Trends in Plant Science

The Algal Revolution

--Manuscript Draft--

Manuscript Number:	PLANTS-D-17-00020R1
Article Type:	Review
Corresponding Author:	Debashish Bhattacharya, PhD Rutgers University New Brunswick, New Jersey UNITED STATES
First Author:	Debashish Bhattacharya, PhD
Order of Authors:	Debashish Bhattacharya, PhD
	Juliet Brodie
	Cheong Xin Chan
	Olivier De Clerck
	J. Mark Cock
	Susana M Coelho
	Claire Gachon
	Arthur R Grossman
	Thomas Mock
	John A Raven
	Alison G Smith
	Hwan Su Yoon
Abstract:	Algae are (mostly) photosynthetic eukaryotes that occupy multiple branches of the tree of life, and are vital for planet function and health. This review highlights a transformative period in studies of the evolution and functioning of this extraordinary group of organisms and their potential for novel applications, wrought by high-throughput 'omic' and reverse genetic methods. It covers the origin and diversification of algal groups, explores advances in understanding the link between phenotype and genotype, considers algal sex determination, and reviews progress in understanding the roots of algal multicellularity. Experimental evolution studies to determine how algae evolve in changing environments are highlighted, as is their potential as production platforms for compounds of commercial interest such as biofuel precursors, nutraceuticals, or therapeutics.

Review

The Algal Revolution

Juliet Brodie,[1,†] Cheong Xin Chan,[2] Olivier De Clerck,[3] J. Mark Cock,[4] Susana M. Coelho,[4] Claire Gachon,[5] Arthur R. Grossman,[6] Thomas Mock,[7] John A. Raven,[8,9] Alison G. Smith,[10] Hwan Su Yoon,[11] Debashish Bhattacharya[12, †,*]

[1] Natural History Museum, Department of Life Sciences, London SW7 5BD, United Kingdom
[2] Institute for Molecular Bioscience, The University of Queensland, Brisbane QLD 4072, Australia
[3] Research Group Phycology, Ghent University, Krijgslaan 281, S8, 9000 Gent, Belgium
[4] CNRS, Sorbonne Université, UPMC University Paris 06, Algal Genetics Group, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, CS 90074, Roscoff F-29688, France
[5] Scottish Association for Marine Science, Scottish Marine Institute, PA37 1QA Oban, United Kingdom
[6] Department of Plant Biology, The Carnegie Institution, Stanford, CA 94305, USA
[7] School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom
[8] Division of Plant Sciences, University of Dundee at the James Hutton Institute, Dundee DD2 5DA, United Kingdom (permanent address)
[9] School of Plant Biology, University of Western Australia, Crawley, WA 6009, Australia
[10] Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom
[11] Department of Biological Sciences, Sungkyunkwan University, Suwon 440-746, Korea
[12] Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08901, USA
[†] These authors contributed equally to this work.
[*]For correspondence: d.bhattacharya@rutgers.edu (D. Bhattacharya)
Laboratory website: http://dblab.rutgers.edu/

Twitter handle: DebashB

Keywords: Archaeplastida, genomics, origin of multicellularity, plastid endosymbiosis, systems biology

Abstract

Algae are (mostly) photosynthetic eukaryotes that occupy multiple branches of the tree of life, and are vital for planet function and health. This review highlights a transformative period in studies of the evolution and functioning of this extraordinary group of organisms and their potential for novel applications, wrought by high-throughput 'omic' and reverse genetic methods. It covers the origin and diversification of algal groups, explores advances in understanding the link between phenotype and genotype, considers algal sex determination, and reviews progress in understanding the roots of algal multicellularity. Experimental evolution studies to determine how algae evolve in changing environments are highlighted, as is their potential as production platforms for compounds of commercial interest such as biofuel precursors, nutraceuticals, or therapeutics.

The diversity and ecological importance of algae

Algae represent a vast array of photosynthetic and non-photosynthetic eukaryotes of ancient origin scattered throughout the tree of life (Fig. 1A; **Box 1**), but less than 10% have been formally described [1]. Algae are critical to the health of our planet because they are dominant in the oceans, which cover about 71% of the Earth's surface where they provide core ecosystem services and produce about one-half of the oxygen that we breathe [2]. Algae are also integral to human activities; they can serve as a food source or in some cases represent a significant nuisance or health hazard (**Box 2**). These organisms occupy a vast range of habitats from desert crusts to coastal and oligotrophic oceans where massive algal blooms can extend over thousands of km^2. Furthermore, the body mass of algae can differ by 20 orders of magnitude (Fig. 1B). *Ostreococcus*, a green microalga has a volume of <1 μm^3 making it the smallest known free-living eukaryote [3], the giant kelp (*Macrocystis pyrifera*) can grow up to 45 m in length [4], and the siphonous green alga *Caulerpa* has cells that grow to a meter in length and differentiate into distinct 'organs' [5]. The brown and red macroalgae are two of the few groups of organisms that have made the transition from unicellularity to complex multicellularity (**Box 1**; and see below). In addition to the photosynthetic lineages that exploit sunlight to reduce CO_2 to organic carbon, there exist a myriad of mixotrophs and facultative heterotrophs, some of which have lost the plastid and evolved into obligate heterotrophs such as *Plasmodium* and *Toxoplasma* that parasitize other organisms [6-8]. The study of algae is undergoing a revolution, largely due to the development of increasingly sophisticated 'omics' (e.g., genomics, proteomics, epigenomics, metagenomics), bioinformatics, systems biology, and novel reverse genetic approaches such as the use of CRISPR-Cas9 for editing their genomes [9-12]. In contrast to the routine application of genomic editing methods to land plant species, their development in different algal groups such as diatoms, haptophytes, and red algae is highly challenging because these taxa are ancient splits in the eukaryotic tree of life. For example, tools such as transformation protocols, transgene promoters, and nuclear localization signals usually have to be developed specifically for each lineage. Nonetheless, modern high-throughput methods have allowed researchers to generate testable hypotheses that set the stage for years of detailed biochemical, cell biological, and genetic experimentation. Here, we document the transformative impact of modern omics and reverse genetics research on our understanding of algae.

Exploring the origins and functions of algal genes

To understand the success of algae, it is important to identify the major genetic processes that have shaped their genomes and how these processes impacted phenotypes. However, disentangling the effects of genotype versus phenotype on algal evolution and fitness is challenging because both are linked and highly plastic. In this section, we discuss the interplay between genotype and phenotype, and review new tools for functional analysis in algae (**Box 3**).

Endosymbiotic gene transfer

The most significant introduction of foreign genes into the nuclear genome of all photosynthetic organisms was the result of endosymbiotic gene transfer (**EGT**; see Glossary) [13], which enriched the nuclear genome with cyanobacterial genes originating from the primary plastid endosymbiont (Fig. 1C and **Box 1**). Most EGT-derived gene products are targeted back to the organelle (via identifiable N-terminal targeting sequences called transit peptides) where they express their original or related function, whereas others evolved novel roles in the host cell [14, 15]. The plastid-destined EGT proteins encode a range of functions, many of which are associated with photosynthesis, metabolite biosynthesis, transcription and translation, plastid biogenesis and macromolecular complex assembly, and the regulation of photosynthetic activity [16, 17]. EGT-derived cyanobacterial genes may also have fused with various DNA sequences from the host or from various prokaryotes to create the so-called symbiogenetic genes (**S-genes;** see Glossary) [18]. S-genes play key roles in algal and plant metabolism, including responses to oxidative stress, phototropism, and adaptation to N limitation [18]. In the following section, we focus on the search for functions encoded by EGT-derived genes using the model green alga, *Chlamydomonas reinhardtii* and the so-called **GreenCut** (see Glossary).

The GreenCut proteins

Comparative genomics using the *C. reinhardtii* genome and genomes from other green lineage organisms, including the terrestrial embryophytes *Arabidopsis thaliana*, *Physcomitrella patens*, *Oryza sativa*, *Populus trichocarpa*, as well as *Ostreococcus* species (*O. tauri*, *O. lucimarinus*, or *Ostreococcus* sp. RCC809) revealed a set of proteins designated the GreenCut. These genes are absent from, or are highly diverged in heterotrophic (non-photosynthetic) organisms [16]. The latest GreenCut (designated GreenCut2) includes a set of 597 *C. reinhardtii* proteins

[16, 17], of which almost one-half have unknown functions, although many have domains that provide some information about function. Whereas a subset of these proteins are exclusively conserved in green lineage organisms (Conserved in Green Lineage, CGL), others are also found in at least one red alga (Conserved in PLantae, CPL), at least one diatom (Conserved in Green Lineage and Diatoms, CGLD), or in the Plantae (in the Archaeplastida) and diatoms (Conserved in the PLantae and Diatoms, CPLD) [19]. Orthologs of many GreenCut proteins are also present in cyanobacteria. The use of the GreenCut suite of proteins assumes that the functions of member proteins of any specific ortholog group are conserved among green lineage organisms.

Approximately 70% of GreenCut proteins are predicted to be targeted to chloroplasts where they function in photosynthesis and several other metabolic pathways, chloroplast biogenesis/assembly, and the regulation of photosynthetic activity [16, 17]. A recently characterized GreenCut protein, CGL71, resides in thylakoid membranes, has a tetratricopeptide repeat domain, and has been shown to be required for accumulation/maintenance of photosystem I (PSI) [20-22]. A *C. reinhardtii* mutant that is null for CGL71 exhibits very low levels of PSI, cannot grow as a photoautotroph, and when grown heterotrophically, is high light sensitive. Amazingly, this mutant can be largely rescued (ca. 70% recovery of PSI activity) when it is maintained under hypoxic conditions. These results suggest that CGL71 functions under atmospheric O_2 conditions to prevent oxidative disruption of the maturation/assembly of PSI, which might reflect the need for PSI to associate with O_2 sensitive iron-sulfur (Fe-S) clusters. These findings have implications for the evolution of the Earth's atmosphere. Oxygenic photosynthesis evolved roughly 2.5 BYA when the atmosphere was largely anoxic. Therefore, even an O_2-sensitive cofactor/complex that became integral to the ancestral photosynthetic electron transport system would have likely remained stable in the early Earth's atmosphere. However, as the atmosphere transitioned to oxic conditions as a consequence of photosynthetic O_2 evolution, specific proteins, like CGL71, may have evolved to assist in assembling and stabilizing complexes containing O_2-sensitive cofactors. The need for mechanisms to prevent oxic disruption of both assembly intermediates and mature macromolecular complexes extends beyond photosynthetic processes; i.e., it would be critical for any complex/activity in the cell that requires an O_2 sensitive cofactor.

Algae in a changing environment

Microalgae have proven to be ideal models for experimental evolution approaches due to their ability to adapt and/or acclimate to changing environmental conditions. Their short generation times and asexual mode of reproduction makes it possible to cultivate them for hundreds to thousands of generations within a relatively short time. Despite these advantages, the large number of interacting variables associated with increasing atmospheric CO_2 and temperature makes it a challenge to design short-term studies to probe the acclimation of algae to these factors, and longer-term experimental evolution studies of genetic change. Accordingly, most experimental work has involved only a single variable: increased CO_2 [23, 24]. One of the longest experiments (4 years, 2100 generations [24]) examined the response of a high-CO_2-adapted population of the bloom-forming *Emiliania huxleyi* after it was returned to its original CO_2 condition and concluded that phytoplankton may evolve complex phenotypic plasticity. An instructive example for algal systems focused on N_2 fixation by the marine cyanobacterium *Trichodesmium* that was exposed to elevated CO_2 (750 p.p.m.) for 4.5 years (about 850 generations). These cultures showed significantly higher N_2 fixation rates and growth rates under P-limited conditions, as well as shifts in the diel occurrence of peak N_2 fixation [25]. The results were consistent with those for short-term (two week) incubations in elevated CO_2. These effects were maintained even when the cultures were returned to the ancestral CO_2 levels (380 p.p.m.) for 2 years. However, analysis of the proteome and enzyme activities did not reveal the basis of these changes [25]. These results might indicate a high level of phenotypic plasticity and surprisingly, genetic fixation of the adaptive phenotypes after a relatively short exposure to high CO_2. To test this hypothesis would require identifying genetic and epigenetic processes of adaptation, which are currently not well understood in cyanobacteria and algae [26].

Padfield et al. [27] investigated the changes in the metabolism of freshwater *Chlorella vulgaris* over 100 generations of exposure to a higher temperature, but without demonstrating that the effects were due to adaptation rather than acclimation. Schlüter et al. [28] examined the impact of increased temperature and increased CO_2, separately and together, on *E. huxleyi* and showed that there was no interaction of CO_2 and temperature. Bermúdez et al. [29] investigated the nutritional value of marine foods using a fully factorial design (three CO_2 concentrations, two temperatures) and reported adverse effects of high CO_2 on the synthesis of essential amino acids and polyunsaturated fatty acids in *Cylindrotheca fusiformis* over 250 generations. Innovative

work by Schaum and Collins [30], Schaum et al. [31], and Doblin and Sebille [32] provide hope for predicting the adaptation of phytoplankton strains to changes in CO_2 and temperature levels.

Transcriptomic data are available for some acclimation and adaptation studies [33], but there is very little genomic analysis coupled to adaptation experiments, except for the work done by Perrineau et al. [34, 35] on the evolution of laboratory cultures of *C. reinhardtii*. As long as we do not understand how the genetic repertoire of algae underpins the response to changing environmental conditions, the evolutionary mechanisms remain elusive. For instance, the polar diatom *Fragilariopsis cylindrus* has evolved to cope with what can be considered one of the most extreme environmental changes in nature: the transformation from a liquid (sea water) into a partly habitat (sea ice) [36]. The genome sequence of this diatom revealed that almost a quarter of the genes had markedly divergent alleles that were differentially expressed under changing environmental conditions. Metatranscriptomes from natural communities of *F. cylindrus* were dominated by these divergent alleles, providing evidence that the alleles are important for cellular function and were selected by the polar environment as an evolutionary mechanism underpinning the success of *F. cylindrus* under highly variable conditions. Whether these diverged alleles will provide an advantage for coping with global warming remains unknown, but with an effective population size (N_e) ≈ 16.5 x 10^7, it might be said that there is an allele for every occasion.

Algal sex determination

Sexual cycles are an ancestral feature of eukaryotes, and genomic analyses demonstrate the presence of meiotic genes in all supergroups, including those containing algae [37]. The systems that regulate mating (mating type loci and sex chromosomes) are more recent in origin, remarkably diverse, and have emerged independently and repeatedly during evolution [38].

In genetically controlled sexual systems, mating types or sexes are determined by defined non-recombining chromosomal regions that can be as small as a single locus or as large as an entire chromosome. Analysis of green and brown (stramenopile) algal lineages has led to novel and important contributions to our understanding of how sexes have emerged and how sex-determining mechanisms evolve. One particularly interesting group is the volvocine algae, which possess a broad range of sexual systems ranging from unicellular, isogamous species such as *C. reinhardtii*, which have two equal-sized gametes of plus and minus mating types, to

multicellular, oogamous species such as *Volvox carteri,* with sperm-producing males and egg-producing females. Comparative genomic analyses of volvocine species have shed new light on the origin of male and female sexes, providing clear evidence that the sexes emerged from mating types in this lineage [39-41]. Surprisingly, the emergence of a sexual system with male sperm and female eggs in *V. carteri* does not appear to have involved the recruitment of additional genes (influencing gamete size for example) into the sex-determining genomic region, as was previously predicted [42]. Rather, sexual dimorphism can arise from isogamy largely via adaptations of the master sex-determining gene (*MID*) itself. These changes appear to re-wire regulatory networks, such that in an isogamous organism, the result is a simple system that determines mating type. In contrast, oogamous organism such as *Volvox* require more complex developmental programs that lead to the determination of spermatogenesis or oogenesis [39].

The brown algae represent a key group for studying the evolution of sexes because they too exhibit a broad range of different sexual characteristics (e.g., isogamy/ anisogamy, sex determination in either the diploid or the haploid phase of the life cycle, and varying levels of sexual dimorphism) [43]. Sexual systems that function during the haploid phase of the life cycle (so-called UV sexual systems, where U and V refer to the female and male sex chromosomes, respectively) are of particular interest because they exhibit novel evolutionary characteristics compared to the better studied diploid phase systems (XY and ZW sex chromosomes) [38, 43, 44]. Analysis of the UV sex chromosomes of the brown alga *Ectocarpus* provided the first detailed genetic description of a sex-determination system for a multicellular species outside the opisthokont and green plant lineages [44]. These data supported theoretical predictions about the evolutionary dynamics of sex chromsomes that could be tested empirically. For example, it has been proposed that purifying selection during the haploid phase of the life cycles should prevent degeneration of both the U and V sex-determining regions. Analysis of the *Ectocarpus* sex-determining region did not suggest marked degeneration but there was, nonetheless, evidence for some minor genic erosion. Theoretical models also predict that the presence of sexually antagonistic genes in the recombining regions of sex chromosomes may drive expansion of the non-recombining sex-determining region [45]. The small size of the *Ectocarpus* sex-determining region is consistent with this prediction, given that this alga exhibits a low degree of sexual dimorphism and has few sex-biased genes, indicative of a low level of sexual antagonism [46]. The recombining regions of the *Ectocarpus* sex chromosome, known as pseudoautosomal

regions, also have exceptional evolutionary and structural features. These regions are enriched in orphan genes that may be selectively maintained because of their important roles in sporophytes [47]. Ongoing analysis of the diverse algal sexual systems across the brown algal tree of life is expected to bring new insights not only to the understanding of the mechanics of sex chromosome function, but also to the interplay between the sexual system and sex chromosome evolution.

Advances in algal biotechnology

With the availability of >30 microalgal genomes and >500 transcriptomes [48], there are significant opportunities to exploit algal metabolism for biotechnological purposes such as biofuels, pharmaceuticals, nutraceuticals and biomaterials. For the purpose of this review, we will only highlight some of these areas because the field of algal biotechnology is rapidly expanding with many studies being done on a variety of algal species. The reason why algae are promising organisms for biotechnology is again rooted in their evolution. Their genetic diversity results in biochemical diversity, which offers opportunities to discover novel metabolic pathways and novel active molecules to serve many different biotechnological purposes [49]. Moreover, being photosynthetic means they offer an advantage over the more traditional bacterial or yeast hosts that require inputs of fixed carbon, and so in principle algae are more sustainable. However, this is not guaranteed, and any industrial process requires careful life cycle assessment to establish its level of sustainability [50]. The fact that many algae can grow mixotrophically or heterotrophically would potentially provide an alternative if growth in photobioreactors is unsustainable or too constrained by the footprint of any commercial venture.

Algae have been the subject of much investigation for biofuel production because many of them accumulate triacylglycerides (TAGs), which can be used as a feedstock for biodiesel [51], although in many cases the accumulation of TAG only occurs in response to nutrient deprivation (e.g., nitrogen [N]). As a result, cell growth is inhibited leading to poor overall productivity, which is one of the major challenges that need to be overcome to make algal biofuels a commercial reality [52, 53]. Recent work shows that microalgae encode multiple genes for enzymes that catalyze the last step in the metabolic pathway for the production of TAG, the addition of a third acyl group onto diacylglycerol [54]. By analyzing RNAseq data it was possible to find which of these genes were upregulated upon N-starvation, and therefore

potentially involved in the increase in TAG synthesis during starvation. In *C. reinhardtii* only one of the five genes encoding diacylglycerol acyltransferase type 2 (DGTT1) had this characteristic [55]. However, overexpression of this and two other DGAT2 genes in *C. reinhardtii* under the control of the strong light-responsive *PSAD* promoter had no effect on total lipid or TAG levels, likely indicating tight regulation of this pathway. A similar experiment in the diatom *Phaeodactylum tricornutum* found that expression of Pt*DGAT2A* under control of the light-responsive *FCPC* promoter somewhat increased neutral lipid levels, but also had an effect on the proportion of unsaturated fatty acids in all cellular glycerolipids, not just TAGs [56]. This illustrates a common observation, which is the generation of unexpected consequences from the introduction of genes for metabolic enzymes.

More detailed systems-level expression has started to identify key factors important in the cellular response to alterations in the C:N ratio. Transcriptomic analysis of N-starved *P. tricornutum* cells provided evidence that N limitation led to a remodeling of intermediary metabolism that shifted the flux of photosynthetically assimilated carbon from amino acid biosynthesis towards lipids, and helped to conserve N further by recycling products of protein degradation through the urea cycle [57]. Boyle et al. [58] analyzed *C. reinhardtii* RNAseq data and found that a transcript for a SQUAMOSA promoter-binding protein domain transcription factor increased just prior to the transcript encoding DGTT1. Mutants in this gene, named *NRR1* for nitrogen response regulator, showed much lower TAG accumulation. A slightly different approach was taken with *P. tricornutum*. Promoters of genes that were activated early and strongly upregulated upon N deprivation contain certain overrepresented motifs. A RING-domain protein (RGQ1) identified by yeast 1-hybrid analysis that binds these N-deprivation motifs acts as a transcription regulator [59]. These insights suggest a number of strategies that might provide effective targets for manipulation, such as the reduction of lipid catabolism to engineer the synthesis of TAG without compromising growth. An example of this approach was the knockdown of lipid catabolism, specifically lipases that catalyze the release of free fatty acids (FAs) from lipids, and therefore increase lipid accumulation. Targeted knockdown of a multifunctional lipase/phospholipase/ acyltransferase increased lipid yields without affecting growth in the diatom *Thalassiosira pseudonana* [60]. As also observed in *C. reinhardtii*, overexpressing genes encoding enzymes of TAG biosynthesis was less successful for stimulating lipid accumulation [60].

Recently, the focus on algal bioproducts has turned away from low-value, high volume biofuels, to high value products such as nutraceuticals (e.g., vitamins, pigments, antioxidants), omega-3 fatty acids, or other novel chemicals [61, 62]. The diversity of algal species means that many novel pathways remain to be identified. For example *Botryococcoccus braunii*, a colonial green alga secretes copious amounts of various straight-chain and branched hydrocarbons between cells in the colony but grows extremely slowly, reducing its potential as a production strain. However, genes thought to encode specific enzymes involved in hydrocarbon synthesis such as squalene synthase-like enzymes (SSL) and triterpenoid methyltransferase (TMT) for the synthesis of botryococcenes, C30-C37 triterpenoids typical of *B. braunii* Race B, were identified by sequence similarity, and their identities have been validated by expression in yeast [63]. Similarly, a recent analysis of *B. braunii* Race L identified a gene encoding an SSL involved in the synthesis of the C40 tetraterpenoid lycopadiene [64], and again the activity of the encoded protein was verified in yeast. These resources offer the means to reconstitute a novel microalgal pathway in a heterologous host, and indeed up to 0.5 mg g^{-1} fresh weight of botryococcenes were produced in tobacco plants into which SSL and TMT genes were introduced, although there were adverse effects on plant growth and morphology [65].

Biomineralizing coccolithophores and diatoms have been widely explored for nanotechnology purposes such as drug delivery, nano-sensors, solar technology, microfluidics, catalyst production and biosensing [66-69]. The structural and physical properties of the biomineralized cell walls as in the frustule of diatoms underpin these applications. Purified frustules are used as filter material and their replicas as biosensors. Genetically engineered frustules of *T. pseudonana* that displayed an immunoglobulin G (IgG)-binding domain on their surface had antibodies attached to selectively target and kill cancer cells [69]. Treatment with the drug-loaded frustules led to tumor growth regression in mice. Despite the variety of uses offered by algal cell walls, the main hindrance in being able to exploit fully their structural and physical properties lies in a lack of knowledge about the genes and proteins required for their formation. Furthermore, the functions of those genes and proteins that have been identified thus far are primarily derived from biochemical *in vitro* studies with recombinant proteins. Thus, *in vivo* studies utilizing reverse genetic approaches are needed to reveal their biological functions and therefore to unravel molecular processes that are responsible for the formation of the morphologically complex algal cell walls. The recent establishment of CRISPR/Cas9 to edit the

genome of the model diatom *T. pseudonana* [12] is the first step towards elucidating the biogenesis and structural and physical properties of the diatom frustule.

Concurrently, the ability to introduce transgenes into microalgae means that it is possible to consider the development of microalgal platforms for industrial production, not just of endogenous molecules but also of a range of non-native compounds, ranging from therapeutics (e.g., plant natural products, vaccines) to platform chemicals used for plastics [70, 71]. As microbes, microalgae can be cultivated in enclosed photobioreactors rather than in open ponds that are susceptible to dynamic environmental conditions, introduction of contaminating eukaryotes and prokaryotes, and predation. Increasingly sophisticated molecular tools are being developed particularly for *C. reinhardtii*, *P. tricornutum*, *T. pseudonana*, and *Nannochloropsis* species. These include a wide range of vectors, promoters, and targeting sequences for transgene expression [72, 73], as well as ways to edit their genomes using TALEN [74] and CRISPR-Cas9 [9-12]. Increasingly, genome sequence and transcriptomic data are being mined to identify new regulatory sequences, and this approach also provides information that can be used to manipulate other algae [75-77]. At the same time, synthetic biology approaches (Fig. 2) that exploit engineering design principles are increasingly being applied to algae [78]. The application of the Design-Build-Test cycle coupled with high-throughput methods and automation will speed up identification of parts, verification of gene function, and analysis of mutant phenotypes, to make manipulation much easier. Nevertheless, significant knowledge gaps need to be filled between omics output and assigning gene functions and building metabolic and regulatory networks that will ultimately lead to a systems-level understanding of algal biology.

Concluding remarks

Algae are extraordinary organisms that exhibit wide diversity in morphology, physiology, gene content, and sexual systems. They have independently 'discovered' multicellularity (both simple and complex) on several occasions and offer a wide range of biotechnological opportunities to produce high value commercial products. From sustaining many ecosystems through primary production and providing an array of human foods, algae are increasingly being targeted for omics approaches to elucidate their diverse properties. Apart from commercial or academic concerns, the impacts of a warming climate on algal health and the role of these taxa as biomarkers of environmental change are also of paramount importance. The coming years will

prove pivotal for algal biologists and the broader public alike as the full array of transformative scientific methods is brought to bear upon these organisms.

Acknowledgements

This manuscript is an outcome of a symposium hosted in June 2016 by The Royal Society entitled, "Into the genome: advances in the world of algal genomics," in Chicheley Hall, Buckinghamshire, United Kingdom. The symposium organizers, J.B. and D.B. are grateful to the Royal Society for supporting this event. The University of Dundee is a registered Scottish charity, No. 015096.

References

1. de Vargas, C. et al. (2015) Eukaryotic plankton diversity in the sunlit ocean. Science 348, 1261605
2. Barsanti, L. and Gualtieri, P. (2014) *Algae: Anatomy, Biochemistry, and Biotechnology*. CRC Press
3. Courties, C. et al. (1994) Smallest eukaryotic organism. Nature 370, 255-255
4. Steneck, R.S. et al. (2002) Kelp forest ecosystems: biodiversity, stability, resilience and future. Environ. Conserv. 29, 436-459
5. Jacobs, W.P. (1994) *Caulerpa*. Sci. Am. 271, 100-105
6. Armstrong, E. et al. (2000) The first record of *Nitzschia alba* from UK coastal waters with notes on its growth potential. J. Mar. Biol. Assoc. U. K. 80, 355-356
7. Horiguchi, T. (2015) Diversity and phylogeny of marine parasitic dinoflagellates. In *Marine Protists: Diversity and Dynamics* (Ohtsuka, S. et al., eds) pp. 397-419, Springer Japan
8. Stoecker, D.K. et al. (2017) Mixotrophy in the marine plankton. Annu. Rev. Mar. Sci. 9, 311-335
9. Wang, Q. et al. (2016) Genome editing of model oleaginous microalgae *Nannochloropsis* spp. by CRISPR/Cas9. Plant J. 88, 1071-1081
10. Shin, S.E. et al. (2016) CRISPR/Cas9-induced knockout and knock-in mutations in *Chlamydomonas reinhardtii*. Sci. Rep. 6, 27810
11. Nymark, M. et al. (2016) A CRISPR/Cas9 system adapted for gene editing in marine algae. Sci. Rep. 6, 24951
12. Hopes, A. et al. (2016) Editing of the urease gene by CRISPR-Cas in the diatom *Thalassiosira pseudonana*. Plant Methods 12, 49
13. Martin, W. and Herrmann, R.G. (1998) Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol. 118, 9-17
14. Deusch, O. et al. (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. Mol. Biol. Evol. 25, 748-761
15. Martin, W. et al. (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc. Natl. Acad. Sci. U. S. A. 99, 12246-12251

16. Karpowicz, S.J. et al. (2011) The GreenCut2 resource, a phylogenomically derived inventory of proteins specific to the plant lineage. J. Biol. Chem. 286, 21427-21439
17. Wittkopp, T.M. et al. (2016) The GreenCut - functions and relationships of proteins conserved in green lineage organisms. In *Chloroplasts: Current Research and Future Trends* (Kirchhoff, H., ed) pp. 241-278, Caister Academic Press
18. Méheust, R. et al. (2016) Protein networks identify novel symbiogenetic genes resulting from plastid endosymbiosis. Proc. Natl. Acad. Sci. U. S. A. 113, 3579-3584
19. Merchant, S.S. et al. (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science 318, 245-251
20. Kim, J. et al. (2016) Flux balance analysis of primary metabolism in the diatom *Phaeodactylum tricornutum*. Plant J. 85, 161-176
21. Stöckel, J. et al. (2006) The evolutionarily conserved tetratrico peptide repeat protein pale yellow green7 is required for photosystem I accumulation in *Arabidopsis* and copurifies with the complex. Plant Physiol. 141, 870-878
22. Wilde, A. et al. (2001) Characterization of the cyanobacterial ycf37: mutation decreases the photosystem I content. Biochem. J. 357, 211-216
23. Reusch, T.B. and Boyd, P.W. (2013) Experimental evolution meets marine phytoplankton. Evolution 67, 1849-1859
24. Schlüter, L. et al. (2016) Long-term dynamics of adaptive evolution in a globally important phytoplankton species to ocean acidification. Sci. Adv. 2, e1501660
25. Hutchins, D.A. et al. (2015) Irreversibly increased nitrogen fixation in *Trichodesmium* experimentally adapted to elevated carbon dioxide. Nat. Commun. 6, 8155
26. Mock, T. et al. (2016) Bridging the gap between omics and earth system science to better understand how environmental change impacts marine microbes. Glob Chang Biol 22, 61-75
27. Padfield, D. et al. (2015) Rapid evolution of metabolic traits explains thermal adaptation in phytoplankton. Ecol. Lett. 19, 133-142
28. Schlüter, L. et al. (2014) Adaptation of a globally important coccolithophore to ocean warming and acidification. Nat. Clim. Change 4, 1024-1030
29. Bermúdez, R. et al. (2015) Long-term conditioning to elevated pCO_2 and warming influences the fatty and amino acid composition of the diatom *Cylindrotheca fusiformis*. PLoS ONE 10, e0123945
30. Schaum, C.E. and Collins, S. (2014) Plasticity predicts evolution in a marine alga. Proc. R. Soc. B 281, 20141486
31. Schaum, C.E. et al. (2016) Environmental stability affects phenotypic evolution in a globally distributed marine picoplankton. ISME J. 10, 75-84
32. Doblin, M.A. and van Sebille, E. (2016) Drift in ocean currents impacts intergenerational microbial exposure to temperature. Proc. Natl. Acad. Sci. U. S. A. 113, 5700-5705
33. Lohbeck, K.T. et al. (2014) Gene expression changes in the coccolithophore *Emiliania huxleyi* after 500 generations of selection to ocean acidification. Proc. R. Soc. B 281, 20140003
34. Perrineau, M.M. et al. (2014) Using natural selection to explore the adaptive potential of *Chlamydomonas reinhardtii*. PLoS ONE 9, e92533
35. Perrineau, M.M. et al. (2014) Evolution of salt tolerance in a laboratory reared population of *Chlamydomonas reinhardtii*. Environ. Microbiol. 16, 1755-1766

36. Mock, T. et al. (2017) Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. Nature, DOI:10.1038/nature20803 (ahead of print)
37. Speijer, D. et al. (2015) Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. Proc. Natl. Acad. Sci. U. S. A. 112, 8827-8834
38. Bachtrog, D. et al. (2014) Sex determination: why so many ways of doing it? PLoS Biol. 12, e1001899
39. Geng, S. et al. (2014) Evolution of sexes from an ancestral mating-type specification pathway. PLoS Biol. 12, e1001904
40. Hamaji, T. et al. (2016) Sequence of the *Gonium pectorale* mating locus reveals a complex and dynamic history of changes in volvocine algal mating haplotypes. G3 (Bethesda) 6, 1179-1189
41. Umen, J.G. (2011) Evolution of sex and mating loci: an expanded view from volvocine algae. Curr. Opin. Microbiol. 14, 634-641
42. Charlesworth, B. (1978) The population genetics of anisogamy. J. Theor. Biol. 73, 347-357
43. Luthringer, R. et al. (2014) Sexual dimorphism in the brown algae. Perspect. Phycol. 1, 11-25
44. Ahmed, S. et al. (2014) A haploid system of sex determination in the brown alga *Ectocarpus* sp. Curr. Biol. 24, 1945-1957
45. Jordan, C.Y. and Charlesworth, D. (2012) The potential for sexually antagonistic polymorphism in different genome regions. Evolution 66, 505-516
46. Lipinska, A.P. et al. (2015) Development of PCR-based markers to determine the sex of kelps. PLoS ONE 10, e0140535
47. Luthringer, R. et al. (2015) The pseudoautosomal regions of the U/V sex chromosomes of the brown alga *Ectocarpus* exhibit unusual features. Mol. Biol. Evol. 32, 2973-2985
48. Keeling, P.J. 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, e1001889
49. Kim, S.K. and Chojnacka, K., eds (2015) *Marine Algae Extracts : Processes, Products, and Applications, Volumes 1 and 2*. Wiley-VCH Verlag GmbH & Co.
50. Kazamia, E. and Smith, A.G. (2014) Assessing the environmental sustainability of biofuels. Trends Plant Sci. 19, 615-618
51. Rodolfi, L. et al. (2009) Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. Biotechnol. Bioeng. 102, 100-112
52. Georgianna, D.R. and Mayfield, S.P. (2012) Exploiting diversity and synthetic biology for the production of algal biofuels. Nature 488, 329-335
53. Scott, S.A. et al. (2010) Biodiesel from algae: challenges and prospects. Curr. Opin. Biotechnol. 21, 277-286
54. Chen, J.E. and Smith, A.G. (2012) A look at diacylglycerol acyltransferases (DGATs) in algae. J. Biotechnol. 162, 28-39
55. Miller, R. et al. (2010) Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. Plant Physiol. 154, 1737-1752
56. Niu, Y.F. et al. (2013) Improvement of neutral lipid and polyunsaturated fatty acid biosynthesis by overexpressing a type 2 diacylglycerol acyltransferase in marine diatom *Phaeodactylum tricornutum*. Mar. Drugs 11, 4558-4569

57. Levitan, O. et al. (2015) Remodeling of intermediate metabolism in the diatom *Phaeodactylum tricornutum* under nitrogen stress. Proc. Natl. Acad. Sci. U. S. A. 112, 412-417
58. Boyle, N.R. et al. (2012) Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in *Chlamydomonas*. J. Biol. Chem. 287, 15811-15825
59. Matthijs, M. et al. (2016) Profiling of the early nitrogen stress response in the diatom *Phaeodactylum tricornutum* reveals a novel family of RING-domain transcription factors. Plant Physiol. 170, 489-498
60. Trentacoste, E.M. et al. (2013) Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth. Proc. Natl. Acad. Sci. U. S. A. 110, 19748-19753
61. Adarme-Vega, T.C. et al. (2014) Towards sustainable sources for omega-3 fatty acids production. Curr. Opin. Biotechnol. 26, 14-18
62. Scranton, M.A. et al. (2015) *Chlamydomonas* as a model for biofuels and bio-products production. Plant J. 82, 523-531
63. Niehaus, T.D. et al. (2011) Identification of unique mechanisms for triterpene biosynthesis in *Botryococcus braunii*. Proc. Natl. Acad. Sci. U. S. A. 108, 12260-12265
64. Thapa, H.R. et al. (2016) A squalene synthase-like enzyme initiates production of tetraterpenoid hydrocarbons in *Botryococcus braunii* Race L. Nat. Commun. 7, 11198
65. Jiang, Y. and Yu, D. (2016) The WRKY57 transcription factor affects the expression of jasmonate ZIM-domain genes transcriptionally to compromise *Botrytis cinerea* resistance. Plant Physiol. 171, 2771-2782
66. Viji, S. et al. (2014) Diatom-based label-free optical biosensor for biomolecules. Appl. Biochem. Biotechnol. 174, 1166-1173
67. Ren, F. et al. (2013) Enhancing surface plasmon resonances of metallic nanoparticles by diatom biosilica. Opt. Express 21, 15308-15313
68. Li, A. et al. (2016) Towards uniformly oriented diatom frustule monolayers: experimental and theoretical analyses. Microsyst. Nanoeng. 2, 16064
69. Delalat, B. et al. (2015) Targeted drug delivery using genetically engineered diatom biosilica. Nat. Commun. 6, 8791
70. Gimpel, J.A. et al. (2015) In metabolic engineering of eukaryotic microalgae: potential and challenges come with great diversity. Front. Microbiol. 6, 1376
71. Guarnieri, M.T. and Pienkos, P.T. (2015) Algal omics: unlocking bioproduct diversity in algae cell factories. Photosynth. Res. 123, 255-263
72. Rasala, B.A. et al. (2014) Enhanced genetic tools for engineering multigene traits into green algae. PLoS ONE 9, e94028
73. Scaife, M.A. et al. (2015) Establishing *Chlamydomonas reinhardtii* as an industrial biotechnology host. Plant J. 82, 532-546
74. Daboussi, F. et al. (2014) Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. Nat. Commun. 5, 3831
75. Thiriet-Rupert, S. et al. (2016) Transcription factors in microalgae: genome-wide prediction and comparative analysis. BMC Genomics 17, 282
76. Vieler, A. et al. (2012) Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. PLoS Genet. 8, e1003064

77. Yao, L. et al. (2015) RNA-Seq transcriptomic analysis with Bag2D software identifies key pathways enhancing lipid yield in a high lipid-producing mutant of the non-model green alga *Dunaliella tertiolecta*. Biotechnol. Biofuels 8, 191
78. Scaife, M.A. and Smith, A.G. (2016) Towards developing algal synthetic biology. Biochem. Soc. Trans. 44, 716-722
79. Adl, S.M. et al. (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J. Eukaryot. Microbiol. 52, 399-451
80. Ball, S.G. et al. (2013) Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis? Plant Cell 25, 7-21
81. Price, D.C. et al. (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 335, 843-847
82. Yoon, H.S. et al. (2004) A molecular timeline for the origin of photosynthetic eukaryotes. Mol. Biol. Evol. 21, 809-818
83. Marin, B. et al. (2007) The ancestor of the *Paulinella* chromatophore obtained a carboxysomal operon by horizontal gene transfer from a *Nitrococcus*-like gamma-proteobacterium. BMC Evol. Biol. 7, 85
84. Nowack, E.C.M. et al. (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. Curr. Biol. 18, 410-418
85. Reyes-Prieto, A. et al. (2010) Differential gene retention in plastids of common recent origin. Mol. Biol. Evol. 27, 1530-1537
86. Delaye, L. et al. (2016) How really ancient is *Paulinella chromatophora*? PLoS Currents Tree of Life, March 15
87. Nowack, E.C.M. et al. (2016) Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. Proc. Natl. Acad. Sci. U. S. A. 113, 12214-12219
88. Bhattacharya, D. et al. (2012) Single cell genome analysis supports a link between phagotrophy and primary plastid endosymbiosis. Sci. Rep. 2, 356
89. Matt, G. and Umen, J. (2016) *Volvox*: A simple algal model for embryogenesis, morphogenesis and cellular differentiation. Dev. Biol. 419, 99-113
90. Cock, J.M. et al. (2010) The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. Nature 465, 617-621
91. Cock, M.J. and Collén, J. (2015) Independent emergence of complex multicellularity in the brown and red Algae. In *Evolutionary Transitions to Multicellular Life: Principles and mechanisms* (Ruiz-Trillo, I. and Nedelcu, M.A., eds) pp. 335-361, Springer Netherlands
92. Knoll, A.H. (2011) The multiple origins of complex multicellularity. Annu. Rev. Earth Planet. Sci. 39, 217-239
93. Niklas, K.J. and Newman, S.A. (2013) The origins of multicellular organisms. Evol. Dev. 15, 41-52
94. Ye, N. et al. (2015) *Saccharina* genomes provide novel insight into kelp biology. Nat. Commun. 6, 6986
95. Collén, J. et al. (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. Proc. Natl. Acad. Sci. U. S. A. 110, 5247-5252
96. Macaisne, N. et al. (2017) The *Ectocarpus IMMEDIATE UPRIGHT* gene encodes a member of a novel family of cysteine-rich proteins that have an unusual distribution across the eukaryotes. Development, DOI:10.1242/dev.141523 (ahead of print)

97. Dillehay, T.D. et al. (2008) Monte Verde: seaweed, food, medicine, and the peopling of South America. Science 320, 784-786
98. Tseng, C.K. and Chang, C.F. (1984) Chinese seaweeds in herbal medicine. Hydrobiologia 116, 152-154
99. FAO (2014) *The State of World Fisheries and Aquaculture*. Food and Agriculture Organization of the United Nations
100. Duarte, C.M. et al. (2007) Rapid domestication of marine species. Science 316, 382-383
101. Cottier-Cook, E.J. et al. (2016) *Safeguarding the future of the global seaweed aquaculture industry*. United Nations University (INWEH) and Scottish Association for Marine Science Policy Brief
102. Smetacek, V. and Zingone, A. (2013) Green and golden seaweed tides on the rise. Nature 504, 84-88
103. Varela-Álvarez, E. et al. (2012) Mediterranean species of *Caulerpa* are polyploid with smaller genomes in the invasive ones. PLoS ONE 7, e47728
104. Loureiro, R. et al. (2015) Seaweed cultivation: potential and challenges of crop domestication at an unprecedented pace. New Phytol. 206, 489-492
105. Botebol, H. et al. (2015) Central role for ferritin in the day/night regulation of iron homeostasis in marine phytoplankton. Proc. Natl. Acad. Sci. U. S. A. 112, 14652-14657
106. Kilian, O. et al. (2011) High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis* sp. Proc. Natl. Acad. Sci. U. S. A. 108, 21265-21269
107. Lozano, J.C. et al. (2014) Efficient gene targeting and removal of foreign DNA by homologous recombination in the picoeukaryote *Ostreococcus*. Plant J. 78, 1073-1083
108. Minoda, A. et al. (2004) Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, *Cyanidioschyzon merolae* 10D. Plant Cell Physiol. 45, 667-671

Figure legends

Figure 1. A) The major eukaryotic supergroups (kingdoms) showing the polyphyletic origins of algae in the tree of life. *Incertae sedis* (unknown affiliations) are shown with the dotted lines. **Archaeplastida** (see Glossary) are shown in violet text, whereas algae with (solely, or primarily) red algal-derived plastids are shown with the red text and those with green algal derived plastids in green text. The plastid in photosynthetic *Paulinella* species (indicated in blue) is an example of an independent cyanobacterial primary endosymbiosis. Image adapted from Kim et al. [20]. B) Moving clockwise from top left: a *Sargassum fusiforme* (stramenopile) farm in South Korea; the invasive green alga (Viridiplantae) *Caulerpa taxifolia*; *Saccharina japonica* (stramenopile); an underwater kelp 'forest' of *Undaria pinnatifida* (stramenopile); (top) dividing cells of *Paulinella microporus* KR01 (Rhizaria; image prepared by Jong Im Kim); (bottom) light microscope image showing the single sheet of cells that comprises the *Porphyra umbilicalis*

(Rhodophyta) blade; an underwater tropical forest of the stramenopile *Sargassum* (credit: R. Terada); (below) examples of harmful algal blooms and their biotic interactions. Shown are two unreported protistan parasites infecting the toxic diatom (stramenopile) *Pseudo-Nitszchia* sp. (scale bars = 10 μm; images prepared by Andrea Garvetto). Images of seaweed farms were provided by the National Institute of Fisheries Science in South Korea. C) Major sources of foreign genes in the nuclear genome of the Archaeplastida ancestor that underwent primary plastid endosymbiosis. EGT refers to genes derived via intracellular gene transfer from the cyanobacterial endosymbiont. HGT refers to a variety of genes derived from multiple non-plastid sources, including symbionts, prey, viruses (represented as the blue and magenta ovals), or other sources of DNA that were present in the cell or entered it from the environment.

Figure 2. A synthetic biology approach to engineering microalgae. Using defined parts (e.g., promoters, terminators) with predictable behavior, and a standardized way to combine them, many permutations can be tested for optimal expression of the gene of interest through the Design-Build-Test-Learn cycle [78].

BOX 1. ***Algae in the tree of life and the evolution of multicellularity***

Primary endosymbiosis

The primary plastid characteristic of the Archaeplastida (land plants, glaucophyte, green and red algae) can be traced back to a single primary endosymbiosis of a cyanobacterium that occurred in the Archaeplastida common ancestor [79-81] about 1.6 billion years ago [82]. Thereafter, the plastid in red and green algae spread to other taxa through secondary or additional rounds of eukaryote-eukaryote endosymbioses (Fig. 1A).

Paulinella **primary endosymbiosis**

There is only one other known example of a primary plastid endosymbiosis: in the photosynthetic *Paulinella* species [83-85]. This plastid (known as a chromatophore) originated 90-140 million years ago [86] and is derived from an α-cyanobacterium with Form IA Rubisco, unlike the Form IB Rubisco derived by EGT from the β-cyanobacterial ancestor of the archaeplastidial plastid [13]. The chromatophore genome has undergone genome reduction [87] but is ca. 1 Mbp in size, which is about 5-10 times larger than other plastid genomes. Genomic

and transcriptomic data demonstrate the existence of extensive horizontal gene transfer (HGT) from non-endosymbiont bacterial sources (Fig. 1C). Many of these nuclear HGTs complement functions lost in the chromatophore due to genome reduction [87]. This result suggests that phagotrophy, which typified the heterotrophic ancestor of photosynthetic *Paulinella* species [88] was crucial during the early phases of primary endosymbiosis to allow gene acquisition.

Evolution of multicellularity

Algae play an important role in improving our understanding of multicellularity. In particular, the volvocine algae represent one of the best model systems for understanding the key initial transition from unicellularity to multicellularity [89]. These organisms have provided several compelling examples of how genes that were already present in the unicellular ancestor were co-opted for functions related to multicellularity.

Transitions from unicellularity to multicellularity have occurred many times but only a limited number of groups have evolved complex multicellularity. Animals, plants, fungi and brown and red seaweeds, which all have organized macroscopic body plans consisting of multiple cell types are generally considered to have made this transition [90-93]. Complex multicellularity evolved independently in each of these groups providing an ideal situation to apply comparative approaches to understand this important evolutionary process. Developmental processes are well understood in land plants, however the two macroalgal groups remain poorly studied at the molecular level. The availability of genome sequence information from both brown [90, 94] and red macroalgae [95] has allowed some comparative analyses to be done, but experimental investigation of the molecular basis of developmental processes in these algae is essential. The recent demonstration that key developmental genes can be identified in the brown algal model *Ectocarpus* using a forward genetic approach [96] represents a first step towards the emergence of experimental macroalgal developmental biology.

BOX 2. ***Algae in human affairs***

Human consumption of seaweed dates back at least 14,000 years in South America [97], with medicinal uses recorded in Chinese and Japanese literature from 1500 to 2000 years ago [98]. Algae are increasingly used in an array of manufactured foods and feed, as a source for hydrocolloids of commercial importance and in food processing, pharmaceutical, cosmetic,

biofuels and biomaterials industries. With a value of $5.6 billion USD in 2014, the seaweed market is currently the fastest growing aquaculture sector (8% year^{-1}) and represents 25% of global aquaculture production (27 million tonnes [Mt] in 2014) [99]. Five major groups of algae contribute to 97% of production: eucheumatoids (*Eucheuma* spp., *Kappaphycus*) accounted for about 11 Mt, the kelp *Saccharina japonica* for 7.6 Mt, the red alga *Gracilaria* spp. for 3.7 Mt, wakame *Undaria pinnatifida* for 2.4 Mt, and laver (*Porphyra* spp. and *Pyropia*) for 1.8 Mts. Major seaweed production areas include China (49%), Indonesia (37%), the Philippines (6%), and South Korea (4%), followed by Malaysia, North Korea, and Japan. However, cultivation of macro- and microalgae is developing quickly on all continents and consequently the number of farmed species is rapidly rising. This domestication follows a similar trend for animal marine species [100], and poses significant challenges to the sustainable exploitation and conservation of marine genetic resources [101]. Some microalgae form blooms that are toxic for other organisms, in particular fish and humans, and uncontrolled growth of seaweeds (e.g., *Ulva* spp.) as a response to eutrophication along coastlines worldwide is a major cause of reduced ecosystem functioning [102]. In addition, non-native seaweeds can cause dramatic ecological shifts in marine ecosystems, for example the green seaweed *Caulerpa* spp. in the Mediterranean Sea [103].

Future domestication of algae

Sustainable exploitation and domestication of algae will require greatly improved knowledge of the ecological and molecular factors controlling growth and reproduction [104]. For general applications, such as lipids for biofuel, or higher-value compounds for a wider range of nutraceuticals, therapeutics or even bulk chemicals, microalgal production needs to be reproducible at scale under ambient conditions [52]. Similarly, for seaweeds, the challenge is to balance growth and reproduction with culture loss through diseases and pests. Identification of ecotypic genetic variation and a comprehensive understanding of molecular mechanisms controlling such intraspecific variation (e.g., quantitative trait loci), and the introduction of agrigenomics tools, will be instrumental to facilitate successful algal domestication.

BOX 3. ***New tools for functional analysis in algae***

Genome-enabled studies have revealed the significance of specific transcription factors, small RNAs, epigenetics, and transposable elements in algal plasticity. Linking hypotheses generated from genomic analyses to specific activities/phenotypes is facilitating the development of a new range of tools.

CRISPR/Cas9: RNA-guided gene-editing tool

CRISPR/Cas9 gene editing is currently available for the algae *C. reinhardtii*, *Phaeodactylum tricornutum*, *Nannochloropsis* sp. and *Thalassiosira pseudonana* [9-12] and high frequency gene targeting by homologous recombination has been reported for the haploid genomes of *Cyanidioschizon merolae*, *Ostreococcus tauri*, and *Nannochloropsis* sp. [105-108]. These tools, because they can target genes or loci precisely, will enable insight into how (epi) genes impact phenotypic expression in algae. Furthermore, the role of gene redundancy (e.g., isoforms from different endosymbionts) can be dissected either by knocking out specific family members, or increasing dosage by adding various synthetic isoforms. This approach would shed light on how different endosymbionts and their gene repertoires contribute to the adaptation and evolution of extant algae in addition to the functionality associated with genome plasticity. Genome editing with CRISPR/Cas9 can target more than a single gene/locus at a time. Thus by simultaneously editing a large number of nuclear genes encoding plastid proteins the contributions of "green" vs. "red" genes for key metabolic processes such as carbon fixation, lipid and polysaccharide synthesis/storage, and macronutrient uptake/assimilation could be determined.

Trends

- Application of modern omic and genetic methods has significantly advanced our understanding of the origin, evolution, and metabolic potential of unicellular and multicellular algae as well as their diverse modes of sexual reproduction.

- The GreenCut proteins, a conserved gene set in the Viridiplantae are primarily plastid targeted and play key roles in the function and regulation of photosynthesis, including the maintenance of photosynthetic reaction complexes.

- Lab evolution experiments demonstrate strong adaptability of microalgae to environmental changes that are associated with climate change, although it is unclear if these results will hold in natural ecosystems.

- Development of algae as ‘cell factories’ promises to allow the production not only of endogenous molecules but also non-native compounds useful in therapeutics and in the production of plastics.

Outstanding Questions

- What role did EGT play in expanding the genetic inventory of algae and plants and how can we determine the roles of these foreign genes in extant cells?

- What do the GreenCut proteins teach us about how ancient Viridiplantae dealt with the rise in atmospheric oxygen precipitated by the great oxygenation event, the subsequent spread of eukaryotic photosynthetic lineages, and the stresses posed by hosting an oxygen-evolving organelle?

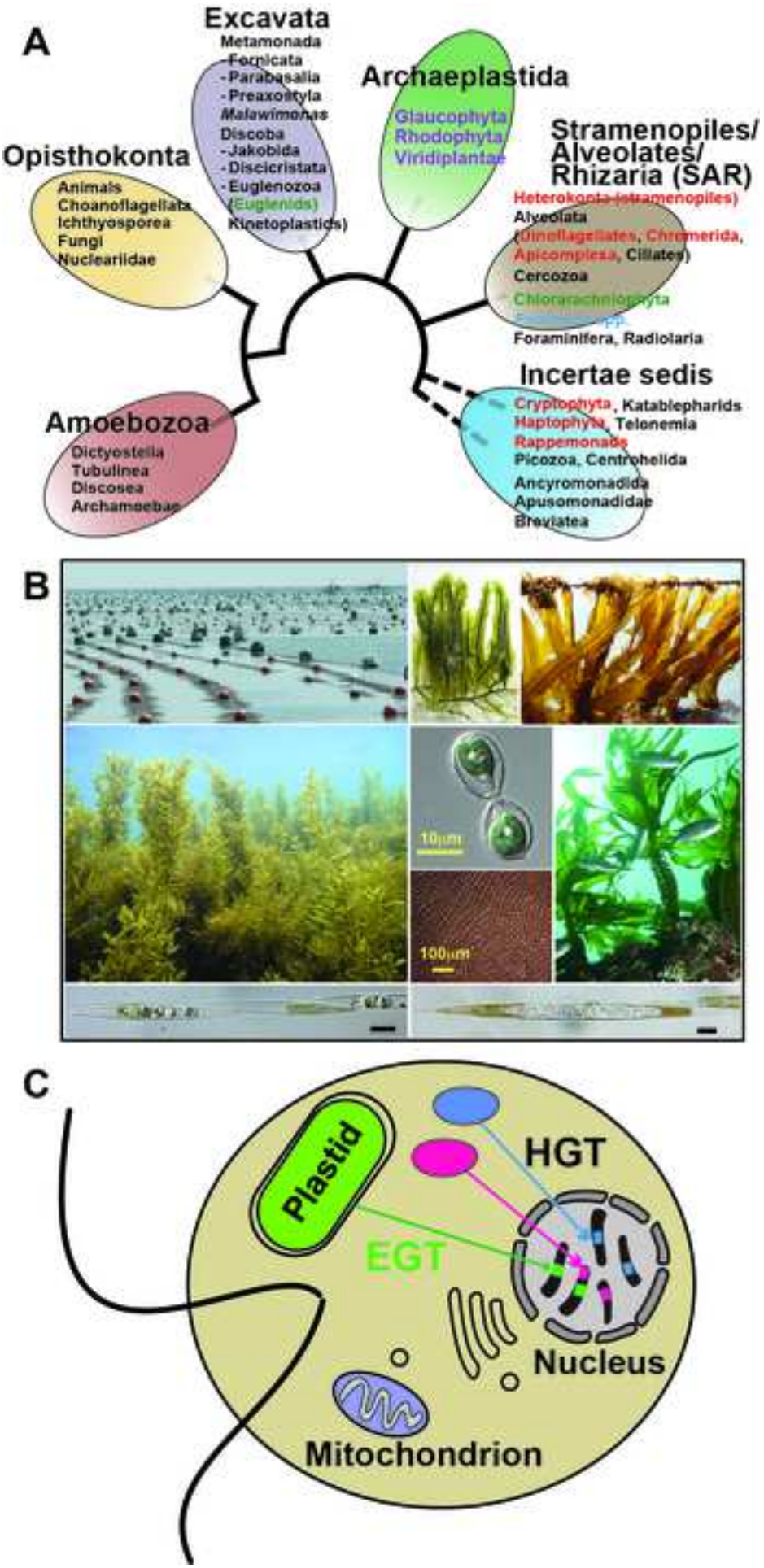
- Will the fluctuating environmental conditions forecasted for the coming century impact the health and distribution of aquatic microalgae, and which types of acclimatory and adaptive mechanisms do these taxa possess to deal with increasing temperatures and lower pH levels?

- What is the basis of sex determination in algae and seaweeds and do these mechanisms, such as the existence of sex-determining regions or sex chromosomes, follow the same patterns of evolution as described in well-studied, classical eukaryotic model systems?

- What are the unique and most promising aspects of, and the constraints on using algae as ‘cell factories’ and what types of advances in genetic manipulation techniques and genomic methods are needed to make these taxa more useful to the biotechnology industry?

Figure

Click here to download Figure Figure 1.tif



Figure

Click here to download Figure Figure 2.tif

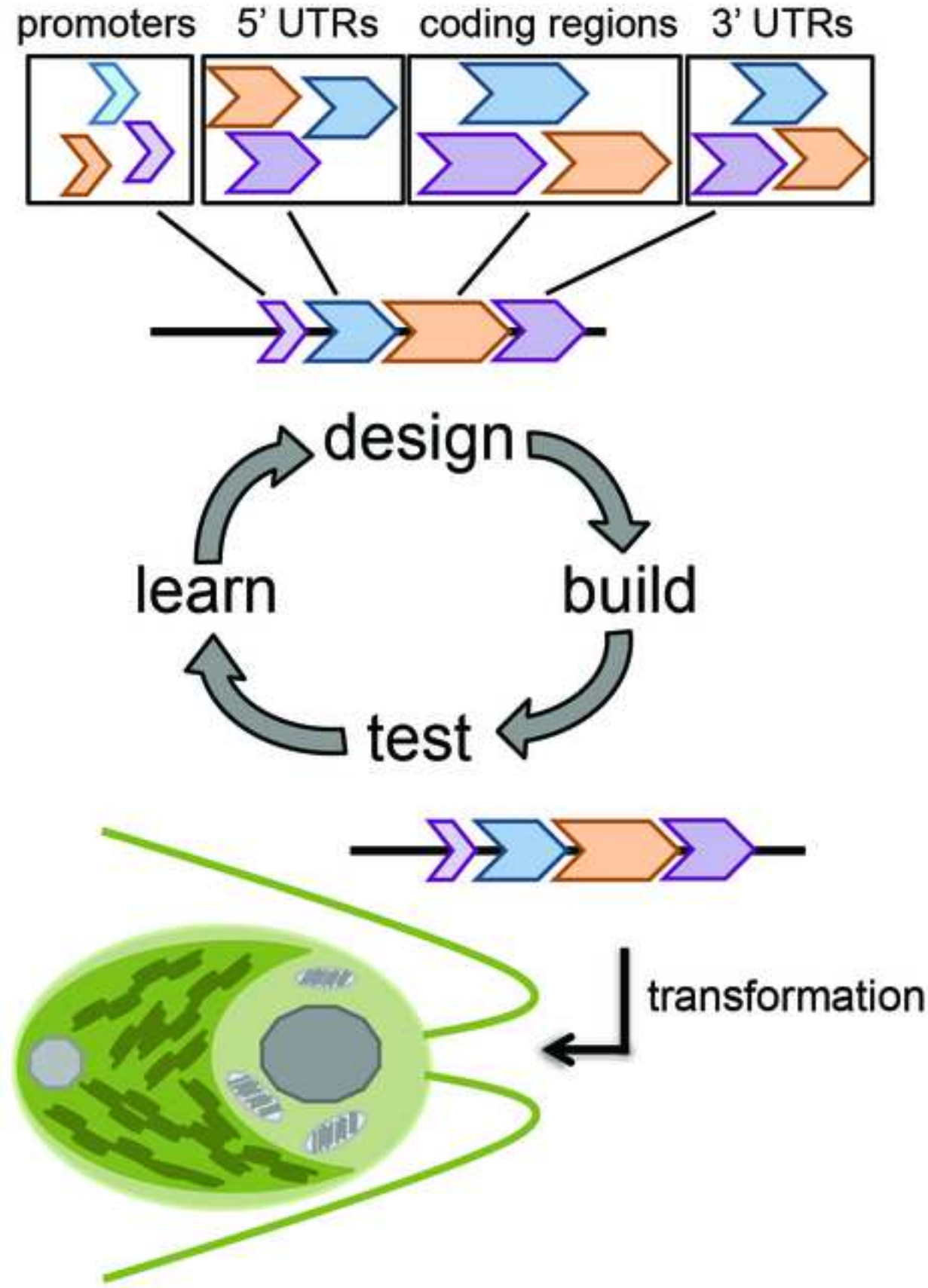
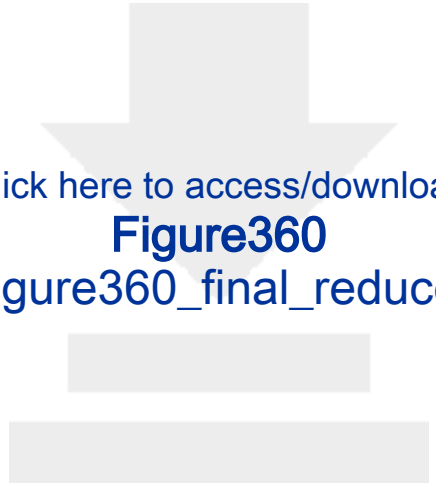


Figure360



Click here to access/download

Figure360

TIPS_figure360_final_reduced.mov