



Review

# An Insight into the Algal Evolution and Genomics

Amna Komal Khan <sup>1</sup>, Humera Kausar <sup>1</sup>, Syyada Samra Jaferi <sup>1</sup>, Samantha Drouet <sup>2</sup>,  
Christophe Hano <sup>2</sup> , Bilal Haider Abbasi <sup>3</sup> and Sumaira Anjum <sup>1,\*</sup>

<sup>1</sup> Department of Biotechnology, Kinnaird College for Women, Lahore 54000, Pakistan; aaykay28@gmail.com (A.K.K.); humera.kausar@kinnaird.edu.pk (H.K.); syyada.samra@kinnaird.edu.pk (S.S.J.)

<sup>2</sup> Laboratoire de Biologie des Ligneux et des Grandes Cultures (LBLGC), INRAE USC1328, Université d'Orléans, 28000 Chartres, France; samantha.drouet@univ-orleans.fr (S.D.); hano@univ-orleans.fr (C.H.)

<sup>3</sup> Department of Biotechnology, Quaid-i-Azam University, Islamabad 45320, Pakistan; bhabbasi@qau.edu.pk

\* Correspondence: sumaira.anjum@kinnaird.edu.pk; Tel.: +92-300-6957038

Received: 21 September 2020; Accepted: 5 November 2020; Published: 6 November 2020



**Abstract:** With the increase in biotechnological, environmental, and nutraceutical importance of algae, about 100 whole genomic sequences of algae have been published, and this figure is expected to double in the coming years. The phenotypic and ecological diversity among algae hints at the range of functional capabilities encoded by algal genomes. In order to explore the biodiversity of algae and fully exploit their commercial potential, understanding their evolutionary, structural, functional, and developmental aspects at genomic level is a pre-requisite. So forth, the algal genomic analysis revealed us that algae evolved through endosymbiotic gene transfer, giving rise to around eight phyla. Amongst the diverse algal species, the unicellular green algae *Chlamydomonas reinhardtii* has attained the status of model organism as it is an ideal organism to elucidate the biological processes critical to plants and animals, as well as commercialized to produce range of bio-products. For this review, an overview of evolutionary process of algae through endosymbiosis in the light of genomics, as well as the phylogenomic, studies supporting the evolutionary process of algae was reviewed. Algal genomics not only helped us to understand the evolutionary history of algae but also may have an impact on our future by helping to create algae-based products and future biotechnological approaches.

**Keywords:** algae; bio-products; bioengineering; *Chlamydomonas reinhardtii*; evolution; genomics; lignin; phenylpropanoids; phylogenetics; phyla

## 1. Introduction

The application of genomic approaches, such as genome sequencing and gene function analysis, to algal research has caused a step change in our understanding of algal biology, ecology, and evolution. The rise of next-generation sequencing (NGS) technologies at a lower cost has led to the acquisition of significant genomic data on algae since 1990s [1]. Whole genomes of large-sized organisms can be sequenced more easily along with the availability of a greater number of bioinformatics tools for data assembly and annotation, easing-up the sequence analyses and prediction of the encoded biological function [2]. Genomics is a field concerned with defining sequences of genes and genomes and analysis of those sequences to understand the organization, architecture, expression, evolution, and function of genes and genomes [3]. With the increased biotechnological, environmental, and nutraceutical importance of algae, about 100 whole genomic sequences of algae have been published, and this figure is expected to double in the coming years [4].

Algae diversity range from tiny unicellular microalgae to giant seaweeds which can grow over 50 m long, abundantly found in virtually every ecosystem [5]. These ancient organisms are defined as unicellular or multicellular photosynthetic organisms regarded as polyphyletic due to morphological resemblance with plants but do not share a common ancestor [6]. Similar to plants, algae are autotrophs, photosynthesize, and produce the same storage compounds and defense strategies. On the contrary, algae do not possess specialized root, stem, nor vascular bundles, and they lack diploid embryo stage. The characteristic features of algae include autotrophic chlorophyll-bearing thalloid plant body, lacks sterile tissue around its reproductive structures, and zygote development is by mitosis or meiosis but not via embryo formation [7]. Algae are not only diverse in their size but also in their ecological distribution, cellular biology, photosynthetic pigments, structural and reserve polysaccharides, and evolutionary origin of heterogeneous algae include prokaryotic and eukaryotic species [8].

Algae are polyphyletic as their origin cannot be traced back to single common hypothetical ancestor. The symbiogenic events believed to have occurred more than 1.5 billion years ago were the reason algae came into existence, and, because of so much time gap, it is difficult to track the key events that resulted in algae occurrence [9]. It is thought that they came into existence when a photosynthetic cyanobacteria invaded a unicellular eukaryote giving rise to double-membranous primary plastid. This is known as primary endosymbiosis and as a result of this endosymbiosis Green algae (Chlorophytes and other land plants), Red algae (Rhodophytes), and Glaucophytes came into existence [10]. Another secondary endosymbiosis occurred when other heterotrophic eukaryotes engulfed red and green algae, due to which many other algal species came into existence, forming three or four membranes of secondary plastid [11]. The molecular sequence analysis suggests that there exist around eight to nine major phyla (divisions) of algae. These are the cyanobacteria (chloroxybacteria) and the eukaryotic phyla Dinophyta (dinoflagellates), Glaucophyta, (glaucophytes), Cryptophyta (cryptomonads), Euglenophyta (euglenoids), Ochrophyta (a diverse array of tiny flagellates, chrysophyceans, diatoms, brown algae, and many other groups), Haptophyta (haptophytes), Rhodophyta (red algae), and Chlorophyta (green algae) [12].

Algae are not only the primary producers in food chain and contributing enormously to various processes of ecosystem, they are also economically very important. Algae are a rich source of pharmaceutically important primary and secondary metabolites exhibiting anti-viral, anti-bacterial, and anti-cancerous activities [13,14]. Algae is also being exploited as an alternate renewable source of energy in the form of biofuels, bio-oils, biodiesel, bio-hydrogen, bio-methane, bio-ethanol, and bio-butanol [15]. Algae is a rich source of vitamin, pigments, carotenoids, fertilizer, fiber, agar, alginate, and numerous other products [16]. Interestingly, algae are consumed as food in the form of soups, and salads around the world. The famous algal-food wakame, nori, or wrap for sushi is rich in nutrition. Moreover, they are used as animal feed to improve animal health and quality of animal meat [17]. These are the uses of some of the algal species, while more than 30,000 algal species have already been described, and other algal groups still await discovery [18]. The phenotypic and ecological diversity among algae hints the range of functional capabilities encoded by algal genomes. In order to explore the biodiversity of algae and fully exploit their commercial potential, understanding their evolutionary, structural, functional, and developmental aspects at genomic level is a pre-requisite.

## 2. Insights to Algal Evolution

There is a general consensus that first photosynthetic eukaryote emerged from a heterotrophic eukaryote which captured a coccoid cyanobacterium about 1.6 billion years ago. Gradually, the cyanobacterium was enslaved, and the endosymbiont's genes transferred to the host nucleus as a new organelle called as primary plastid. This event is named as primary endosymbiosis [19]. The genome of cyanobacteria underwent evolution, resulting in either loss or transfer of genes to the host nucleus in a process termed as endosymbiotic gene transfer (EGT). As a result, only little portion of genome retained within primary plastid, including some genes encoding for photosynthesis [20].

The primary endosymbiosis gave rise to three groups, having primary plastids, forming a major group of autotrophic eukaryotes Archaeplastida: Glaucophytes, Red algae, and Green algae. Genomes of primary plastids consist of highly conserved genes employed extensively to study the evolution of diverse Archaeplastida groups [21].

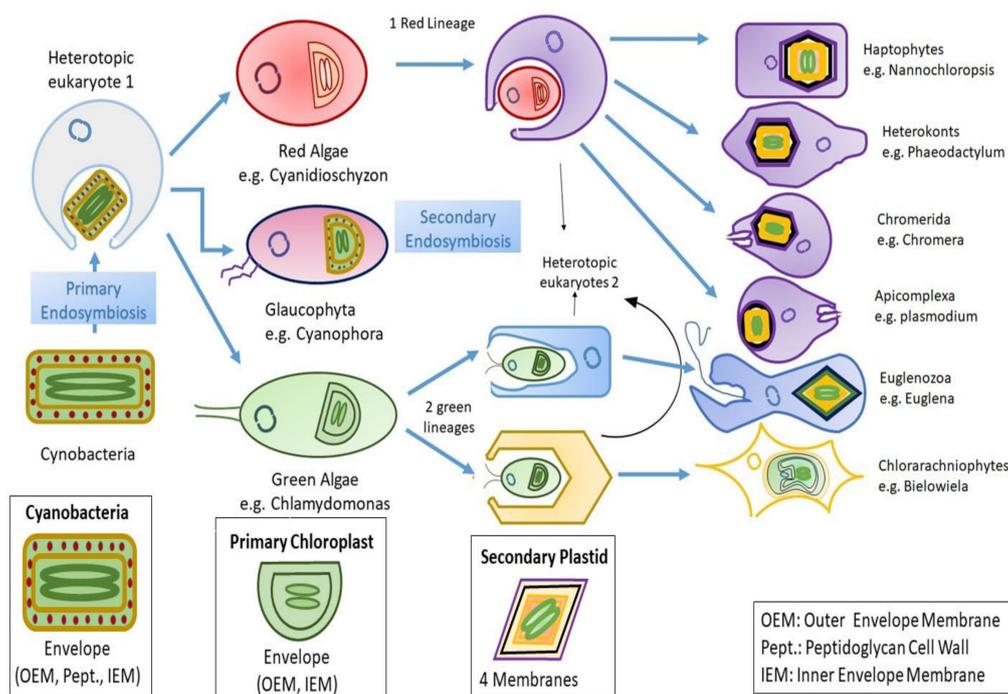
The Red algal lineage or Rhodophyta consist of 5000–6000 multicellular eukaryotic marine algal species. Phylum Rhodophyta is comprised of two sub-phyla Rhodophytina and Cyanidiophytina with seven classes [22]. Rhodophyta have evolved with plastids lacking chlorophyll accessory pigments, rather, containing chlorophyll-a, phycocyanin, and phycoerytherin, so light is captured by phycobiliproteins [23]. Secondly, there is complete absence of centrioles and flagella [24]. The genome in the red algae are compact in nature and have high gene capacity. The mitochondrial genome in Florideophyceae are considered highly conserved despite the loss of 32 genes, which suggest endosymbiotic gene transfer [25]. Genomic and transcriptomic analysis of four different classes of Red algae (Compsopogonophyceae, Stylonematophyceae, Rhodellophyceae, and Porphyridiophyceae) reveals significant evolutionary insights. The proteins from the plastid genome of 37 red algae were translated and analyzed, which supported the idea that all four classes show monophyly and are deeply rooted in red algal phylogenetic tree [26].

Green algae are termed as Chlorophyceae, and they are characterized by the presence of double membranous chloroplasts, stacked thylakoids, and photosynthetic pigments: chlorophyll-*a* and *b* [27]. Within their plastids they have starch molecules as polysaccharide reserve, a characteristic similar to green plants. The *Chlamydomonas reinhardtii* mutants showed starch biosynthesis similar to starch biosynthesized by maize endosperm [28]. Chlorophyceae contains five lineages: Chlamydomonadales, Chaetopeltidales, Oedogoniales, Sphaerpleales, and Chaetophorales [29]. They emerged about 470 million years ago, which marks one of the important evolutions of time, adapted all the photic zones of Earth, and paved way for the evolution of other life [30].

The glaucophytes (glaucocystophytes) are a group of freshwater microalgae which contain blue-green plastids often called as cyanelles. In five genera of glaucophytes, about thirteen species have been described. Glaucophyte plastids are similar to cyanobacteria, and different from plastids of other algae, in having a thin peptidoglycan wall and contain only phycobilins and chlorophyll *a* [31]. Glaucophyte is a poorly studied genera probably due to their rareness in nature and restricted habitat. Genomic data and a better understanding of the phylogenetic position of glaucophytes will provide valuable insights into the endosymbiotic origin and evolution of plastids in eukaryotes [32].

A phylogenomic study revealed a sister-group relationship between green or red algae and glaucophytes with a bootstrap value  $\geq 90\%$  [33]. The comparison of genome of diatom with other organism revealed that many genes present in its genome were obtained by the horizontal gene transfer (HGT) from bacteria. Moreover, viruses as gene transfer agents are of considerable importance in the marine environments [34]. Along with achieving new genes with HGT and EGT, gene loss, fusion and duplication have also added to size of algal genome. There is a great difference in the genome size among eukaryotes because of genome organization, like gene distribution. For example, *Ostreococcus tauri*, contain 8166 genes located on 20 chromosomes and spanning a genome size of 1.54 kb/gene [35].

Algae outside Archaeplastida have their plastids originating from acquisition of primary plastids during secondary endosymbiosis. This gave rise to haptophytes, cryptophytes, dinoflagellates, heterokonts, and other photosynthetic eukaryotes [36]. The existence of secondary endosymbiosis was first indicated by the occurrence of more than two envelope membranes around the secondary plastids. As evident from the presence of four membrane bounding plastids in chlorarachniophytes, cryptomonads, ochrophytes, and haptophytes, dinoflagellates and euglenoids appear to have lost one of these membranes possessing three [37]. Moreover, nucleomorph (remnant of endosymbiont nucleus) can be found within chlorarachniophytes and cryptophytes [38]. The process of algal evolution through endosymbiosis is depicted in Figure 1.



**Figure 1.** Diagrammatic representation of endosymbiosis: Primary endosymbiosis occurs when a heterotopic eukaryote engulfs a prokaryote (gram negative cyanobacteria), which leads to emergence of three lineages bearing primary chloroplast with two membranes that it vertically inherited from the cyanobacteria. Later on, the independent secondary endosymbiosis of green algae (eukaryote) by two unknown heterotopic eukaryotes that leads to emergence of Euglenozoa and Chlorarachniophyte. Red algae also go through similar mechanism of secondary endosymbiosis. The consequent secondary plastid contains four membranes, out of which two are the ones that were inherited from the cyanobacteria during primary endosymbiosis, and the other two are still unknown.

To add to the complexity of algal evolution, there seems to exist tertiary endosymbiosis too. Some dinoflagellates have gained their plastid either by replacing the red-alga-derived plastid with a green-alga-derived plastid, or by capturing secondary endosymbiotic cryptophyte, haptophyte or heterokont alga in tertiary endosymbiosis [39]. Genomic data is rapidly accumulating to help us further understand the evolutionary process of photosynthetic eukaryotes, which appears to be a complex and haphazard process.

### 3. Algal Genomes

The last decade has witnessed intense progress in algal genome sequencing with many whole-genome sequences (WGS) of representative algae available for major lineages. Green microalgae are single-celled green algae which inhabit a diverse ecosystem. The unicellular green algae *Chlamydomonas reinhardtii* has attained the status of model organism as it is ideal organism to elucidate the biological processes critical to plants and animals [40]. It is a haploid organism, consist of 110 Mb sized nuclear genome spanning over 17 chromosomes, a 16 kb mitochondrial genome, and chloroplast genome of 203 kb, encodes for 14,000 proteins [41]. Comparative genomic analysis of *Chlamydomonas* genome can help trace the green plant or plant-animal common ancestor. The BLASTp (Basic Local Alignment Search Tool for searching protein collections) comparison between *Chlamydomonas* proteome with human and *Arabidopsis* proteome revealed more similarity with *Arabidopsis*, whereas basal body and flagellar proteome showed similarity with human homologs [42].

Diatoms (Bacillariophyta) are found abundantly in every aquatic photic zone, emerging between 225–200 million years ago [43,44]. WGS of several representative diatoms have become available which includes *Thalassiosira oceanica*, *Thalassiosira pseudonana*, *Cyclotella cryptica*, *Phaeodactylum tricorutum*,

*Fragilariopsis cylindrus*, *Fistulifera solaris*, *Pseudo-nitzschia multistriata*, and *Synedra acus* [45]. In the phylogeny-based HGT study on all sequenced diatoms genomes, 1979 genes were found to be of horizontal origin in diatoms. Furthermore, five vitamin B12 biosynthesis pathway genes were detected as HGT, imparting a competitive advantage to diatoms in cobalt depleted environment, as cobalt is essential co-factor for diatoms [46]. In an iron-scarce environment, diatoms have the ability to reduce their cellular iron requirement and replaces plastid-localized ferredoxin with flavodoxin, using flavin as co-factor. This feature is similar to cyanobacteria and most other algae [47,48]. *Osterococcus tauri* and related species are a phytoplankton, and their smallest eukaryotes, known having a diameter of  $\approx 1\mu\text{M}$ , play a significant role in global primary productivity, food chain, and biogeochemical cycles [49]. The ability of *O. tauri* to thrive oceanic environment of low iron availability is a good model to study genetic adaptation of algae in ocean environment. The RNA sequencing of *O. tauri* under iron limitation showed that iron uptake and metabolism are tightly coordinated with diurnal cycles and differ fundamentally from *C. reinhardtii* iron metabolism [50].

The genome analysis also reveals insights into the genome mutations and its consequences of protein, as well as environmental, survival. Recently, transcriptome analysis of *C. reinhardtii* revealed cadmium tolerance genes and pathways on exposure to  $0.5\ \mu\text{M}$  cadmium for over 420 days. Genome analysis suggested a mutation and differential expression of plasma membrane calcium-transporting ATPase encoding gene (CHLREDRAFT\_189266), calmodulin binding protein encoding gene (CSE1) and calcium ion binding genes (CHLREDRAFT\_187187, CHLREDRAFT\_191203) in mutant and wild-type strains [51]. Moreover, differential expression of genes associated with cell cycle, protein synthesis and protein kinase-based phosphorylation in *C. reinhardtii* was observed under cold stress [52]. Such studies are important in enhancing our understanding of evolution of algae under harsh environment and comparing with other organisms, like plant stress tolerance.

Genomes of algae inhabiting extreme environments helps us understand the adaptive evolution of extremophile algae from their neutrophilic ancestors. *Chlamydomonas eustigma* is an acidophilic green algae and the comparative genome analyses between *C. eustigma* and *C. reinhardtii* (neutrophilic relative) revealed higher expression of  $\text{H}^+$ -ATPase and heat-shock proteins, as well as loss of metabolic pathways, that acidify cytosol by HGT in acidophile. Similar features are found in red and green acidophilic algae, hinting at a common evolutionary mechanism [53]. *Chlamydomonas sp.* ICE-L and *Tetraabaena socialis* are two Antarctic psychrophilic green algae thrive in extreme polar environment. Their phylogenetic analysis showed proofs of positive evolution and complex modifications in genes encoding for photosynthesis and protein synthesis. These adaptations include highly stable photosynthetic apparatus, enhanced membrane fluidity, antioxidant strategies, and proteins for cold response [54].

Cyanidiophytina red algae, sister to mesophilic red algae (the Rhodophytina), is a unicellular taxa able to thrive at pH 0–4 and temperatures extremes up to  $56\ ^\circ\text{C}$ . *Galdieria sulphuraria* is a hot-spring dwelling cryptoendolithic encompass 13.7 Mbp genomic sequence, demonstrated HGT from prokaryotic sources, and extensive gene loss in common ancestor [55].

The small genome of unicellular red algae *Cyanidioschyzon merolae* provides useful information on basic and essential genes required for the survival of autotrophic eukaryotes. The unique features of 20 chromosomal *C. merolae* genome were presence of very few introns, three copies of ribosomal DNA involved in the maintenance of nucleolus, unique telomeric repeats for chromosomal ends, and a small number of genes [56]. Moreover, the mosaic origin of Calvin cycle was identical in *C. merolae* and *A. thaliana*, supporting the single primary endosymbiotic plastid concept [57].

Studies based on comparative transcriptomic analysis of single genes also provide useful insights of algal evolution. Han and colleagues studied the origin and evolution of xylosyltransferases (XylTs), which are polysaccharide biosynthesis enzymes, in 22 sequenced primary endosymbionts (glaucomphytes, red algae, and green algae), as well as in secondary endosymbionts (diatoms, brown algae, Eustigmatophyceae, Cryptophyta, and Pelagophyceae). The study revealed that XylTs are distributed in Choanoflagellata and Echinodermata, suggesting their origin from last common eukaryotic ancestor

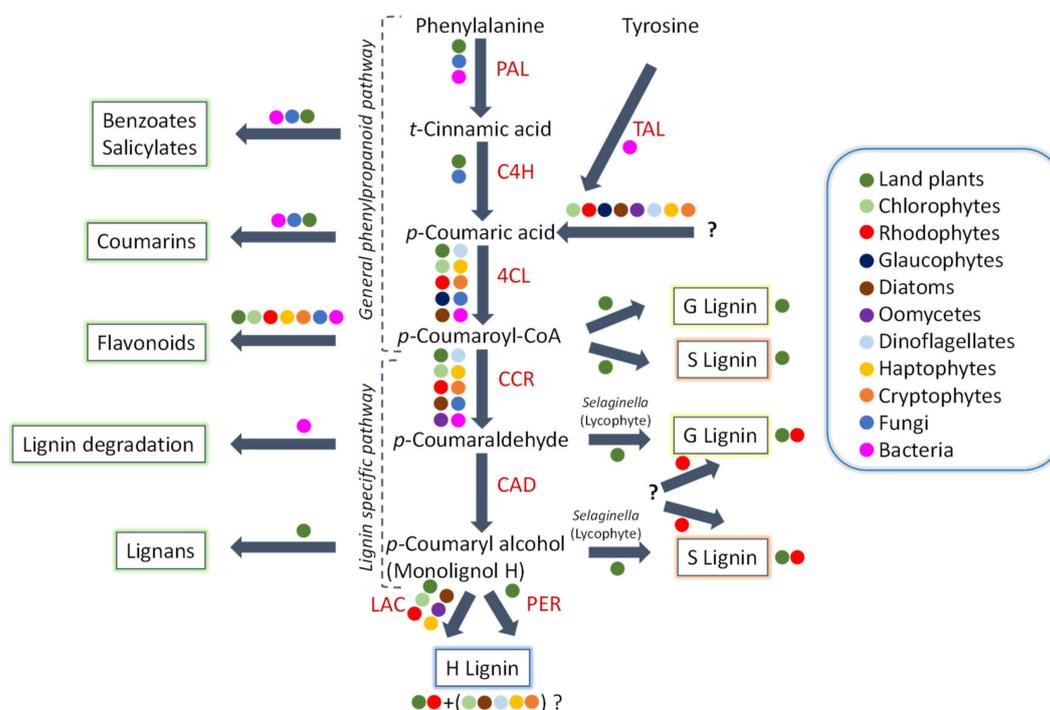
instead of being plant-specific [58]. Similarly, GRAS (Gibberellic-acid insensitive (GAI), Repressor of GAI (RGA) and Scarecrow (SCR)) and PYR/PYL/RCAR (Pyrabactin Resistance1 (PYR1)/PYR1-like (PYL)/regulatory components of ABA receptors (RCAR)) are genes which help zygnematophyceae and embryophytes survive biotic and abiotic stress. The two *Mesotaenium endlicherianum* and *Spirogloea muscicola* zygnematophyceae species were analyzed for their genome, revealed the origin of stress genes through HGT from soil inhabiting bacteria in common ancestor of zygnematophyceae and embryophytes [59].

Studying plastid genome of algae reveals information about algal evolution through endosymbiosis. *Emiliana huxleyi* is a marine calcifying phytoplankton belonging to coccolithophorids [60]. Its plastid DNA is densely packed and smaller than other red plastid lineages. The comparison of gene content, gene cluster, and gene function with plastid genomes of primary plastid and secondary plastid containing algae was drawn. Results disclosed that *Emiliana huxleyi* is related more to chlorophyll *c*-containing cryptophytes and heterokonts, and supports the descent of chromophyte from red algae [61]. Plastid genome comparison of microalgae *Gracilaria chilensis* with 17 red algal plastid sequences suggested separation of red algae into different groups, where Florideophyceae formed a large branch and Gracilariales as its sub-branch [62]. *Helicosporidium* is a non-photosynthetic alga, in which the plastid genome is evolved to have lost genes for major metabolic functions. Comparison with green algae and other distantly related non-photosynthetic plastids shows possibility of convergent evolution with apicomplexa [63].

In-depth knowledge of algal genomes can provide a thorough insight into the evolution of metabolic pathways critical to understanding the processes of evolution in the green lineage. One good example is the phenylpropanoid pathway including the lignification process. About 475 million years ago, the aquatic ancestral terrestrial plants adapted to form lignified cell walls which is widely regarded as one of the key innovations in the process of evolution [64]. Lignin plays a pivotal role in transport of water and structural support of plants by strengthening the secondary cell walls of xylem tissues and creates a dense matrix. The precursors of lignin may also have antimicrobial properties and may offer protection when a plant is damaged [65]. Such developmentally specialized cell walls were thought to be restricted to vascular plants as adaptations to terrestrial habitats within the algal lineage. The discovery of lignin or its precursor in different algae led to ponder over the evolution of its biosynthetic pathway [66]. The genes encoding for enzymes involved in the synthesis of *p*-coumaryl alcohol, which is a simplest lignin monomer (H units) have been reported in a range of photosynthetic eukaryotes. Moreover, presence of different lignin units (i.e., G, S, and H units, Figure 2) have been found in the cell walls of calcified red alga *Calliarthron cheilosporioides* (Rhodophyta), as well as in other algae [67]. In algae, putative key genes involved in the biosynthesis pathway of this simple lignin precursor have been identified in the haptophyte *Emiliana huxleyi*. The presence of *p*-coumaric acid and flavonoids have been also reported in diatom (*Phaedactylum tricorutum*) [68]. The potential of diatoms and haptophytes to biosynthesize monolignols suggests the possible evolution of this biosynthetic pathway in ancient oceans and later predating the origin of land plants, despite the fact that they are evolutionary distant from each other and plants, as well [67].

Interestingly, there are three main enzymes (cinnamoyl-CoA reductase (CCR), 4-coumarate:CoA ligase (4CL), and cinnamyl alcohol dehydrogenase (CAD)) involved in the biosynthesis pathway of *p*-coumaryl alcohol which are able to produce sufficient level of *p*-coumaryl alcohol from *p*-coumaric acid [67]. The phylogenetic analysis of the genes encoding these three enzymes reveals ancient origin of this metabolic pathway as their homologous enzymes are present in red algae, green algae, dinoflagellates, glaucophytes, cryptophytes, haptophytes, and diatoms alongside land plants. During the evolutionary process, these lignin biosynthesis enzymes evolved to become multifunctional or paralogues with divergent functions, such as a CAD enzyme, which catalyzes the last step in monolignol biosynthesis. Since red algae are polyphyletic, there might be some paralogs weakly clustering with homologs from heterotrophic eukaryotes or bacteria, whereas the *p*-coumaryl alcohol genes seem to be present in shared ancestor of green algae and plants. However, the presence

of pathway in the archaeplastid ancestor itself is difficult to predict before glaucophytes diverged. Further genomic analysis reveals the EGT of *p*-coumaryl alcohol biosynthesis pathway from the ancestors of green or red algae or themselves [67]. Further studies on the lignin or monolignols would help in filling knowledge gaps to understanding the crucial roles of these metabolites in terrestrial adaptation in the green lineage.



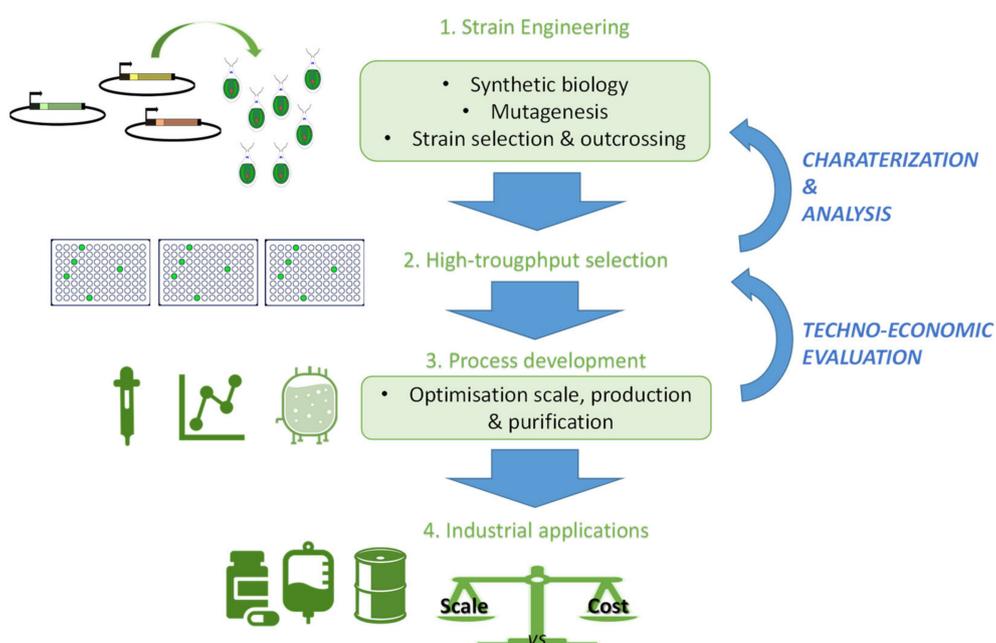
**Figure 2.** The lignin biosynthesis pathway. Colored dots represent the presence of a given enzyme in a specific taxonomic group. The abbreviations used for the enzyme are: phenylalanine ammonia-lyase; cinnamate 4-hydroxylase (C4H); 4-coumarate:CoA ligase (4CL); cinnamoyl-CoA reductase (CCR); cinnamyl alcohol dehydrogenase (CAD); peroxidase; and laccase (LAC). Question marks (?) signify that the enzyme responsible for the particular function or its substrate is still uncertain, or that the occurrence of a particular compound in a given taxonomic group is suspected but has not been proven. The figure was adapted from Reference [13,49,67].

#### 4. Conclusions and Future Prospects

The rapid increase in the number of sequenced genomes of algal species is opening doors of knowledge for scientists and giving us a new perspective to understand these very diverse and complex photosynthetic eukaryotes. Through phylogenomics and whole genome sequencing, we are able to recognize functional potential of individual algae, their adaptability to each ecological niche, and the movement of these genes among different organism through span of evolution.

The simplicity of algal genome, fast growth, direct synthesis of specific products, and the ability to fix CO<sub>2</sub> by capturing solar energy, algae serves an exceptional candidate as cellular factories for the production of value-added pharmaceuticals, nutraceuticals, and bioenergy. The molecular tools, such as sequencing technologies and genome editing, are empowering the research on algae and development of algal bioproducts [68]. Genetic modification of metabolic pathways in algae can help enhance yield of valuable products [69]. *C. reinhardtii* appears as one such algae which have a prominent role in basic research on algae that has allowed scientists to understand and genetically engineer algal metabolism. *C. reinhardtii* is among the first engineered algae to be studied at commercial scale; therefore, it helped to understand and improve the production of biofuels and bio-products in algae [69]. It was engineered for heterologous production of diterpenoids, which have health promoting benefits [70]. Algal genomics has also provided wealth of information regarding genes

involved in lipid biosynthesis. This information has helped in strain engineering to enhance lipid content for biofuel development [71]. However, the naturally low oil content, weak growth in outdoor ponds and sensitivity to high solar irradiation are some of the limitations in this regard which can be overcome with genetic engineering [69] (Figure 3). Given the generally regarded as safe (GRAS) status of *C. reinhardtii*, it can also be used as an effective production system for recombinant proteins at a lower cost and in larger quantities than in other eukaryotic production models, such as Chinese Hamster Ovary (CHO) cells [69,72]. To date, different therapeutic recombinant proteins have been produced in algae with proper post-translation modifications, most of them produced in *C. reinhardtii*, such as the erythropoietin hormone [73,74], or the anti-cancerous TNF (tumor necrosis factor)-related apoptosis inducing ligand [75].



**Figure 3.** The key steps in bioengineered green algae production (adapted from Scranton et al. [69]).

To summarize, it is clear that the algal genomics has not only helped us in understanding evolutionary history of algae, but it has impacted our future by helping in development of algal based products and future biotechnology approaches.

**Author Contributions:** Conceptualization, S.A, B.H.A., and C.H.; methodology, A.K.K., H.K., and S.S.J.; software, S.A. and C.H.; validation, B.H.A., A.K.K., and C.H.; formal analysis, S.A., H.K., and C.H.; investigation, S.S.J., H.K., S.D., A.K.K., C.H., B.H.A., and S.A.; resources, S.A. and C.H.; data curation, B.H.A. and A.K.K.; writing-original draft preparation, H.K., S.A., and A.K.K.; writing-review and editing, S.A., C.H., and A.K.K.; visualization, A.K.K., S.D., C.H., and B.H.A.; supervision, S.A. and C.H.; project administration, S.A.; funding acquisition, C.H., S.A., and B.H.A.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** S.A. acknowledges the administration of Kinnaird College for Women, Lahore for providing support. B.H.A. and C.H. acknowledge Le Studium-Institute for Advanced Studies, Loire Valley, Orléans, France.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Gilbert, J.A.; Dupont, C.L. Microbial metagenomics: Beyond the genome. *Ann. Rev. Mar. Sci.* **2011**, *3*, 347–371. [[CrossRef](#)] [[PubMed](#)]
2. Kong, J.; Huh, S.; Won, J.I.; Yoon, J.; Kim, B.; Kim, K. GAAP: A Genome Assembly + Annotation Pipeline. *Biomed. Res. Int.* **2019**, *2019*, 4767354. [[CrossRef](#)] [[PubMed](#)]
3. Blaby-Haas, C.E.; Merchant, S.S. Comparative and Functional Algal Genomics. *Annu. Rev. Plant Biol.* **2019**, *70*, 605–638. [[CrossRef](#)] [[PubMed](#)]
4. Tirichine, L.; Bowler, C. Decoding algal genomes: Tracing back the history of photosynthetic life on Earth. *Plant J.* **2011**, *66*, 45–57. [[CrossRef](#)] [[PubMed](#)]
5. Adelheid, S. *The Structure and Reproduction of the Algae Vol. I*; Cambridge University Press: London, UK, 1967; Volume 7, pp. 168–169.
6. Bajpai, P. Characteristics of Algae. In *Third Generation Biofuels*; Springer: Singapore, 2019; pp. 11–15. [[CrossRef](#)]
7. Andersen, R. Diversity of eukaryotic algae. *Biodivers. Conserv.* **1992**, *1*, 267–292. [[CrossRef](#)]
8. Sukenik, A.; Zohary, T.; Padisák, J. Cyanoprokaryota and other prokaryotic algae. In *Encyclopedia of Inland Waters*; Elsevier Inc.: Amsterdam, The Netherlands, 2009; pp. 138–148.
9. Reyes-Prieto, A.; Weber, A.P.; Bhattacharya, D. The origin and establishment of the plastid in algae and plants. *Annu. Rev. Genet.* **2007**, *41*, 147–168. [[CrossRef](#)]
10. Keeling, P.J.; Burger, G.; Durnford, D.G.; Lang, B.F.; Lee, R.W.; Pearlman, R.E.; Roger, A.J.; Gray, M.W. The tree of eukaryotes. *Trends Ecol. Evol.* **2005**, *20*, 670–676. [[CrossRef](#)]
11. Keeling, P.J. Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* **2004**, *91*, 1481–1493. [[CrossRef](#)]
12. Baweja, P.; Sahoo, D. Classification of algae. In *The Algae World*; Dinabandhu, S., Seckbach, J., Eds.; Springer: New York, NY, USA, 2015; pp. 31–55.
13. Stevenson, C.S.; Capper, E.A.; Roshak, A.K.; Marquez, B.; Eichman, C.; Jackson, J.R.; Mattern, M.; Gerwick, W.H.; Jacobs, R.S.; Marshall, L.A.; et al. The identification and characterization of the marine natural product scytonemin as a novel antiproliferative pharmacophore. *J. Pharm. Exp. Therap.* **2002**, *303*, 858–866. [[CrossRef](#)] [[PubMed](#)]
14. Tsai, C.-C.; Emau, P.; Jiang, Y.; Agy, M.B.; Shattock, R.J.; Schmidt, A.; Morton, W.R.; Gustafson, K.R.; Boyd, M.R. Cyanovirin-N inhibits AIDS virus infections in vaginal transmission models. *Aids Res. Hum. Retrovir.* **2004**, *20*, 11–18. [[CrossRef](#)]
15. Hannon, M.; Gimpel, J.; Tran, M.; Rasala, B.; Mayfield, S. Biofuels from algae: Challenges and potential. *Biofuels* **2010**, *1*, 763–784. [[CrossRef](#)] [[PubMed](#)]
16. Sharma, P.; Sharma, N. Industrial and biotechnological applications of algae: A review. *J. Adv. Plant Biol.* **2017**, *1*, 1. [[CrossRef](#)]
17. Scieszka, S.; Klewicka, E. Algae in food: A general review. *Crit. Rev. Food Sci. Nutr.* **2019**, *59*, 3538–3547. [[CrossRef](#)] [[PubMed](#)]
18. Bowler, C.; Allen, A.E. The contribution of genomics to the understanding of algal evolution. In *Unravelling the Algae: The Past, Present, Future of Algal Systematics*; CRC Press: Boca Raton, FL, USA, 2007; p. 331.
19. Yoon, H.S.; Hackett, J.D.; Ciniglia, C.; Pinto, G.; Bhattacharya, D. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **2004**, *21*, 809–818. [[CrossRef](#)]
20. Timmis, J.N.; Ayliffe, M.A.; Huang, C.Y.; Martin, W. Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **2004**, *5*, 123–135. [[CrossRef](#)]
21. Mackiewicz, P.; Gagat, P. Monophyly of Archaeplastida supergroup and relationships among its lineages in the light of phylogenetic and phylogenomic studies. Are we close to a consensus? *Acta Soc. Bot. Pol.* **2014**, *83*, 263–280. [[CrossRef](#)]
22. Yoon, H.S.; Müller, K.; Sheath, R.; Ott, F.; Bhattacharya, D. Defining the major lineages of red algae (Rhodophyta). *J. Phycol.* **2006**, *42*, 482–492. [[CrossRef](#)]
23. Glazer, A.N. Structure and molecular organization of the photosynthetic accessory pigments of cyanobacteria and red algae. *Mol. Cell. Biochem.* **1977**, *18*, 125–140. [[CrossRef](#)]

24. Yoon, H.S.; Nelson, W.; Lindstrom, S.C.; Boo, S.M.; Pueschel, C.; Qiu, H.; Bhattacharya, D. Rhodophyta. In *Handbook of the Protists*; Archibald, J.M., Simpson, A.G.B., Slamovits, C.H., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 89–133.
25. Salomaki, E.D.; Lane, C.E. Red Algal Mitochondrial Genomes Are More Complete than Previously Reported. *Genome Biol. Evol.* **2017**, *9*, 48–63. [[CrossRef](#)]
26. Moreira, D.; López-García, P. Evolution: King-Size Plastid Genomes in a New Red Algal Clade. *Curr. Biol.* **2017**, *27*, R651–R653. [[CrossRef](#)]
27. Tanaka, A.; Tanaka, R. Chapter Six-The biochemistry, physiology, and evolution of the chlorophyll cycle. In *Advances in Botanical Research*; Grimm, B., Ed.; Academic Press: Cambridge, MA, USA, 2019; Volume 90, pp. 183–212.
28. Busi, M.V.; Barchiesi, J.; Martín, M.; Gomez-Casati, D.F. Starch metabolism in green algae. *Starch* **2014**, *66*, 28–40. [[CrossRef](#)]
29. Lemieux, C.; Vincent, A.T.; Labarre, A.; Otis, C.; Turmel, M. Chloroplast phylogenomic analysis of chlorophyte green algae identifies a novel lineage sister to the Sphaeropleales (Chlorophyceae). *BMC Evol. Biol.* **2015**, *15*, 264. [[CrossRef](#)] [[PubMed](#)]
30. Domozych, D.; Ciancia, M.; Fangel, J.; Mikkelsen, M.; Ulvskov, P.; Willats, W. The Cell Walls of Green Algae: A Journey through Evolution and Diversity. *Front. Plant Sci.* **2012**, *3*, 1–7. [[CrossRef](#)]
31. Jackson, C.; Clayden, S.; Reyes-Prieto, A. The Glaucophyta: The blue-green plants in a nutshell. *Acta Soc. Bot. Pol.* **2015**, *84*, 149–165. [[CrossRef](#)]
32. Figueroa-Martinez, F.; Jackson, C.; Reyes-Prieto, A. Plastid Genomes from Diverse Glaucophyte Genera Reveal a Largely Conserved Gene Content and Limited Architectural Diversity. *Genome Biol. Evol.* **2018**, *11*, 174–188. [[CrossRef](#)] [[PubMed](#)]
33. Price, D.C.; Chan, C.X.; Yoon, H.S.; Yang, E.C.; Qiu, H.; Weber, A.P.M.; Schwacke, R.; Gross, J.; Blouin, N.A.; Lane, C.; et al. Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants. *Science* **2012**, *335*, 843–847. [[CrossRef](#)]
34. Bowler, C.; Allen, A.E.; Badger, J.H.; Grimwood, J.; Jabbari, K.; Kuo, A.; Maheswari, U.; Martens, C.; Maumus, F.; Otilar, R.P. The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature* **2008**, *456*, 239. [[CrossRef](#)] [[PubMed](#)]
35. Derelle, E.; Ferraz, C.; Rombauts, S.; Rouzé, P.; Worden, A.Z.; Robbens, S.; Partensky, F.; Degroeve, S.; Echeynié, S.; Cooke, R. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11647–11652. [[CrossRef](#)]
36. McFadden, G.I. Primary and secondary endosymbiosis and the origin of plastids. *J. Phycol.* **2001**, *37*, 951–959. [[CrossRef](#)]
37. Archibald, J. Nucleomorph genomes: Structure, function, origin and evolution. *BioEssays* **2007**, *29*, 392–402. [[CrossRef](#)]
38. Burki, F.; Kaplan, M.; Tikhonenkov, D.V.; Zlatogursky, V.; Minh, B.Q.; Radaykina, L.V.; Smirnov, A.; Mylnikov, A.P.; Keeling, P.J. Untangling the early diversification of eukaryotes: A phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. *Proc. R. Soc. B Biol. Sci.* **2016**, *283*, 20152802. [[CrossRef](#)]
39. Yoon, H.S.; Hackett, J.D.; Van Dolah, F.M.; Nosenko, T.; Lidie, K.L.; Bhattacharya, D. Tertiary Endosymbiosis Driven Genome Evolution in Dinoflagellate Algae. *Mol. Biol. Evol.* **2005**, *22*, 1299–1308. [[CrossRef](#)]
40. Shrager, J.; Hauser, C.; Chang, C.-W.; Harris, E.H.; Davies, J.; McDermott, J.; Tamse, R.; Zhang, Z.; Grossman, A.R. Chlamydomonas reinhardtii genome project. A guide to the generation and use of the cDNA information. *Plant Physiol.* **2003**, *131*, 401–408. [[CrossRef](#)] [[PubMed](#)]
41. Smith, D.R.; Lee, R.W. Nucleotide diversity of the Chlamydomonas reinhardtii plastid genome: Addressing the mutational-hazard hypothesis. *BMC Evol. Biol.* **2009**, *9*, 120. [[CrossRef](#)] [[PubMed](#)]
42. Merchant, S.S.; Prochnik, S.E.; Vallon, O.; Harris, E.H.; Karpowicz, S.J.; Witman, G.B.; Terry, A.; Salamov, A.; Fritz-Laylin, L.K.; Maréchal-Drouard, L. The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science* **2007**, *318*, 245–250. [[CrossRef](#)]
43. Janech, M.; Krell, A.; Mock, T.; Raymond, J. Ice-binding proteins from sea ice diatoms (Bacillariophyceae). *J. Phycol.* **2006**, *42*, 410–416. [[CrossRef](#)]

44. Nakov, T.; Beaulieu, J.M.; Alverson, A.J. Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (Diatoms, Bacillariophyta). *New Phytol.* **2018**, *219*, 462–473. [[CrossRef](#)]
45. Falciatore, A.; Jaubert, M.; Bouly, J.-P.; Bailleul, B.; Mock, T. Diatom Molecular Research Comes of Age: Model Species for Studying Phytoplankton Biology and Diversity. *Plant Cell* **2020**, *32*, 547–572. [[CrossRef](#)]
46. Vancaester, E.; Depuydt, T.; Osuna-Cruz, C.M.; Vandepoele, K. Systematic and functional analysis of horizontal gene transfer events in diatoms. *bioRxiv* **2020**. [[CrossRef](#)]
47. Pierella Karlusich, J.J.; Carrillo, N. Evolution of the acceptor side of photosystem I: ferredoxin, flavodoxin, and ferredoxin-NADP(+) oxidoreductase. *Photosynth. Res.* **2017**, *134*, 235–250. [[CrossRef](#)] [[PubMed](#)]
48. Groussman, R.D.; Parker, M.S.; Armbrust, E.V. Diversity and evolutionary history of iron metabolism genes in diatoms. *PLoS ONE* **2015**, *10*, e0129081. [[CrossRef](#)]
49. Palenik, B.; Grimwood, J.; Aerts, A.; Rouzé, P.; Salamov, A.; Putnam, N.; Dupont, C.; Jorgensen, R.; Derelle, E.; Rombauts, S.; et al. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7705–7710. [[CrossRef](#)]
50. Lelandais, G.; Scheiber, I.; Paz-Yepes, J.; Lozano, J.-C.; Botebol, H.; Pilátová, J.; Žárský, V.; Léger, T.; Blaiseau, P.-L.; Bowler, C.; et al. *Ostreococcus tauri* is a new model green alga for studying iron metabolism in eukaryotic phytoplankton. *BMC Genom.* **2016**, *17*, 319. [[CrossRef](#)]
51. Yu, Z.; Zhang, T.; Zhu, Y. Whole-genome re-sequencing and transcriptome reveal cadmium tolerance related genes and pathways in *Chlamydomonas reinhardtii*. *Ecotoxicol. Environ. Saf.* **2020**, *191*, 110231. [[CrossRef](#)]
52. Li, L.; Peng, H.; Tan, S.; Zhou, J.; Fang, Z.; Hu, Z.; Gao, L.; Li, T.; Zhang, W.; Chen, L. Effects of early cold stress on gene expression in *Chlamydomonas reinhardtii*. *Genomics* **2020**, *112*, 1128–1138. [[CrossRef](#)]
53. Hirooka, S.; Hirose, Y.; Kanesaki, Y.; Higuchi, S.; Fujiwara, T.; Onuma, R.; Era, A.; Ohbayashi, R.; Uzuka, A.; Nozaki, H.; et al. Acidophilic green algal genome provides insights into adaptation to an acidic environment. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8304–E8313. [[CrossRef](#)]
54. Zhang, Z.; Qu, C.; Yao, R.; Nie, Y.; Xu, C.; Miao, J.; Zhong, B. The Parallel Molecular Adaptations to the Antarctic Cold Environment in Two Psychrophilic Green Algae. *Genome Biol. Evol.* **2019**, *11*, 1897–1908. [[CrossRef](#)]
55. Qiu, H.; Price, D.C.; Weber, A.P.M.; Reeb, V.; Chan Yang, E.; Lee, J.M.; Kim, S.Y.; Yoon, H.S.; Bhattacharya, D. Adaptation through horizontal gene transfer in the cryptoendolithic red alga *Galdieria phlegrea*. *Curr. Biol.* **2013**, *23*, R865–R866. [[CrossRef](#)]
56. Nozaki, H.; Takano, H.; Misumi, O.; Terasawa, K.; Matsuzaki, M.; Maruyama, S.; Nishida, K.; Yagisawa, F.; Yoshida, Y.; Fujiwara, T.; et al. A 100%-complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae*. *BMC Biol.* **2007**, *5*, 28. [[CrossRef](#)] [[PubMed](#)]
57. Matsuzaki, M.; Misumi, O.; Shin-i, T.; Maruyama, S.; Takahara, M.; Miyagishima, S.-y.; Mori, T.; Nishida, K.; Yagisawa, F.; Nishida, K.; et al. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* **2004**, *428*, 653–657. [[CrossRef](#)]
58. Han, W.; Fan, X.; Teng, L.; Kaczurowski, M.J.S.; Zhang, X.; Xu, D.; Yin, Y.; Ye, N. Identification, classification, and evolution of putative xylosyltransferases from algae. *Protoplasma* **2019**, *256*, 1119–1132. [[CrossRef](#)]
59. Cheng, S.; Xian, W.; Fu, Y.; Marin, B.; Keller, J.; Wu, T.; Sun, W.; Li, X.; Xu, Y.; Zhang, Y.; et al. Genomes of Subaerial Zygnematophyceae Provide Insights into Land Plant Evolution. *Cell* **2019**, *179*, 1057–1067.e14. [[CrossRef](#)]
60. McClelland, H.L.O.; Barbarin, N.; Beaufort, L.; Hermoso, M.; Ferretti, P.; Greaves, M.; Rickaby, R.E.M. Calcification response of a key phytoplankton family to millennial-scale environmental change. *Sci. Rep.* **2016**, *6*, 34263. [[CrossRef](#)]
61. Puerta, M.V.S.; Bachvaroff, T.R.; Delwiche, C.F. The Complete Plastid Genome Sequence of the Haptophyte *Emiliana huxleyi*: A Comparison to Other Plastid Genomes. *DNA Res.* **2005**, *12*, 151–156. [[CrossRef](#)]
62. Wang, G.; Liu, N.; Li, Y.; Zhang, L.; Meinita, M.D.N.; Chen, W.; Liu, T.; Chi, S. The complete plastid genome and phylogenetic analysis of *Gracilaria chilensis*. *Mitochondrial DNA Part B* **2020**, *5*, 1282–1283. [[CrossRef](#)]
63. De Koning, A.P.; Keeling, P.J. The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured. *BMC Biol.* **2006**, *4*, 12. [[CrossRef](#)]
64. Kenrick, P.; Crane, P. The origin and early evolution of plants on Land. *Nature* **1997**, *389*, 33–39. [[CrossRef](#)]

65. Tronchet, M.; Balague, C.; Kroj, T.; Jouanin, L.; Roby, D. Cinnamyl alcohol dehydrogenases-C and D, key enzymes in lignin biosynthesis, play an essential role in disease resistance in Arabidopsis. *Mol. Plant Pathol.* **2010**, *11*, 83–92. [[CrossRef](#)]
66. Martone, P.T.; Estevez, J.M.; Lu, F.; Ruel, K.; Denny, M.W.; Somerville, C.; Ralph, J. Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. *Curr. Biol.* **2009**, *19*, 169–175. [[CrossRef](#)]
67. Labeeuw, L.; Martone, P.T.; Boucher, Y.; Case, R.J. Ancient origin of the biosynthesis of lignin precursors. *Biol. Direct.* **2015**, *10*, 23. [[CrossRef](#)]
68. Dahlin, L.; Guarnieri, M. Recent advances in algal genetic tool development. *Curr. Biotechnol.* **2016**, *5*, 192–197. [[CrossRef](#)]
69. Scranton, M.A.; Ostrand, J.T.; Fields, F.J.; Mayfield, S.P. Chlamydomonas as a model for biofuels and bio-products production. *Plant J.* **2015**, *82*, 523–531. [[CrossRef](#)]
70. Lauersen, K.J.; Wichmann, J.; Baier, T.; Kampranis, S.C.; Pateraki, I.; Møller, B.L.; Kruse, O. Phototrophic production of heterologous diterpenoids and a hydroxy-functionalized derivative from Chlamydomonas reinhardtii. *Metab. Eng.* **2018**, *49*, 116–127. [[CrossRef](#)]
71. Radakovits, R.; Jinkerson, R.E.; Darzins, A.; Posewitz, M.C. Genetic engineering of algae for enhanced biofuel production. *Eukaryot. Cell* **2010**, *9*, 486–501. [[CrossRef](#)]
72. Manuell, A.L.; Beligni, M.V.; Elder, J.H.; Siefker, D.T.; Tran, M.; Weber, A.; McDonald, T.L.; Mayfield, S.P. Robust expression of a bioactive mammalian protein in Chlamydomonas chloroplast. *Plant Biotechnol. J.* **2007**, *5*, 402–412. [[CrossRef](#)] [[PubMed](#)]
73. Eichler-Stahlberg, A.; Weisheit, W.; Ruecker, O.; Heitzer, M. Strategies to facilitate transgene expression in Chlamydomonas reinhardtii. *Planta* **2009**, *229*, 873–883. [[CrossRef](#)]
74. Walsh, G. Biopharmaceutical benchmarks. *Nat. Biotechnol.* **2000**, *18*, 831–833. [[CrossRef](#)]
75. Yang, Z.; Chen, F.; Li, D.; Zhang, Z.; Liu, Y.; Zheng, D.; Wang, Y.; Shen, G. Expression of human soluble TRAIL in Chlamydomonas reinhardtii chloroplast. *Chin. Sci. Bull.* **2006**, *51*, 1703–1709. [[CrossRef](#)]

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).