different biophysical properties of PHYA and PHYB have been demonstrated (13) that presumably extend further to other members of this family and, by extension, likely exist in algal versions. A major challenge ahead is to characterize not only their global function in unicellular marine green algae, and the transcription factors with which they interact, but also modulation of their activation and downstream signaling under the massive range of light and temperature conditions experienced in the marine environment.

While the wavelengths perceived by algal phytochromes have been characterized (35, 142), the role(s) of phytochrome in algae are still mysterious. In prasinophytes, phytochrome proteins have been shown to localize to the nucleus, but the transcription factors with which they presumably interact have not been identified. Secondly, over a sampling of diel time points, researchers have shown that the gene exhibits the highest relative expression in the morning just before light appears, but protein localization to the nucleus occurs as the cell begins to replicate, later in the light period (35). Recombinant expression of phytochrome from another prasinophyte enabled production of proteins with light-regulated autophosphorylation activity, which was induced by the C-terminal histidine kinase modules (**Figure 5**), demonstrating that this alga has light-activated histidine kinases, similar to land plants (179). Notably, many studies examining plant phytochromes highlight their roles in triggering cascading responses to light and thermal stress and the interaction of these stressors (19, 32, 38, 70, 91, 138) and even a role for PHYB in promoting iron uptake (59). In the ocean, light and temperature can fluctuate rapidly, particularly in mixing scenarios, fronts, and other aspects of fluid environments, and iron is known to be a limiting growth factor in open ocean waters. Together with hypotheses that can be developed from plant phytochrome functions, advances in prasinophyte model systems for genetic manipulation (40) provide a springboard for establishing key phytochrome roles in marine taxa, possibly reflecting back to ancestral functions.

The transcriptional and protein dynamics of phytochrome observed for *Micromonas pusilla* in laboratory experiments (35) provide first steps toward targeting downstream interactions and cellular responses. However, the role of phytochromes in regulation and ecology most likely will not be fully understood until they are studied in natural marine habitats. This is likely the case not just for phytochrome but also for other algal photoreceptors, and the first metatranscriptomic field studies are only now becoming available. A caveat is that metatranscriptomic responses can be difficult to interpret without genomes from organisms present at the respective time and site, with the slow rate of sample collection and processing (such that expression might change), and with limits to the depth of sequencing (given that transcripts from many different organisms are present in each sample). Pronounced diel differential expression alongside extensive phylogenetic diversity has been documented for another class of photoreceptors, light-oxygen-voltage (LOV) domain–containing proteins (62), in a time-resolved study using robotic sampling in the eastern North Pacific Ocean (85). A follow-up study has confirmed similar patterns in a broader expanse of the Pacific Ocean (23). Nevertheless, the ultimate role of these proteins, as well as that of PHY, remains unclear in marine algae from both the green lineage and other eukaryotic supergroups.

Raman scattering: a phenomenon that occurs when photons interact with matter (e.g., seawater) and are re-emitted at different wavelengths

Metatranscriptomics: a method of characterizing the holistic gene expression profile of a microbial assemblage; difficult to interpret due to community complexity

THE REMARKABLY VARIED JOURNEY FROM CYANOBACTERIAL CELL WALL TO VIRIDIPLANTAE PLASTIDS

Murein polymerase/ transpeptidase: bidomain protein involved in peptidoglycan biosynthesis, composed of a GT51 peptidoglycan synthase that polymerizes sugar chains and a penicillin-binding protein transpeptidase that forms peptide cross-links

The most prominent feature retained from the cyanobacterial endosymbiont is, of course, the machinery for photosynthesis. Some of the genes involved are encoded by the chloroplast genome, and others have been relocated to the nucleus in a process termed endosymbiotic gene transfer (EGT) (1). Proteins encoded by such transferred genes must transit across the chloroplast envelope after translation in order to be integrated into their respective photosystems. The envelope presumably derives from the primary endosymbiont's cell wall, which (in cyanobacteria) has a peptidoglycan layer between the inner and outer membrane. However, the peptidoglycan biosynthesis pathway central to the cell wall of both cyanobacteria and other bacteria was long thought to have been lost during the evolution of plastids in many plant lineages.

Peptidoglycan biosynthesis by bacteria is well characterized, in part due to it being the specific target of many medical therapies (antibiotics) against bacterial pathogens of animals, such as the long-known penicillin (6). A core of 11 enzymes is involved in peptidoglycan biosynthesis in bacteria, including cyanobacteria. Seven core enzymes (MurA–MurF and MraY) are responsible for converting UDP-N-acetyl-D-glucosamine (UDP-GlcNAc) to acetylmuramyl-pentapeptide-pyrophosphoryl-undecaprenol or lipid I (37) (**Figure 6a**). Lipid II (GlcNAc-N-acetylmuramyl-pentapeptide-pyrophosphoryl-undecaprenol) is then formed by the elongation of lipid I's carbohydrate unit with one GlcNAc by MurG, which is a glycosyltransferase belonging to CAZy family GT28. The integral membrane protein MurJ has been identified as the lipid flippase in *Escherichia coli* and appears to be essential for peptidoglycan synthesis (139, 144). The nascent peptidoglycan chain is polymerized by a murein polymerase/transpeptidase comprising a CAZy family GT51 sugar polymerase domain and a penicillin-binding protein (PBP) domain, which cross-links the polysaccharide chains using peptides into the three-dimensional structure of the cell wall central layer (145) that appears as an electron-dense structure that is ~10-nm thick in cyanobacteria (103).

Among eukaryotes, glaucophyte algae, which belong to the Archaeplastida supergroup that also contains Viridiplantae, have long been known to maintain a visible peptidoglycan layer in their chloroplast wall (83). Evidence for peptidoglycan retention was lacking for the Viridiplantae for many years, and in early analyses of the *Chlamydomonas reinhardtii* (120) and *O. tauri* (30), genes for this pathway were not reported, possibly because it had not been observed in *Arabidopsis* and was assumed to no longer be relevant or present. Thus, until more recent studies, the mechanisms for plastid division or wall formation in the Viridiplantae previously appeared to have been dramatically changed from those of glaucophytes and cyanobacteria (83).

The view that peptidoglycan was absent from plants changed dramatically when the pathway was detected in the genomes of the bryophyte *Physcomitrella patens* (68, 111), the lycophyte *Selaginella moellendorffii* (155), and prasinophyte alga *Micromonas pusilla* (162). Comparative genomics revealed that these organisms have the canonical or nearly complete peptidoglycan biosynthesis pathway that is observed in cyanobacteria (**Figure 6b**). Treatment with antibiotics that act on various components of this pathway modified plastid characteristics in both *Micromonas* and the moss *Physcomitrella* (66, 162). However, until the elegant studies of Hirano et al. (66), who manipulated *Physcomitrella*, peptidoglycan had not been visualized in any plant. So far, it is still visually unobserved in prasinophytes and the prasinodermophytes wherein the complete pathway was also recently reported (100).

In many other Viridiplantae, the peptidoglycan pathway is very much reduced (**Figure 6b**), including all sequenced members of the chlorophytes, except the multicellular *U. mutabilis*. It is also highly reduced in most flowering plants (155). Unfortunately, our understanding of loss patterns of peptidoglycan in plant lineages is muddy due to publications involving inaccurate gene searches or literature reviews, such that its absence in algae was overstated and/or incorrect.
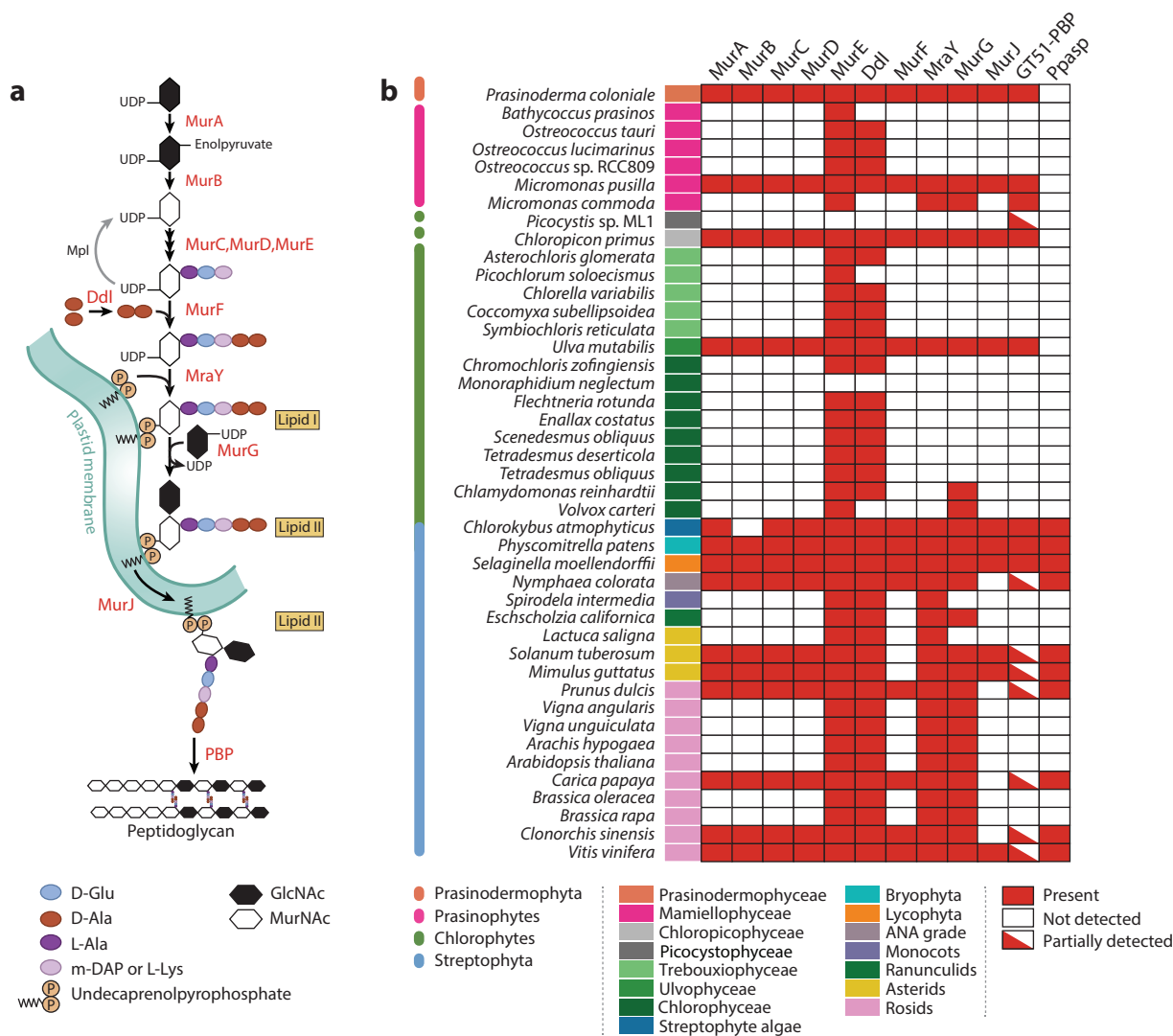
Figure 6

The patchiness of Viridiplantae peptidoglycan biosynthesis pathways. (*a*) Proteins involved in the peptidoglycan biosynthetic pathway and final cross-linking steps (162), updated using data from Reference 37. (*b*) Peptidoglycan pathway proteins in the Viridiplantae, with MurE, Ddl, MraY, and MurG present throughout. The full peptidoglycan pathway complement, with the exception of GT51-PBP, occurs in several major plant lineages. The final column indicates the gene for Ppasp, identified in Reference 162. Red boxes mean that the gene is present in the genome and white that the gene is not detected. Boxes that are half red and half white indicate partial detection with only the transglycosylase GT51 domain in the N terminus detected, without the C-terminal PBP transpeptidase. Note that the full pathway has been found in four prasinophyte lineages based on transcriptome sequencing (162) and hence is not displayed in this genome-sequencing-based figure. Additionally, gymnosperms with the complete pathway (105) were not included due to their large and still unresolved genome sequences and protein models. Abbreviations: ANA, Amborellales, Nymphaeales, Austrobaileyaceae; GlcNAc, *N*-acetylglucosamine; GT51, glycosyltransferase family 51; MurNAc, *N*-acetylmuramic acid; PBP, penicillin-binding protein; Ppasp, plant peptidoglycan–associated streptophyte protein.

The current status of knowledge on peptidoglycan biosynthesis in green organisms indicates that members of several prasinophyte groups, as well as the prasinodermophyte *Prasinoderma coloniale*, alongside two *Micromonas* species represent a more ancestral viridiplant state than do the most reduced Mamiellophyceae (i.e., *Ostreococcus* and *Bathycoccus*) and chlorophyte algae such as *Chlamydomonas reinhardtii* (120, 162) (**Figure 6b**). With respect to streptophytes, retention levels are varied. Like some prasinophytes and the sequenced prasinodermophyte, the streptophyte alga *Chlorokybus atmophyticus*, bryophytes, and lycophytes all have the complete pathway—but with a twist (highlighted below). Perhaps more exciting, the canonical pathway was also recently reported in another streptophyte seed lineage, the gymnosperms (105), specifically *Picea abies* and *Pinus taeda*. Moreover, some members of the ANA grade, asterids, and core rosids have retained nearly all proteins in this pathway, typically lacking just one enzyme, MurF or MurJ, and also lacking the PBP transpeptidase domain in the murein polymerase/transpeptidase. For others, retention involves just four proteins that appear to have been conserved throughout streptophyte evolution and possibly operate in non-peptidoglycan-forming pathways. *A. thaliana* is a typical example of these taxa, with retention of MurE, MraY, MurG, and Ddl enzymes (111). Strikingly, selective losses of other components have occurred in multiple independent events within the flowering plants (**Figure 6b**).

Given the conserved nature of these losses, with the same enzymes being lost in independent events, it is important that we reconceptualize the concepts of loss and reduction and consider possible redirections and modifications of function. One of the most exciting developments in understanding how functions may have changed follows from results of knockout studies. Knockout of *MURE* in *Physcomitrella* was shown to disrupt plastid division, resulting in a macrochloroplast. However, in *Arabidopsis* it inhibited chloroplast development, resulting in a white seedling phenotype; moreover, expression of *AtMURE* in *Physcomitrella* could not rescue the macrochloroplast phenotype, suggesting functional divergence of moss and *Arabidopsis* MurE (49). Cross-species complementation assays used the *MURE* gene from a gymnosperm with an otherwise uncharacterized peptidoglycan pathway, *Larix gmelinii* (larch), which occupies an intermediate but still early streptophyte divergence position (**Figure 1a**), and resulted in complete rescue of *Arabidopsis MURE* mutants but, again, did not rescue the moss macrochloroplast phenotype. These results were taken to suggest that in seed plants MurE has the same function but functions differently in earlier branching streptophytes (105). Further research would help us understand and identify the functional differences in the retention of the full and almost full peptidoglycan biosynthesis pathway, as well as those with lineages that have only the four conserved genes. Remarkably, this variation level exists even between two members of the same prasinophyte genus, *Micromonas pusilla* and *Micromonas commoda* (**Figure 6b**).

For flowering plants retaining the full peptidoglycan pathway, one of the final steps, the transpeptidase function encoded in PBP, has seemingly been lost; only the glycosyltransferase domain is detected (**Figure 6b**). This raises the possibility that a new PBP-like enzyme class exists that lacks transpeptidation activity, meaning that the cross-linkage of glycan chains either does not take place or is realized by another protein. Interestingly, in bacteria, one of the final steps is binding of peptidoglycan by a LysM domain. In plants, LysM domain–containing proteins are generally thought to be involved in sensing bacterial peptidoglycan (171) and chitin (7) by recognizing *N*-acetylglucosamine moieties (12, 147). In streptophytes with a complete or nearly complete peptidoglycan biosynthesis pathway (**Figure 6b**), a protein has been found that contains a C-terminal LysM domain (Pfam 01476) and is absent from streptophytes with a reduced peptidoglycan pathway (as well as from chlorophytes and prasinophytes, regardless of the reduction or completeness of the pathway) (162). This peptidoglycan pathway–associated streptophyte protein (PPASP) is

predicted to form an α-helical transmembrane region with the LysM domain on the outside and typically has a conserved hydrophobic N-terminal 21-amino-acid domain. Our analysis of newer genome data shows that *Chlorokybus atmophyticus*, bryophytes, and lycophytes have both the murein polymerase/transpeptidase and PPASP. Presence/absence patterns of the murein polymerase/transpeptidase and PPASP across the broader Viridiplantae suggest that PPASP evolved or was acquired around the time that streptophytes diverged from the ancestor shared with prasinodermophyte, chlorophyte, and prasinophyte algae (**Figure 6**). PPASP would then have been differentially lost in multiple streptophyte lineages, along with other peptidoglycan pathway components. Functional characterization of PPASP will be important for testing the hypothesis (162) that it has a role in peptidoglycan formation or binding. The surprise here is that despite the tremendous wealth of information on functions within the chloroplast (such as the function of photosystem components), there are still questions regarding the functional implications and potential co-option that has occurred for Viridiplantae members with highly reduced peptidoglycan pathways as well as those with nearly but not quite complete pathways. This lack of understanding extends across extant plants, prasinophytes, chlorophytes, and prasinodermophytes, despite the fundamental role that peptidoglycan has in bacterial, and presumably plastidial, division and cell wall structure.

CRAZY FOR CARBS: CAZyme PROFILES ACROSS THE VIRIDIPLANTAE

Plant-derived carbohydrates serve as one of the main sources of energy for many heterotrophic life forms and are exceptionally diverse. Unlike the peptide bonds that linearly link the 20 amino acids, carbohydrates have more monomer diversity (over 70 monosaccharides are known) and can be linked together via many of their constitutive hydroxyl groups. They can also be attached to different noncarbohydrate molecules, generating a large diversity of structures that can be synthesized and broken down selectively (88, 90). Each single glycosidic bond—synthesized or degraded—requires its own specialized enzyme. Thus, the diversity of CAZymes in the Viridiplantae parallels that of their carbohydrate substrates and includes enzymes such as those required for the synthesis of polysaccharides like cellulose (in some) and starch.

Sequence-based classification of CAZymes was initiated in 1991, starting with the class of glycoside hydrolases (GHs) (64) and recognized families that correlated with the structure of the enzymes; the classification had predictive power because the catalytic mechanism and orientation of the glycosidic bond undergoing catalysis could be extrapolated reliably. After the initial work on GHs, the family classification was extended to GTs (16), polysaccharide lyases (106), and carbohydrate-binding modules (CBMs) (10). Grouping enzymes into families based on the substrates they act upon has helped reveal the evolution of CAZyme specificities and honed attention to functional diversification.

The genome of *A. thaliana* is the first to have been formally searched for encoded CAZymes (65). The general conclusions of this study remain correct today—specifically, that for several reasons (25), land plants generally encode more CAZymes than other eukaryotes, an idea upheld through inspection of prasinodermophytes, prasinophytes, and chlorophytes (**Figure 7**). That said, since then a number of novel activities and families have been found. The CAZy database has a section of sequences in the nonclassified section that are only distantly related to bona fide GH or GT families. Despite considerable progress in the 20 years since completion of the *A. thaliana* genome, examination of the CAZy nonclassified section suggests that there are still many *A. thaliana* CAZymes to discover, particularly GTs. This is true not just for *Arabidopsis* but also for the genomes of other well-studied plants (e.g., *Brassica* and *Vigna*), prasinophytes (e.g., *Bathycoccus*, *Micromonas*, and *Ostreococcus*), and prasinodermophytes (i.e., *Prasinoderma coloniale*). The CAZy

Figure 7

CAZymes in the Viridiplantae, including presence and copy numbers of CAZy classes in representative plant genomes. CAZy family and class definitions follow the CAZy Database (107). Shades of red in the heatmap indicate presence and copy number in each genome of CAZyme classes, with their respective family numbers indicated at the left. Within families, CAZyme classes are ordered bottom to top, based on the number of species in which a respective class was present. The vertical bar below the species name indicates the lineage based on the same grouping in **Figure 1**, with squares above the species name indicating the taxonomic groups. It should be noted that other plant genomes have been analyzed previously but are not represented here, for instance, Poplar (86). Furthermore, there are many unclassified CAZymes that are not presented and await further studies to establish function. Abbreviations: ANA, Amborellales, Nymphaeales, Austrobaileyaceae; CAZyme, carbohydrate-active enzyme.
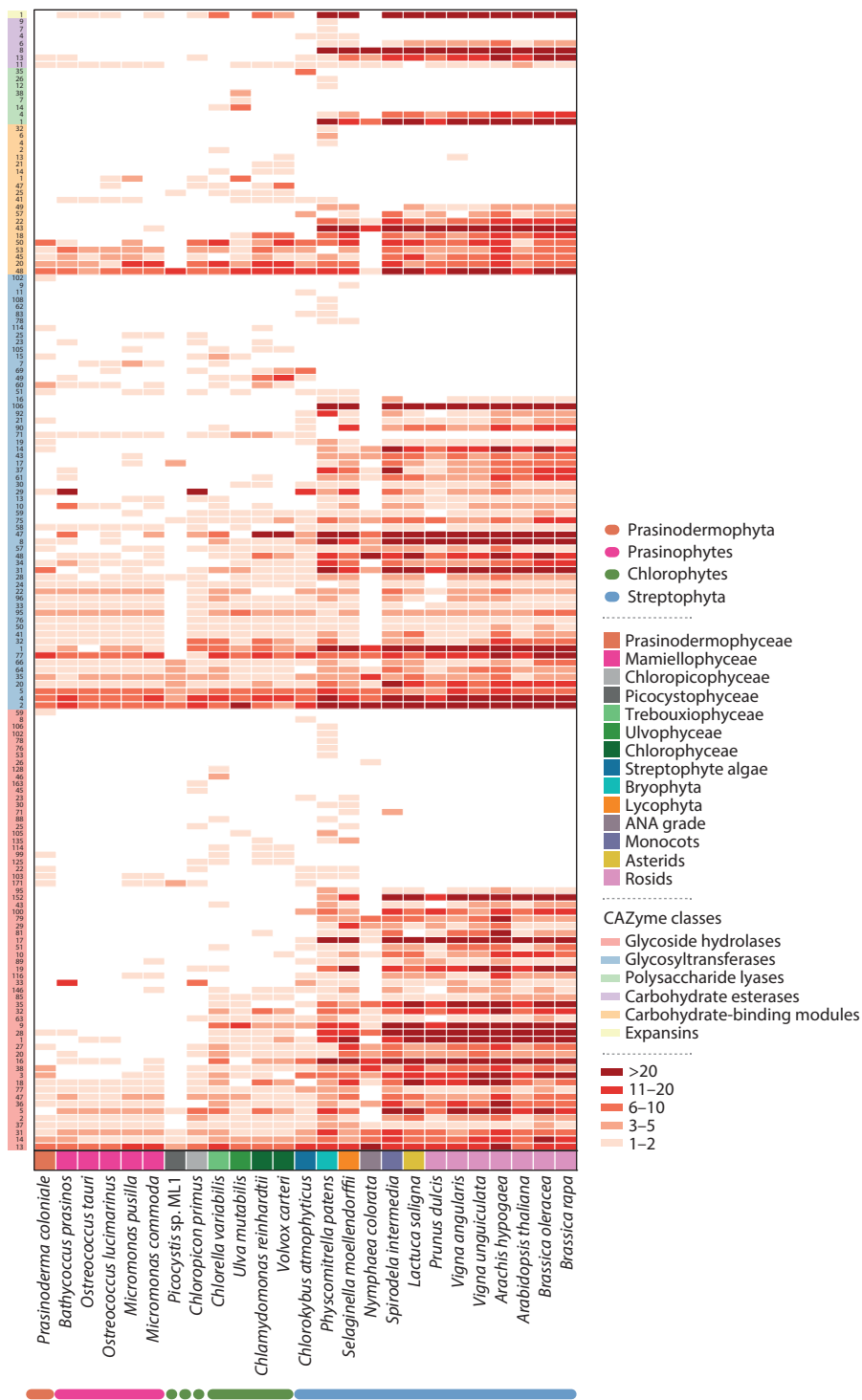
database has been updated, analyzed, and made available with 10 new organisms spanning green algae to basal streptophytes in the process of performing this review (**Figure 7**). As of June 2021, the CAZy database (**http://www.cazy.org**) has provided open access to the CAZyme profile of 18 analyzed Viridiplantae genomes, released as finished entries by GenBank, and to updated analysis of *Micromonas pusilla*, using a finished genome assembly and model set (162).

The importance of GTs in plant metabolism is reflected by the presence of starch in the Viridiplantae plastids (14) (e.g., **Figure 3d**) and the ubiquitous presence of starch-related CAZymes (GH13, GH14, GH31, GT5, GT35, CBM48; see **Figure 7**). Studies of *Chlamydomonas reinhardtii* have helped identify aspects of starch synthesis (e.g., 158), and analysis of the pathway in prasinophytes revealed similar complexity to that of vascular plants, establishing its presence in the green algal progenitor (31, 140). Linked to starch metabolism are the metabolites sucrose and trehalose, which act as signaling molecules, osmoprotectants, or carbon shuttle molecules (43). Their importance is reflected by the omnipresence of the required synthase enzymes (GT4 and GT20, respectively) across the Viridiplantae (**Figure 7**). For *O. tauri*, trehalose has been proposed to serve as an energy and carbon buffer molecule linked to starch degradation, or potentially to signaling during the dark period, based on a metabolomics study of growth under a diel light cycle (67). In contrast, starch content in *A. thaliana* has been shown to be independent of circadian cycles, although the transcription of genes encoding the enzymes involved in synthesis and degradation depends on the presence or absence of light (152), suggesting dissimilar mechanisms of starch metabolism regulation in algae and vascular plants.

Perhaps more surprising is that the GT2 family, comprising cellulose and chitin synthases, is also found across the Viridiplantae, based on the taxonomic sampling herein (**Figure 7**). Of course, cellulose is an important component of land plant cell walls, giving stability and enabling vertical growth (146, 153), but its presence in unicellular algae is more variable. Cellulose is a cell wall component in many core chlorophyte and streptophyte algae, but not in *Chlamydomonas*. Most recently, cellulose was proposed to fulfill similar functions in streptophyte algae as in vascular plants, based on a transcriptomic study of the streptophyte alga *Zygnema circumcarinatum* under osmotic stress (45). However, cellulose has not been observed in prasinophyte cell walls (33); cell wall composition of prasinodermophytes is not yet known. Moreover, *Micromonas* and *Ostreococcus* are so-called naked prasinophytes, lacking even the scales that are found on other prasinophytes. Thus, the presence of GT2 genes in their genomes (72, 121, 122) suggests either a more ancestral role or a different functional role than in extant chlorophytes and streptophytes.

Thus, among the already-classified CAZymes there is still much to learn about how they manifest in the cell, and there are still many CAZyme functions yet to be discovered. Other examples include enrichment of sialic acid–related CAZymes in the prasinophytes *Bathycoccus prasinos* and *Chloropicon primus*, specifically GH33 (sialidases/neuraminidases) and GT29 (sialyltransferases; see **Figure 7**), compared to other Viridiplantae (96, 127, 162). For *Bathycoccus prasinos*, it was speculated that these enzymes are putative participants in scale biosynthesis (127), based on the chemical structure scales in the chlorophyte *Scherffelia dubia* being similar to sialic acids (119). However, scales have not been observed in the three-layer cell wall of *Chloropicon primus* (108), and its GH33 and GT29 genes appear to be silent or expressed at very low levels under the growth conditions tested (96). Thus, it remains unclear whether scales are produced at some as-yet-unknown life stage in *Chloropicon primus*, as is seen for *Chlorokybus atmophyticus* zoospores during asexual reproduction (169), or used for another function, or if the proposed role in scale formation might be erroneous. Deciphering the functions of orthologous GH33 and GT29 genes in streptophytes should help simplify studies of the plethora of such genes in *Bathycoccus prasinos* and *Chloropicon primus*.

The *Bathycoccus prasinos* genome also has an expansion of GT10 genes (galactoside/glycoprotein α-ʟ-fucosyltransferases) compared to all other Viridiplantae, which is remarkable given that the land plants have much bigger genomes (**Figure 1***b*) but contain fewer GT10 genes than *Bathycoccus* (**Figure 7**). Such genes appear to be present in Streptophyta and prasinophytes (except *Micromonas pusilla*), but they are almost completely lacking in prasinodermophytes and chlorophytes, with just one detected in genomes of chlorophytes *Chlorella variabilis* and *Chlamydomonas reinhardtii*. These results suggest that an as-yet-unknown fucosylated structure exists in *Bathycoccus* (162), or potentially complex undiscovered protein glycosylations, as reported for some *Arabidopsis* GT10 genes (5, 172).

GT49 and GT47 exemplify the differential distributions of various CAZy families in the Viridiplantae. The complex *N*-glycan structures in *Chlamydomonas* (116, 164) likely explain the almost exclusive presence of GT49 genes (β-*N*-acetyl glucosaminyltransferases) in *Chlamydomonas* and *Volvox* (**Figure 7**), and their virtual absence from prasinodermophytes, prasinophytes, and streptophytes. By contrast, GT47 profiles exhibit similar numbers in *Chlamydomonas*, *Volvox* and streptophytes, whereas fewer or none are found in other chlorophytes, prasinophytes, and prasinodermophytes (**Figure 7**). In *Arabidopsis*, 4 of 39 predicted GT47 enzymes have been characterized and are either linked to pectin (61, 74) or xyloglucan (73, 112) biosynthesis, the latter being thought to have evolved after the divergence of chlorophyte and streptophyte algae (72, 121, 122). GT47 function in *Chlamydomonas* and *Volvox* therefore remains elusive but could be connected to glycosylation of the hydroxyproline-rich glycoproteins (HRGPs), constituents of *Chlamydomonas* and plant cell walls (124). At least one *Arabidopsis* GT47 was shown to be active on extensins, a family of streptophyte cell wall HRGPs (125). Additionally, GT77 and GT95 genes are connected to HRGP/extensin glycosylation in land plants (130, 165, 166), and the GT77 are abundant in the genome (175). Given these differential patterns, it should be possible to identify appropriate algal workhorses for heterologous expression of proteins and enzymes (which exist in multiple copies in plants) for functional studies that have proved challenging in other model systems (44).

The CAZyme profiles treated in a phylogenetic context reveal 24 families that are largely exclusive to land plants, 9 others that appear to be absent from prasinophytes sequenced to date, and 21 that are present across all Viridiplantae but enriched in land plants. Expansins, although technically not CAZymes, are included in the latter group. These proteins are important in cell wall turnover, where they contribute to its loosening (115). Nevertheless, expansins are found even in naked prasinophytes, albeit few, pointing again to the importance of addressing both evolutionary patterns and functional diversification. Most of the other families are linked to hemicellulose (xylan, xyloglucan, pectin) synthesis and modification, which is no surprise since these are almost exclusively a feature of land plant cell walls and some streptophyte algae (33).

CONCLUSION

The Viridiplantae are truly remarkable for persisting as important primary producers in the ocean since the advent of photosynthetic eukaryotes and for bringing photosynthesis to freshwater and land. Collectively, Viridiplantae today are responsible for much more than half of global CO_2 uptake from the atmosphere, taking the contributions of aquatic and terrestrial greens together. These organisms have survived major environmental perturbations, including warming and cooling events, likely using molecular mechanisms that are conserved across the lineage even today. It is clear that a major push is required to delineate factors that allow unicellular algae to thrive in the oceans, as well as those key to niche partitioning. These studies are now possible given the emerging model systems for several prasinophyte algae—including widespread species that have

Niche partitioning: the process by which competing taxa differently utilize resources or space, allowing for coexistence within an environment

small genomes with low gene redundancy (40). Over the past decades, model streptophyte plants such as *Physcomitrella patens* and *Arabidopsis thaliana*, both of which are quite derived among the Viridiplantae as a whole, have shed light on many biological processes. By extending the scope of comparisons to include early-diverging Viridiplantae, such as prasinophytes, alongside prasinodermophytes and chlorophytes, functional roles of proteins that may be retained in other plant lineages can be identified and contribute to our understanding of processes that are essential to the long-term evolution and survival of plants on land and in the sea.

SUMMARY POINTS

1. Viridiplantae emerged after the primary endosymbiosis event, wherein a cyanobacterium was incorporated as a plastid, and today comprise the Chlorophyta (i.e., the primarily freshwater chlorophyte algae and marine prasinophyte algae), Streptophyta, and Prasinodermophyta. Collectively, marine green algae and land plants are responsible for well over half of global primary production.

2. Phytochromes steer major aspects of streptophyte development and thermal and light responses. They are present in most prasinophytes, but not chlorophytes. Those in prasinophytes have a similar core domain structure to plant phytochromes, but do not have complex gene family expansion. However, while they are known to perceive blue-shifted wavelengths, their cellular roles in algae are still unclear.

3. Cyanobacterial peptidoglycan biosynthesis has been fully retained in various Viridiplantae lineages, including some prasinophytes, prasinodermophytes, chlorophytes, and streptophytes. In other lineages within these major groupings, a conserved reduced set has been retained. This highlights the likelihood of at least two differentiated functionalities that each span diverse but divergent Viridiplantae lineages.

4. The main groupings of Viridiplantae algae each retain different sets of orthologs shared with streptophytes; hence, they capture different ancestral aspects of the green lineage and, as model systems, speak to disparate but complementary aspects of plant biology.

5. CAZyme diversity mirrors the large variety of monosaccharides that exist in nature, with many families being observed across all Viridiplantae, especially GTs, which are involved in cell wall and starch metabolism. The handful of folds that GTs adopt and the availability of algorithms able to detect distant similarities or even able to accurately predict protein folds (77) help identify sequences that may be glycosyltransferases. However, the massive differences in cell wall glycan composition and other carbohydrate transformations in the plant lineage render the identification of the molecular function of these candidate biosynthetic CAZymes extremely difficult in spite of the large advances in their functional characterization over the last 20 years.

6. Unicellular marine prasinophytes that grow rapidly and synchronously represent excellent model systems for plant biology owing to their small genomes with few duplications, few repeats, and limited gene family expansion. Likewise, hypotheses can be developed from plant model systems on roles of orthologous proteins in algae, serving as a springboard for closing the massive gap in understanding environmental factors and cellular responses for major groups of aquatic primary producers.

7. Under changing climate conditions, organisms such as green algae must respond to dramatic shifts in temperature and light conditions, and may do so with some alacrity due to their fast reproduction times. Studies of these organisms can enhance our understanding of both ancestral algae and the role of putative homologs in multicellular lineages, as well as potential insights for genetic manipulation or assisted evolution efforts.

DISCLOSURE STATEMENT

ACKNOWLEDGMENTS

LITERATURE CITED

1. Archibald JM. 2015. Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* 25(19):R911–21

2. Azam F, Malfatti F. 2007. Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* 5(10):782–91

3. Bachy C, Charlesworth CJ, Chan AM, Finke JF, Wong C-H, et al. 2018. Transcriptional responses of the marine green alga *Micromonas pusilla* and an infecting prasinovirus under different phosphate conditions. *Environ. Microbiol.* 20(8):2898–912

4. Bachy C, Yung CCM, Needham DM, Gazitúa MC, Roux S, et al. 2021. Viruses infecting a warm water picoeukaryote shed light on spatial co-occurrence dynamics of marine viruses and their hosts. *ISME J.* 15:3129–47

5. Bakker H, Schijlen E, de Vries T, Schiphorst WECM, Jordi W, et al. 2001. Plant members of the α1→3/4-fucosyltransferase gene family encode an α1→4-fucosyltransferase, potentially involved in Lewis^a biosynthesis, and two core α1→3-fucosyltransferases. *FEBS Lett.* 507(3):307–12

6. Baquero F, Levin BR. 2021. Proximate and ultimate causes of the bactericidal action of antibiotics. *Nat. Rev. Microbiol.* 19(2):123–32

7. Bateman A, Bycroft M. 2000. The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD). *J. Mol. Biol.* 299(4):1113–19

8. Blaby-Haas CE, Merchant SS. 2019. Comparative and functional algal genomics. *Annu. Rev. Plant Biol.* 70:605–38

9. **Bolaños LM, Karp-Boss L, Choi CJ, Worden AZ, Graff JR, et al. 2020. Small phytoplankton dominate western North Atlantic biomass. *ISME J.*14:1663–74**

10. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* 382(Part 3):769–81

11. Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, et al. 2017. Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* 171(2):287–304.e15

9. Discovered that tiny prasinophyte algae are key components of an annual Atlantic spring bloom responsible for supporting many fisheries and carbon export.

12. Buist G, Steen A, Kok J, Kuipers OP. 2008. LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol. Microbiol.* 68(4):838–47

13. Burgie ES, Gannam ZTK, McLoughlin KE, Sherman CD, Holehouse AS, et al. 2021. Differing biophysical properties underpin the unique signaling potentials within the plant phytochrome photoreceptor families. *PNAS* 118(22):e2105649118

14. Busi MV, Barchiesi J, Martín M, Gomez-Casati DF. 2014. Starch metabolism in green algae. *Starch* 66(1–2):28–40

15. Butcher RW. 1952. Contributions to our knowledge of the smaller marine algae. *J. Mar. Biol. Ass. U. K.* 31(1):175–91

16. Campbell JA, Davies GJ, Bulone V, Henrissat B. 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* 326(Part 3):929–39

17. Chavez FP, Messié M, Pennington JT. 2011. Marine primary production in relation to climate variability and change. *Ann. Rev. Mar. Sci.* 3:227–60

18. Chen Y-R, Su Y-S, Tu S-L. 2012. Distinct phytochrome actions in nonvascular plants revealed by targeted inactivation of phytobilin biosynthesis. *PNAS* 109(21):8310–15

19. Cheng M-C, Kathare PK, Paik I, Huq E. 2021. Phytochrome signaling networks. *Annu. Rev. Plant Biol.* 72:217–44

20. Chisholm SW. 1992. Phytoplankton size. In *Primary Productivity and Biogeochemical Cycles in the Sea*, ed. PG Falkowski, AD Woodhead, K Vivirito, pp. 213–37. Boston, MA: Springer

21. Chrétiennot-Dinet M-J, Courties C, Vaquer A, Neveux J, Claustre H, et al. 1995. A new marine picoeucaryote: *Ostreococcus tauri* gen. et sp. nov. (Chlorophyta, Prasinophyceae). *Phycologia* 34(4):285–92

22. Clayton S, Lin Y-C, Follows MJ, Worden AZ. 2017. Co-existence of distinct *Ostreococcus* ecotypes at an oceanic front. *Limnol. Oceanogr.* 62:75–88

23. Coesel SN, Durham BP, Groussman RD, Hu SK, Caron DA, et al. 2021. Diel transcriptional oscillations of light-sensitive regulatory elements in open-ocean eukaryotic plankton communities. *PNAS* 118(6):e2011038118

24. Corlett RT. 2016. Plant diversity in a changing world: status, trends, and conservation needs. *Plant Divers.* 38(1):10–16

25. Coutinho PM, Stam M, Blanc E, Henrissat B. 2003. Why are there so many carbohydrate-active enzyme-related genes in plants? *Trends Plant Sci.* 8(12):563–65

26. Delaux P-M, Radhakrishnan GV, Jayaraman D, Cheema J, Malbreil M, et al. 2015. Algal ancestor of land plants was preadapted for symbiosis. *PNAS* 112(43):13390–95

27. Delwiche CF, Cooper ED. 2015. The evolutionary origin of a terrestrial flora. *Curr. Biol.* 25(19):R899–910

28. Demir-Hilton E, Sudek S, Cuvelier ML, Gentemann CL, Zehr JP, Worden AZ. 2011. Global distribution patterns of distinct clades of the photosynthetic picoeukaryote *Ostreococcus*. *ISME J.* 5(7):1095–107

29. Demory D, Baudoux A-C, Monier A, Simon N, Six C, et al. 2019. Picoeukaryotes of the *Micromonas* genus: sentinels of a warming ocean. *ISME J.* 13:132–46

30. Derelle E, Ferraz C, Rombauts S, Rouzé P, Worden AZ, et al. 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *PNAS* 103(31):11647–52

31. Deschamps P, Moreau H, Worden AZ, Dauvillée D, Ball SG. 2008. Early gene duplication within Chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. *Genetics* 178(4):2373–87

32. Devireddy AR, Liscum E, Mittler R. 2020. Phytochrome B is required for systemic stomatal responses and reactive oxygen species signaling during light stress. *Plant Physiol.* 184(3):1563–72

33. Domozych DS, Ciancia M, Fangel JU, Mikkelsen MD, Ulvskov P, Willats WGT. 2012. The cell walls of green algae: a journey through evolution and diversity. *Front. Plant Sci.* 3:82

34. Doney SC, Fabry VJ, Feely RA, Kleypas JA. 2009. Ocean acidification: the other CO_2 problem. *Annu. Rev. Mar. Sci.* 1:169–92

35. Duanmu D, Bachy C, Sudek S, Wong C-H, Jiménez V, et al. 2014. Marine algae and land plants share conserved phytochrome signaling systems. *PNAS* 111(44):15827–32

36. Dutcher SK. 2020. Asymmetries in the cilia of *Chlamydomonas*. *Philos. Trans. R. Soc. B* 375:20190153

29. Combined growth experiments and theoretical ecology to characterize the differentiated thermal responses of several *Micromonas* species.

35. Demonstration of an operational conserved phytochrome signaling mechanism between prasinophytes and plants alongside phylogenetic analysis showing a eukaryotic origin of the Archaeplastida phytochromes.

37. Egan AJF, Errington J, Vollmer W. 2020. Regulation of peptidoglycan synthesis and remodelling. *Nat. Rev. Microbiol.* 18(8):446–60

38. Ejaz M, Bencivenga S, Tavares R, Bush M, Sablowski R. 2021. *ARABIDOPSIS THALIANA HOMEO-BOX GENE 1* controls plant architecture by locally restricting environmental responses. *PNAS* 118(17):e2018615118

39. Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238

40. Introduces first steps toward genetic manipulation methods for nascent model organisms, including the widespread marine prasinophytes *Micromonas commoda*, *Bathycoccus prasinos*, and *Ostreococcus lucimarinus*.

40. Faktorová D, Nisbet RER, Robledo JAF, Casacuberta E, Sudek L, et al. 2020. Genetic tool development in marine protists: emerging model organisms for experimental cell biology. *Nat. Methods* 17(5):481–94

41. Falciatore A, Bowler C. 2005. The evolution and function of blue and red light photoreceptors. *Curr. Top. Dev. Biol.* 68:317–50

42. Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281:237–40

43. Figueroa CM, Lunn JE. 2016. A tale of two sugars: trehalose 6-phosphate and sucrose. *Plant Physiol.* 172(1):7–27

44. Fischl R, Bertelsen K, Gaillard F, Coelho S, Michel G, et al. 2016. The cell-wall active mannuronan C5-epimerases in the model brown alga *Ectocarpus*: from gene context to recombinant protein. *Glycobiology* 26(9):973–83

45. Fitzek E, Orton L, Entwistle S, Grayburn WS, Ausland C, et al. 2019. Cell wall enzymes in *Zygnema circumcarinatum* UTEX 1559 respond to osmotic stress in a plant-like fashion. *Front. Plant Sci.* 10:732

46. Fortunato AE, Jaubert M, Enomoto G, Bouly J-P, Raniello R, et al. 2016. Diatom phytochromes reveal the existence of far-red-light-based sensing in the ocean. *Plant Cell* 28(3):616–28

47. Fujiwara MT, Sanjaya A, Itoh RD. 2019. *Arabidopsis thaliana* leaf epidermal guard cells: a model for studying chloroplast proliferation and partitioning in plants. *Front. Plant Sci.* 10:1403

48. Gaarder T. 1933. Untersuchungen über Produktions- und Lebensbedingungen in norwegischen Austern-Pollen. *Bergens Museums Arbok Naturv.* 3

49. Garcia M, Myouga F, Takechi K, Sato H, Nabeshima K, et al. 2008. An *Arabidopsis* homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development. *Plant J.* 53(6):924–34

50. Geng Y, Cai C, McAdam SAM, Banks JA, Wisecaver JH, Zhou Y. 2021. A de novo transcriptome assembly of *Ceratopteris richardii* provides insights into the evolutionary dynamics of complex gene families in land plants. *Genome Biol. Evol.* 13(3):evab042

51. Gibbons IR. 1981. Cilia and flagella of eukaryotes. *J. Cell Biol.* 91(3):107–24

52. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, et al. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40(Database issue):D1178–86

53. Grigoriev IV, Hayes RD, Calhoun S, Kamel B, Wang A, et al. 2021. PhycoCosm, a comparative algal genomics resource. *Nucleic Acids Res.* 49(D1):D1004–11

54. Grossman A, Sanz-Luque E, Yi H, Yang W. 2019. Building the GreenCut2 suite of proteins to unmask photosynthetic function and regulation. *Microbiology* 165(7):697–718

55. Gruber N, Gloor M, Mikaloff Fletcher SE, Doney SC, Dutkiewicz S, et al. 2009. Oceanic sources, sinks, and transport of atmospheric CO_2. *Global Biogeochem. Cycles* 23(1):GB1005

56. Guasto JS, Rusconi R, Stocker R. 2012. Fluid mechanics of planktonic microorganisms. *Annu. Rev. Fluid Mech.* 44:373–400

57. Guillou L, Eikrem W, Chrétiennot-Dinet M-J, Le Gall F, Massana R, et al. 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(2):193–214

58. Guiry MD. 2012. How many species of algae are there? *J. Phycol.* 48(5):1057–63

59. Guo Z, Xu J, Wang Y, Hu C, Shi K, et al. 2021. The phyB-dependent induction of HY5 promotes iron uptake by systemically activating *FER* expression. *EMBO Rep.* 22(7):e51944

60. Hamilton M, Mascioni M, Hehenberger E, Bachy C, Yung C, et al. 2021. Spatiotemporal variations in Antarctic protistan communities highlight phytoplankton diversity and seasonal dominance by a novel cryptophyte lineage. *mBio* 12:e02973-21

61. Harholt J, Jensen JK, Sørensen SO, Orfila C, Pauly M, Scheller HV. 2006. ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in *Arabidopsis*. *Plant Physiol.* 140(1):49–58

62. Hart JE, Gardner KH. 2021. Lighting the way: recent insights into the structure and regulation of phototropin blue light receptors. *J. Biol. Chem.* 296:100594

63. Hasegawa T, Miyashita H, Kawachi M, Ikemoto H, Kurano N, et al. 1996. *Prasinoderma coloniale* gen. et sp. nov., a new pelagic coccoid prasinophyte from the western Pacific Ocean. *Phycologia* 35(2):170–76

64. Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280(2):309–16

65. Henrissat B, Coutinho PM, Davies GJ. 2001. A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. *Plant Mol. Biol.* 47(1–2):55–72

66. Hirano T, Tanidokoro K, Shimizu Y, Kawarabayasi Y, Ohshima T, et al. 2016. Moss chloroplasts are surrounded by a peptidoglycan wall containing D-amino acids. *Plant Cell* 28(7):1521–32

67. Hirth M, Liverani S, Mahlow S, Bouget F-Y, Pohnert G, Sasso S. 2017. Metabolic profiling identifies trehalose as an abundant and diurnally fluctuating metabolite in the microalga *Ostreococcus tauri*. *Metabolomics* 13(6):68

68. Homi S, Takechi K, Tanidokoro K, Sato H, Takio S, Takano H. 2009. The peptidoglycan biosynthesis genes *MurA* and *MraY* are related to chloroplast division in the moss *Physcomitrella patens*. *Plant Cell Physiol.* 50(12):2047–56

69. Ibarbalz FM, Henry N, Brandão MC, Martini S, Busseni G, et al. 2019. Global trends in marine plankton diversity across kingdoms of life. *Cell* 179(5):1084–97.e21

70. Ikeda H, Suzuki T, Oka Y, Gustafsson ALS, Brochmann C, et al. 2021. Divergence in red light responses associated with thermal reversion of phytochrome B between high- and low-latitude species. *New Phytol.* 231(1):75–84

71. Ishizaki K, Nishihama R, Yamato KT, Kohchi T. 2016. Molecular genetic tools and techniques for *Marchantia polymorpha* research. *Plant Cell Physiol.* 57(2):262–70

72. Jensen JK, Busse-Wicher M, Poulsen CP, Fangel JU, Smith PJ, et al. 2018. Identification of an algal xylan synthase indicates that there is functional orthology between algal and plant cell wall biosynthesis. *New Phytol.* 218(3):1049–60

73. Jensen JK, Schultink A, Keegstra K, Wilkerson CG, Pauly M. 2012. RNA-Seq analysis of developing nasturtium seeds (*Tropaeolum majus*): identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. *Mol. Plant.* 5(5):984–92

74. Jensen JK, Sørensen SO, Harholt J, Geshi N, Sakuragi Y, et al. 2008. Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in *Arabidopsis*. *Plant Cell* 20(5):1289–302

75. Jeong J, Choi G. 2013. Phytochrome-interacting factors have both shared and distinct biological roles. *Mol. Cells* 35(5):371–80

76. Jouenne F, Eikrem W, Le Gall F, Marie D, Johnsen G, Vaulot D. 2011. *Prasinoderma singularis* sp. nov. (Prasinophyceae, Chlorophyta), a solitary coccoid Prasinophyte from the South-East Pacific Ocean. *Protist* 162(1):70–84

77. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* 596:583–89

78. Karlusich JJP, Ibarbalz FM, Bowler C. 2020. Exploration of marine phytoplankton: from their historical appreciation to the omics era. *J. Plankton Res.* 42(6):595–612

79. Karsten U, Holzinger A. 2014. Green algae in alpine biological soil crust communities: acclimation strategies against ultraviolet radiation and dehydration. *Biodivers. Conserv.* 23(7):1845–58

80. Kathare PK, Huq E. 2021. Light-regulated pre-mRNA splicing in plants. *Curr. Opin. Plant Biol.* 63:102037

81. Keeling PJ, Burki F. 2019. Progress towards the Tree of Eukaryotes. *Curr. Biol.* 29(16):R808–17

82. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, et al. 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLOS Biol.* 12(6):e1001889

83. Kies L. 1984. Cytological aspects of blue-green algal endosymbiosis. In *Compartments in Algal Cells and Their Interaction*, ed. W Weissner, DG Robinson, RC Starr, pp. 191–99. Berlin: Springer

70. Shows differences in responses connected to thermal stability of a physiologically active form of phytochrome (Pfr) that has diverged between sister species of Brassicaceae (the cabbage family) inhabiting ecosystems at different latitudes.

84. Knight-Jones EW, Walne PR. 1951. *Chromulina pusilla* Butcher, a dominant member of the ultraplankton. *Nature* 167(4246):445–46

85. Kolody BC, McCrow JP, Ziegler Allen L, Aylward FO, Fontanez KM, et al. 2019. Diel transcriptional response of a California Current plankton microbiome to light, low iron, and enduring viral infection. *ISME J.* 13:2817–33

86. Kumar V, Hainaut M, Delhomme N, Mannapperuma C, Immerzeel P, et al. 2019. Poplar carbohydrate-active enzymes: whole-genome annotation and functional analyses based on RNA expression data. *Plant J.* 99(4):589–609

87. Kuroiwa H, Sugai M, Kuroiwa T. 1988. Behavior of chloroplasts and chloroplast nuclei during spermatogenesis in the fern, *Pteris vittata* L. *Protoplasma* 146(2):89–100

88. Laine RA. 1994. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the *Isomer Barrier* to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4(6):759–67

89. Lang D, Ullrich KK, Murat F, Fuchs J, Jenkins J, et al. 2018. The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* 93(3):515–33

90. Lapébie P, Lombard V, Drula E, Terrapon N, Henrissat B. 2019. Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nat. Commun.* 10:2043

91. Lee S, Zhu L, Huq E. 2021. An autoregulatory negative feedback loop controls thermomorphogenesis in *Arabidopsis*. *PLOS Genet.* 17(6):e1009595

92. Leliaert F. 2019. Green algae: Chlorophyta and Streptophyta. In *Encyclopedia of Microbiology*, ed. TM Schmidt, pp. 457–68. Oxford: Academic. 4th ed.

93. Leliaert F, Smith DR, Moreau H, Herron MD, Verbruggen H, et al. 2012. Phylogeny and molecular evolution of the green algae. *Crit. Rev. Plant Sci.* 31(1):1–46

94. Leliaert F, Tronholm A, Lemieux C, Turmel M, DePriest MS, et al. 2016. Chloroplast phylogenomic analyses reveal the deepest-branching lineage of the Chlorophyta, Palmophyllophyceae class. nov. *Sci. Rep.* 6:25367

95. Leliaert F, Verbruggen H, Zechman FW. 2011. Into the deep: new discoveries at the base of the green plant phylogeny. *Bioessays* 33(9):683–92

96. Lemieux C, Turmel M, Otis C, Pombert J-F. 2019. A streamlined and predominantly diploid genome in the tiny marine green alga *Chloropicon primus*. *Nat. Commun.* 10:4061

97. Lewis LA, McCourt RM. 2004. Green algae and the origin of land plants. *Am. J. Bot.* 91(10):1535–56

98. Li F-W, Brouwer P, Carretero-Paulet L, Cheng S, de Vries J, et al. 2018. Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nat. Plants* 4:460–72

99. Li F-W, Melkonian M, Rothfels CJ, Villarreal JC, Stevenson DW. 2015. Phytochrome diversity in green plants and the origin of canonical plant phytochromes. *Nat. Commun.* 6:7852

100. **Li L, Wang S, Wang H, Sahu SK, Marin B, et al. 2020. The genome of *Prasinoderma coloniale* unveils the existence of a third phylum within green plants. *Nat. Ecol. Evol.* 4(9):1220–31**

101. Li WKW, McLaughlin FA, Lovejoy C, Carmack EC. 2009. Smallest algae thrive as the Arctic Ocean freshens. *Science* 326(5952):539

102. Liang Y, Wang J, Zheng J, Gong Z, Li Z, et al. 2021. Genome-wide comparative analysis of heat shock transcription factors provides novel insights for evolutionary history and expression characterization in cotton diploid and tetraploid genomes. *Front. Genet.* 12:658847

103. Liberton M, Berg RH, Heuser J, Roth R, Pakrasi HB. 2006. Ultrastructure of the membrane systems in the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. *Protoplasma* 227(2):129

104. Limardo AJ, Sudek S, Choi CJ, Poirier C, Rii YM, et al. 2017. Quantitative biogeography of picoprasinophytes establishes ecotype distributions and significant contributions to marine phytoplankton. *Environ. Microbiol.* 19(8):3219–34

105. **Lin X, Li N, Kudo H, Zhang Z, Li J, et al. 2017. Genes sufficient for synthesizing peptidoglycan are retained in gymnosperm genomes, and MurE from *Larix gmelinii* can rescue the albino phenotype of *Arabidopsis MurE* Mutation. *Plant Cell Physiol.* 58(3):587–97**

106. Lombard V, Bernard T, Rancurel C, Brumer H, Coutinho PM, Henrissat B. 2010. A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem. J.* 432(3):437–44

100. Genome analysis of *Prasinoderma coloniale*, previously considered a prasinophyte, establishes the early divergence of Prasinodermophyta and recognizes it as a Viridiplantae phylum.

105. Functional exploration of the gymnosperm and *Arabidopsis* peptidoglycan pathway member MurE, establishing modification of functional roles in different Viridiplantae lineages.

107. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42(Database issue):D490–95

108. Lopes dos Santos A, Pollina T, Gourvil P, Corre E, Marie D, et al. 2017. Chloropicophyceae, a new class of picophytoplanktonic prasinophytes. *Sci. Rep.* 7(1):14019

109. López-García P, Eme L, Moreira D. 2017. Symbiosis in eukaryotic evolution. *J. Theor. Biol.* 434:20–33

110. Lovejoy C, Vincent WF, Bonilla S, Roy S, Martineau M-J, et al. 2007. Distribution, phylogeny, and growth of cold-adapted picoprasinophytes in Arctic seas. *J. Phycol.* 43:78–89

111. Machida M, Takechi K, Sato H, Chung SJ, Kuroiwa H, et al. 2006. Genes for the peptidoglycan synthesis pathway are essential for chloroplast division in moss. *PNAS* 103(17):6753–58

112. Madson M, Dunand C, Li X, Verma R, Vanzin GF, et al. 2003. The *MUR3* gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *Plant Cell* 15(7):1662–70

113. Makita Y, Suzuki S, Fushimi K, Shimada S, Suehisa A, et al. 2021. Identification of a dual orange/far-red and blue light photoreceptor from an oceanic green picoplankton. *Nat. Commun.* 12:3593

114. Margulis L, Corliss JO, Melkonian M, Chapman DJ, eds. 1990. *Handbook of Protoctista: The Structure, Cultivation, Habitats and Life Histories of the Eukaryotic Microorganisms and Their Descendants Exclusive of Animals, Plants and Fungi*. Boston, MA: Jones & Bartlett

115. Marowa P, Ding A, Kong Y. 2016. Expansins: roles in plant growth and potential applications in crop improvement. *Plant Cell Rep.* 35(5):949–65

116. Mathieu-Rivet E, Scholz M, Arias C, Dardelle F, Schulze S, et al. 2013. Exploring the *N*-glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Mol. Cell. Proteom.* 12(11):3160–83

117. McDonald SM, Plant JN, Worden AZ. 2010. The mixed lineage nature of nitrogen transport and assimilation in marine eukaryotic phytoplankton: a case study of *Micromonas*. *Mol. Biol. Evol.* 27(10):2268–83

118. McFadden GI. 2014. Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 6(4):a016105

119. Melkonian M, Preisig HR. 1986. A light and electron microscopic study of *Scherffelia dubia*, a new member of the scaly green flagellates (Prasinophyceae). *Nord. J. Bot.* 6(2):235–56

120. Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, et al. 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318(5848):245–50

121. Mikkelsen MD, Harholt J, Ulvskov P, Johansen IE, Fangel JU, et al. 2014. Evidence for land plant cell wall biosynthetic mechanisms in charophyte green algae. *Ann. Bot.* 114(6):1217–36

122. Mikkelsen MD, Harholt J, Westereng B, Domozych D, Fry SC, et al. 2021. Ancient origin of fucosylated xyloglucan in charophycean green algae. *Commun. Biol.* 4:754

123. Miyashita H, Ikemoto H, Kurano N, Miyachi S, Chihara M. 1993. *Prasinococcus capsulatus* gen. et sp. nov., a new marine coccoid prasinophyte. *J. Gen. Appl. Microbiol.* 39(6):571–82

124. Mnich E, Bjarnholt N, Eudes A, Harholt J, Holland C, et al. 2020. Phenolic cross-links: building and de-constructing the plant cell wall. *Nat. Prod. Rep.* 37(7):919–61

125. Møller SR, Yi X, Velásquez SM, Gille S, Hansen PLM, et al. 2017. Identification and evolution of a plant cell wall specific glycoprotein glycosyl transferase, ExAD. *Sci. Rep.* 7:45341

126. Monier A, Worden AZ, Richards TA. 2016. Phylogenetic diversity and biogeography of the Mamiellophyceae lineage of eukaryotic phytoplankton across the oceans. *Environ. Microbiol. Rep.* 8(4):461–69

127. Moreau H, Verhelst B, Couloux A, Derelle E, Rombauts S, et al. 2012. Gene functionalities and genome structure in *Bathycoccus prasinos* reflect cellular specializations at the base of the green lineage. *Genome Biol.* 13:R74

128. Niklas KJ, Cobb ED, Matas AJ. 2017. The evolution of hydrophobic cell wall biopolymers: from algae to angiosperms. *J. Exp. Bot.* 68(19):5261–69

129. Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, et al. 2018. The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell* 174(2):448–64.e24

130. Ogawa-Ohnishi M, Matsushita W, Matsubayashi Y. 2013. Identification of three hydroxyproline O-arabinosyltransferases in *Arabidopsis thaliana*. *Nat. Chem. Biol.* 9(11):726–30

131. One Thousand Plant Transc. Initiat. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574:679–85

122. Identified fucosylated xyloglucan in a streptophyte alga, banishing the long-believed dogma that it evolved in multicellular land plants.

132. O'Toole ET. 2010. *Chlamydomonas* cryopreparation methods for the 3-D analysis of cellular organelles. *Methods Cell Biol*. 96:71–91

133. Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, et al. 2013. Pattern and synchrony of gene expression among sympatric marine microbial populations. *PNAS* 110(6):E488–97

134. Palenik B, Grimwood J, Aerts A, Rouze P, Salamov A, et al. 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *PNAS* 104:7705–10

135. Petriman N-A, Lorentzen E. 2020. Structural insights into the architecture and assembly of eukaryotic flagella. *Microb. Cell Fact*. 7(11):289–99

136. Pimm SL, Joppa LN. 2015. How many plant species are there, where are they, and at what rate are they going extinct? *Ann. Mo. Bot. Gard*. 100(3):170–76

137. Possart A, Xu T, Paik I, Hanke S, Keim S, et al. 2017. Characterization of phytochrome interacting factors from the moss *Physcomitrella patens* illustrates conservation of phytochrome signaling modules in land plants. *Plant Cell* 29(2):310–30

138. Prerostova S, Dobrev PI, Knirsch V, Jarosova J, Gaudinova A, et al. 2021. Light quality and intensity modulate cold acclimation in *Arabidopsis*. *Int. J. Mol. Sci*. 22(5):2736

139. Qiao Y, Srisuknimit V, Rubino F, Schaefer K, Ruiz N, et al. 2017. Lipid II overproduction allows direct assay of transpeptidase inhibition by β-lactams. *Nat. Chem. Biol*. 13(7):793–98

140. Ral J-P, Derelle E, Ferraz C, Wattebled F, Farinas B, et al. 2004. Starch division and partitioning. A mechanism for granule propagation and maintenance in the picophytoplanktonic green alga *Ostreococcus tauri*. *Plant Physiol*. 136(2):3333–40

141. Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, et al. 2008. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319(5859):64–69

142. Rockwell NC, Duanmu D, Martin SS, Bachy C, Price DC, et al. 2014. Eukaryotic algal phytochromes span the visible spectrum. *PNAS* 111:3871–76

143. Rockwell NC, Lagarias JC. 2020. Phytochrome evolution in 3D: deletion, duplication, and diversification. *New Phytol*. 225(6):2283–300

144. Ruiz N. 2008. Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in *Escherichia coli*. *PNAS* 105(40):15553–57

145. Sauvage E, Kerff F, Terrak M, Ayala JA, Charlier P. 2008. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev*. 32(2):234–58

146. Saxena IM, Brown RM Jr. 2005. Cellulose biosynthesis: current views and evolving concepts. *Ann. Bot*. 96(1):9–21

147. Schlöffel MA, Käsbauer C, Gust AA. 2019. Interplay of plant glycan hydrolases and LysM proteins in plant—bacteria interactions. *Int. J. Med. Microbiol*. 309(3):252–57

148. Shi XL, Marie D, Jardillier L, Scanlan DJ, Vaulot D. 2009. Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PLOS ONE* 4(10):e7657

149. Simmons MP, Bachy C, Sudek S, van Baren MJ, Sudek L, et al. 2015. Intron invasions trace algal speciation and reveal nearly identical Arctic and Antarctic *Micromonas* populations. *Mol. Biol. Evol*. 32(9):2219–35

150. Simmons MP, Sudek S, Monier A, Limardo AJ, Jimenez V, et al. 2016. Abundance and biogeography of picoprasinophyte ecotypes and other phytoplankton in the eastern North Pacific Ocean. *AEM* 82(6):1693–705

151. Simon N, Foulon E, Grulois D, Six C, Desdevises Y, et al. 2017. Revision of the genus *Micromonas* Manton et Parke (Chlorophyta, Mamiellophyceae), of the type species *M. pusilla* (Butcher) Manton & Parke and of the species *M. commoda* van Baren, Bachy and Worden and description of two new species based on the genetic and phenotypic characterization of cultured isolates. *Protist* 168(5):612–35

152. Smith SM, Fulton DC, Chia T, Thorneycroft D, Chapple A, et al. 2004. Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. *Plant Physiol*. 136(1):2687–99

153. Somerville C. 2006. Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol*. 22:53–78

154. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–13

155. Takano H, Takechi K. 2010. Plastid peptidoglycan. *Biochim. Biophys. Acta Gen. Subj.* 1800(2):144–51

156. Teves ME, Nagarkatti-Gude DR, Zhang Z, Strauss JF. 2016. Mammalian axoneme central pair complex proteins: broader roles revealed by gene knockout phenotypes. *Cytoskeleton* 73(1):3–22

157. Thomas S, Kumar R, Sharma K, Barpanda A, Sreelakshmi Y, et al. 2021. iTRAQ-based proteome profiling revealed the role of Phytochrome A in regulating primary metabolism in tomato seedling. *Sci. Rep.* 11:7540

158. Toyokawa C, Yamano T, Fukuzawa H. 2020. Pyrenoid starch sheath is required for LCIB localization and the CO_2-concentrating mechanism in green algae. *Plant Physiol.* 182(4):1883–93

159. Tragin M, Lopes dos Santos A, Christen R, Vaulot D. 2016. Diversity and ecology of green microalgae in marine systems: an overview based on 18S rRNA gene sequences. *Perspect. Phycol.* 3:141–54

160. Tragin M, Vaulot D. 2018. Green microalgae in marine coastal waters: The Ocean Sampling Day (OSD) dataset. *Sci. Rep.* 8:14020

161. **Umen JG. 2014. Green algae and the origins of multicellularity in the plant kingdom. *Cold Spring Harb. Perspect. Biol.* 6(11):a016170**

162. **van Baren MJ, Bachy C, Reistetter EN, Purvine SO, Grimwood J, et al. 2016. Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC Genom.* 17:267**

163. Van Vlierberghe M, Philippe H, Baurain D. 2021. Broadly sampled orthologous groups of eukaryotic proteins for the phylogenetic study of plastid-bearing lineages. *BMC Res. Notes* 14:143

164. Vanier G, Lucas P-L, Loutelier-Bourhis C, Vanier J, Plasson C, et al. 2017. Heterologous expression of the *N*-acetylglucosaminyltransferase I dictates a reinvestigation of the *N*-glycosylation pathway in *Chlamydomonas reinhardtii*. *Sci. Rep.* 7:10156

165. Velasquez SM, Marzol E, Borassi C, Pol-Fachin L, Ricardi MM, et al. 2015. Low sugar is not always good: impact of specific *O*-glycan defects on tip growth in Arabidopsis. *Plant Physiol.* 168(3):808–13

166. Velasquez SM, Ricardi MM, Dorosz JG, Fernandez PV, Nadra AD, et al. 2011. O-glycosylated cell wall proteins are essential in root hair growth. *Science* 332(6036):1401–3

167. Vergin KL, Beszteri B, Monier A, Thrash JC, Temperton B, et al. 2013. High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J.* 7(7):1322–32

168. Viprey M, Guillou L, Ferréol M, Vaulot D. 2008. Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. *Environ. Microbiol.* 10(7):1804–22

169. **Wang S, Li L, Li H, Sahu SK, Wang H, et al. 2020. Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nat. Plants* 6(2):95–106**

170. Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, et al. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *PNAS* 111(45):E4859–68

171. Willmann R, Lajunen HM, Erbs G, Newman M-A, Kolb D, et al. 2011. *Arabidopsis* lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *PNAS* 108(49):19824–29

172. Wilson IBH, Rendić D, Freilinger A, Dumić J, Altmann F, et al. 2001. Cloning and expression of cDNAs encoding α1,3-fucosyltransferase homologues from *Arabidopsis thaliana*. *Biochim. Biophys. Acta Gen. Subj.* 1527(1):88–96

173. Worden AZ. 2006. Picoeukaryote diversity in coastal waters of the Pacific Ocean. *AME* 43:165–75

174. Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ. 2015. Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. *Science* 347(6223):1257594

175. Worden AZ, Lee J-H, Mock T, Rouzé P, Simmons MP, et al. 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324(5924):268–72

176. Worden AZ, Nolan JK, Palenik B. 2004. Assessing the dynamics and ecology of marine picophytoplankton: the importance of the eukaryotic component. *Limnol. Oceanogr.* 49(1):168–79

161. Review of the independent transitions from unicellularity to multicellularity in two well-studied groups of green algae within the chlorophytes and streptophyte algae.

162. Showed that the complete peptidoglycan pathway is present not only in some prasinophytes but also in some plant lineages (but greatly reduced in others); identified the PPASP protein [putatively replacing penicillin-binding protein (PBP)], which is present only in plants with the pathway.

169. Plant terrestrialization steps reconsidered in the light of genome sequencing and analysis of the streptophyte algae *Mesostigma viride* and *Chlorokybus atmophyticus*.

177. Xu T, Yuan J, Hiltbrunner A. 2020. PHYTOCHROME INTERACTING FACTORs in the moss *Physcomitrella patens* regulate light-controlled gene expression. *Physiol. Plant.* 169(3):467–79

178. Yau S, Krasovec M, Benites LF, Rombauts S, Groussin M, et al. 2020. Virus-host coexistence in phytoplankton through the genomic lens. *Sci. Adv.* 6(14):eaay2587

179. Yeh K-C, Lagarias JC. 1998. Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry. *PNAS* 95(23):13976–81

Contents

Errata

An online log of corrections to *Annual Review of Plant Biology* articles may be found at
http://www.annualreviews.org/errata/arplant