

# Algal genomics perspective: the pangenome concept beyond traditional molecular phylogeny and taxonomy

JunMo Lee\*

Department of Oceanography, Kyungpook National University, Daegu 41566, Republic of Korea

\*Correspondent: junmolee@knu.ac.kr

Algal genomics approaches provide a massive number of genome/transcriptome sequences and reveal the evolutionary history vis-à-vis primary and serial endosymbiosis events that contributed to the biodiversity of photosynthetic eukaryotes in the eukaryote tree of life. In particular, phylogenomic methods using several hundred or thousands of genes have provided new insights into algal taxonomy and systematics. Using this method, many novel insights into algal species diversity and systematics occurred, leading to taxonomic revisions. In addition, horizontal gene transfers (HGTs) of functional genes have been identified in algal genomes that played essential roles in environmental adaptation and genomic diversification. Finally, algal genomics data can be used to address the pangenome, including core genes shared among all isolates and partially shared strain-specific genes. However, some aspects of the pangenome concept (genome variability of intraspecies level) conflict with population genomics concepts, and the issue is closely related to defining species boundaries using genome variability. This review suggests a desirable future direction to merge algal pangenomics and population genomics beyond traditional molecular phylogeny and taxonomy.

Keywords: algal genomics, biodiversity, genome variability, molecular phylogeny, molecular taxonomy, pangenome

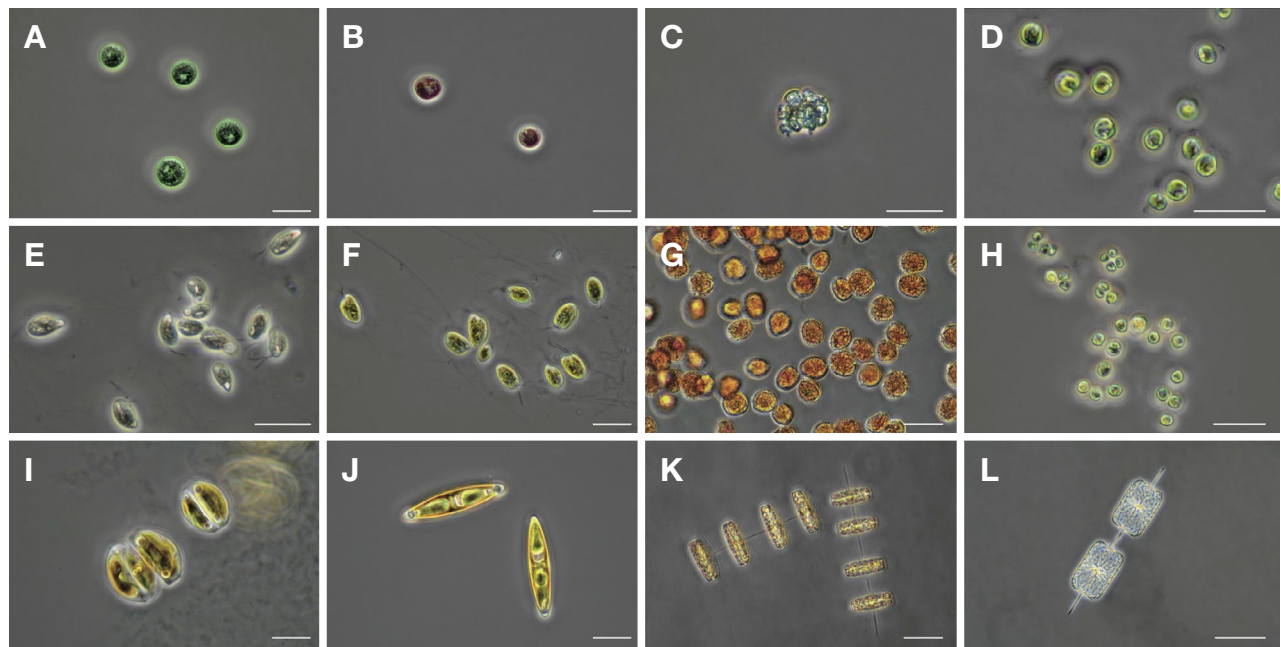
© 2021 National Institute of Biological Resources  
DOI:10.12651/JSR.2021.10.2.142

## INTRODUCTION

Algae are photosynthetic eukaryotes, including macrophytes and unicellular microalgae distributed in aquatic ecosystems (Simpson *et al.*, 2017). Diverse algal species inhabit a variety of environments, including common mesophilic environments and some extreme environments (e.g., high acidity, low temperature, and high salinity places; Rossoni *et al.*, 2019; Zhang *et al.*, 2020). Algal photosynthesis contributes globally to carbon fixation and primary production, which ecological roles (e.g., primary production in polar regions, symbiotic microalgae in corals, and photosynthesis under low light conditions) are extremely important in interactions with other eukaryotes (Gradinger, 2009; Brodie *et al.*, 2017; Mercado *et al.*, 2020). As photosynthetic organelles, plastids were derived from a primary endosymbiosis event, in which an ancient heterotrophic eukaryote engulfed the cyanobacterial symbiont and then integrated it as an intracellular organelle (Bhattacharya *et al.*, 2004; Timmis *et al.*, 2004; Keeling, 2010; Ponce-Toledo *et al.*, 2017). The primary endosymbiosis is believed to have occurred at least >1 billion years ago (BYA) up to 1.6 BYA from a single common ancestor,

which is supported by molecular timeline methods with fossil records (Butterfield, 2000; Yoon *et al.*, 2004; Parfrey *et al.*, 2011; Bengtson *et al.*, 2017; Sánchez-Baracaldo *et al.*, 2017; Gibson *et al.*, 2018; Tang *et al.*, 2020; Strasser *et al.*, 2021). The ancient endosymbiosis group diverged into the current three major photosynthetic lineages (i.e., primary endosymbiosis group) known as Archaeplastida including Rhodophyta (red algae), Viridiplantae (also known as Chloroplastida; green algae, and land plants), and Glaucophyta (glaucophyte algae) (Adl *et al.*, 2005; Rodríguez-Ezpeleta *et al.*, 2005; Chan *et al.*, 2011; Price *et al.*, 2012). Currently, non-photosynthetic *Rhododelphis* has been reported as a sister group of red algae within Archeplastida (Gawryluk *et al.*, 2019).

Through the primary endosymbiosis, a large proportion (>90%) of genes in the ancient cyanobacterial symbiont genome has been lost, but some of them were transferred into the eukaryotic host nuclear genome, which is referred to as endosymbiotic gene transfer (EGT) (Timmis *et al.*, 2004; de Vries *et al.*, 2016; Lee *et al.*, 2016a). Interestingly, three Archaeplastida lineages underwent differential rates of EGT and gene gain/loss processes. Red algae showed the most extensive gene inventory in



**Fig. 1.** Phase-contrast microscopy images of diverse algal taxa. A. *Rhodella maculata* CCMP736 (Rhodophyta). B. *Dixoniella grisea* CCMP1916 (Rhodophyta). C. *Emiliana huxleyi* (Haptophyta). D. *Diacronema lutheri* LIMS-PS-0073 (Haptophyta). E. *Proteomonas sulcata* (Cryptophyta). F. *Rhinomonas nottbecki* (Cryptophyta). G. *Coolia monotis* (Alveolata). H. *Sungminbooa australiensis* (Pelagophyceae; Stramenopiles). I. *Halamphora pseudohyalina* (Bacillariophyceae; Stramenopiles). J. *Navicula avium* (Bacillariophyceae; Stramenopiles). K. *Thalassiosira gravida* (= *T. rotula*; Bacillariophyceae; Stramenopiles). L. *Ditylum sol* (Bacillariophyceae; Stramenopiles). Multifocus light microscopy images were merged, and white balances were properly adjusted by Adobe Photoshop and Illustrator (scale bars: A–F, and H–J = 15  $\mu$ m; G, and K = 40  $\mu$ m; L = 100  $\mu$ m).

plastid genomes but the smallest number of EGT-derived nuclear-encoded genes (e.g., 75% in plastid genomes, 1% in nuclear genomes, and outright gene loss). Thus, red algal photosynthetic machineries are mainly encoded in the plastid genome rather than in the nuclear genome. In contrast, Viridiplantae shows the largest number of EGT-derived nuclear-encoded genes (e.g., 50% in plastid genomes, 25% in nuclear genomes, and outright gene loss) among Archaeplastida lineages (Lee *et al.*, 2016a). Although the common ancient EGTs of several plastid core functions are present in Archaeplastida species genomes (e.g., TIC/TOC complex, cytochrome, ferredoxin, and ATPase complexes), the different evolutionary trajectories construct the different types of photosynthetic machineries between green and red lineages, such as light-harvesting complexes (i.e., chlorophyll-*b* and phycobilisome, respectively), and their pigment compositions (Bryant and Canniffe, 2018; Giovagnetti and Ruban, 2018; Lee *et al.*, 2019). From the primary endosymbiosis lineages, plastids of red and green algal ancestors were transferred again to other eukaryotes, establishing both secondary and tertiary (or serial) endosymbiosis lineages. The plastids of stramenopiles (Stramenopila; e.g., diatoms and brown algae), alveolates (e.g., dinoflagellates and apicomplexans), haptophytes, and cryptophytes were derived from the red algal ancestor, whereas the chlorarachnio-

phytes and euglenids were originated from a green algal ancestor (Bhattacharya *et al.*, 2004; Keeling, 2010; Bhattacharya and Price, 2020; Sibbald and Archibald, 2020). Algal lineages show huge biodiversity (>20,000 species records) in aquatic environments with morphological and physiological diversities (Fig. 1; Guiry and Guiry, 2021). In terms of genetic diversity, independent functional gene divergences (e.g., phycobilisome linker proteins in red algae) and horizontal gene transfer events (e.g., ice-binding proteins in green alga *Chlamydomonas* sp. and mercuric reductases in extremophilic red algae), from broad taxonomic ranges to algae, contributed to increase genetic diversity in their genomes, and some of the genes led to the acquisition of innovative adaptive functions (Lee *et al.*, 2016b; 2019; Rossoni *et al.*, 2019; Zhang *et al.*, 2020; Van Etten and Bhattacharya, 2020). Furthermore, another independent primary endosymbiosis occurred in the rhizarian amoeba *Paulinella*, which contains alpha-cyanobacterium-derived photosynthetic organelle (called chromatophore), with a different evolutionary history in the genomes compared to other primary endosymbiosis lineages (Nowack *et al.*, 2008; Reyes-Prieto *et al.*, 2010; Lhee *et al.*, 2021). Algal genomics has contributed to the discovery of the complicated evolutionary history of algal lineages within the eukaryotic tree of life (Keeling, 2010; Bhattacharya and Price, 2020; Sibbald and Archibald,

2020).

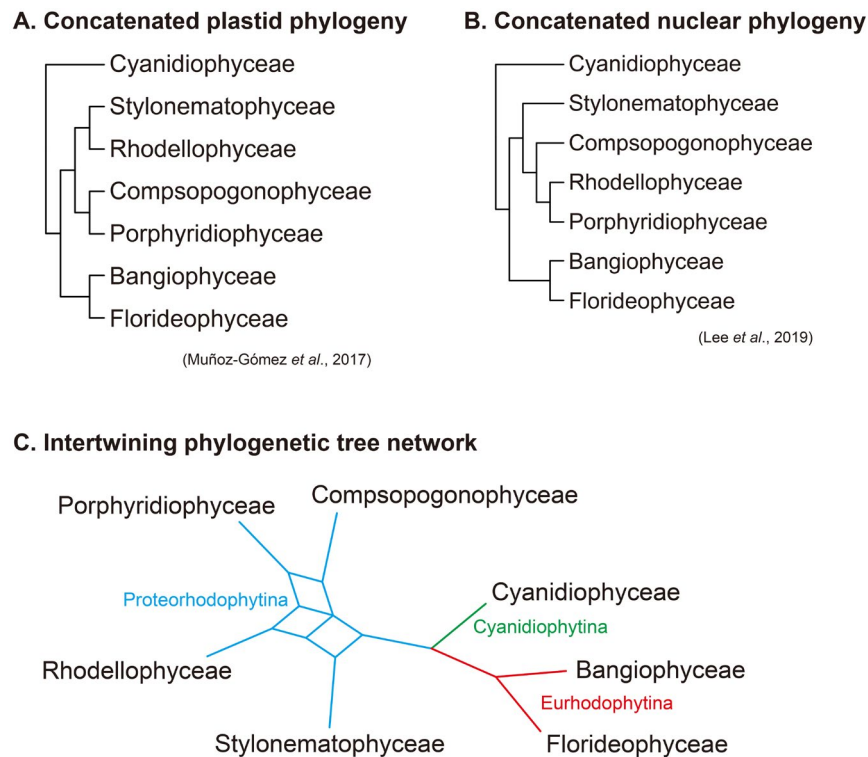
However, independently evolved gene families, recent rapid radiations, natural hybridization, and introgression can make incongruent evolutionary relationships within the lineage compared to previously studied species-trees in diverse taxa (Shaw, 2002; Fehrer *et al.*, 2007; Toews and Brelsford, 2012; Jarvis *et al.*, 2014; Kumar *et al.*, 2017; Lee *et al.*, 2018a; Russell *et al.*, 2021). Because genomics approaches enable the use of thousands of genes to construct multigene phylogenies, this method can help to recognize such complicated evolutionary histories (i.e., phylogenetic relationship) and discover what is the majority among the different phylogenetic or taxonomic scenarios (Keeling and Burki, 2019; Lee *et al.*, 2019; Bhattacharya and Price, 2020). Horizontal gene transfers (HGTs) can introduce diverse functions in algal genomes, and the HGT-derived algal pangenomes, including both core genome shared by all isolates and partially shared strain-specific genes of the species (Tettelin *et al.*, 2005), are related to environmental adaptations (Rossoni *et al.*, 2019; Zhang *et al.*, 2020). Due to the falling cost of genome sequencing, more algal genomes and transcriptomes at species/population-level will be available in the near future. These data will provide better information on genome variability between species or populations. Comprehensive understanding using population genomics and pangenome data would bring a new concept for the integrated algal species boundaries.

### ALGAL GENOMICS FOR MOLECULAR PHYLOGENY AND TAXONOMY

Algal taxonomy was initially established by morphological (by light and electron microscopic observations) and physiological features (e.g., the composition of pigments). Beyond the algal alpha taxonomy, diverse methodological approaches have been developed, such as characteristics-based cladistics and PCR-based molecular phylogeny (Gabrielson *et al.*, 1985; Freshwater *et al.*, 1994; Ragan *et al.*, 1994; Yoon *et al.*, 2002). In particular, the advances in molecular biology (i.e., determination of DNA sequences and molecular phylogenetic methods) enable the easy recognition of diverse cryptic species, whereas morphological characteristics cannot be distinguished; this has led to the discovery of several hidden species' diversity (Zucarello and West, 2003; Saunders and Lehmkuhl, 2005; Medlin *et al.*, 2007; Payo *et al.*, 2013; Díaz-Tapia *et al.*, 2018). Through the molecular phylogenetic approaches, algal systematics is being gradually re-established (Yoon *et al.*, 2006; Medlin *et al.*, 2007; Yoon *et al.*, 2010; Graf *et al.*, 2020). However, incongruent phylogenetic relationships were frequently identified at all taxonomic levels when analyzed by two or more independent gene align-

ments or different combinations of gene concatenations. Several technical issues (e.g., insufficient data, model selection, and taxon sampling bias) could make such phylogenetic incongruence (Theriot *et al.*, 2010; Sun *et al.*, 2015; Hu *et al.*, 2020). Still, in many cases, even more data and broad taxon samples with appropriate analytical methods cannot establish a clear consensus among the datasets because the conflicts are potentially caused by diverse evolutionary histories of genes (e.g., convergent evolution, rapid diversification, hybridization, and incomplete lineage sorting) (Wendel and Doyle, 1998; Medlin *et al.*, 2007).

In the end, genomics and transcriptomics approaches were introduced to molecular phylogeny fields to resolve such phylogenetic incongruences because these methods allow the generation of substantial gene datasets and help researchers to reach a clear consensus (Rokas *et al.*, 2003; Rudd, 2003). Recently, diverse and useful high-throughput sequencing technologies have been developed that can be applied in many algal studies (Song *et al.*, 2016; Oliveira *et al.*, 2018). Through the methodological improvements with a reasonable cost, diverse algal genomes and transcriptomes were generated. Especially, organelle genomes (i.e., mitochondria and plastids) were sequenced from diverse algal lineages, including broader taxon sampling. As a result, the phylogenomic analysis based on up to over a hundred organelle genes provided highly resolved phylogenetic relationships at order and family (or higher taxonomic) level (Costa *et al.*, 2016; Lee *et al.*, 2016a; 2018; Díaz-Tapia *et al.*, 2017; 2019; Kim *et al.*, 2017; Muñoz-Gómez *et al.*, 2017; Iha *et al.*, 2018). In addition, several approaches for marker gene selection (Janouškovec *et al.*, 2013; Gulbrandsen *et al.*, 2021) and binning methods (e.g., TIGER; Cummins and McInerney, 2011; Lee *et al.*, 2018a) have also been used to construct multigene phylogenies. Through these studies, the field of algal systematics became more rigorous. However, conflicts still occur between gene-trees and species-trees, and several hidden evolutionary histories were still discovered whether multigene phylogeny was constructed from several genes or from hundreds of genes. For example, a glaucophyte genus *Cyanophora* shows a strong monophyly in the concatenated multigene phylogenies using 134 plastid and 34 mitochondrial genes, but its interspecies relationships differ from these organelle genomes (Russell *et al.*, 2021). Coralline red algae also showed two different phylogenetic relationships with strong bootstrap supporting values at the ordinal level, constructed by concatenated multigene phylogenies of 195 plastid and 22 mitochondrial genes, respectively (Lee *et al.*, 2018a). In addition, individual phylogenies of the coralline red algal plastid genes were divided into three different phylogenetic divergence patterns with fully supported bootstrap values in all branches that are believed to be derived from



**Fig. 2.** The red algal phylogenomic approaches. A. Concatenated multigene phylogeny using 170 plastid genes (Muñoz-Gómez *et al.*, 2017). B. Concatenated multigene phylogeny using 4,777 nuclear genes (Lee *et al.*, 2019). C. Intertwining phylogenetic network tree of red algal plastid and nuclear multigene phylogenies.

incomplete lineage sorting (ILS) in early diverged coral-line orders (Lee *et al.*, 2018a). ILS generally refers to the conflict between a single gene phylogeny and population or species level phylogeny caused by the retention of genetic polymorphism, and this leads to serious difficulties for phylogenetic inference (Maddison *et al.*, 2006). The reticulate gene phylogenies could be derived not only from commonly shared genes (e.g., cases of ILS) but also from foreign genetic materials. HGTs, which indicate the introduction of novel gene sources from taxonomically distinct organisms (e.g., bacteria, and other algal lineages) to algal genomes, could create reticulate gene phylogenies (Van Etten and Bhattacharya, 2020). Therefore, to analyze a clear phylogenetic relationship, whether target genes are derived from a common ancestor or from HGTs, it has to be accurately recognized by individual gene phylogeny with homologous target genes.

As the generation of genome sequencing data in the algal research field is rapidly growing, phylogenomic approaches using diverse algal genomes and transcriptomes have been widely applied. The analysis results provide helpful information to recognize what is the consensus (i.e., species-tree or true evolutionary history) among the diverse phylogenetic relationships (Price *et al.*, 2012; Collén *et al.*, 2013; Keeling *et al.*, 2014; Brawley *et al.*,

2017; Lee *et al.*, 2018a; 2019; Rossoni *et al.*, 2019; Bhattacharya and Price, 2020; Zhang *et al.*, 2020). However, because some meaningful biological processes (e.g., ILS) could be easily ignored in the species-tree constructed using concatenated multigene data, several different phylogenetic approaches for constructing the species-tree have been suggested, such as the coalescence tree, network-based tree, or their combinations, that enable the visualization of several major (or all) phylogenetic variations (Huson and Bryant, 2006; Xi *et al.*, 2014; Edge and Coop, 2019; Mao *et al.*, 2020). For example, the concatenated multigene phylogeny using red algal 170 plastid genes showed a clear interclass relationship (Fig. 2A; Muñoz-Gómez *et al.*, 2017), but the concatenated multigene phylogeny of 4,777 nuclear genes showed a different interclass relationship (Fig. 2B; Lee *et al.*, 2019). Nonetheless, interclass relationships in both phylogenies were strongly supported. The phylogenetic conflict between plastid and nuclear genomes may be derived by complex evolutionary histories in subphylum Proteorhodophytina, identified by the intertwining phylogenetic tree network (Fig. 2C). The two different topologies are merged and visualized by the method of intertwining phylogenetic trees and networks (Schliep *et al.*, 2017) with the phangorn package in R (<https://github.com/KlausVigo/phangorn>).

Modern phylogenomic approaches allow us to identify unified (e.g., species-tree), complex (e.g., network), and hidden (e.g., conflict with each strong support) phylogenetic relationships. However, it is still controversial to answer what is the best method of phylogenetic analysis. Therefore, multiple or sequential methodological approaches of phylogenetic analysis are required to accurately interpret algal phylogeny and evolution, although the concatenation method is still helpful for constructing a consensus species-tree (Pirie, 2015).

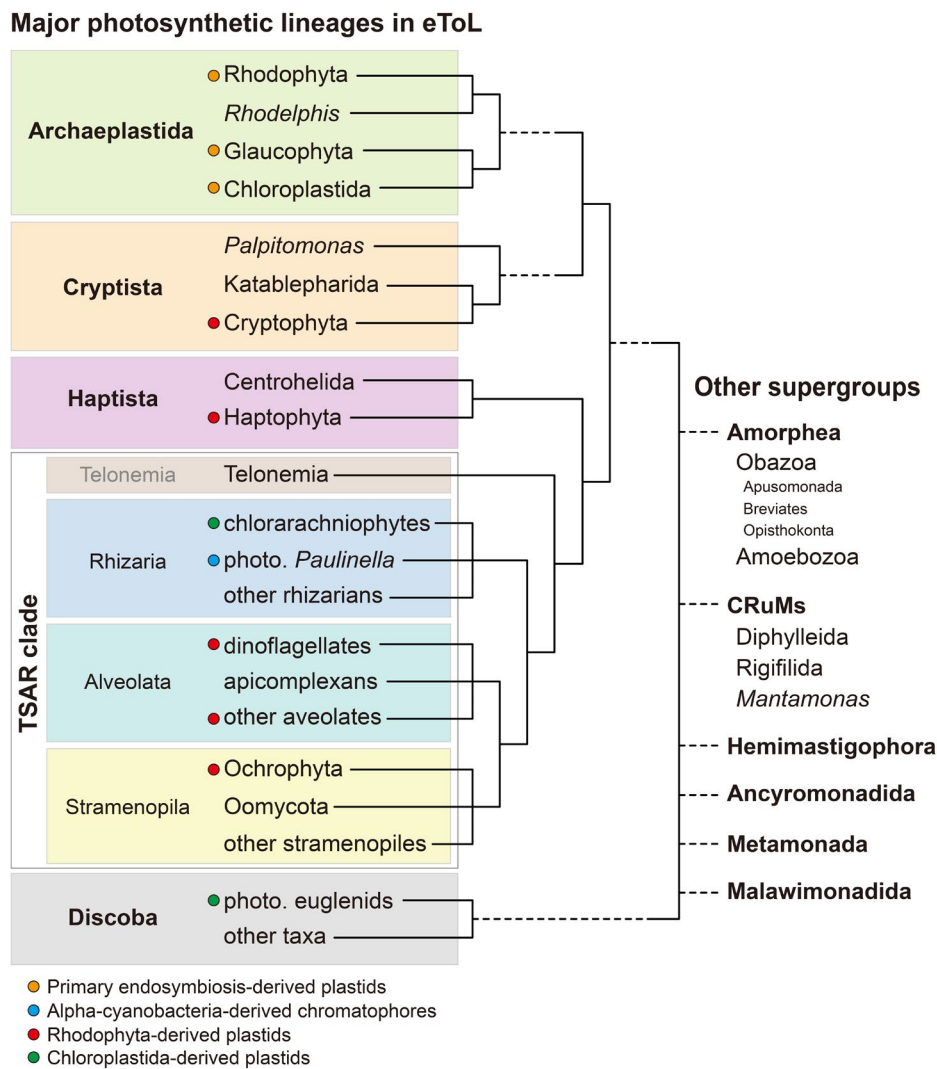
### ALGAL GENOMICS FOR BIODIVERSITY WITHIN EUKARYOTIC TREE OF LIFE

The primary endosymbiosis event gave rise to a large number of algal species that was recently reported as 7,314 red algae, 11,667 green lineages (i.e., non-land plants), and 25 glaucophyte algae species (March 2021; Guiry and Guiry, 2021). After the establishment of the primary endosymbiosis group, both the red and green algal plastids were also transferred to other eukaryotes through secondary and tertiary endosymbiosis, and then diverse photosynthetic eukaryotes have emerged, such as stramenopiles (e.g., diatoms and brown algae), alveolates (e.g., dinoflagellates and apicomplexans), haptophytes, cryptophytes, chlorarachniophytes, and euglenids (Keeling *et al.*, 2014; Keeling and Burki, 2019; Bhattacharya and Price, 2020; Sibbald and Archibald, 2020). These “secondary/tertiary” endosymbiosis lineages show considerable biodiversity (>20,000 species records) compared to algal descendants (excluding land plants) of primary endosymbiosis lineages (Guiry and Guiry, 2021). In particular, diatoms (Fig. 1), which are the most widespread marine phytoplankton lineage belonging to stramenopiles, show the most extensive biodiversity (>16,000 species) among photosynthetic eukaryotes, followed by dinoflagellates (>3,000 species), brown algae (>2,000 species), and euglenozoans (>1,500 species) (Guiry and Guiry, 2021). In addition, algal lineages constitute a large portion within the eukaryote Tree of Life (eToL). Currently, the eToL has been profoundly rearranged through phylogenomic approaches and the discoveries of several novel eukaryotic lineages, resulting in the suggestion of the supergroup model of the eToL (Fig. 3; Burki *et al.*, 2019; Sibbald and Archibald, 2020). The supergroup model includes Archaeplastida (Rhodophyta, *Rhodolphis*, Viridiplantae/Chloroplastida, and Glaucophyta), TSAR (Telonemia, Stramenopila, Alveolata, and Rhizaria), Haptista (Haptophyta, Centrohelida, and Rappemonads), Cryptista (Cryptophyta, Katablepharida, and *Palpitomonas*), Amorphea (Amoebozoa, and Obazoa including Opisthokonta, Breviataes, and Apusomonadida), CRuMs (Diphyllleida, Rigifilida, and *Mantamonas*), Discoba, Metamonada, Hemimastigophora, and several orphan

taxa (Malawimonadida, Ancyromonadida, Ancoracysta, and Picozoa) (Burki *et al.*, 2019). In the supergroup model, diverse photosynthetic algal lineages are mainly occupied as follows: Archaeplastida (Rhodophyta, Viridiplantae/Chloroplastida, and Glaucophyta), TSAR (chlorarachniophytes and photosynthetic *Paulinella* in Rhizaria, Alveolata, and Stramenopila), Haptista (haptophytes), Cryptista (cryptophytes), and Discoba (photosynthetic euglenids) (Fig. 3; Burki *et al.*, 2019; Sibbald and Archibald, 2020). Studies on algal genomics (i.e., algal genome/transcriptome data and applications of phylogenomics) have strongly contributed to constructing the eToL, and the result enables recognizing huge-scaled algal biodiversity within the eToL (Fig. 3). However, there are still controversies in the construction of true evolutionary histories. For example, Cryptista, including a red algal plastid-derived lineage (cryptophytes), usually shows a monophyletic relationship with Archaeplastida whether the branch of Cryptista is a sister clade of Archaeplastida or monophyly with one of Archaeplastida lineages. However, Haptista, including another red algal plastid-derived lineage (haptophytes), usually shows a monophyletic relationship with the TSAR clade. Still, it is controversial in the eToL because the Haptista clade could also show a monophyletic relationship with the Archaeplastida + Cryptista clade, depending on phylogenomic approaches (Burki *et al.*, 2012; Yabuki *et al.*, 2014; Strasser *et al.*, 2019; Bhattacharya and Price, 2020). In contrast with the phylogenetic controversy, the *rpl36* gene in the plastid genomes of photosynthetic haptophytes and cryptophytes is commonly replaced as bacterial-derived *rpl36*. This condition supports that plastids in haptophytes and cryptophytes are regarded as sister groups (Rice and Palmer, 2006). The plastid division machinery, *minD* and *minE* genes, encode in plastid genomes of haptophytes and cryptophytes but not in the plastid genomes of red algae, stramenopiles, and alveolates (de Vries and Gould, 2018). Although these commonly shared features are present in the plastid genomes of haptophytes and cryptophytes, their paraphyletic relationship may be derived from the chimeric origins of ochrophytes (photosynthetic stramenopiles) and haptophytes (Dorrell *et al.*, 2017). The useful information (e.g., phylogeny and genome variation) of algal genome sequences will provide new insights into algal biodiversity and its evolutionary trajectories within the eToL (Fig. 3).

### THE PANGENOME CONCEPT AND ALGAL TAXONOMY

Currently, the prokaryote-derived HGTs affected diverse functions in algal genomes (e.g., polysaccharide biosynthesis in brown algae and diatoms), leading to the

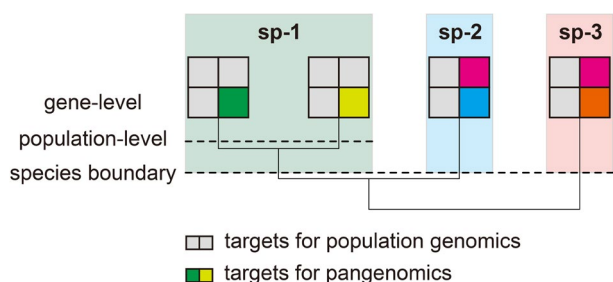


**Fig. 3.** Major photosynthetic algal lineages in the eukaryote Tree of Life (eToL). The eToL is reconstructed based on previous studies (Burki *et al.*, 2019; Keeling and Burki, 2019; Strassert *et al.*, 2019; Bhattacharya and Price, 2020; Sibbald and Archibald, 2020).

pangenome concept in algal genomics (Fan *et al.*, 2020). The pangenome refers to the set of all genes and genomic contents for all the strains of a particular species (typically bacteria), including a core genome shared by all isolates, plus partially shared strain-specific genes derived by lateral gene transfers (Tettelin *et al.*, 2005). Algal genomics studies have revealed that algal genomes also contain strain-specific variations in genomes and metabolic repertoires (Read *et al.*, 2013; Fan *et al.*, 2020). Indeed, the prokaryotic HGT-derived functional genes in algal genomes play ecologically and evolutionarily important roles, such as environmental adaptations (e.g., heavy metal detoxification, xenobiotic detoxification, and reactive oxygen species scavenging) and genome diversification (Rossoni *et al.*, 2019; Fan *et al.*, 2020). Through the HGTs, whether this is gene-scaled or genome-scaled (i.e., endosymbiosis), the algal species/genome complexity and diversity in the

eToL are consistently increasing up to date.

Algal pangenomes (i.e., genome variability of intraspecies level) may be derived by diverse evolutionary processes like HGTs (Rossoni *et al.*, 2019; Fan *et al.*, 2020), sexual genetic recombination (Umen and Coelho, 2019), the activity of transposable elements within the genome (Bennetzen and Wang, 2014; Lee *et al.*, 2018b), and external mutagens (e.g., UV light; Godfroy *et al.*, 2015). The algal genome variability, including genic region variations, can be easily detected by the resequencing method, which is generally used for population genomics (Read *et al.*, 2013; Graf *et al.*, 2021). Then, how can we distinguish between genome variability of population-level and pangenomes? How can we define species boundaries and novel species based on algal genome sequences? These questions would not simply be issues in studies for lineage-level or deep phylogeny analysis, but there are no



**Fig. 4.** The pangenome concept based on a comparison of gene inventory. Colored squares indicate commonly shared or newly acquired genes between species or populations.

clear answers yet in species/population-level studies in the algal genomics field because there is little study for algal pangenome (e.g., *Emiliana huxleyi*; Read *et al.*, 2013; Fan *et al.*, 2020; Golicz *et al.*, 2020; Richard *et al.*, 2020). The concept of population genomics (genetics) could conflict with the pangenome concept because the HGT-derived gene introductions are hard to share with allelic changes in population genomics (Shapiro, 2017). In other words, the pangenome concept generally focuses on the presence/absence of genes in the genomes of intra-species, but in the population genomics concept, the commonly shared genes and their mutations are important in intraspecies (Fig. 4). Therefore, there is ambiguity in demarcating the boundary of gene evolution between population genomics and pangenomics. At present, the issue of including the species concept is actively discussed in the bacterial genomics field, but the conclusion still remains challenging (Shapiro, 2017; Bobay, 2020; Azarian, *et al.*, 2020), although many strain-specific bacterial genomes were reported on the NCBI genome database (<https://www.ncbi.nlm.nih.gov/genome/>). Besides, genomics approaches in diverse taxa tried to define species boundaries: bacteria (Jain *et al.*, 2018; Matteoli *et al.*, 2020; Willis and Woodhouse, 2020), fungi (Matute and Sepúlveda, 2019; Ono *et al.*, 2020), and animals (MacGuigan *et al.*, 2017; Barth *et al.*, 2020). To date, however, many novel species (or variety/forma) have been described based on morphological characteristics and genetic variations in current algal research. Because algal taxonomy and systematics were initially established according to morphological characteristics and molecular markers, it is difficult to accept a bonafide species boundary based solely on algal genome variability (i.e., pangenomics). To establish the modern algal taxonomy, including the pangenome concept, firstly, many strain (or population)-specific algal genomes are required because these will provide a consensus (or new) concept between pangenomics and population genomics (Read *et al.*, 2013; Graf *et al.*, 2021). Secondly, studies on HGT mechanisms in eukaryotic genomes are required (e.g., plasmid-mediated HGTs; Lee *et al.*,

2016b) because it may cause algal pangenomes and diverse innovative evolutions. Lastly, single-cell genomics studies are required because these approaches could provide genome information at the individual-based level, even if for unculturable natural populations (Yoon *et al.*, 2011; Seeleuthner *et al.*, 2018). Multilateral analytic approaches will bring a new era of algal genomics and taxonomy, although more algal genome sequences are still required to clarify algal taxonomy, systematics, and evolutionary history.

## ACKNOWLEDGEMENTS

The author thanks for technical supports of light microscope works to Hyun Shik Cho (Marine Ecological Genomics Lab., Department of oceanography, Kyungpook National University, Daegu).

## REFERENCES

- Adl, S.M., A.G. Simpson, M.A. Farmer, R. Andersen, O.R. Anderson, J.R. Barta, S.S. Bowser, G. Brugerolle, R.A. Fensome, S. Fredericq, T.Y. James, S. Karpov, P. Kugrens, J. Krug, C.E. Lane, L.A. Lewis, J. Lodge, D.H. Lynn, D.G. Mann, R.M. McCourt, L. Mendoza, O. Moestrup, S.E. Mozley-Standridge, T.A. Nerad, C.A. Shearer, A.V. Smirnov, F.W. Spiegel and M.F. Taylor. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology* 52(5):399-451.
- Azarian, T., I.T. Huang and W.P. Hanage. 2020. Structure and dynamics of bacterial populations: pangenome ecology. In: H. Tettelin and D. Medini (eds.), *The pangenome: diversity, dynamics and evolution of genomes*, Springer, Cham, Switzerland. pp. 115-128.
- Barth, J.M.I., C. Gubili, M. Matschiner, O.K. Tørresen, S. Watanabe, B. Egger, Y.S. Han, E. Feunteun, R. Sommaruga, R. Jehle and R. Schabetsberger. 2020. Stable species boundaries despite ten million years of hybridization in tropical eels. *Nature Communications* 11:1433.
- Bengtson, S., T. Sallstedt, V. Belivanova and M. Whitehouse. 2017. Three-dimensional preservation of cellular and subcellular structures suggests 1.6 billion-year-old crown-group red algae. *PloS Biology* 15(3):e2000735.
- Bennetzen, J.L. and H. Wang. 2014. The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annual Review of Plant Biology* 65:505-530.
- Bhattacharya, D., H.S. Yoon and J.D. Hackett. 2004. Photosynthetic eukaryotes unite: endosymbiosis connects the dots. *BioEssays* 26:50-60.
- Bhattacharya, D. and D.C. Price. 2020. The algal tree of life from a genomics perspective. In: A.W.D. Larkum, A.R.

- Grossman and J.A. Raven (eds.), *Photosynthesis in algae: biochemical and physiological mechanisms*, Springer, Cham, Switzerland. pp. 11-24.
- Bobay, L.M. 2020. The prokaryotic species concept and challenges. In: H. Tettelin and D. Medini (eds.), *The pangenome: diversity, dynamics and evolution of genomes*, Springer, Cham, Switzerland. pp. 21-49.
- Brawley, S.H., N.A. Blouin, E. Ficko-Blean, G.L. Wheeler, M. Lohr, H.V. Goodson, J.W. Jenkins, C.E. Blaby-Haas, K.E. Helliwell, C.X. Chan, T.N. Marriage, D. Bhattacharya, A.S. Klein, Y. Badis, J. Brodie, Y. Cao, J. Collén, S.M. Dittami, C.M.M. Gachon, B.R. Green, S.J. Karpowicz, J.W. Kim, U.J. Kudahl, S. Lin, G. Michel, M. Mittag, B.J.S.C. Olson, J.L. Pangilinan, Y. Peng, H. Qiu, S. Shu, J.T. Singer, A.G. Smith, B.N. Sprecher, V. Wagner, W. Wang, Z.Y. Wang, J. Yan, C. Yarish, C. Zäuner-Riek, Y. Zhuang, Y. Zou, E.A. Lindquist, J. Grimwood, K.W. Barry, D.S. Rokhsar, J. Schmutz, J.W. Stiller, A.R. Grossman and S.E. Prochnik. 2017. Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta). *Proceedings of the National Academy of Sciences of the USA* 114(31): E6361-6370.
- Brodie, J., S.G. Ball, F.Y. Bouget, C.X. Chan, O. De Clerck, J.M. Cock, C. Gachon, A.R. Grossman, T. Mock, J.A. Raven, M. Saha, A.G. Smith, A. Vardi, H.S. Yoon and D. Bhattacharya. 2017. Biotic interactions as drivers of algal origin and evolution. *New Phytologist* 216(3):670-681.
- Bryand, D.A. and D.P. Canniffe. 2018. How nature designs light-harvesting antenna systems: design principles and functional realization in chlorophototrophic prokaryotes. *Journal of Physics B: Atomic, Molecular and Optical Physics* 51(3):033001.
- Burki, F., N. Okamoto, J.F. Pombert and P.J. Keeling. 2012. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. *Proceedings of the Royal Society B: Biological Sciences* 279(1736):2246-2254.
- Burki, F., A.J. Roger, M.W. Brown and A.G.B. Simpson. 2019. The new tree of eukaryotes. *Trends in Ecology & Evolution* 31(1):43-55.
- Butterfield, N.J. 2000. *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* 26(3):386-404.
- Chan, C.X., E.C. Yang, T. Banerjee, H.S. Yoon, P.T. Martone, J.M. Estevez and D. Bhattacharya. 2011. Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. *Current Biology* 21(4): 328-333.
- Collén, J., B. Porcel, W. Carré, S.G. Ball, C. Chaparro, T. Tonon, T. Barbeyron, G. Michel, B. Noel, K. Valentin, M. Elias, F. Artiguenave, A. Arun, J.M. Aury, J.F. Barbosa-Neto, J.H. Bothwell, F.Y. Bouget, L. Brillet, F. Cabello-Hurtado, S. Capella-Gutiérrez, B. Charrier, L. Cladière, J.M. Cock, S.M. Coelho, C. Colleoni, M. Czjzek, C.D. Silva, L. Delage, F. Denoëud, P. Deschamps, S.M. Dittami, T. Gabaldón, C.M.M. Gachon, A. Groisillier, C. Hervé, K. Jabbari, M. Katinka, B. Kloareg, N. Kowalczyk, K. Labadie, C. Leblanc, P.J. Lopez, D.H. McLachlan, L. Meslet-Cladière, A. Moustafa, Z. Nehr, P.N. Collén, O. Panaud, F. Partensky, J. Poulain, S.A. Rensing, S. Rousvoal, G. Samson, A. Symeonidi, J. Weissenbach, A. Zambounis, P. Wincker and C. Boyen. 2013. Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proceedings of the National Academy of Sciences of the USA* 110(13):5247-5252.
- Costa, J.F., S.M. Lin, E.C. Macaya, C. Fernández-García and H. Verbruggen. 2016. Chloroplast genomes as a tool to resolve red algal phylogenies: a case study in the Nemaliales. *BMC Evolutionary Biology* 16:205.
- Cummins, C.A. and J.O. McInerney. 2011. A method for inferring the rate of evolution of homologous characters that can potentially improve phylogenetic inference, resolve deep divergence and correct systematic biases. *Systematic Biology* 60(6):833-844.
- de Vries, J., A. Stanton, J.M. Archibald and S.B. Gould. 2016. Streptophyte terrestrialization in light of plastid evolution. *Trends in Plant Science* 21(6):467-476.
- de Vries, J. and S.B. Gould. 2018. The monoplastidic bottleneck in algae and plant evolution. *Journal of Cell Science* 131(2):jcs203414.
- Díaz-Tapia, P., C.A. Maggs, J.A. West and H. Verbruggen. 2017. Analysis of chloroplast genomes and a supermatrix inform reclassification of the Rhodomelaceae (Rhodophyta). *Journal of Phycology* 53(5):920-937.
- Díaz-Tapia, P., C.A. Maggs, E.C. Macaya and H. Verbruggen. 2018. Widely distributed red aglae often represent hidden introductions, complexes of cryptic species or species with strong phylogeographic structure. *Journal of Phycology* 54(6):829-839.
- Díaz-Tapia, P., M.M. Pasella, H. Verbruggen and C.A. Maggs. 2019. Morphological evolution and classification of the red algal order Ceramiales inferred using plastid phylogenomics. *Molecular Phylogenetics and Evolution* 137:76-85.
- Dorrell, R.G., G. Gile, G. McCallum, R. Méheust, E.P. Bapteste, C.M. Klinger, L. Brillet-Guéguen, K.D. Freeman, D.J. Richter and C. Bowler. 2017. Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome. *eLife* 6:e23717.
- Edge, M.D. and G. Coop. 2019. Reconstructing the history of polygenic scores using coalescent trees. *Genetics* 211:235-262.
- Fan, X., H. Qiu, W. Han, Y. Wang, D. Xu, X. Zhang, D. Bhattacharya and N. Ye. 2020. Phytoplankton pangenome reveals extensive prokaryotic horizontal gene transfer of diverse functions. *Science Advances* 6(18):eaba0111.
- Fehrer, J., B. Gemeinholzer and J.B.S. Chrték. 2007. Incongruent plastid and nuclear DNA phylogenies reveal anc-



- ient intergeneric hybridization in *Pilosella hawkweeds* (Hieracium, Cichorieae, Asteraceae). *Molecular Phylogenetics and Evolution* 42(2):347-361.
- Freshwater, D.W., S. Fredericq, B.S. Butler, M.H. Hommersand and M.W. Chase. 1994. A gene phylogeny of the red algae (Rhodophyta) based on plastid *rbcL*. *Proceedings of the National Academy of Sciences of the USA* 91(15): 7281-7285.
- Gabrielson, P.W., D.J. Garbary and R.F. Scagel. 1985. The nature of the ancestral red alga: inferences from a cladistic analysis. *BioSystems* 18(3-4):335-346.
- Gawryluk, R.M.R., D.V. Tikhonenkov, E. Hehenberger, F. Husnik, A.P., Mylnikov and P.J. Keeling. 2019. Non-photosynthetic predators are sister to red algae. *Nature* 572(7768):240-243.
- Gibson, T.M., P.M. Shih, V.M. Cumming, W.W. Fischer, P.W. Crockford, M.S.W. Hodgskiss, S. Wörndle, R.A. Creaser, R.H. Rainbird, T.M. Skulski and G.P. Halverson. 2018. Precise age of *Bangiomorpha pubescens* dates the origin of eukaryotic photosynthesis. *Geology* 46(2):135-138.
- Giovagnetti, V. and A.V. Ruban. 2018. The evolution of the photoprotective antenna proteins in oxygenic photosynthetic eukaryotes. *Biochemical Society Transaction* 46(5): 1263-1277.
- Godfroy, O., A.F. Peters, S.M. Coelho and J.M. Cock. 2015. Genome-wide comparison of ultraviolet and ethyl methanesulphonate mutagenesis methods for the brown alga *Ectocarpus*. *Marine Genomics* 24:109-113.
- Golicz, A.A., P.E. Bayer, P.L. Bhalla, J. Batley and D. Edwards. 2020. Pangenomics comes of age: from bacteria to plant and animal applications. *Trends in Genetics* 36(2):132-145.
- Gradinger, R. 2009. Sea-ice algae: major contributors to primary production and algal biomass in the Chukchi and Beaufort Seas during May/June 2002. *Deep-Sea Research II* 56(17):1201-1212.
- Graf, L., E.C. Yang, K.Y. Han, F.C. Küpper, K.M. Benes, J.K. Oyadomari, R.J.H. Herbert, H. Verbruggen, R. Wetherbee, R.A. Andersen and H.S. Yoon. 2020. Multigene phylogeny, morphological observation and re-examination of the literature lead to the description of the Phaeosacciophyceae classis nova and four new species of the Heterokontophyta SI clade. *Protist* 171(6):125781.
- Graf, L., Y. Shin, J.H. Yang, J.W. Choi, I.K. Hwang, W. Nelson, D. Bhattacharya, F. Viard and H.S. Yoon. 2021. A genome-wide investigation of the effect of farming and human-mediated introduction on the ubiquitous seaweed *Undaria pinnatifida*. *Nature Ecology & Evolution* 5(3):360-368.
- Guiry, M.D. and G.M. Guiry. 2021. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. <https://www.algaebase.org>; searched on March 2021.
- Gulbrandsen, Ø.S., I.J. Andresen, A.K. Krabberød, J. Bråte and K. Shalchian-Tabrizi. 2021. Phylogenomic analysis restructures the ulvophyceae. *Journal of Phycology* (accepted article).
- Hu, Y., W. Xing, Z. Hu and G. Liu. 2020. Phylogenetic analysis and substitution rate estimation of colonial volvocine algae based on mitochondrial genomes. *Genes* 11(1):115.
- Huson, D.H. and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(3):254-267.
- Iha, C., C.J. Grassa, G.D.M. Lyra, C.C. Davis, H. Verbruggen and M.C. Oliveira. 2018. Organellar genomics: a useful tool to study evolutionary relationships and molecular evolution in Gracilariaceae (Rhodophyta). *Journal of Phycology* 54(6):775-787.
- Jain, C., L.M. Rodriguez-R, A.M. Phillippy, K.T. Konstantinidis and S. Aluru. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications* 9:5114.
- Janouškovec, J., S.L. Liu, P.T. Martone, W. Carré, C. Leblanc, J. Collén and P.J. Keeling. 2013. Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLoS ONE* 8(3):e59001.
- Jarvis, E.D., S. Mirarab, A.J. Aberer, B. Li, P. Houde, C. Li, S.Y.W. Ho, B.C. Faircloth, B. Nabholz, J.T. Howard, A. Suh, C.C. Weber, R.R. da Fonseca, J. Li, F. Zhang, H. Li, L. Zhou, N. Narula, L. Liu, G. Ganapathy, B. Boussau, M.S. Bayzid, V. Zavidovych, S. Subramanian, T. Gabaldón, S. Capella-Gutiérrez, J. Huerta-Cepas, B. Rekepalli, K. Munch, M. Schierup, B. Lindow, W.C. Warren, D. Ray, R.E. Green, M.W. Bruford, X. Zhan, A. Dixon, S. Li, N. Li, Y. Huang, E.P. Derryberry, M.F. Bertelsen, F.H. Sheldon, R.T. Brumfield, C.V. Mello, P.V. Lovell, M. Wirthlin, M.P.C. Schneider, F. Prosdocimi, J.A. Samaniego, A.M.V. Velazquez, A. Alfaro-Núñez, P.F. Compos, B. Petersen, T. Sicheritz-Ponten, A. Pas, T. Bailey, P. Scofield, M. Bunce, D.M. Lambert, Q. Zhou, P. Perelman, A.C. Driskell, B. Shapiro, Z. Xiong, Y. Zeng, S. Liu, Z. Li, B. Liu, K. Wu, J. Xiao, X. Yinqi, Q. Zheng, Y. Zhang, H. Yang, J. Wang, L. Smeds, F.E. Rheindt, M. Braun, J. Fjeldsa, L. Orlando, F.K. Barker, K.A. Jønsson, W. Johnson, K.P. Keopfli, S. O'Brien, D. Haussler, O.A. Ryder, C. Rahbek, E. Willerslev, G.R. Graves, T.C. Glenn, J. McCormack, D. Burt, H. Ellegren, P. Alström, S.V. Edwards, A. Stamatakis, D.P. Mindell, J. Cracraft, E.L. Braun, T. Warnow, W. Jun, M.T.P. Gilbert and G. Zhang. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346(6215):1320-1331.
- Keeling, P.J. 2010. The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1541):729-748.
- Keeling, P.J. and F. Burki. 2019. Progress towards the tree of eukaryotes. *Current Biology* 29(16):R808-R817.
- Keeling, P.J., F. Burki, H.M. Wilcox, B. Allam, E.E. Allen, L.A. Amaral-Zettler, E.V. Armbrust, J.M. Archibald, A.K. Bharti, C.J. Bell, B. Beszteri, K.D. Bidle, C.T. Cameron, L. Campbell, D.A. Caron, R.A. Cattolico, J.L. Collier, K. Coyne, S.K. Davy, P. Deschamps, S.T. Dyhrman, B. Ed-

- varlsen, R.D. Gates, C.J. Gobler, S.J. Greenwood, S.M. Guida, J.L. Jacobi, K.S. Jakobsen, E.R. James, B. Jenkins, U. John, M.D. Johnson, A.R. Juhl, A. Kamp, L.A. Katz, R. Kiene, A. Kudryavtsev, B.S. Leander, S. Lin, C. Lovejoy, D. Lynn, A. Marchetti, G. McManus, A.M. Nedelcu, S. Menden-Deuer, C. Miceli, T. Mock, M. Montresor, M.A. Moran, S. Murray, G. Nadathur, S. Nagai, P.B. Ngam, B. Palenik, J. Pawlowski, G. Petroni, G. Piganeau, M.C. Posewitz, K. Rengefors, G. Romano, M.E. Rumpho, T. Rynearson, K.B. Schilling, D.C. Schroeder, A.G.B. Simpson, C.H. Slamovits, D.R. Smith, G.J. Smith, S.R. Smith, H.M. Sosik, P. Stief, E. Theriot, S.N. Twary, P.E. Umale, D. Vaultot, B. Wawrik, G.L. Wheeler, W.H. Wilson, Y. Xu, A. Zingone and A.Z. Worden. 2014. The marine microbial eukaryote transcriptome sequencing project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PloS Biology* 12(6):e1001889.
- Kim, J.I., C.E. Moore, J.M. Archibald, D. Bhattacharya, G. Yi, H.S. Yoon and W. Shin. 2017. Evolutionary dynamics of cryptophyte plastid genomes. *Genome Biology and Evolution* 9(7):1859-1872.
- Kumar, V., F. Lammers, T. Bidon, M. Pfenninger, L. Kolter, M.A. Nilsson and A. Janke. 2017. The evolutionary history of bears is characterized by gene flow across species. *Scientific Reports* 7:46487.
- Lee, J.M., C.H. Cho, S.I. Park, J.W. Choi, H.S. Song, J.A. West, D. Bhattacharya and H.S. Yoon. 2016a. Parallel evolution of highly conserved plastid genome architecture in red seaweeds and seed plants. *BMC Biology* 14:75.
- Lee, J.M., K.M. Kim, E.C. Yang, K.A. Miller, S.M. Boo, D. Bhattacharya and H.S. Yoon. 2016b. Reconstructing the complex evolutionary history of mobile plasmids in red algal genomes. *Scientific Reports* 6:23744.
- Lee, J.M., H.J. Song, S.I. Park, Y.M. Lee, S.Y. Jeong, T.O. Cho, J.H. Kim, H.G. Choi, C.G. Choi, W.A. Nelson, S. Fredericq, D. Bhattacharya and H.S. Yoon. 2018a. Mitochondrial and plastid genomes from coralline red algae provide insights into the incongruent evolutionary histories of organelles. *Genome Biology and Evolution* 10(11):2961-2972.
- Lee, J.M., E.C. Yang, L. Graf, J.H. Yang, H. Qiu, U. Zelzion, C.X. Chan, T.G. Stephens, A.P.M. Weber, G.H. Boo, S.M. Boo, K.M. Kim, Y. Shin, M. Jung, S.J. Lee, H.S. Yim, J.H. Lee, D. Bhattacharya and H.S. Yoon. 2018b. Analysis of the draft genome of the red seaweed *Gracilariopsis chorda* provides insights into genome size evolution in Rhodophyta. *Molecular Biology and Evolution* 35(8):1869-1886.
- Lee, J.M., D. Kim, D. Bhattacharya and H.S. Yoon. 2019. Expansion of phycobilisome linker gene families in mesophilic red algae. *Nature Communications* 10:4823.
- Lhee, D., J.M. Lee, K. Ettahi, C.H. Cho, J.S. Ha, Y.F. Chan, U. Zelzion, T.G. Stephens, D.C. Price, A. Gabr, E.C.M. Nowack, D. Bhattacharya and H.S. Yoon. 2021. Amoeba Genome Reveals Dominant Host Contribution to Plastid Endosymbiosis. *Molecular Biology and Evolution* 38(2):344-357.
- MacGuigan, D.J., A.J. Geneva and R.E. Glor. 2017. A genomic assessment of species boundaries and hybridization in a group of highly polymorphic anoles (*distichus* species complex). *Ecology and Evolution* 7(11):3657-3671.
- Maddison, W.P. and L.L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55(1):21-30.
- Mao, Y., S. Hou, J. Shi and E.P. Economo. 2020. TREEasy: an automated workflow to infer gene trees, species trees, and phylogenetic networks from multilocus data. *Molecular Ecology Resources* 20:832-840.
- Matteoli, F.P., H. Passarelli-Araujo, F. Pedrosa-Silva, F.L. Olivares and T.M. Venancio. 2020. Population structure and pangenome analysis *Enterobacter bugandensis* uncover the presence of blaCTX-M-55, blaNDM-5 and blaIMI-1, along with sophisticated iron acquisition strategies. *Genomics* 112(2):1182-1191.
- Matute, D.R. and V.E. Sepúlveda. 2019. Fungal species boundaries in the genomics era. *Fungal Genetics and Biology* 131:103249.
- Medlin, L.K., K. Metfies, U. John and J.L. Olsen. 2007. Algal molecular systematics: a review of the past and prospects for the future. In: J. Brodie and J. Lewis (eds.), *Unravelling the algae: the past, present, and future of algal systematics*, Taylor and Francis, Boca Raton, Florida. pp. 341-353.
- Mercado, J.M., S. Salles and D. Cortés. 2020. Ecophysiology of marine algae: factors regulating primary production and growth in natural environments. In: S.K. Kim (ed.), *Encyclopedia of marine biotechnology*, John Wiley & Sons. pp. 341-357.
- Muñoz-Gómez, S.A., F.G. Mejía-Franco, K. Durnin, M. Colp, C.J. Grisdale, J.M. Archibald and C.H. Slamovits. 2017. The new red algal subphylum Proteorhodophytina comprises the largest and most divergent plastid genomes known. *Current Biology* 27(11):1677-1684.
- Nguyen, L.T., H.A. Schmidt, A. Von Haeseler and B.Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32(1):268-274.
- Nowack, E.C.M., M. Melkonian and G. Glöckner. 2008. Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Current Biology* 18(6):410-418.
- Oliveira, M.C., S.I. Repetti, C. Iha, C.J. Jackson, P. Díaz-Tapia, K.M.F. Lubiana, V. Cassano, J.F. Costa, M.C.M. Cremen, V.R. Marcelino and H. Verbruggen. 2018. High-throughput sequencing for algal systematics. *European Journal of Phycology* 53(3):256-272.
- Ono, J., D. Greig and P.J. Boynton. 2020. Defining and disrupting species boundaries in *Saccharomyces*. *Annual Review of Microbiology* 74:477-495.
- Parfrey, L.W., D.J.G. Lahr, A.H. Knoll and L.A. Katz. 2011.

- Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proceedings of the National Academy of Sciences of the USA* 108(33):13624-13629.
- Payo, D.A., F. Leliaert, H. Verbruggen, S. D'hondt, H.P. Calumpong and O. De Clerck. 2013. Extensive cryptic species diversity and fine-scale endemism in the marine red alga *Portieria* in the Philippines. *Proceedings of the Royal Society B: Biological Sciences* 280(1753):20122660.
- Pirie, M.D. 2015. Phylogenies from concatenated data: is the end nigh? *Taxon* 64(3):421-423.
- Ponce-Toledo, R., P. Deschamps, P. López-García, Y. Zivanovic, K. Benzerara and D. Moreira. 2017. An early-branching freshwater cyanobacterium at the origin of plastids. *Current Biology* 27(3):386-391.
- Price, D.C., C.X. Chan, H.S. Yoon, E.C. Yang, H. Qiu, A.P.M. Weber, R. Schwacke, J. Gross, N.A. Blouin, C. Lane, A. Reyes-Prieto, D.G. Durnford, J.A.D. Neilson, B.F. Lang, G. Burger, J.M. Steiner, W. Löffelhardt, J.E. Meuser, M.C. Posewitz, S. Ball, M.C. Arias, B. Henrissat, P.M. Coutinho, S.A. Rensing, A. Symeonidi, H. Doddapaneni, B.R. Green, V.D. Rajah, J. Boore and D. Bhattacharya. 2012. *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science* 355(6070):843-847.
- Ragan, M.A., C.J. Bird, E.L. Rice, R.R. Gutell, C.A. Murphy and R.K. Singh. 1994. A molecular phylogeny of the marine red algae (Rhodophyta) based on the nuclear small-subunit rRNA gene. *Proceedings of the National Academy of Sciences of the USA* 91(15):7276-7280.
- Read, B.A., J. Kegel, M.J. Klute, A. Kuo, S.C. Lefebvre, F. Maumus, C. Mayer, J. Miller, A. Monier, A. Salamov, J. Young, M. Aguilar, J.M. Claverie, S. Frickenhaus, K. Gonzalez, E.K. Herman, Y.C. Lin, J. Napier, H. Ogata, A.F. Sarno, J. Shmutz, D. Schroeder, C. de Vargas, F. Verret, P. von Dassow, K. Valentin, Y. Van de Peer, G. Sheeler, *Emiliana huxleyi* Annotation Consortium, J.B. Dacks, C.F. Delwiche, S.T. Dyhrman, G. Glöckner, U. John, T. Richards, A.Z. Worden, X. Zhang and I.V. Grigoriev. 2013. Pan genome of the phytoplankton *Emiliana* underpins its global distribution. *Nature* 499(7457):209-213.
- Reyes-Prieto, A., H.S. Yoon, A. Moustafa, E.C. Yang, R.A. Andersen, S.M. Boo, T. Nakayama, K.I. Ishida and D. Bhattacharya. 2010. Differential gene retention in plastids of common recent origin. *Molecular Biology and Evolution* 27(7):1530-1537.
- Rice, D.W. and J.D. Palmer. 2006. An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. *BMC Biology* 4:31.
- Richard, G.F. 2020. Eukaryotic pangenomes. In: H. Tettelin and D. Medini (eds.), *The pangenome: diversity, dynamics and evolution of genomes*, Springer, Cham, Switzerland. pp. 253-291.
- Rodríguez-Ezpeleta, N., H. Brinkmann, S.C. Burey, B. Roure, G. Burger, W. Löffelhardt, H.J. Bohnert, H. Philippe and B.F. Lang. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Current Biology* 15(14):1325-1330.
- Rokas, A., B.L. Williams, N. King and S.B. Carroll. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425(6960):798-804.
- Rossoni, A.W., D.C. Price, M. Seger, D. Lyska, P. Lammers, D. Bhattacharya and A.P.M. Weber. 2019. The genomes of polyextremophilic cyanidiales contain 1% horizontally transferred genes with diverse adaptive functions. *eLife* 8:e45017.
- Rudd, S. 2003. Expressed sequence tags: alternative or complement to whole genome sequences? *Trends in Plant Science* 8(7):1360-1385.
- Russell, S., C. Jackson and A. Reyes-Prieto. 2021. High sequence divergence but limited architectural rearrangements in organelle genomes of *Cyanophora* (Glaucophyta) species. *Journal of Eukaryotic Microbiology* 68:e12831.
- Sánchez-Baracaldo, P., J.A. Raven, D. Pisani and A.H. Knoll. 2017. Early photosynthetic eukaryotes inhabited low-salinity habitats. *Proceedings of the National Academy of Sciences of the USA* 114(37):E7737-7745.
- Saunders, G.W. and K.V. Lehmkuhl. 2005. Molecular divergence and morphological diversity among four cryptic species of *Plocamium* (Plocamiales, Florideophyceae) in northern Europe. *European Journal of Phycology* 40(3):293-312.
- Seeleuthner, Y., S. Mondy, V. Lombard, Q. Carradec, E. Pelletier, M. Wessner, J. Leconte, J.F. Mangot, J. Poulain, K. Labadie, R. Logares, S. Sunagawa, V. de Berardinis, M. Salanoubat, C. Dimier, S. Kandels-Lewis, M. Picheral, S. Searson, Tara Oceans Coordinators, S. Pesant, N. Poulton, R. Stepanauskas, P. Bork, C. Bowler, P. Hingamp, M.B. Sullivan, D. Iudicone, R. Massana, J.M. Aury, B. Henrissat, E. Karsenti, O. Jaillon, M. Sieracki, C. de Vargas and P. Wincker. 2018. Single-cell genomics of multiple uncultured stramenopiles reveals underestimated functional diversity across oceans. *Nature Communications* 9:310.
- Schliep, K., A.J. Potts, D.A. Morrison and G.W. Grimm. 2017. Intertwining phylogenetic trees and networks. *Methods in Ecology and Evolution* 8(10):1212-1220.
- Shapiro, B.J. 2017. The population genetics of pangenomes. *Nature Microbiology* 2(12):1574-1574.
- Shaw, K.L. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Sciences of the USA* 99(25):16122-16127.
- Sibbald, S.J. and J.M. Archibald. 2020. Genomic insights into plastid evolution. *Genome Biology and Evolution* 12(7):978-990.
- Simpson, A.G.B., C.H. Slamovits and J.M. Archibald. 2017. Protist diversity and eukaryote phylogeny. In: J.M. Archibald, A.G.B. Simpson and C.H. Slamovits (eds.), *Handbook of the protists* (2nd ed.), Springer International Pub-

- lishing, Cham, Switzerland. pp. 1-21.
- Song, H.J., J.M. Lee, L. Graf, M. Rho, H. Qiu, D. Bhattacharya and H.S. Yoon. 2016. A novice's guide to analyzing NGS-derived organelle and metagenome data. *Algae* 31(2):137-154.
- Strassert, J.F.H., M. Jamy, A.P. Mylnikov, D.V. Tikhonenkov and F. Burki. 2019. New phylogenomic analysis of the enigmatic phylum Telonemia further resolves the eukaryote tree of life. *Molecular Biology and Evolution* 36(4):757-765.
- Strassert, J.F.H., I. Irisarri, T.A. Williams and F. Burki. 2021. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nature Communications* 12:1879.
- Sun, M., D.E. Soltis, P.S. Soltis, X. Zhu, J.G. Burleigh and Z. Chen. 2015. Deep phylogenetic incongruence in the angiosperm clade *Rosidae*. *Molecular Phylogenetics and Evolution* 83:156-166.
- Tang, Q., K. Pang, X. Yuan and S. Xiao. 2020. A one-billion-year-old multicellular chlorophyte. *Nature Ecology & Evolution* 4(4):543-549.
- Tettelin, H., V. Massignani, M.J. Cieslewicz, C. Donati, D. Medini, N.L. Ward, S.V. Angiuoli, J. Crabtree, A.L. Jones, A.S. Durkin, R.T. DeBoy, T.M. Davidsen, M. Mora, M. Scarselli, I.M. Ros, J.D. Peterson, C.R. Hauser, J.P. Sundaram, W.C. Nelson, R. Madupu, L.M. Brinkac, R.J. Dodson, M.J. Rosovitz, S.A. Sullivan, S.C. Daugherty, D.H. Haft, J. Selengut, M.L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G. Dimitrov, K. Watkins, K.J.B. O'Connor, S. Smith, J.L. Telford, M.R. Wessels, R. Rappuoli and C.M. Fraser. 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proceedings of the National Academy of Sciences of the USA* 102(39):13950-13955.
- Theriot, E.C., M. Ashworth, E. Ruck, T. Nakov and R.K. Jansen. 2010. A preliminary multigene phylogeny of the diatoms (Bacillariophyta): challenges for future research. *Plant Ecology and Evolution* 143(3):278-296.
- Timmis, J.N., M.A. Ayliffe, C.Y. Huang and W. Martin. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Reviews Genetics* 5(2):123-135.
- Toews, D.P.L. and A. Brelford. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* 21(16):3907-3930.
- Umen, J. and S. Coelho. 2019. Algal sex determination and the evolution of anisogamy. *Annual Review of Microbiology* 73:267-291.
- Van Etten, J. and D. Bhattacharya. 2020. Horizontal gene transfer in eukaryotes: not if, but how much? *Trends in Genetics* 36(12):915-925.
- Wendel, J.F. and J.J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. In: D.E. Soltis, P.S. Soltis and J.J. Doyle (eds.), *Molecular Systematics of Plants II*, Springer, Boston, MA. pp. 265-296.
- Willis, A. and J.N. Woodhouse. 2020. Defining cyanobacterial species: diversity and description through genomics. *Critical Reviews in Plant Science* 39(2):101-124.
- Xi, Z., L. Liu, J.S. Rest and C.C. Davis. 2014. Coalescent versus concatenation methods and the placement of *Amborella* as sister to water lilies. *Systematic Biology* 63(6):919-932.
- Yabuki, A., R. Kamikawa, S.A. Ishikawa, M. Kolisko, E. Kim, A.S. Tanabe, K. Kume, K.I. Ishida and Y. Inagaki. 2014. *Palpitomonas bilix* represents a basal cryptist lineage: insight into the character evolution in Cryptista. *Scientific Reports* 4:4641.
- Yang, E.C., K.M. Kim, S.Y. Kim, J.M. Lee, G.H. Boo, J.H. Lee, W.A. Nelson, G. Yi, W.E. Schmidt, S. Fredericq, S.M. Boo, D. Bhattacharya and H.S. Yoon. 2015. Highly conserved mitochondrial genomes among multicellular red algae of the Florideophyceae. *Genome Biology and Evolution* 7(8):2394-2406.
- Yoon, H.S., J.D. Hackett and D. Bhattacharya. 2002. A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. *Proceedings of the National Academy of Sciences of the USA* 99(18):11724-11729.
- Yoon, H.S., J.D. Hackett, C. Ciniglia, G. Pinto and D. Bhattacharya. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Molecular Biology and Evolution* 21(5):809-818.
- Yoon, H.S., K.M. Müller, R.G. Sheath, F.D. Ott and D. Bhattacharya. 2006. Defining the major lineages of red algae (Rhodophyta). *Journal of Phycology* 42(2):482-492.
- Yoon, H.S., G.C. Zuccarello and D. Bhattacharya. 2010. Evolutionary history and taxonomy of red algae. In: J. Seckbach and D.J. Chapman (eds.), *Red algae in the genomic age*, Springer, Dordrecht. pp. 27-42.
- Yoon, H.S., D.C. Price, R. Stepanauskas, V.D. Rajah, M.E. Sieracki, W.H. Wilson, E.C. Yang, S. Duffy and D. Bhattacharya. 2011. Single-cell genomics reveals organismal interactions in uncultivated marine protists. *Science* 332(6030):714-717.
- Zhang, Z., C. Qu, K. Zhang, Y. He, X. Zhao, L. Yang, Z. Zheng, X. Ma, X. Wang, W. Wang, K. Wang, D. Li, L. Zhang, X. Zhang, D. Su, X. Chang, M. Zhou, D. Gao, W. Jiang, F. Leliaert, D. Bhattacharya, O. De Clerck, B. Zhong and J. Miao. 2020. Adaptation to extreme Antarctic environments revealed by the genome of a sea ice green alga. *Current Biology* 30(17):3330-3341.
- Zuccarello, G.C. and J.A. West. 2003. Multiple cryptic species: molecular diversity and reproductive isolation in the *Bostrychia radicans*/*B. moritziana* complex (Rhodomelaceae, Rhodophyta) with focus on North American isolates. *Journal of Phycology* 39(5):948-959.

Submitted: April 5, 2021  
 Revised: April 20, 2021  
 Accepted: April 21, 2021