Evolutionary history of the main extracellular matrix polysaccharides in brown algae

Lisa Mazéas[†,1], Ahlem Bouguerba-Collin[†,1], J. Mark Cock[1], France Denoeud[2], Olivier Godfroy[1], Loraine Brillet-Guéguen[1,3], Tristan Barbeyron[1], Agnieszka P. Lipinska[4], Ludovic Delage[1], Erwan Corre[3], Elodie Drula[5,6,7], Bernard Henrissat[5,6,8], Mirjam Czjzek[1], Nicolas Terrapon[5,6], Cécile Hervé[1,*]

[1]Sorbonne Université, CNRS, Integrative Biology of Marine Models Laboratory, Station Biologique de Roscoff, Roscoff, France

[2]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Université Evry, Université Paris-Saclay, Evry, 91057, France

[3]CNRS, Sorbonne Université, FR2424, ABiMS-IFB, Station Biologique, Roscoff, France

[4]Department of Algal Development and Evolution, Max Planck Institute for Biology, 72076, Tübingen, Germany.

[5]Aix Marseille Univ, CNRS, UMR 7257 AFMB, Marseille, France

[6]INRAE, USC 1408 AFMB, Marseille, France

[7]INRAE, Aix-Marseille Univ., UMR1163 BBF, Marseille, France

[8]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark

[†]These authors contributed equally

*Corresponding author at: UMR8227, *Station Biologique de Roscoff, Roscoff, France*

E-mail address: cecile.herve@sb-roscoff.fr (C. Hervé)

Abstract (max 200 words)

Brown algae belong to the Stramenopiles phylum and are phylogenetically distant from plants and other multicellular organisms. This independent evolutionary history has shaped brown algae with numerous metabolic characteristics specific to this group, including the synthesis of peculiar polysaccharides contained in their extracellular matrix (ECM). Alginates and fucose-containing sulphated polysaccharides (FCSP), the latter including fucans, are the main components of ECMs. However, the metabolic pathways of these polysaccharides remain poorly described due to a lack of genomic data. An extensive genomic dataset has been recently released for brown algae and their close sister species. We performed an expert annotation of key genes involved in ECM-carbohydrate metabolisms, combined with comparative genomics, phylogenetics analyses, and protein modelling. Our analysis indicates that the gene families involved in both the synthesis and degradation of alginate were acquired by the common ancestor of brown algae and their closest sister species *Schizocladia ischiensis*, and subsequently expanded in brown algae. The pathway for the biosynthesis of fucans still remains biochemically unresolved and we identify the most likely fucosyltransferase genes that may harbour a fucan synthase activity in brown algae. Our analysis questions the possible occurrence of FCSPs outside brown algae, notably within their closest sister taxon.

Keywords: extracellular matrix, brown algae, carbohydrate-active enzymes, alginates, fucans, glycosyltransferases, sulfotransferases

Introduction

Brown algae correspond to the class Phaeophyceae within the Stramenopile lineage (1). Many of these organisms are key components of extensive coastal ecosystems that provide high value ecosystem services, including the sequestration of several megatons of carbon per year. They represent a rich but underutilized biomass that will likely provide solutions to address global challenges, such as climate change, food shortages, and rising demands for natural bioactive compounds (2,3). These organisms are also impacted by climate change and shifts in their habitats and communities have been reported (4,5). They are key components of highly dynamic ecosystems, and to thrive in such habitats they have evolved plasticity in terms of metabolic responses, including at the level of their extracellular matrices (ECMs), or cell walls (6).

Brown algae evolved complex multicellularity independently from animals, fungi and land plants, and their metabolic networks have been shaped by secondary endosymbiosis and a number of horizontal gene transfers (HGTs) (6-8). This has resulted in the emergence of a range of unique metabolic features, including the generation of specific glycans in their ECMs (9,10). Alginates and fucose-containing sulfated polysaccharides (FCSPs) are the main components of the ECM, representing up to 45% of algal dry weight. Alginates are linear polymers made solely of two 1,4-linked epimers: β-D-mannuronic acid (M) and α-L-guluronic acid (G). FCSPs include both fucans and fucoidans (9). Fucans are highly sulfated polysaccharides with a backbone structure based on sulfated L-fucose residues, on which additional branches of sulfated fucose, galactose and glucuronate can occur. Fucoidans are a set of heterogeneous polymers with non-fucose backbones (e.g. chains of galactose, mannose or glucuronate), and branches of sulfated fucose (9).

The metabolic pathways for both alginates and sulfated fucans have been predicted (11,12). All enzymes biochemically characterized so far are involved either in the synthesis of sugar precursors or in the remodelling of alginates (12). GDP-mannose is an essential activated sugar used in the alginate biosynthetic pathway, and in the production of fucans through its conversion to GDP-fucose (12). The three enzymatic steps that lead to the production of GDP-mannose from fructose-6 phosphate are catalysed by mannose-6-phosphate isomerase (MPI), bifunctional phosphomannomutase (PMM)/phosphoglucomutase (PGM) and mannose-1-phosphate guyanylyltransferase (MPG). The later enzyme also has MPI activity (12,13). As in other eukaryotes and in most bacteria, the precursor GDP-fucose is likely produced in brown algae from GDP-mannose (the *de novo* pathway) and from fucose (the salvage pathway) (12). In the *de novo* pathway, GDP-mannose is first converted into GDP-4-keto-6-deoxymannose by GDP-mannose 4,6-dehydratase (GM46D), and subsequently converted to GDP-fucose by GDP-fucose synthase (GFS). In the salvage pathway, cytosolic fucose is first phosphorylated by a fucokinase (FK)

and then condensed to GTP by a GDP-fucose pyrophosphorylase (GFPP) to form GDP-fucose. Once GDP-mannose has been synthesized, its oxidation to GDP-mannuronic acid will result in the elongation of an alginate polymer in the form of polymannuronate. Oxidation of GDP-mannose is catalyzed by GDP-mannose dehydrogenase (GMD).

Once these activated sugars (e.g. GDP-mannose, GDP-fucose) have been generated, the elongation or grafting of the glucan chains is the most significant step in ECM synthesis, but, for brown algae, the genes involved have not been identified (12). Carbohydrate active enzymes (CAZymes) are major actors in polysaccharide metabolism (14,15). They are classified into different families in the CAZy database (www.cazy.org, (14)) based on their amino acid sequence similarities. CAZy classification includes glycosyltransferases (GTs), glycoside hydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), enzymes with Auxiliary Activities (AAs) and associated modules such as Carbohydrate-Binding Modules (CBMs). Additional important enzymes involved in the remodeling of glycans in brown algae include mannuronan C5-epimerases which convert M to G residues in alginates, and the sulfotransferases (STs) and sulfatases (SAs), which mediate the sulfation and desulfation of FCSPs, respectively (7,12). The SAs have been classified into families in the SulfAtlas database based on amino acid sequence similarities (https://sulfatlas.sb-roscoff.fr, (16,17)).

Sequencing a large number of genomes of brown algae and their sister groups, allows the emergence and diversification of key biological traits within the brown algae to be investigated, particularly traits associated with the transition to complex multicellularity, such as the elaboration of an adherent ECM. Brown seaweeds have remained poorly described in terms of genome sequencing due, in part, to difficulties with extracting nucleic acids. The Phaeoexplorer project has recently generated 60 new genomes corresponding to 40 species of brown algae and four close sister species (https://phaeoexplorer.sb-roscoff.fr, (18)). The 40 brown algal species include representatives of 16 families, spanning all the major orders of the Phaeophyceae. The sister species notably include the sister taxon *Schizocladia ischiensis*, the closest taxon to the brown algae (1,18,19). Here, we have investigated the evolutionary history of carbohydrate metabolism gene families in brown algae by a combination of expert annotations and evolutionary analysis of CAZymes (14), STs and SAs (16), as well as protein modelling of the most promising candidate genes. When relevant, we also performed genomic analyses of additional eukaryotes such as echinoderms or diatoms. This study offers the first representative view of the core CAZome in brown algae, derived from an extensive resource of sequences from distinct taxonomic origins. Our analysis indicates that the abilities to both synthesize and degrade alginate were gained prior to the diversification of the brown algae. We also identify the most likely fucan synthase genes in brown algae, an enzymatic step that still remains biochemically unresolved.

Material and methods

Identification of carbohydrate-related proteins

The CAZyme predicted proteins were identified by human curators of the CAZy database (http://www.cazy.org/, (14)). Curation consisted in the expert validation of the predicted modularity for each protein, using a combination of BLASTp pairwise alignments against both characterized members and previously curated sequences integrated in the CAZy database, as well as HMMER v3.4 (www.hmmer.org; (20)) comparisons to in-house HMM profiles of the families, and subfamilies when available. Proteins obtaining 100% coverage without gaps in pairwise alignments against a previously curated sequence were not manually inspected, allowing to decrease the load of human curation by processing genomes sequentially. The ST predicted proteins from various Stramenopiles were identified in the UniProt database based on the presence of a sulfotransferase-specific Pfam signature. Sequences of brown algal genomes not present in Uniprot were obtained by BLASTp searches in the Phaeoexplorer database (https://phaeoexplorer.sb-roscoff.fr, (18)) using the sequences found in the UniProt database as query. All other genes were identified by BLASTp searches of brown algal enzymes in the Phaeoexplorer database, using characterized brown algal genes as the queries, when known (MPI, PMM, GMD, UGD, ManC5-E, DEHU reductase). The homology threshold was chosen as ~35% pairwise sequence identity. In addition to the identified brown algal sequences, representative and characterized members of the targeted families were retrieved from the CAZy database (for CAZymes) and from public databases.

Phylogenetic analyses

Protein sequences were aligned with the MAFFT program using the FFT-NS-i iterative refinement method (https://mafft.cbrc.jp/alignment/server, (21)). In the case of multimodular proteins, the targeted catalytic domains were extracted prior to their alignment. The multiple alignments were curated manually using the Jalview program (https://www.jalview.org, (22)) to retain only the informative sites. Prior to carrying out comprehensive phylogenetics, a preliminary analysis was carried out for each protein family to select the optimal model of sequence evolution using the IQ-TREE progam (http://iqtree.cibiv.univie.ac.at/; (23)). Phylogenetic trees were subsequently constructed using the maximum likelihood approach with the RAxML program ((24), available at https://usegalaxy.fr). The reliability of the trees was tested with an MR-based Bootstopping criterion which generated resamplings of the datasets (range of 350 to 1000). The trees were edited in Mega 11.0.13 (https://www.megasoftware.net/; (25)) and graphically shaped with Adobe Illustrator CS6. For all phylogenies, only bootstrap values above 60% are shown. The brown algal sequences

are indicated with a color code in relation to their taxonomy as indicated by the key legend. Sequences from the closest sister taxa *S. ischiensis* are further indicated by the red circle. In all instances, the green squares indicate sequences which have been biochemically characterized.

Heatmaps based on Blast Score Ratios

To calculate the Blast Score Ratio (BSR), all pairs of sequences were compared using the NCBI BLASTp tool. The BSR, which corresponds to the bitscore normalized to the length of the protein query, was then calculated from the BLAST result following the procedure described in (26). Heat maps representing these BSR results and associated dendrograms were constructed using the ggheatmap function in the ggplot2 package in RStudio.

Extracted alignments and AlphaFold2 structural modeling

Some specific catalytic regions were extracted from alignments (Figures 2C, 5B, 8C). Alignments of full-length proteins were made using the MAFFT program and edited in the ESPript 3.0 program (https://espript.ibcp.fr; (27)). The color code of the amino acid residues indicates their biochemical character. Regions displaying high sequence similarity are boxed. The green squares indicate sequences which have been biochemically characterized. AlphaFold v2 (28) was used to predict 3D protein structures from primary amino acid sequences. The PyMOLv2.5.5 Molecular Graphics program (http://www.pymol.org/) was used to create and the display structural images.

Syntenic region analysis

Orthogroups, derived from predicted proteins of the 60 strains sequences in the Phaeoexplorer project (https://phaeoexplorer.sb-roscoff.fr, (18)) were searched for genes of interest. Microsynteny analysis was carried out on a subset of species where the target orthologues resided on a singular chromosome/scaffold, using McscanX (29), following the pipeline described at https://github.com/tanghaibao/jcvi/wiki/Mcscan. Gene annotations were then processed using AGAT (https://github.com/NBISweden/AGAT) and LASTZ (30) tools (C-score cutoff =0.7 to infer synteny blocks between species.

Results

CAZyme profiles in brown algae

Expert annotation of carbohydrate metabolism gene families was performed for all the major orders of brown algae (Phaeophyceae) and their close sister groups. Regarding CAZymes, we have already reported a quite constant number of genes (237 genes on average) for the 7 orders of brown algae (Figure 1A) (18). This number is similar in the closest sister taxon *S. ischiensis* (233 genes) while it is twice more than the more distantly related unicellular alga *Heterosigma akashiwo* (120 genes). The brown algal and sister taxa CAZymes are distributed across 69 families (including 37 GT, 21 GH and 1 PL families) with, on average, three genes per family (while reaching up to 16 genes for GT2 and GT4 and 19 genes for the PL41) (Figure 1B). The GT31 and GT47 families which are abundant in Viridiplantae (essentially β-galactosyl- or β-glucuronyltransferases, (31,32)) are also among the largest GT families in brown algae (average of 9 genes per genome for both families). We also detected in brown algae the occurrence of unexpected CAZyme families such as GT23 (α-1,6-L-fucosyltransferase, (33,34)), usually found in metazoan, GT49 (α-glucuronyltransferase), found in sac fungi, GT60 (α-N-acetyl-glucosaminyltransferase), found in amoeba plus chlorophyte green algae. Regarding glycosyl hydrolases, the most populated families were GH16, GH47, GH81 and GH114 (with 6, 5, 4 and 5 genes per genome respectively). The GH16 and GH81 enzymes are likely involved in the metabolism of the storage polysaccharide laminarin in brown algae, with β-1,3-glucanase activities already referenced in those families. GH47 genes are frequently found in eukaryotic genomes and are associated with α-1,2-mannosidase activity. Outside brown algae, GH114 are found in chlorophytes, fungi and bacteria, with a reported α-1,4-galactosaminidase activity (35,36). Finally PL41, which is, so far, the only PL family reported in brown algae and with alginate lyase activity (37), contains a high number of genes.

The genes involved in alginate metabolism in brown algae are also present in its closest sister taxon *S. ischiensis*

The alginate and the *de novo* fucan biosynthetic pathways share common steps leading to the generation of GDP-mannose as a precursor, prior to its conversion to GDP-fucose in the case of fucans. The two enzymes involved in these enzymatic conversions, MPI/MPG and PMM/PGM, have been characterized for *Saccharina japonica* (13,38). Analysis of their phylogeny indicated that these enzymes are widely present in brown algae but also more generally within the Stramenopiles, with several occurrences in the closest sister taxa and diatoms (Supplementary Figure S1A,B). GDP-mannose dehydrogenase (GMD) is essential to provide GDP-mannuronic acid for the initial assembly of alginate in its mannuronan form

by a mannuronan synthase. The latter putatively belongs to the large and multispecific GT2 family, although no functional validation has been obtained so far to identify specific candidate genes, making phylogenetic studies of this family highly speculative. The alginate obtained undergoes further modelling through the action of mannuronan C5-epimerases (ManC5-E). Both GMD and ManC5-E activities have been described, including in brown algae (13,39-42).

GMD belong to a protein superfamily also including and uridine diphosphate-glucose dehydrogenases (UGD). In order to explore the sequence diversity within this family, we gathered GMD and UGD protein sequences, using representative sets of brown algal sequences and additional characterized members from public databases. After clustering their BLASTp+ Score Ratio (BSR), we generated heatmaps that clearly illustrates that UGD and GMD form two distinct clusters (Supplementary Figure S2A). Within the first cluster, the UGD sequences from brown algae and the sister taxon *S. ischiensis*, share more than 90% pairwise protein sequence identity. In the second cluster, GMD sequences from brown algae and the sister taxon *S. ischiensis*, share more than 77% pairwise protein sequence identity, while they share less than 25% identity with the UGD sequences, indicating a large evolutive distance between these two families. The BSR-based heatmap indicates a distant cluster of sequences from *Actinobacteria*, which share around 30% identity with both GMD and UGD brown algal sequences. An additional BLAST search indicated that the brown algal GMDs share stronger sequence identities (40%) with a protein belonging to *Candidatus woesebacteriota* (MBP7967557.1), an uncultivated bacterium from a putative novel taxon (43). The phylogeny suggests that the common ancestor of *S. ischiensis* and brown algae probably gained the GMD sequence via horizontal gene transfer (HGT), and that the gene was duplicated in all brown algal groups (Figure 2A, Supplementary Figure S2B). The mannuronan C5 epimerases are a large multigenic family in brown algae, with an average of 27 members identified per genome. Eleven homologues were identified in *S. ischiensis* (Figure 2B, Supplementary Figure S3). Unlike aforementioned families, screening the other Stramenopile genomes available (*i.e*. other brown algal sister taxa, diatoms and oomycetes), did not identify any ManC5-E sequences. Three ManC5-E orthologue groups were identified (Figure 2B, Supplementary Figure S3). All the ManC5-E sequences possess the DPHD motif, typical of the ManC5-E catalytic site (Figure 2C), indicating that these sequences are genuine ManC5-Es. Marked amplifications of the gene family were detected in the Fucales, Desmarestiales and Laminariales. The closest bacterial representatives are from *Actinobacteria* with ~30% protein pairwise identity (Figure 2B, Supplementary Figure S3).

In addition to alginate synthesis, brown algae also possess a metabolic pathway for alginate degradation, the main actors being the PL41s. This is one of the largest degrading CAZyme family in brown algae (Figure 1B), with an average of 19 genes per genome, and

can be classified into at least four orthologous groups. PL41 duplication events have occurred in several brown algal orders including the Ectocarpales and the Laminariales. *S. ischiensis* displays three PL41 sequences while no homologues were found in other Stramenopiles (Figure 3A, Supplementary Figure S4A). PL41s are however not limited to brown algae, with bacterial members in some *Gammaproteobacteria*, *Actinobacteria* and *Alphaproteobacteria*, but with an average of one gene copy per genome (34-40% pairwise identify with the PL41 domain of brown algae). The oligoalginates generated by PL41s have been shown to be further degraded into 2-keto-3-deoxy-gluconate (KGD) by a DEHU reductase in the brown alga *S. japonica* (44). DEHU homologues were found in most brown algal genomes but not all sister taxa. This family of genes is found in other eukaryote genomes (Supplementary Figure S5). Overall, the broad distribution of this family of enzymes across eukaryotic organisms and the lack of homologues in some brown algae, argue against a strong specificity for alginate degradation in brown algae.

GH88 family members were found in all the brown algal genomes analysed (Figure 3B, Supplementary Figure S6). This CAZyme family includes characterized members which specifically act on oligosaccharides with an unsaturated uronic acid moiety (Δ), typically introduced by the lytic mechanism of lyases, such as heparan lyases (45) or ulvan lyases (46). As the PL41 alginate lyases are the only known PL family described so far in brown algae, the GH88 family would be an ideal candidate for further hydrolysis of oligosaccharides generated by PL41 action. The GH88 family in brown algae shares a common ancestor with that of its sister taxon *S. ischiensis* and the family has amplified in brown algae (Figure 3B, Supplementary Figure S6), especially within the Fucales where it is a large family (around 16 genes per genome).

Prediction of candidate genes for the synthesis of sulfated fucans

All the candidate genes putatively involved in the synthesis of GDP-fucose await functional validation in brown algae. Genes encoding the *de novo* pathway enzymes GM46D and GFS were identified in all the orders of brown algae analysed, with close homologues in sister taxa, including *S. ischiensis* (Supplementary Figure S7A,B). In contrast, genes encoding fucokinase (FK) and GDP-fucose pyrophophorylase (GFPP) from the salvage pathway were identified in the Phaeophyceae, but not in *S. ischiensis* (Supplementary Figure S7C,D). These two enzymes are usually found as a bimodular protein (FK-GFPP) in brown algae. They are found in a gene cluster previously reported in three Ectocarpales, which also encodes a sulfotransferase (ST) and a hydrolase (47). This microsyntenic region was found in other members of the Ectocarpales (Figure 4) and is conserved in other brown algal orders (Laminariales, Fucales, Desmarestiales), albeit with a stronger distance between the

contained genes near the root (*Dictyota dichotoma*). The synteny was not observed with the Discosporangiales and more distant taxa (sister taxa).

The most important set of enzymes involved in the synthesis of fucans are the fucosyltransferases (FucTs), which transfer GDP-fucose to oligofucans to elongate the chains. Their exact nature is still unknown. The CAZy database includes several FucTs that have been identified and these are classified into a variety of GT families. Our previous analyses of a restricted number of brown algal genomes, indicated that our FucT candidate genes for fucan synthesis may belong to either the GT10, GT23, GT41 and GT74 families (12). The extensive Phaeoexplorer dataset allowed us to performed a more in-depth analyses of these CAZyme families in brown algae, and to refine the identification of putative fucan synthases (Supplementary Table S1).

GT10s are referenced as α-1,3 or α-1,4-FucTs in the CAZy database, and the activities described relate to the transfer of an N-acetyl glucosamine (GlcNAc) residue (48,49). The fucose residues in fucans are α-1,3 and/or α-1,4-linked. It is thus tempting to hypothesise that the GT10 family is involved in their synthesis. However, there is only one GT10 gene per brown algal genome on average (Figure 5A, Supplementary Figure S8, Supplementary Table S1). Considering the importance of sulfated fucans in ECMs of all brown algae, we would expect the corresponding FucT gene to be present in multiple copies. As such this GT10 family is not the most favorable candidate for their biosynthesis. Moreover, a close homologue of brown algal GT10s is found in the Pelagophyceae, which are known not to contain fucans. The phylogeny further indicates that the brown algal sequences are closely related to plant GT10s, which are involved in protein glycosylation, where the α-1-3 FucT activity takes place on a GlcNAc residue, directly linked to an asparagine (50). While the activity of brown algal GT10s on oligofucans is unlikely, the GDP-fucose binding site is nonetheless highly conserved (Figure 5B,C), indicative of FucT activity. Comparisons were made with the only known structure of a GT10 from bacteria (51), and additional animal GT10 sequences for which the donor binding site has been studied (52,53). Both the alignment (Figure 5B) and the structure predictions (Figure 5C) indicate the conservation of key residues within the donor site in the brown algal sequences, while the acceptor site exhibits considerable variation (Figure 5D,E).

GT41 is a multispecific family in the CAZy database, with N-acetylglucosaminyl-transferase and protein O-fucosyltransferase activity, the latter being restricted, so far, to the green lineage (54). Brown algal genomes harbor ten GT41 genes on average, with homologues found in all ochrophytes, usually with several copies per genome. The GT41 phylogeny indicates that brown algal GT41s are organized into numerous orthologous groups, with *S. ischiensis* frequently found at the base of these algal groups. None of these orthologues are closely related to plant O-FucT (Figure 6A, Supplementary Figure S9). A

significant number of the orthologous groups are distant from the orthologous groups that contain the GlcNAc characterized activity, and could indicate new activities in the family. Overall, the phylogeny shows that the evolution of O-FucT activity has been restricted to the green lineage so far. Finally, the presence of multiple GT41 homologues in ochrophytes other than brown algae, does not favour the hypothesis that this class of enzymes contains fucan synthases.

α-1-2 fucosyltransferase activity has been reported in the GT74 family, which is also represented in brown algal genomes. This family has a restricted occurrence in eukaryotes, being only characterized in the amoeba *Dictyostelium discoideum*, where it participates in the glycosylation of the Skp-1 protein, involved in the ubiquitination regulatory system (55). The *D. discoideum* enzyme has a bimodular structure, the GT74 module being appended to a GT2 domain with β 1,3-galactosyltransferase activity. Bimodularity was also observed in the brown algal sequences, although the GT74 domain was found at the N-terminus, with the GT2 domain at the C-terminus, similarly to the situation found in some bacteria and the inverse of the situation in amoeba. Phylogenetic analysis indicated a common origin of GT74 genes from brown algae and other Stramenopiles (Figure 7, Supplementary Figure S10), although some genomes lacked this gene (*i.e.* some diatoms and brown algal sister taxa). Most brown algae have one GT74 gene. Again, considering the preponderance of sulfated fucans in ECMs of brown algae, and that this gene occurs in Stramenopiles known not to produce fucans, the GT74 family is not the best candidate to contain fucan synthases.

Finally, the GT23 family is present in all brown algae and is known to contain fucosyltransferases in other eukaryotes. GT23 genes are usually found in metazoa, where α-1,6 FucT activity has been reported and the structure of the human isoform FUT8 has been characterised (33,34). Uncharacterized homologues have also been found in bacteria, fungi and the Chlorophyta but this family is very rare in these lineages. On average, seven GT23 genes were found per genome in brown algae (Supplementary Table S1). Characteristic GT23 features are conserved in the brown algal sequences but the proteins are only distantly related to other sequences in the family. A BSR-based heatmap indicated two main clusters (Figure 8A), with the brown algal sequences representing an independent cluster, sharing on average less than 30% pairwise identity with the second cluster that contains the metazoan sequences. The phylogeny of the brown algal GT23 sequences shows that they do not strictly distribute in relation to taxonomy and this may indicate the presence of orthologue groups (Figure 8B, Supplementary Figure S11). All these GT23 sequences share, on average, 37% pairwise protein identity. Some duplication events have occurred in the Fucales and Laminariales. Overall, this analysis indicated that brown algae contain GT23 sequences that may belong to one or to several sub-classes within the GT23 family. The diversity of brown algal GT23 proteins may indicate functionally and/or structurally distinct

proteins. This diversity of proteins could potentially reflect the variety of positions to which a fucose is transferred in fucans and more generally in FCSPs (variety of monosaccharides, carbon position, linkage type). As such, the diverse GT23 proteins may reflect the FCSP diversity and the complex enzymatic machinery required for their synthesis.

As the brown algal GT23 sequences seem rather distant from other GT23s, and knowing that some of the latter have been biochemically and structurally characterized, we examined further the potential catalytic sites of the brown algal sequences. A sequence alignment was made for a selection of GT23 proteins from brown algae and the sister taxon *S. sichensis*, against the human FUT8 sequence (33,34). Despite the low sequence identity between our algal sequences and FUT8 (~22% on average), the donor site where the GDP-fucose binds is highly conserved (Figure 8C). This situation was even more apparent when the predicted structure of a representative brown algal protein was superposed on the FUT8 structure, also featuring the hydrolyzed GDP (Figure 8D). The catalytic residues His363, Arg365, Asp368, Lys369 and Ser469, involved in GDP-fucose fixation, are all conserved. This observation indicates that GT23 algal sequences are genuine fucosyltransferases. In contrast the acceptor site shows poor conservation between the algal and human GT23 (Figure 8E). The metazoan GT23s transfer fucose to the innermost GlcNAc residue of N-glycans (34,56). The Src homology 3 (SH3) domain at the C-terminus seems to be involved in protein dimerization, subcellular localization and to some extent, acceptor fixation (57). The SH3 domain is absent from brown algal sequences, and the N-terminal terminus is poorly predicted by AlphaFold2 (and therefore not shown in Figure 8E; the coverage being 89-458 on a total of 458 amino acid sequence). In the *S. latissima* GT23 model, several aromatic residues can be found in close vicinity to the GDP-fucose binding site and might participate in acceptor binding (Figure 8E, yellow residues). Overall, these observations indicate that GT23s in brown algae transfer the fucose residue onto an acceptor molecule distinct from those used by mammalian GT23s. One could speculate that brown algal GT23s catalyse transfer onto oligosaccharidic precursors of FCSPs (*e.g.* FCSOs).

Sulfated fucans are glycans largely restricted to brown algae, exception made of some marine echinoderms (i.e. body wall of sea cucumbers, egg jelly coat of sea urchins) where these molecules are known to be transiently produced and to occur with more linear and regular structures (9,58,59). We therefore also explored the CAZyme content of publicly available echinoderm genomes. We report the presence of the GT10, GT23, GT41 families (Supplementary Table S1). GT10 members are particularly abundant in these organisms, with on average 79 members, compared to 4 GT23 members. The GT23 members in these systems are not very similar to brown algal members of the family and are more closely associated with animal orthologues (Figure 8A). We also screened diatom genomes as recent literature suggest the presence of fucans in these organisms (60). We did not detect

any GT23 members in this case and only an average of 3 GT10 members per genome (Supplementary Table S1).

FCSPs, and especially fucans, are known to be heavily sulfated in brown algae. The FCSPs are likely initially polymerized in the state of neutral glycans by GTs, prior to their sulfation by STs. GTs and STs form large multigenic families in brown algae and brown algal STs can be clustered into at least 15 clades (12). In order to investigate the origin of brown algal STs within the Stramenopiles, we analysed the ST gene content of all the Phaeoexplorer brown algal and sister taxa genomes, along with genomes of other Stramenopiles. A phylogenetic tree built using these sequences (Figure 9A, Supplementary Figure S12) confirmed the occurrence of ~10 major orthologous groups in brown algae. All of these clades could include members with activities on polysaccharides, apart from clade 1, which is expected to be active on non-carbohydrate substrates (7,11,12). Clade 12 includes genes that are located within the syntenic region that also contains the dual FK-GFPP enzymes for the synthesis of the GDP-fucose precursor (Figure 4). This ST and FK-GFPP gene cluster is present in most brown algal genomes, including early diverging ones, suggesting a role for the STs in the sulfation of fucans or fucoidans. In brown algae, clades 4 and 10 include a high number of duplicated genes. Clade 4 includes members from other ochrophyta species but clade 10 appears to be specific to brown algae. In order to study the possible relationship of echinoderm STs with those of brown algae, especially those belonging to clade 10, we clustered these protein sequences using the BLASTP+ Score Ratio (BSR) approach (Figure 9B). The BSR-based heatmap indicates a closer relationship of the brown algal ST sequences from group 10 to echinoderm counterparts than to vertebrate STs, albeit with a pairwise identity score of below 30%. Note that no echinoderm STs have been biochemically characterized so far, thus the conclusion of their possible involvement in fucan sulfation is highly speculative. Regarding the sulfatase genes, as previously reported, brown algae and *S. ischiensis* contain 7 members per genome on average and they all belong to the S1_2 family (18). As for the STs, no biochemical activities have been determined so far for the corresponding enzymes.

Discussion

CAZyme numbers do not correlate with developmental complexity in brown algae

Our analyses provide the first representative view of the core CAZome in brown algae, derived from an exhaustive resource of sequences of distinct taxonomic origins, covering 16 Phaeophyceae families. Prior to the acquisition of large genomic datasets for these organisms, it had been suggested that some CAZyme families, especially GTs, were amplified in algae with more complex tissues (e.g. the kelp *Saccharina japonica*) relative to filamentous forms (e.g. *Ectocarpus* species 7) (61). However, this statement has been gradually invalidated as more genomes have become available (12), and the analysis reported here confirms that the number of total CAZymes in individual brown algal genomes is not correlated with either tissue complexity or taxonomic origin. The filamentous sister taxon *S. ischiensis* only harbors a slightly lower number of CAZymes and the other sister taxa analyzed possessed half as many as *S. ischiensis*. The latter case may indicate the absence of a dense and carbohydrate-rich ECM surrounding these unicellular organisms. One difficulty regarding 'exotic' genomes as compared to referenced/model species, lays in the automatic structural annotation of the genes, which may lead to the prediction of fragmented genes. In brown algae, the accounting of domain duplicates in predicted proteins and/or gene fragments by different authors, sometimes in the same organism, can lead to distinct conclusions (12,61,62). One should proceed with caution when drawing conclusions from the analysis of datasets coming from distinct sources. In our specific case, while we cannot address the issue regarding prediction of gene fragments at the whole genome scale, all species have been treated equally from sample preparation to structural and expert annotation, thus limiting the risks in giving biased statements.

Regarding the CAZyme families shared by all brown algae, GT2, GT23, GT31, GT47 and PL41 have a relatively high number of genes per genome, which would be coherent with them being involved in the assembly and remodelling of cell wall polysaccharides. The GT2 family is a large and multispecific family (e.g. includes proteins with different biochemical activities). It is generally one of the largest GT families in eukaryotes, and this is also the case in brown algae. Around two thirds of the GT2 sequences in brown algae are putative cellulose synthase (CESAs) and cellulose synthase-like (CSL) proteins (11), with some being possibly involved in the synthesis of mixed linkage glucans (63). Other GT2s related to dolichyl-phosphate β-glucosyltransferases are involved in protein N-linked glycosylation. The remaining GT2 members do not relate to any specific GT2 homologue with a known activity, and may include putative mannuronan synthases (11). Note that brown algal GT2 genes do not share homology with the Alg8 gene, a GT2 member known to be involved in alginate synthesis in alginate-producing bacteria (11).

GT31 is also a large family in other eukaryotes (*e.g.* 26 genes in *Homo sapiens* and 33 genes in *Arabidopsis thaliana*) with essentially glycoprotein galactosyltransferase and N-acetylglucosaminyltransferase activities and glycolipid activity in metazoans (64), and galactosyltransferase activity during the synthesis of arabinogalactan proteins (AGPs) in plants (31). Similarly, GT47 members are involved in the synthesis of glycosaminoglycans in metazoans and have various activities in the synthesis of cell-wall polysaccharides in plants (32). Brown algae have been shown to express chimeric AGPs (65) and it is tempting to speculate that the GT31 members may be involved in the synthesis of these molecules, although a definite chemical and structural analysis of those glycoproteins still needs to be carried out for brown algae. The PL41 family may be the largest, probably monospecific, CAZyme family in brown algae, indicating the importance of alginate metabolism in these organisms.

The abilities to synthesize and degrade alginate in brown algae were gained prior to their diversification

Alginate is a major component of brown algal cell walls where its rheological properties plays important roles in determining the wall's mechanical strength. Our extended analyses confirmed a previous statement (18) that several key gene families involved in alginate synthesis (GMD, ManC5-E) and degradation (PL41, and possibly GH88) were acquired by the common ancestor of brown algae and *S. ischiensis*. This event was followed by a marked amplification of some key families (ManC5-E, PL41) in brown algae relative to its sister taxon.

The ManC5-Es catalyze the final step of alginate synthesis in brown algae: epimerization of mannuronic acid (M) residues into guluronic acid (G) residues, which occurs directly within the polymannuronate polymer. Brown algae are known to contain a large number of ManC5-E genes (12,41,62,66) and this diversity of enzymes presumably yields polymers with a variety of either random or block-wise distribution of G residues, and thus alginates of distinct rheological properties. These modifications are thought to be regulated in response to stress (66), seasonal variations (67) and developmental processes (41). However, this is a one-way process, as the G residues cannot be modified further once generated, thus locking the alginate in a definitive chemical and rheological state. PL41 activity would be relevant in this context, to further modulate the length and the properties of the alginate chains during cell elongation or development. The large size of the PL41 family in brown algae indicates a variety of lytic activities correlated with alginate patterns. As a consequence of these activities, alginates, mannuronan C5-epimerases and alginate lyases, are likely to determine the porosity and mechanical strength of the wall, and may therefore

be seen as functional analogues to pectins, pectin methylesterases and pectin lyases in plants, respectively (68).

Alginate is also synthesized by bacteria belonging to the *Azotobacter* and *Pseudomonas* genera (*Gammaproteobacteria*) (69,70). The multigenic mannuronan C5-epimerase family has been extensively studied in these organisms, including at the biochemical level (71). These bacteria also produce alginate lyases, exclusively of the PL5 (72,73) and PL7 families (74,75). *Pseudomonas aeruginosa* is an opportunistic human pathogen, and a leading cause of chronic infection in the lungs of immunocompromised individuals. *Azotobacter vinelandii* is a nitrogen-fixing bacterium found in soil. The ability of these two bacteria to synthesize alginate gives them several selective advantages, the main one being the implementation of a protective shield against adverse environmental conditions (75,76). With the exception of PMM (AlgC), all the genes required for alginate biosynthesis are clustered in a single operon, including PL5-alginate lyase (AlgL) (77,78). The products derived from the lyase action are not used as a carbon source by the bacteria, thus several functions have been putatively attributed to these lyases (76,77). In *P. aeruginosa* they contribute to cell viability by clearing the periplasmic space of trapped alginate (76,78). In *A. vinelandii* they trigger coat rupture during cyst germination (74). Additional alginate lyase activities have been described in other PL families, essentially from bacteria (PL6, PL15, PL17, PL18, PL31, PL32, PL36, PL39) and some eukaryotes (PL8, PL14) (12). Screening of public databases indicates uncharacterized PL41 homologues in bacteria, essentially from the *Streptomyces* within the *Actinomycetota* (Figure3A, Supplementary Figure S4) but these organisms are not known to contain an alginate biosynthetic pathway. Indeed, in contrast to the *Azotobacter* and *Pseudomonas* genera, most of these bacteria use the products of alginate lyase action as a carbon source (79).

The occurrence of alginate metabolic pathways in *S. ischiensis* correlates with the immunodetection of alginate in this species (19). There is strong evidence that the ManC5-E, PL41 and GH88 genes have been acquired by *S. ischiensis* via horizontal gene transfer (HGT). We have previously proposed that about 1% of brown algal genes are derived from HGTs, with these genes having been principally acquired from bacterial genomes (18). Around 10% of those HGT-derived genes have predicted functions in carbohydrate transport and metabolism (18).

While previous work indicated that brown algal GMD was derived from a HGT of an ancestral GMD sequence from *Actinobacteria* (11), our analyses based on additional algal sequences indicates that this conclusion is not strongly supported, and other origins of horizontal gene transfer cannot be excluded, in particular from other bacterial groups. Surprisingly, the brown algal ManC5-E and PL41 families seem to derived from distinct HGTs (different donor taxa). Thus, the catalytic steps for the synthesis and the degradation

of alginate, may have arisen independently in *S. ischiensis* and brown algae, although within a narrow time-scale. While a global CAZome expansion is not observed in brown algae as compared to *S. ischiensis*, the PL41 family is the exception to this rule, with a greater number observed in brown algae as compared to its sister taxon (i.e. an average of 22 genes vs. 3). The situation is similar for the ManC5-E family which shows great expansion in brown algae (i.e. an average of 27 genes vs. 11). The acquisition of these enzymatic steps, and the expansion of the gene families in brown algae, probably allowed the generation of extensive ECMs with fine-tuned structures and rigidity. While there is no dynamic correlation between the number of CAZyme families and tissue complexity, these extensive ECMs may have represented an important prerequisite for the evolution of developmental complexity, and for the emergence of large, resilient substrate-anchored multicellular organisms, in the hostile coastal environment.

The fucan biosynthetic pathway in brown algae needs to be fully resolved

The set of genes involved in the synthesis of sulfated glycans remains unknown. Access to the large Phaeoexplorer genome dataset allowed us to identify the most promising candidate genes. Based on the large size of the GT23 family in all brown algal genomes, the specific phylogenetic position of the brown algal members of this family compared to biochemically characterized members, and information gained from modelling of GT23 protein structure, this family represents a strong candidate for fucan synthases in brown algae. This hypothesis is also supported by transcriptomic data obtained during the development of the *Fucus* embryo (80), where the expression of GT23 genes is up-regulated during early fucan deposition into the walls. Elongation of the oligosaccharidic moieties of FCSPs is believed to be followed by sulfation, catalysed by STs. The diversity of the brown algal ST orthologue groups may indicate a variety of acceptors, including many monosaccharides and their available hydroxyl positions found in FCSPs. In contrast to the GTs, we found homologues of some brown algal STs in echinoderms and these enzymes are known to synthesize sulfated fucans of regular patterns (58). While the origin of such homology might find its roots near the origin of eukaryotes (LECA), this would imply that the ability to perform fucan elongation and its subsequent sulfation, to have been lost in all eukaryotes, except brown algae and a narrow number of taxa within animals. The current available data does not identify taxa donors to favour a scenario involving HGT in the two groups of organisms. Functional convergence could also explain part of the similarity, yet this needs to be biochemically explored. STs belonging to orthologue group 10 are interesting candidates for the sulfation of fucans but other groups cannot be excluded. For example, orthogroup 12 contains a gene that is frequently associated with FK-GFFP in a syntenic region, and groups 2 and 8 contain genes that are also up-regulated during early development of the *Fucus* embryo (80). These

latter two groups contain orthologues from *S. ischiensis*. At this stage we cannot conclude that the GT23 family, the STs, the SAs and the fucan biosynthetic pathway were obtained independently in brown algae as compared to echinoderms. Functional validation of all of those genes is needed to progress further on this question.

In addition to brown algae and echinoderms, immunodetection procedures and fucose assays have provided evidence for the presence of fucans/fucoidans in the diatoms *Chaetoceros* and *Thalassiosira weissflogii* (81). The composition of these secreted glycans in diatoms differ markedly from those of brown algal fucans/fucoidans (82). The precise chemical structure of the diatom glycans has not been investigated in detail, the structures reported being essentially limited to glucuronomannans (83). It is therefore possible that the anti-fucans/fucoidans antibodies cross-react with additional, yet to be described, glycan structures in these organisms. The BAM2 antibody, which recognises α-(1→3)-fucans featuring 4-O-sulfate esters based on analysis using synthetic oligofucans (60), detects epitopes in diatoms, but precise structural validation of the compounds detected still needs to be carried out. In addition, glycan content has not been explored in detail in Stramenopiles other than brown algae and diatoms. A few studies have reported analyses of total monosaccharide composition in some of these organisms and these analyses either did not detect fucose (*Tribonema minus,* Xanthophyceae) (84), or only detected fucose at a limited level (*Nannochloropsis oceanica* Eustigmatophyceae) (85). These results suggest that FCSPs are unlikely to be constitutively produced by these organisms at high levels. In short, both the lack of biochemical validations of any putative fucan synthase genes, and the fragmented knowledge of FCSP structures outside brown algae, impairs progression on tracking the evolutionary history of fucan synthesis in eukaryotes.

The situation of the sister taxon *S. ischiensis* is intriguing. Alginate has been detected immunologically in this species (19) and all the genes involved in alginate biosynthesis and degradation in brown algae are conserved. We have also identified GT23 members highly similar to their brown algal orthologues (~48% protein identity). The ability of *S. ischiensis* to produce sulfated fucans needs to be established. The outcome would contribute towards testing the hypothesis that the GT23 family contains fucan synthase genes. If future research establishes the occurrence of fucans in *S. ischiensis*, this would indicate an ECM with a glycan composition very similar to that of brown algae. One substantial difference between the Phaeophyceae and the Schizocladiophyceae is the absence of plasmodesmata in *S. ischiensis* (19,86). Plasmodesmata have been shown to be crucial in cell-to-cell signaling in land plants and may play a role in the evolution of tissue complexity (87-89). Thus, it is not only the building up of a consistent and dynamic glycan-rich ECM in brown algae that may have contributed to the evolution of complex multicellular organisms, but also some structural features that function in close vicinity to ECMs such as plasmodesmata.

Acknowledgements

We are grateful to the Roscoff Bioinformatics platform ABiMS (http://abims.sb-roscoff.fr), part of the Institut Français de Bioinformatique (ANR-11-INBS-0013) and BioGenouest network, for providing computing and storage resources.

Author contributions: Software: L.B.G., E.C.; Formal analysis: F.D., O.G. and J.M.C; Investigation: L.M., N.T., A. L. and C.H; Resources: E.C. , J.M.C. and N.T.; Data Curation: T.B. and N.T.; Writing-Original Draft: L.M., A. B.-C. and C.H.; Writing-Review & Editing: all authors; Visualization: L.M., A.B-C., M.C., A.L. and C.H.; Supervision: C.H.; Project administration: F.D., J.M.C. and C.H.; Funding acquisition: J.M.C. and C.H.; Conceptualization: C.H.

Funding

This work was supported by the ANR project BrownSugar (ANR-20-CE44-0011) and the France Génomique National infrastructure project Phaeoexplorer (ANR-10-INBS-09).

Figure legends

Figure 1. Overview of CAZyme families in brown algae. (**A**) Counts of numbers of genes predicted to encode glycosyltransferases (GTs), glycoside hydrolases (GHs), polysaccharide lyases (PLs) and for all CAZymes (GTs, GHs, PLs, Carbohydrate Esterases CEs, Auxiliary Activities AAs, Carbohydrate Binding Modules CBMs), showing numbers of both full-length proteins (dark colours) and fragments (light colours). The data is averaged by taxonomic order with the number of species analysed per order in brackets. (**B**) Number of genes for selected CAZyme families in brown algae (box plots) and the sister species *S. ischiensis* (red lines). Counts include non-fragmentary proteins and fragments. The species analysed are the same as in Supplementary Table S1. A schematic overview regarding abundance of the corresponding families is indicated below the plots for other groups (the reader is directed to the CAZy database for a more accurate counting in these other groups).

Figure 2. Phylogenetic trees of key enzymes involved in the synthesis of alginates. (**A**) Phylogenetic tree of GDP-mannose 6-dehydrogenases (GMD). The uncompressed tree is shown in Supplementary Figure S2B. (**B**) Phylogenetic trees of mannuronan C5-epimerases (ManC5-E). The circle shown as an inset indicates a global phylogeny of a large collection of ManC5-Es from various genomes, in which 3 main orthologous groups can be found. The phylogeny shown on the right side is representative of this global view, the 3 groups being indicated. The uncompressed tree of the inset is shown in Supplementary Figure S3. (**C**) Extract of an alignment of selected ManC5-E sequences from brown algae and the sister species *S. ischiensis*, the bacterial ManC5-E sequence from *Pseudomonas syringae* (pdbcode: 4NK6). The region shown corresponds to the catalytic site as identified in 4NK6, and the secondary structure assignments shown above the sequences are those determined for 4NK6. The red arrowheads below the alignment indicate the conserved catalytic residues His274 and Asp275. The numbering on the left refers to the phylogenetic clusters indicated in (B).

Figure 3. Phylogenetic trees of key enzymes involved in the degradation of alginates. (**A**) Phylogenetic tree of alginate lyases from the PL41 family. The circle shown as an inset indicates a global phylogeny of the large collection of PL41s from various genomes, in which four main orthologue groups can be found. The phylogeny shown on the right is representative of this global view, the four groups being indicated. The uncompressed tree of the inset is shown in Supplementary Figure S4. (**B**) Phylogenetic tree of glycoside hydrolases from the GH88 family. The uncompressed trees are shown in Supplementary Figure S6.

Figure 4. Syntenic region of putative fucan-related genes in brown algal genomes. This region contains five conserved genes encoding: an Ankyrin repeat-containing protein, a GDP-fucose pyrophosphorylase (GFPP), a L-fucokinase (FK), a sulfotransferase (ST), an alpha/beta hydrolase.

Figure 5. Comparison of brown algal GT10s with biochemically characterised orthologues. (**A**) Phylogenetic tree of the GT10 family. This family is known to contain FucT activities for species outside brown algae. The uncompressed tree is shown in Supplementary Figure S8. (**B**) Extract of an alignment of selected GT10 sequences from brown algae against the bacterial GT10 sequence from *Helicobacter pylori* (pdbcode: 2NZW) focusing on the GDP-fucose donor site. Additional animal sequences are also aligned and secondary structure information is shown for 2NZW. The red arrowheads below the alignment indicate the conserved catalytic residues Tyr246, Glu249 and Lys250, which interact with the donor substrate GDP-fucose. The 2NZW sequence is shown in pink, a *Bos taurus* sequence in blue and the *Ectocarpus* species 7 sequence in green, and the same color code is used in C, D and E. (**C**) Overlay of the *H. pylori* GT10 crystallized structure (2NZW; pink) and the best models generated by AlphaFold2 for the *B. taurus* (blue) and brown algal (green) sequences. The cartoon representation focuses on the GDP-fucose binding site where GDP-fucose is apparent in yellow/orange colors. Some conserved catalytic residues, including those discussed in B), are highlighted. (**D**) Larger view of the overlay presented in B) which allows visualization of the acceptor binding site, in addition to the GDP-fucose binding site. Despite apparent structural conservation, the nature of the residues at the pocket surface differ drastically between the 3 sequences. (**E**) Molecular surfaces of 2NZW and the AlphaFold2 predictions for the *B. taurus* and algal sequences based on their electrostatic potentials. While the substrate binding cleft is clearly apparent in 2NZW and in the *B. taurus* sequence (dotted lines), it appears more enclosed in the brown algal sequence (arrow), although the N-terminal domain may form a flexible loop to allow accommodation of a large substrate.

Figure 6. Phylogenetic tree of the GT41 family. Some GT41 family members in non-brown-algal species are known to have FucT activities (green boxes). The uncompressed tree is shown in Supplementary Figure S9.

Figure 7. Phylogenetic tree of the GT74 family. Some GT74 family members in non-brown-algal species are known to have FucT activities (green box). These proteins frequently occur as bimodular forms with a GT2 domain. The phylogeny is based uniquely on the catalytic GT74 domain. The uncompressed tree is shown in Supplementary Figure S10.

Figure 8. Comparison of brown algal GT23s with characterised orthologues. (**A**) Heatmap representing BLASTP+ Score Ratios (BSRs) for GT23 protein sequence pairwise alignments. The colour scale correlates BSRs. Two groups are apparent: one containing metazoan sequences and the other consisting of brown algal GT23s. (**B**) Phylogenetic tree of GT23s from brown algae and the sister taxa *S. ischiensis*. The uncompressed tree is shown in Supplementary Figure S11. (**C**) Extracts of an alignment of selected GT23 sequences from brown algae and the sister taxa *S. ischiensis*, plus the human FUT8 GT23 sequence (pdbcode: 6X5R), focusing on the GDP-fucose donor site as identified in FUT8. Secondary structure predictions for FUT8 are shown above the sequences. The red arrowheads below the alignment indicate the conserved catalytic residues His363, Arg365, Asp368, Lys369 and Ser469. The blue triangles indicate catalytic residues identified in 6X5R which differ in some of the algal sequences, but may retain similar biochemical properties: Asp453 and Gln470. The FUT8 sequence is shown in blue and the *S. latissima* in green, so as to match the color code used in D and E. (**D**) Overlay of the FUT8 crystal structure (6X5R) and the best model generated by AlphaFold2 of a brown algal sequence from *S. latissima* (Accession number: contig302.9313). The cartoon representation focuses on the GDP-fucose binding site where FUT8 is colored in blue and SlGT23A2D in green. The GDP is apparent in blue/orange colors. The conserved catalytic residues discussed in A) are indicated. (**E**) Global view of the overlay presented in D). The conserved catalytic residues Lys369 and E373 are shown in cyan. Aromatic residues present in the *S. latissima* sequence and possibly involved in acceptor fixation are shown in yellow.

Figure 9. Phylogenetic position of STs from brown algae and other eukaryotes. (**A**) Phylogenetic tree of STs from brown algae and other Stramenopiles. The phylogenetic tree is shown in circle format with collapsed clades. Brown algal sequences are indicated in brown. Sequences belonging to sister groups of brown algae Chrysoparadoxophyceae and Schizocladiaphyceae are indicated by red lines, with red circles for the latter. Other Stramenopile sequences are shown in black. Numbers indicate the brown algal clades as identified in (12). The corresponding uncompressed and fully-labelled tree is shown in Supplementary Figure S12. (**B**) Heatmap representing BLASTP+ Score Ratios (BSRs) for ST protein sequence pairwise alignments. The colour code correlates BSRs. Four groups are apparent: two harbouring brown algal STs (orthologues from group 10), one containing metazoan sequences and the last group containing STs from echinoderms, which is more closely related to brown algal sequences than the metazoan group.

Supplementary Figure legends

Supplementary Figure S1. Phylogenetic trees of key enzymes involved in the synthesis of GDP-mannose, a precursor for the synthesis of both alginates and fucans. (**A**) The bifunctional annose-6-phosphate isomerases (MPI)/mannose-1-phosphate guyanylyltransferases (MPG) and (**B**) the bifunctional phosphomannomutases (PMM)/ phosphoglucomutase (PGM).

Supplementary Figure S2. Clustering and phylogenetic position of brown algal proteins within the GDP-mannose/UDP-glucose dehydrogenases (GMD/UGDs) superfamily. (**A**) Heatmap representing BLASTP+ Score Ratios (BSRs) for protein sequence pairwise alignments. The colour code correlates BSRs. Two main groups of Stramenopile sequences are clearly apparent: one containing the GMD sequences and one harbouring the UGD sequences. The bacterial sequences from *Actinobacteria* and *Pseudomonadota* are distantly related to these two groups. (**B**) Uncompressed version of the phylogenetic tree of GMDs shown in Figure 2A of the main manuscript.

Supplementary Figure S3. Uncompressed version of the phylogenetic tree of ManC5-E proteins shown as an inset in Figure 2B of the main manuscript

Supplementary Figure S4. Uncompressed version of the phylogenetic tree of PL41 proteins shown in Figure 3A of the main manuscript

Supplementary Figure S5. Phylogenetic tree of DEHU reductases, known to be active in alginate degradation in brown algae

Supplementary Figure S6. Uncompressed versions of the phylogenetic tree of GH88 proteins shown in Figure 3B of the main manuscript

Supplementary Figure S7. Phylogenetic trees of key enzymes involved in the synthesis of GDP-fucose. Enzymes involved in the synthesis of GDP-fucose through the *de novo* pathway: (**A**) GDP-mannose 4,6-dehydratases (GM46D) and (**B**) the GDP-L-fucose synthetases (GFS); and through the salvage pathway: (**C**) fucokinase (FK) and (**D**) GDP-fucose pyrophosphorylase (GFPP).

Supplementary Figure S8. Uncompressed version of the phylogenetic tree of GT10 proteins shown in Figure 5A of the main manuscript

Supplementary Figure S9. Uncompressed version of the phylogenetic tree of GT41 proteins shown in Figure 6 of the main manuscript

Supplementary Figure S10. Uncompressed version of the phylogenetic tree of GT74 proteins shown in Figure 7 of the main manuscript

Supplementary Figure S11. Uncompressed version of the phylogenetic tree of GT23 proteins shown in Figure 8B of the main manuscript

Supplementary Figure S12. Uncompressed version of the phylogenetic tree of ST proteins shown in Figure 9A of the main manuscript.

Supplementary Table 1. Gene number of a selection of GT families in brown algae and other relevant organisms and known to contain FucT activities in the CAZy database.

All known FucTs are classified in the CAZy GT families as shown. The generated glycosidic linkages identified from characterized enzymes (outside brown algae) are indicated. The average value is shown for the following species: *Ectocarpus* sp.7, *Pylaiella littoralis, Chordaria linearis* (Ectocarpales), *Saccharina latissima, Saccharina japonica, Macrocystis pyrifera* (Laminariales), *Porterinema fluviatiles* (Ralfsiales), *Ascophyllum nodosum, Fucus serratus, Fucus distichus* (Fucales), *Saccorhiza dermatodea, Saccorhiza polyschides* (Tilopteridiales), *Desmarestia herbacea*, *Desmarestia dudresnayi* (Desmarestiales), *Sphacelaria rigidula* (Sphacelariales), *Dictyota dichotoma* (Dictyotales), *Discosporangium mesarthrocarpum, Choristocarpus tenellus* (Discosporangiales), *Schizocladia ischiensis* (Schizocladiophyceae), *Heterosigma akashiwo* (Raphidophyceae), *Phaeodactylum tricornutum CCAP 1055, Thalassiosira pseudonana* CCMP 1335*, Chaetoceros tenuissimus* NIES-3715_426638 (diatoms), *Lytechinus variegatus* NC3_7654, *Strongylocentrotus purpuratus* Spur-01_7668 (sea urchin), *Apostichopus japon* Shaxun_307972 (sea cucumber). ND. Not determined.

GT families	**GT10**	**GT23**	**GT41**	**GT74**	**GT11**	**GT37**	**GT65**	**GT68**
Known FucT activities in CAZy	α-1,3/4-FucTs	α-1,6-FucTs	protein O-FucTs	α-1,2-FucTs	α-1,2/3-FucTs	α-1,2-FucTs	protein O-FucTs	protein O-FucTs
Phaeophyceae								
Ectocarpales [3]	**1**	**9**	**13**	**1**	0	0	0	0
Laminariales [3]	**1**	**9**	**10**	**1**	0	0	0	0
Ralfsiales [1]	**3**	**3**	**12**	**1**	0	0	0	0
Fucales [3]	**1**	**14**	**10**	**1**	0	0	0	0
Tilopteridiales [2]	**0**	**5**	**10**	**2**	0	1	0	0
Desmarestiales [2]	**1**	**11**	**16**	**1**	0	1	0	0
Sphacelariales [1]	**0**	**1**	**6**	**1**	0	0	0	0
Dictyotales [1]	**1**	**4**	**10**	**1**	0	1	0	0
Discosporangiales [2]	**1**	**1**	**5**	**1**	0	1	0	0
Sister groups to Phaeophyceae								
Schizocladiophyceae [1]	**1**	**5**	**11**	**1**	0	1	0	0
Raphidophyceae [1]	**0**	**0**	**8**	**0**	0	0	0	0
Other Stramenopiles								
diatoms [3]	**3**	**0**	**2**	**1**	ND	ND	ND	ND
Echinoderms								
Sea urchin [2]	**60**	**7**	**4**	**0**	ND	ND	ND	ND
Sea cucumber [1]	**97**	**1**	**1**	**0**	ND	ND	ND	ND

References

1. Bringloe, T.T., Starko, S., Wade, R.M., Vieira, C., Kawai, H., De Clerck, O., Cock, J.M., Coelho, S.M., Destombe, C., Valero, M. *et al.* (2020) Phylogeny and Evolution of the Brown Algae. *Critical Reviews in Plant Sciences*, **39**, 281-321.
2. Farghali, M., Mohamed, I.M.A., Osman, A.I. and Rooney, D.W. (2022) Seaweed for climate mitigation, wastewater treatment, bioenergy, bioplastic, biochar, food, pharmaceuticals, and cosmetics: a review. *Environmental Chemistry Letters*.
3. Pessarrodona, A., Assis, J., Filbee-Dexter, K., Burrows, M.T., Gattuso, J.-P., Duarte, C.M., Krause-Jensen, D., Moore, P.J., Smale, D.A. and Wernberg, T. (2022) Global seaweed productivity. *Science Advances*, **8**, eabn2465.
4. Filbee-Dexter, K., Feehan, C.J., Smale, D.A., Krumhansl, K.A., Augustine, S., de Bettignies, F., Burrows, M.T., Byrnes, J.E.K., Campbell, J., Davoult, D. *et al.* (2022) Kelp carbon sink potential decreases with warming due to accelerating decomposition. *PLoS Biol*, **20**, e3001702.
5. Arriaga, O., Wawrzynkowski, P., Ibáñez, H., Muguerza, N., Díez, I., Pérez-Ruzafa, I., Gorostiaga, J.M., Quintano, E. and Becerro, M.A. (2023) Short-term response of macroalgal communities to ocean warming in the Southern Bay of Biscay. *Mar Environ Res*, **190**, 106098.
6. Popper, Z.A., Michel, G., Hervé, C., Domozych, D.S., Willats, W.G.T., Tuohy, M.G., Kloareg, B. and Stengel, D.B. (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. *Annual Review of Plant Biology*, **62**, 567-590.
7. Kloareg, B., Badis, Y., Cock, J.M. and Michel, G. (2021) Role and evolution of the extracellular matrix in the acquisition of complex multicellularity in eukaryotes: a macroalgal perspective. *Genes (Basel)*, **12**.
8. Blaby-Haas, C.E. and Merchant, S.S. (2019) Comparative and Functional Algal Genomics. *Annual Review of Plant Biology*, **70**, 605-638.
9. Deniaud-Bouët, E., Hardouin, K., Potin, P., Kloareg, B. and Hervé, C. (2017) A review about brown algal cell walls and fucose-containing sulfated polysaccharides: Cell wall context, biomedical properties and key research challenges. *Carbohydrate polymers*, **175**, 395-408.
10. Deniaud-Bouët, E., Kervarec, N., Michel, G., Tonon, T., Kloareg, B. and Hervé, C. (2014) Chemical and enzymatic fractionation of cell walls from Fucales: insights into the structure of the extracellular matrix of brown algae. *Annals of Botany*, **114**, 1203-1216.
11. Michel, G., Tonon, T., Scornet, D., Cock, J.M. and Kloareg, B. (2010) The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes. *New Phytol*, **188**, 82-97.
12. Mazéas, L., Yonamine, R., Barbeyron, T., Henrissat, B., Drula, E., Terrapon, N., Nagasato, C. and Hervé, C. (2023) Assembly and synthesis of the extracellular matrix in brown algae. *Seminars in Cell & Developmental Biology*, **134**, 112-124.
13. Chi, S., Liu, T., Wang, X., Wang, R., Wang, S., Wang, G., Shan, G. and Liu, C. (2018) Functional genomics analysis reveals the biosynthesis pathways of important cellular components (alginate and fucoidan) of *Saccharina*. *Current Genetics*, **64**, 259-273.
14. Drula, E., Garron, M.L., Dogan, S., Lombard, V., Henrissat, B. and Terrapon, N. (2022) The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res*, **50**, D571-d577.
15. Carpita, N.C. and McCann, M.C. (2000), *Biochemistry and Molecular Biology of Plants*. B. Buchanan, W. Gruissem and R. Jones, eds. American Society of Plant Physiologists, Rockville, Md.
16. Stam, M., Lelièvre, P., Hoebeke, M., Corre, E., Barbeyron, T. and Michel, G. (2023) SulfAtlas, the sulfatase database: state of the art and new developments. *Nucleic Acids Res*, **51**, D647-d653.
17. Barbeyron, T., Brillet-Guéguen, L., Carré, W., Carrière, C., Caron, C., Czjzek, M., Hoebeke, M. and Michel, G. (2016) Matching the diversity of sulfated biomolecules: creation of a

classification database for sulfatases reflecting their substrate specificity. *PloS one*, **11**, e0164846-e0164846.

18. Denoeud, F., Godfroy, O., Cruaud, C., Heesch, S., Nehr, Z., Tadrent, N., Couloux, A., Brillet-Guéguen, L., Delage, L., Mckeown, D. *et al.* (2024) Evolutionary genomics of the emergence of brown algae as key components of coastal ecosystems. *bioRxiv*, 2024.2002.2019.579948.
19. Kawai, H., Maeba, S., Sasaki, H., Okuda, K. and Henry, E.C. (2003) Schizocladia ischiensis: A New Filamentous Marine Chromophyte Belonging to a New Class, Schizocladiophyceae. *Protist*, **154**, 211-228.
20. Eddy, S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol*, **7**, e1002195.
21. Katoh, K., Rozewicki, J. and Yamada, K.D. (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, **20**, 1160-1166.
22. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189-1191.
23. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. and Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**, 587-589.
24. Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312-1313.
25. Tamura, K., Stecher, G. and Kumar, S. (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, **38**, 3022-3027.
26. Hobbs, E.E.M., Gloster, T.M. and Pritchard, L. (2023) cazy_webscraper: local compilation and interrogation of comprehensive CAZyme datasets. *Microb Genom*, **9**.
27. Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, **42**, W320-W324.
28. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583-589.
29. Tang, H., Bowers, J.E., Wang, X., Ming, R., Alam, M. and Paterson, A.H. (2008) Synteny and collinearity in plant genomes. *Science*, **320**, 486-488.
30. Harris, R. (2007) Improved Pairwise Alignment of Genomic DNA. *Pennsylvania State University* **PhD Thesis**.
31. Geshi, N., Johansen, J.N., Dilokpimol, A., Rolland, A., Belcram, K., Verger, S., Kotake, T., Tsumuraya, Y., Kaneko, S., Tryfona, T. *et al.* (2013) A galactosyltransferase acting on arabinogalactan protein glycans is essential for embryo development in Arabidopsis. *The Plant Journal*, **76**, 128-137.
32. Zhang, L., Prabhakar, P.K., Bharadwaj, V.S., Bomble, Y.J., Peña, M.J. and Urbanowicz, B.R. (2023) Glycosyltransferase family 47 (GT47) proteins in plants and animals. *Essays Biochem*, **67**, 639-652.
33. Boruah, B.M., Kadirvelraj, R., Liu, L., Ramiah, A., Li, C., Zong, G., Bosman, G.P., Yang, J.Y., Wang, L.X., Boons, G.J. *et al.* (2020) Characterizing human α-1,6-fucosyltransferase (FUT8) substrate specificity and structural similarities with related fucosyltransferases. *J Biol Chem*, **295**, 17027-17045.
34. García-García, A., Ceballos-Laita, L., Serna, S., Artschwager, R., Reichardt, N.C., Corzana, F. and Hurtado-Guerrero, R. (2020) Structural basis for substrate specificity and catalysis of α1,6-fucosyltransferase. *Nat Commun*, **11**, 973.
35. Bamford, N.C., Le Mauff, F., Subramanian, A.S., Yip, P., Millán, C., Zhang, Y., Zacharias, C., Forman, A., Nitz, M., Codée, J.D.C. *et al.* (2019) Ega3 from the fungal pathogen Aspergillus fumigatus is an endo-α-1,4-galactosaminidase that disrupts microbial biofilms. *J Biol Chem*, **294**, 13833-13849.

36. Tamura, J.-I., Hasegawa, K., Kadowaki, K., Igarashi, Y. and Kodama, T. (1995) Molecular cloning and sequence analysis of the gene encoding an endo α-1,4 polygalactosaminidase of Pseudomonas sp. 881. *Journal of Fermentation and Bioengineering*, **80**, 305-310.
37. Inoue, A. and Ojima, T. (2019) Functional identification of alginate lyase from the brown alga *Saccharina japonica*. *Scientific Reports*, **9**, 4937.
38. Zhang, P., Shao, Z., Li, L., Liu, S., Yao, J. and Duan, D. (2018) Molecular characterisation and biochemical properties of phosphomannomutase/phosphoglucomutase (PMM/PGM) in the brown seaweed *Saccharina japonica*. *Journal of Applied Phycology*, **30**, 2687-2687-2696.
39. Tenhaken, R., Voglas, E., Cock, J.M., Neu, V. and Huber, C.G. (2011) Characterization of GDP-mannose dehydrogenase from the brown alga *Ectocarpus siliculosus* providing the precursor for the alginate polymer. *Journal of Biological Chemistry*, **286**, 16707-16715.
40. Zhang, P., Shao, Z., Jin, W. and Duan, D. (2016) Comparative characterization of two GDP-mannose dehydrogenase genes from *Saccharina japonica* (Laminariales, Phaeophyceae). *BMC Plant Biology*, **16**, 62.
41. Fischl, R., Bertelsen, K., Gaillard, F., Coelho, S., Michel, G., Klinger, M., Boyen, C., Czjzek, M. and Hervé, C. (2016) The cell-wall active mannuronan C5-epimerases in the model brown alga *Ectocarpus*: From gene context to recombinant protein. *Glycobiology*, **26**, 973-983.
42. Inoue, A., Satoh, A., Morishita, M., Tokunaga, Y., Miyakawa, T., Tanokura, M. and Ojima, T. (2016) Functional heterologous expression and characterization of mannuronan C5-epimerase from the brown alga *Saccharina japonica*. *Algal Research*, **16**, 282-291.
43. Murray, R.G. and Schleifer, K.H. (1994) Taxonomic notes: a proposal for recording the properties of putative taxa of procaryotes. *Int J Syst Bacteriol*, **44**, 174-176.
44. Inoue, A. and Ojima, T. (2021) Functional identification of the 4-deoxy-L-erythro-5-hexoseulose uronate reductase from a brown alga, *Saccharina japonica*. *Biochem Biophys Res Commun*, **545**, 112-118.
45. Cartmell, A., Lowe, E.C., Baslé, A., Firbank, S.J., Ndeh, D.A., Murray, H., Terrapon, N., Lombard, V., Henrissat, B., Turnbull, J.E. *et al.* (2017) How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans. *Proc Natl Acad Sci U S A*, **114**, 7037-7042.
46. Hehemann, J.-H., Boraston, A.B. and Czjzek, M. (2014) A sweet new wave: structures and mechanisms of enzymes that digest polysaccharides from marine algae. *Current Opinion in Structural Biology*, **28**, 77-86.
47. Nishitsuji, K., Arimoto, A., Higa, Y., Mekaru, M., Kawamitsu, M., Satoh, N. and Shoguchi, E. (2019) Draft genome of the brown alga, *Nemacystus decipiens*, Onna-1 strain: Fusion of genes involved in the sulfated fucan biosynthesis pathway. *Scientific Reports*, **9**, 4607-4607.
48. Both, P., Sobczak, L., Breton, C., Hann, S., Nöbauer, K., Paschinger, K., Kozmon, S., Mucha, J. and Wilson, I.B.H. (2011) Distantly related plant and nematode core α1,3-fucosyltransferases display similar trends in structure–function relationships. *Glycobiology*, **21**, 1401-1415.
49. Brito, C., Kandzia, S., Graça, T., Conradt, H.S. and Costa, J. (2008) Human fucosyltransferase IX: specificity towards N-linked glycoproteins and relevance of the cytoplasmic domain in intra-Golgi localization. *Biochimie*, **90**, 1279-1290.
50. Ma, B., Simala-Grant, J.L. and Taylor, D.E. (2006) Fucosylation in prokaryotes and eukaryotes. *Glycobiology*, **16**, 158r-184r.
51. Sun, H.Y., Lin, S.W., Ko, T.P., Pan, J.F., Liu, C.L., Lin, C.N., Wang, A.H. and Lin, C.H. (2007) Structure and mechanism of Helicobacter pylori fucosyltransferase. A basis for lipopolysaccharide variation and inhibitor design. *J Biol Chem*, **282**, 9973-9982.
52. Breton, C., Oriol, R. and Imberty, A. (1998) Conserved structural features in eukaryotic and prokaryotic fucosyltransferases. *Glycobiology*, **8**, 87-94.
53. Oulmouden, A., Wierinckx, A., Petit, J.-M., Costache, M., Palcic, M.M., Mollicone, R., Oriol, R. and Julien, R. (1997) Molecular Cloning and Expression of a Bovine α(1,3)-Fucosyltransferase Gene Homologous to a Putative Ancestor Gene of the Human FUT3-FUT5-FUT6 Cluster*. *Journal of Biological Chemistry*, **272**, 8764-8773.

54. Kumar, S., Wang, Y., Zhou, Y., Dillard, L., Li, F.-W., Sciandra, C.A., Sui, N., Zentella, R., Zahn, E., Shabanowitz, J. *et al.* (2023) Structure and dynamics of the Arabidopsis O-fucosyltransferase SPINDLY. *Nature communications*, 10.1038/s41467-023-37279-1.
55. van Der Wel, H., Morris, H.R., Panico, M., Paxton, T., North, S.J., Dell, A., Thomson, J.M. and West, C.M. (2001) A non-Golgi alpha 1,2-fucosyltransferase that modifies Skp1 in the cytoplasm of Dictyostelium. *J Biol Chem*, **276**, 33952-33963.
56. Yanagidani, S., Uozumi, N., Ihara, Y., Miyoshi, E., Yamaguchi, N. and Taniguchi, N. (1997) Purification and cDNA cloning of GDP-L-Fuc:N-acetyl-beta-D-glucosaminide:alpha1-6 fucosyltransferase (alpha1-6 FucT) from human gastric cancer MKN45 cells. *J Biochem*, **121**, 626-632.
57. Tomida, S., Takata, M., Hirata, T., Nagae, M., Nakano, M. and Kizuka, Y. (2020) The SH3 domain in the fucosyltransferase FUT8 controls FUT8 activity and localization and is essential for core fucosylation. *J Biol Chem*, **295**, 7992-8004.
58. Pomin, V.H. and Mourão, P.A.S. (2008) Structure, biology, evolution, and medical importance of sulfated fucans and galactans. *Glycobiology*, **18**, 1016-1027.
59. Pereira, M.S., Mulloy, B. and Mourão, P.A.S. (1999) Structure and anticoagulant activity of sulfated fucans: comparison between the regular, repetitive, and linear fucans from echinoderms with the more heterogeneous and branched polymers from brown algae *Journal of Biological Chemistry*, **274**, 7656-7667.
60. Crawford, C., Schultz-Johansen, M., Luong, P., Vidal-Melgosa, S., Hehemann, J.-H. and Seeberger, P. (2024) Automated synthesis of fucoidan enables molecular investigations in marine glycobiology. *ChemRxiv*.
61. Liu, T., Wang, X., Wang, G., Jia, S., Liu, G., Shan, G., Chi, S., Zhang, J., Yu, Y., Xue, T. *et al.* (2019) Evolution of complex thallus alga: genome sequencing of *Saccharina japonica*. *Front Genet*, **10**, 378.
62. Ye, N., Zhang, X., Miao, M., Fan, X., Zheng, Y., Xu, D., Wang, J., Zhou, L., Wang, D., Gao, Y. *et al.* (2015) *Saccharina* genomes provide novel insight into kelp biology. *Nat Commun*, **6**.
63. Salmeán, A.A., Duffieux, D., Harholt, J., Qin, F., Michel, G., Czjzek, M., Willats, W.G.T. and Hervé, C. (2017) Insoluble (1→3), (1→4)-β-D-glucan is a component of cell walls in brown algae (Phaeophyceae) and is masked by alginates in tissues. *Scientific Reports*, **7**, 2880.
64. Kolbinger, F., Streiff, M.B. and Katopodis, A.G. (1998) Cloning of a human UDP-galactose:2-acetamido-2-deoxy-D-glucose 3beta-galactosyltransferase catalyzing the formation of type 1 chains. *J Biol Chem*, **273**, 433-440.
65. Hervé, C., Siméon, A., Jam, M., Cassin, A., Johnson, K.L., Salmeán, A.A., Willats, W.G.T., Doblin, M.S., Bacic, A. and Kloareg, B. (2016) Arabinogalactan proteins have deep roots in eukaryotes: identification of genes and epitopes in brown algae and their role in *Fucus serratus* embryo development. *New Phytologist*, **209**, 1428-1441.
66. Tonon, T., Rousvoal, S., Roeder, V. and Boyen, C. (2008) Expression profiling of the mannuronan C5-epimerase multigenic family in the brown alga *Laminaria digitata* (Phaeophyceae) under biotic stress conditions. *Journal of Phycology*, **44**, 1250-1256.
67. Nyvall, P., Corre, E., Boisset, C., Barbeyron, T., Rousvoal, S., Scornet, D., Kloareg, B. and Boyen, C. (2003) Characterization of mannuronan C-5-epimerase genes from the brown alga *Laminaria digitata*. *Plant Physiol*, **133**, 726-735.
68. Sénéchal, F., Wattier, C., Rustérucci, C. and Pelloux, J. (2014) Homogalacturonan-modifying enzymes: structure, expression, and roles in plants. *J Exp Bot*, **65**, 5125-5160.
69. Gawin, A., Tietze, L., Aarstad, O.A., Aachmann, F.L., Brautaset, T. and Ertesvåg, H. (2020) Functional characterization of three Azotobacter chroococcum alginate-modifying enzymes related to the Azotobacter vinelandii AlgE mannuronan C-5-epimerase family. *Scientific Reports*, **10**, 12470.
70. Linker, A. and Jones, R.S. (1966) A new polysaccharide resembling alginic acid isolated from pseudomonads. *J Biol Chem*, **241**, 3845-3851.

71. Bjerkan, T.M., Lillehov, B.E., Strand, W.I., Skjak-Braek, G., Valla, S. and Ertesvag, H. (2004) Construction and analyses of hybrid *Azotobacter vinelandii* mannuronan C-5 epimerases with new epimerization pattern characteristics. *Biochem J*, **381**, 813-821.
72. Ertesvåg, H., Erlien, F., Skjåk-Braek, G., Rehm, B.H. and Valla, S. (1998) Biochemical properties and substrate specificities of a recombinantly produced Azotobacter vinelandii alginate lyase. *J Bacteriol*, **180**, 3779-3784.
73. Schiller, N.L., Monday, S.R., Boyd, C.M., Keen, N.T. and Ohman, D.E. (1993) Characterization of the Pseudomonas aeruginosa alginate lyase gene (algL): cloning, sequencing, and expression in Escherichia coli. *J Bacteriol*, **175**, 4780-4789.
74. Gimmestad, M., Ertesvåg, H., Heggeset, T.M., Aarstad, O., Svanem, B.I. and Valla, S. (2009) Characterization of three new Azotobacter vinelandii alginate lyases, one of which is involved in cyst germination. *J Bacteriol*, **191**, 4845-4853.
75. Yamasaki, M., Moriwaki, S., Miyake, O., Hashimoto, W., Murata, K. and Mikami, B. (2004) Structure and Function of a Hypothetical Pseudomonas aeruginosa Protein PA1167 Classified into Family PL-7: A NOVEL ALGINATE LYASE WITH A β-SANDWICH FOLD*. *Journal of Biological Chemistry*, **279**, 31863-31872.
76. Jain, S. and Ohman, D.E. (2005) Role of an alginate lyase for alginate transport in mucoid Pseudomonas aeruginosa. *Infect Immun*, **73**, 6429-6436.
77. May, T.B. and Chakrabarty, A.M. (1994) Pseudomonas aeruginosa: genes and enzymes of alginate synthesis. *Trends in Microbiology*, **2**, 151-157.
78. Gheorghita, A.A., Wolfram, F., Whitfield, G.B., Jacobs, H.M., Pfoh, R., Wong, S.S.Y., Guitor, A.K., Goodyear, M.C., Berezuk, A.M., Khursigara, C.M. *et al.* (2022) The Pseudomonas aeruginosa homeostasis enzyme AlgL clears the periplasmic space of accumulated alginate during polymer biosynthesis. *Journal of Biological Chemistry*, **298**, 101560.
79. Thomas, F., Barbeyron, T., Tonon, T., Génicot, S., Czjzek, M. and Michel, G. (2012) Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine Flavobacteriia to their independent transfers to marine Proteobacteria and human gut Bacteroides. *Environmental Microbiology*, **14**, 2379-2394.
80. Linardić, M., Cokus, S.J., Pellegrini, M. and Braybrook, S.A. (2020) Growth of the *Fucus* embryo: insights into wall-mediated cell expansion through mechanics and transcriptomics. *bioRxiv*, 2020.2001.2029.925107.
81. Vidal-Melgosa, S., Sichert, A., Francis, T.B., Bartosik, D., Niggemann, J., Wichels, A., Willats, W.G.T., Fuchs, B.M., Teeling, H., Becher, D. *et al.* (2021) Diatom fucan polysaccharide precipitates carbon during algal blooms. *Nature Communications*, **12**, 1150.
82. Gugi, B., Le Costaouec, T., Burel, C., Lerouge, P., Helbert, W. and Bardor, M. (2015) Diatom-Specific Oligosaccharide and Polysaccharide Structures Help to Unravel Biosynthetic Capabilities in Diatoms. *Mar Drugs*, **13**, 5993-6018.
83. Le Costaouëc, T., Unamunzaga, C., Mantecon, L. and Helbert, W. (2017) New structural insights into the cell-wall polysaccharide of the diatom Phaeodactylum tricornutum. *Algal Research*, **26**, 172-179.
84. Huo, S., Wang, H., Chen, J., Hu, X., Zan, X., Zhang, C., Qian, J., Zhu, F., Ma, H. and Elshobary, M. (2022) A preliminary study on polysaccharide extraction, purification, and antioxidant properties of sugar-rich filamentous microalgae Tribonema minus. *Journal of Applied Phycology*, **34**, 2755-2767.
85. Bernaerts, T.M.M., Gheysen, L., Kyomugasho, C., Jamsazzadeh Kermani, Z., Vandionant, S., Foubert, I., Hendrickx, M.E. and Van Loey, A.M. (2018) Comparison of microalgal biomasses as functional food ingredients: Focus on the composition of cell wall related polysaccharides. *Algal Research*, **32**, 150-161.
86. Terauchi, M., Nagasato, C., Kajimura, N., Mineyuki, Y., Okuda, K., Katsaros, C. and Motomura, T. (2012) Ultrastructural study of plasmodesmata in the brown alga Dictyota dichotoma (Dictyotales, Phaeophyceae). *Planta*, **236**, 1013-1026.
87. Sager, R.E. and Lee, J.Y. (2018) Plasmodesmata at a glance. *J Cell Sci*, **131**.

88. German, L., Yeshvekar, R. and Benitez-Alfonso, Y. (2023) Callose metabolism and the regulation of cell walls and plasmodesmata during plant mutualistic and pathogenic interactions. *Plant Cell Environ*, **46**, 391-404.
89. Cook, M.E., Graham, L.E., Botha, C.E.J. and Lavin, C.A. (1997) Comparative ultrastructure of plasmodesmata of Chara and selected bryophytes: toward an elucidation of the evolutionary origin of plant plasmodesmata. *American Journal of Botany*, **84**, 1169-1178.

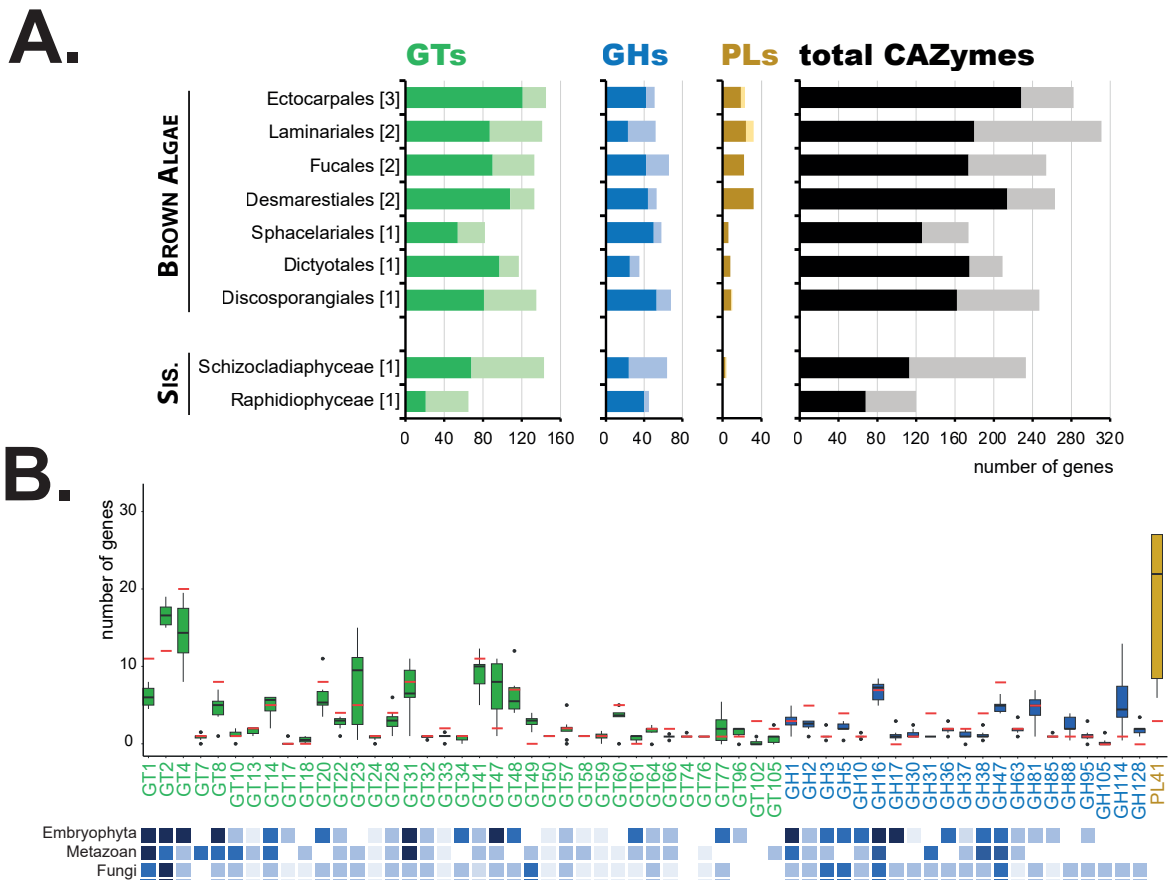

A.
GTs
GHs
PLs
total CAZymes
Brown Algae
Ectocarpales [3]
Laminariales [2]
Fucales [2]
Desmarestiales [2]
Sphacelariales [1]
Dictyotales [1]
Discosporangiales [1]
Sis.
Schizocladiaphyceae [1]
Raphidiophyceae [1]
0 40 80 120 160
0 40 80
0 40
0 40 80 120 160 200 240 280 320
number of genes
B.
number of genes
0 10 20 30
GT1 GT2 GT4 GT7 GT8 GT10 GT13 GT14 GT17 GT18 GT20 GT22 GT23 GT24 GT28 GT31 GT32 GT33 GT34 GT41 GT47 GT48 GT49 GT50 GT57 GT58 GT59 GT60 GT61 GT64 GT66 GT74 GT76 GT77 GT96 GT102 GT105 GH1 GH2 GH3 GH5 GH10 GH16 GH17 GH30 GH31 GH36 GH37 GH38 GH47 GH63 GH81 GH85 GH88 GH95 GH105 GH114 GH128 PL41
Embryophyta
Metazoan
Fungi

Figure 1

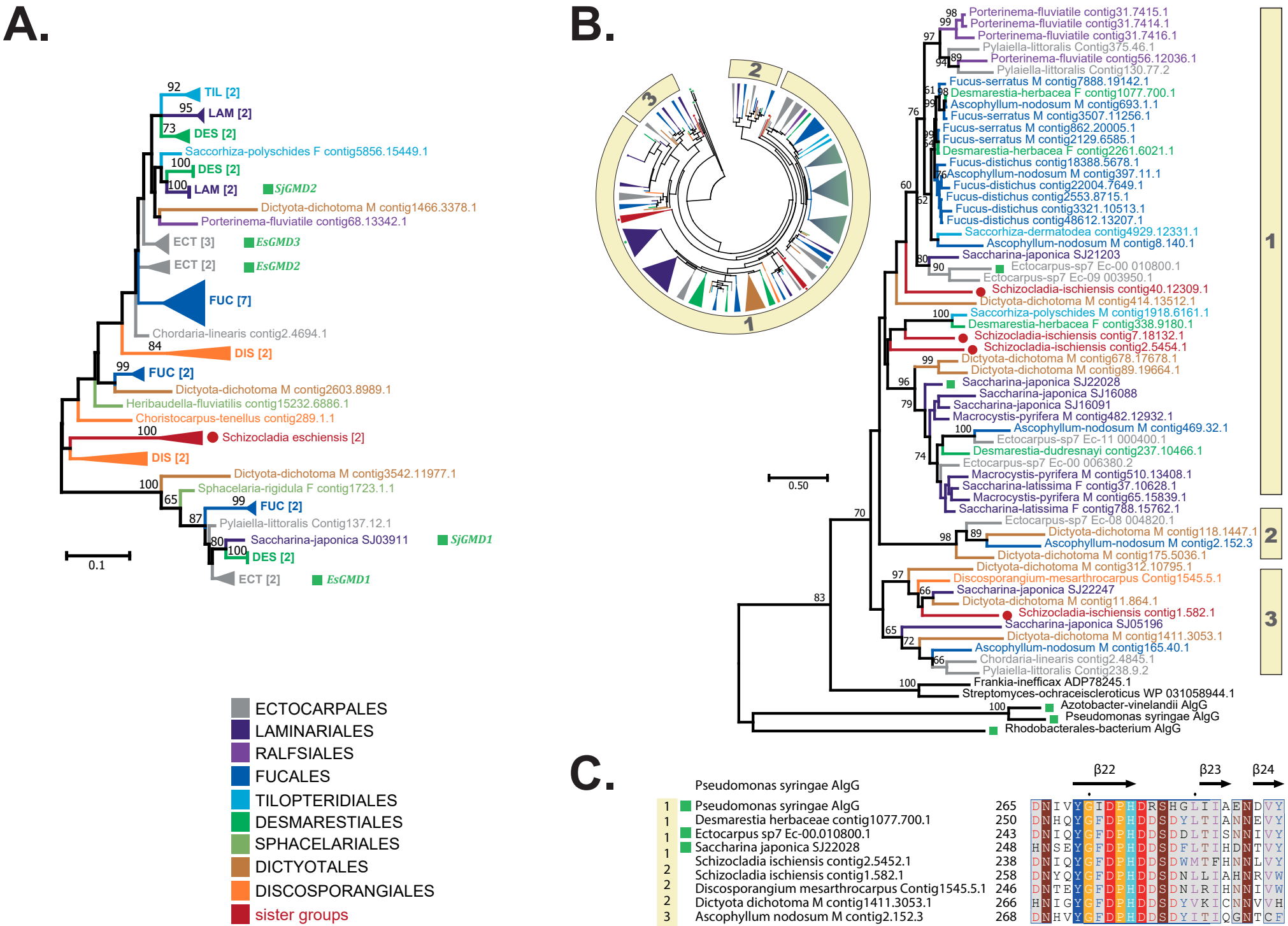

A.
TIL [2]
LAM [2]
DES [2]
Saccorhiza-polyschides F contig5856.15449.1
DES [2]
LAM [2]
SjGMD2
Dictyota-dichotoma M contig1466.3378.1
Porterinema-fluviatile contig68.13342.1
ECT [3]
EsGMD3
ECT [2]
EsGMD2
FUC [7]
Chordaria-linearis contig2.4694.1
DIS [2]
FUC [2]
Dictyota-dichotoma M contig2603.8989.1
Heribaudella-fluviatilis contig15232.6886.1
Choristocarpus-tenellus contig289.1.1
Schizocladia eschiensis [2]
DIS [2]
Dictyota-dichotoma M contig3542.11977.1
Sphacelaria-rigidula F contig1723.1.1
FUC [2]
Pylaiella-littoralis Contig137.12.1
Saccharina-japonica SJ03911
SjGMD1
DES [2]
ECT [2]
EsGMD1
0.1
ECTOCARPALES
LAMINARIALES
RALFSIALES
FUCALES
TILOPTERIDIALES
DESMARESTIALES
SPHACELARIALES
DICTYOTALES
DISCOSPORANGIALES
sister groups
B.
Porterinema-fluviatile contig31.7415.1
Porterinema-fluviatile contig31.7414.1
Porterinema-fluviatile contig31.7416.1
Pylaiella-littoralis Contig375.46.1
Porterinema-fluviatile contig56.12036.1
Pylaiella-littoralis Contig130.77.2
Fucus-serratus M contig7888.19142.1
Desmarestia-herbacea F contig1077.700.1
Ascophyllum-nodosum M contig693.1.1
Fucus-serratus M contig3507.11256.1
Fucus-serratus M contig862.20005.1
Fucus-serratus M contig2129.6585.1
Desmarestia-herbacea F contig2261.6021.1
Fucus-distichus contig18388.5678.1
Ascophyllum-nodosum M contig397.11.1
Fucus-distichus contig22004.7649.1
Fucus-distichus contig2553.8715.1
Fucus-distichus contig3321.10513.1
Fucus-distichus contig48612.13207.1
Saccorhiza-dermatodea contig4929.12331.1
Ascophyllum-nodosum M contig8.140.1
Saccharina-japonica SJ21203
Ectocarpus-sp7 Ec-00 010800.1
Ectocarpus-sp7 Ec-09 003950.1
Schizocladia-ischiensis contig40.12309.1
Dictyota-dichotoma M contig414.13512.1
Saccorhiza-polyschides M contig1918.6161.1
Desmarestia-herbacea F contig338.9180.1
Schizocladia-ischiensis contig7.18132.1
Schizocladia-ischiensis contig2.5454.1
Dictyota-dichotoma M contig678.17678.1
Dictyota-dichotoma M contig89.19664.1
Saccharina-japonica SJ22028
Saccharina-japonica SJ16088
Saccharina-japonica SJ16091
Macrocystis-pyrifera M contig482.12932.1
Ascophyllum-nodosum M contig469.32.1
Ectocarpus-sp7 Ec-11 000400.1
Desmarestia-dudresnayi contig237.10466.1
Ectocarpus-sp7 Ec-00 006380.2
Macrocystis-pyrifera M contig510.13408.1
Saccharina-latissima F contig37.10628.1
Macrocystis-pyrifera M contig65.15839.1
Saccharina-latissima F contig788.15762.1
Ectocarpus-sp7 Ec-08 004820.1
Dictyota-dichotoma M contig118.1447.1
Ascophyllum-nodosum M contig2.152.3
Dictyota-dichotoma M contig175.5036.1
Dictyota-dichotoma M contig312.10795.1
Discosporangium-mesarthrocarpus Contig1545.5.1
Saccharina-japonica SJ22247
Dictyota-dichotoma M contig11.864.1
Schizocladia-ischiensis contig1.582.1
Saccharina-japonica SJ05196
Dictyota-dichotoma M contig1411.3053.1
Ascophyllum-nodosum M contig165.40.1
Chordaria-linearis contig2.4845.1
Pylaiella-littoralis Contig238.9.2
Frankia-inefficax ADP78245.1
Streptomyces-ochraceiscleroticus WP 031058944.1
Azotobacter-vinelandii AlgG
Pseudomonas syringae AlgG
Rhodobacterales-bacterium AlgG
0.50
1
2
3
C.
β22
β23
β24
Pseudomonas syringae AlgG
1 Pseudomonas syringae AlgG 265 DNIVYGIDPHDRSHGLIIAENDVY
1 Desmarestia herbaceae contig1077.700.1 250 DNHQYGFDPHDDSDYLTIANNEVY
1 Ectocarpus sp7 Ec-00.010800.1 243 DNIQYGFDPHDDSDDLTISNNIVY
1 Saccharina japonica SJ22028 248 HNSEYGFDPHDDSDFLTIHDNTVY
Schizocladia ischiensis contig2.5452.1 238 DNIQYGFDPHDDSDWMTFHNNLVY
2 Schizocladia ischiensis contig1.582.1 258 DNYQYGFDPHDDSDNLLIAHNRVW
2 Discosporangium mesarthrocarpus Contig1545.5.1 246 DNTEYGFDPHDDSDNLRIHNNIVW
2 Dictyota dichotoma M contig1411.3053.1 266 HNIGYGFDPHDDSDYVKICNNVVH
3 Ascophyllum nodosum M contig2.152.3 268 DNHVYGFDPHDDSDYITIQGNTCF

Figure 2

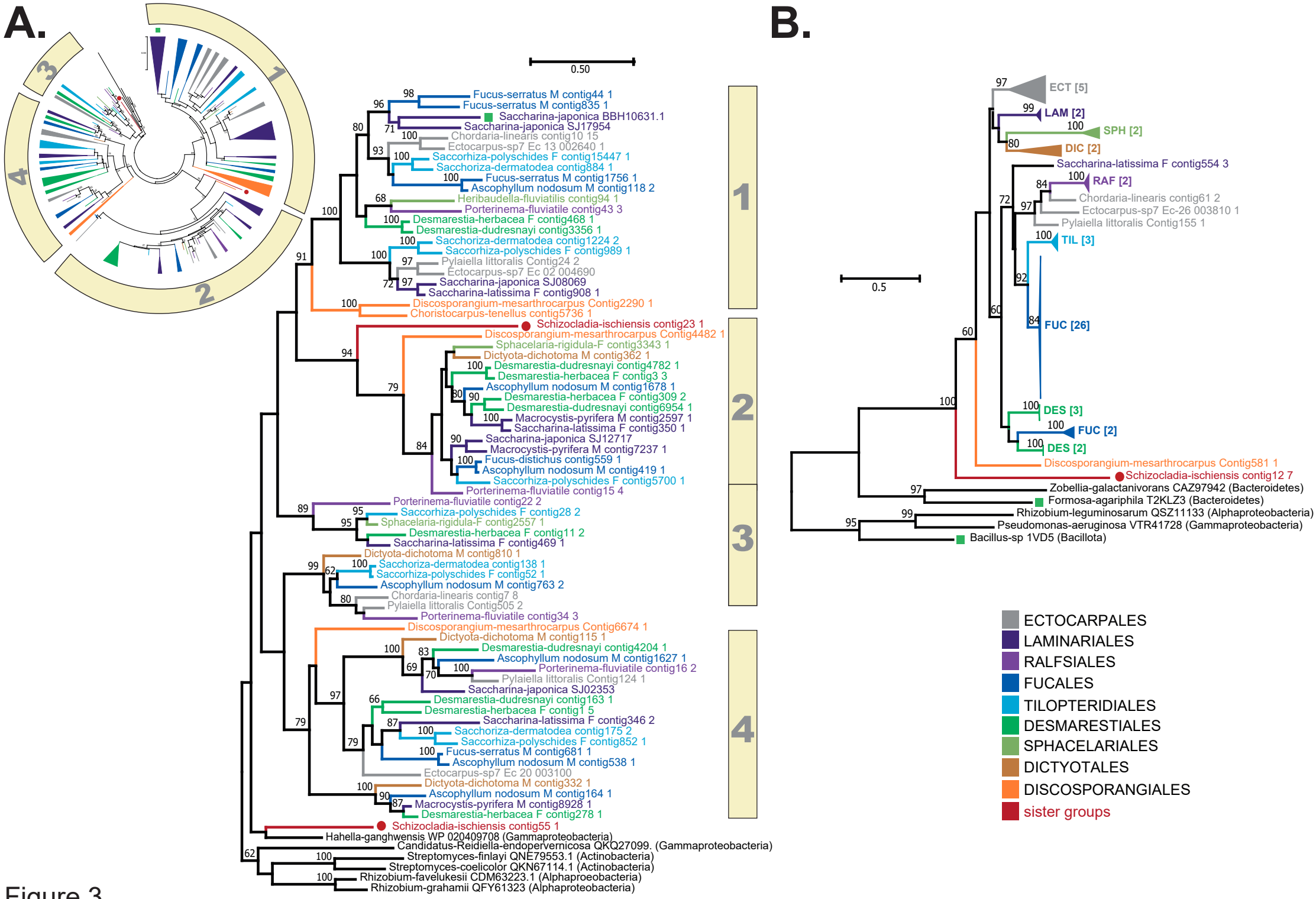

A.
B.
0.50
Fucus-serratus M contig44 1
Fucus-serratus M contig835 1
Saccharina-japonica BBH10631.1
Saccharina-japonica SJ17954
Chordaria-linearis contig10 15
Ectocarpus-sp7 Ec 13 002640 1
Saccorhiza-polyschides F contig15447 1
Sacchoriza-dermatodea contig884 1
Fucus-serratus M contig1756 1
Ascophyllum nodosum M contig118 2
Heribaudella-fluviatilis contig94 1
Porterinema-fluviatile contig43 3
Desmarestia-herbacea F contig468 1
Desmarestia-dudresnayi contig3356 1
Sacchoriza-dermatodea contig1224 2
Saccorhiza-polyschides F contig989 1
Pylaiella littoralis Contig24 2
Ectocarpus-sp7 Ec 02 004690
Saccharina-japonica SJ08069
Saccharina-latissima F contig908 1
Discosporangium-mesarthrocarpus Contig2290 1
Choristocarpus-tenellus contig5736 1
Schizocladia-ischiensis contig23 1
Discosporangium-mesarthrocarpus Contig4482 1
Sphacelaria-rigidula-F contig3343 1
Dictyota-dichotoma M contig362 1
Desmarestia-dudresnayi contig4782 1
Desmarestia-herbacea F contig3 3
Ascophyllum nodosum M contig1678 1
Desmarestia-herbacea F contig309 2
Desmarestia-dudresnayi contig6954 1
Macrocystis-pyrifera M contig2597 1
Saccharina-latissima F contig350 1
Saccharina-japonica SJ12717
Macrocystis-pyrifera M contig7237 1
Fucus-distichus contig559 1
Ascophyllum nodosum M contig419 1
Saccorhiza-polyschides F contig5700 1
Porterinema-fluviatile contig15 4
Porterinema-fluviatile contig22 2
Saccorhiza-polyschides F contig28 2
Sphacelaria-rigidula-F contig2557 1
Desmarestia-herbacea F contig11 2
Saccharina-latissima F contig469 1
Dictyota-dichotoma M contig810 1
Sacchoriza-dermatodea contig138 1
Saccorhiza-polyschides F contig52 1
Ascophyllum nodosum M contig763 2
Chordaria-linearis contig7 8
Pylaiella littoralis Contig505 2
Porterinema-fluviatile contig34 3
Discosporangium-mesarthrocarpus Contig6674 1
Dictyota-dichotoma M contig115 1
Desmarestia-dudresnayi contig4204 1
Ascophyllum nodosum M contig1627 1
Porterinema-fluviatile contig16 2
Pylaiella littoralis Contig124 1
Saccharina-japonica SJ02353
Desmarestia-dudresnayi contig163 1
Desmarestia-herbacea F contig1 5
Saccharina-latissima F contig346 2
Sacchoriza-dermatodea contig175 2
Saccorhiza-polyschides F contig852 1
Fucus-serratus M contig681 1
Ascophyllum nodosum M contig538 1
Ectocarpus-sp7 Ec 20 003100
Dictyota-dichotoma M contig332 1
Ascophyllum nodosum M contig164 1
Macrocystis-pyrifera M contig8928 1
Desmarestia-herbacea F contig278 1
Schizocladia-ischiensis contig55 1
Hahella-ganghwensis WP 020409708 (Gammaproteobacteria)
Candidatus-Reidiella-endopervernicosa QKQ27099. (Gammaproteobacteria)
Streptomyces-finlayi QNE79553.1 (Actinobacteria)
Streptomyces-coelicolor QKN67114.1 (Actinobacteria)
Rhizobium-favelukesii CDM63223.1 (Alphaproeobacteria)
Rhizobium-grahamii QFY61323 (Alphaproteobacteria)
1
2
3
4
0.5
ECT [5]
LAM [2]
SPH [2]
DIC [2]
Saccharina-latissima F contig554 3
RAF [2]
Chordaria-linearis contig61 2
Ectocarpus-sp7 Ec-26 003810 1
Pylaiella littoralis Contig155 1
TIL [3]
FUC [26]
DES [3]
FUC [2]
DES [2]
Discosporangium-mesarthrocarpus Contig581 1
Schizocladia-ischiensis contig12 7
Zobellia-galactanivorans CAZ97942 (Bacteroidetes)
Formosa-agariphila T2KLZ3 (Bacteroidetes)
Rhizobium-leguminosarum QSZ11133 (Alphaproteobacteria)
Pseudomonas-aeruginosa VTR41728 (Gammaproteobacteria)
Bacillus-sp 1VD5 (Bacillota)
ECTOCARPALES
LAMINARIALES
RALFSIALES
FUCALES
TILOPTERIDIALES
DESMARESTIALES
SPHACELARIALES
DICTYOTALES
DISCOSPORANGIALES
sister groups

Figure 3

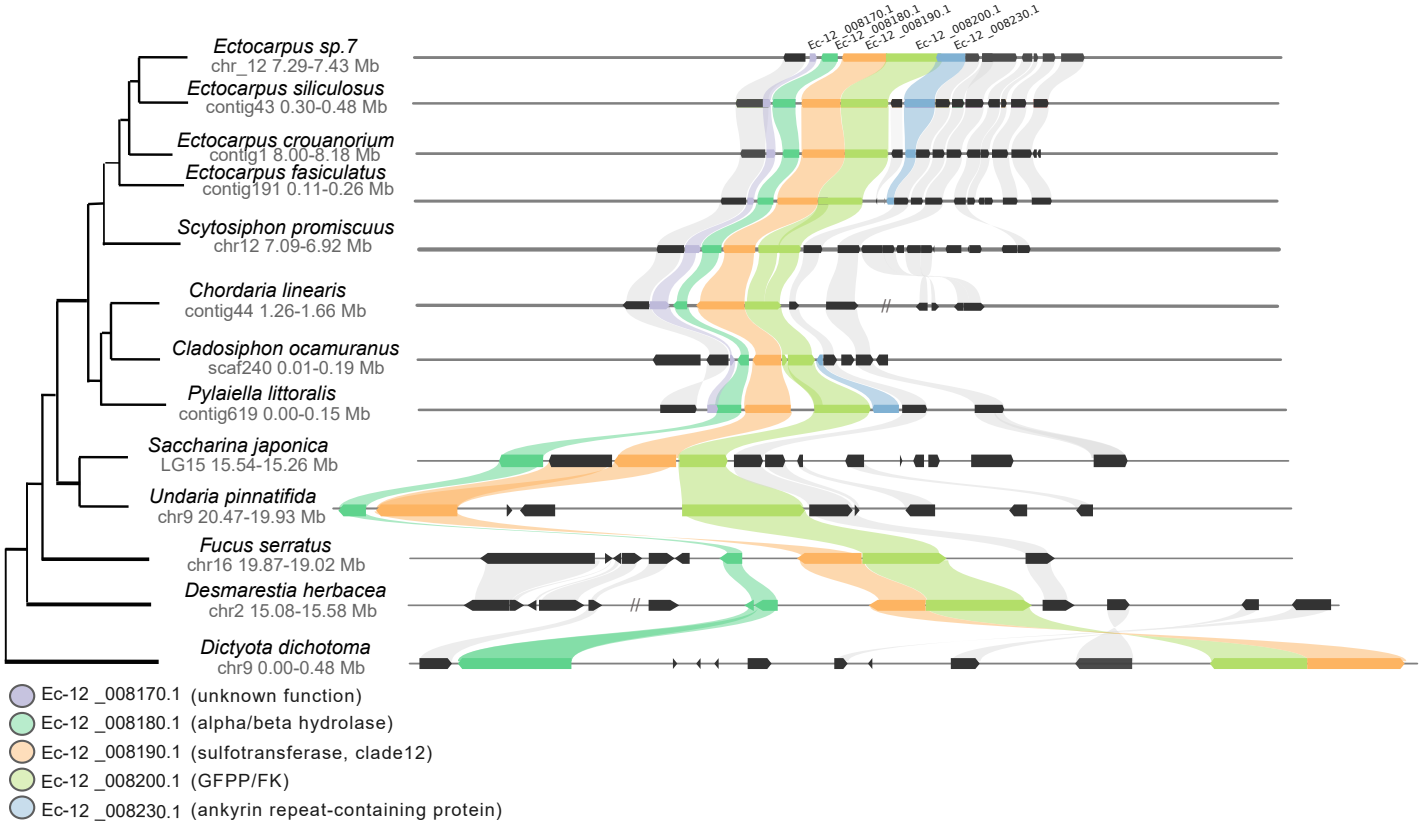
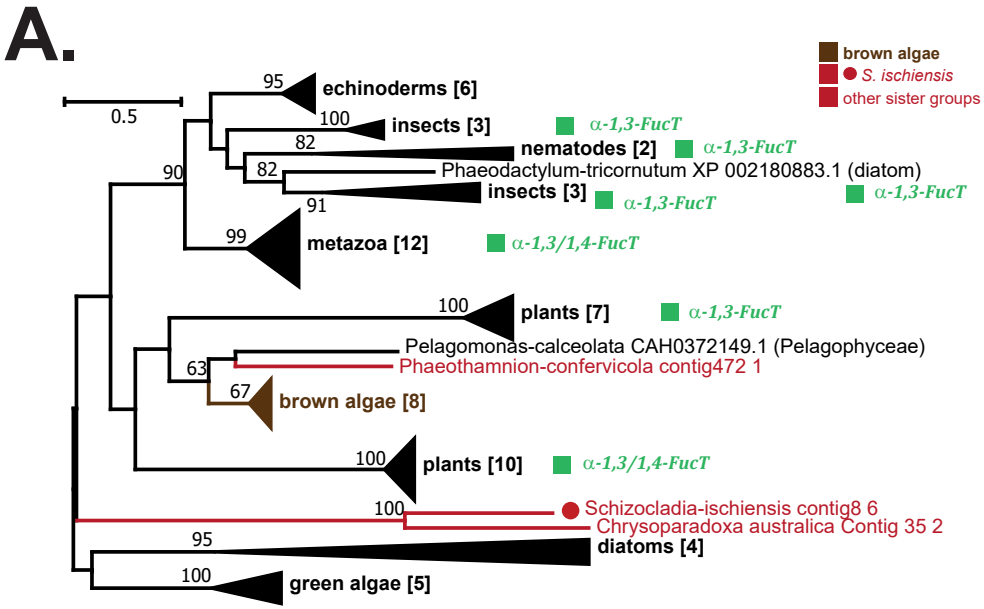
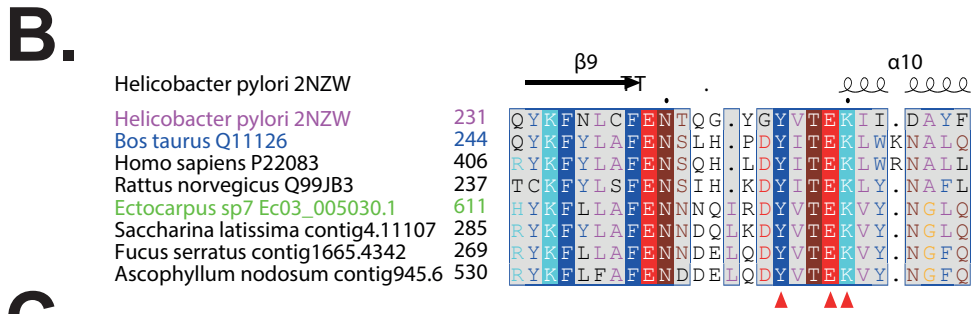
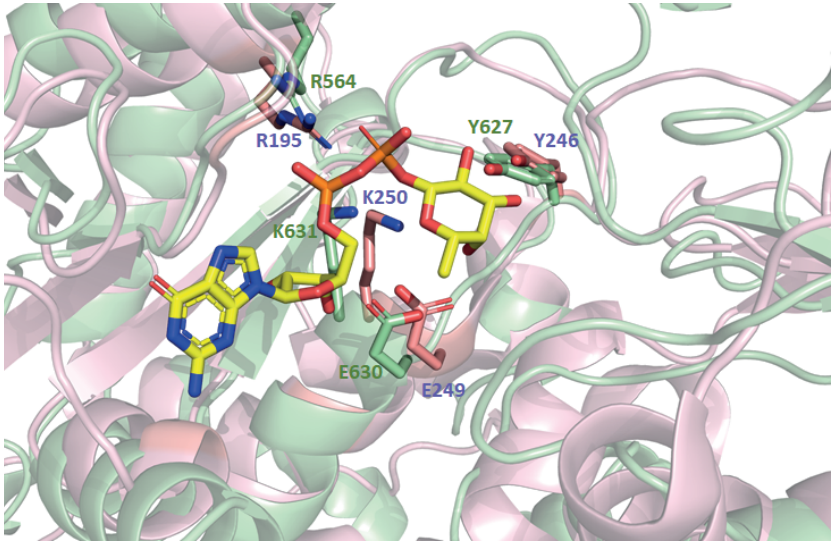
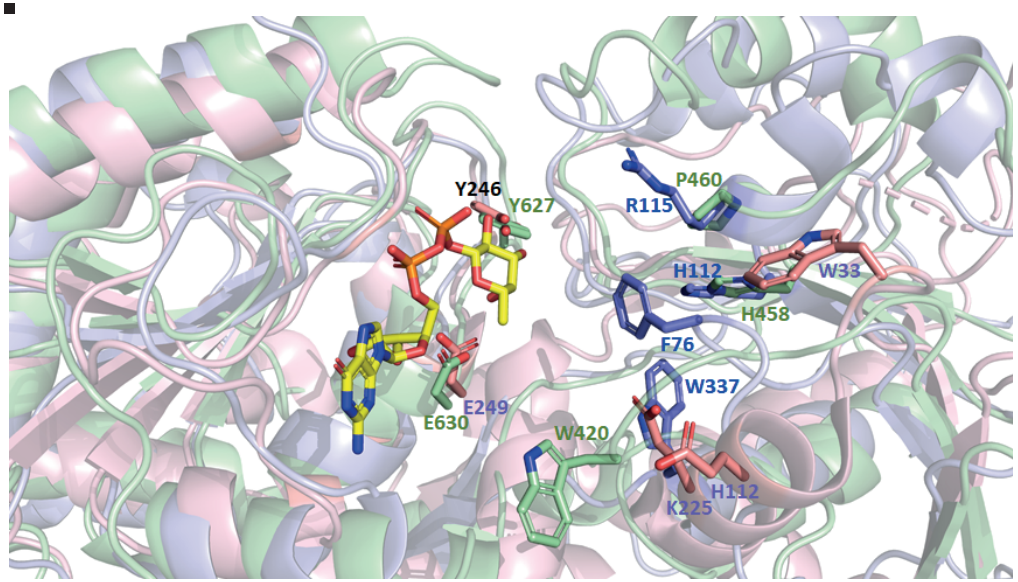

Ec-12_008170.1
Ec-12_008180.1
Ec-12_008190.1
Ec-12_008200.1
Ec-12_008230.1
Ectocarpus sp.7
chr_12 7.29-7.43 Mb
Ectocarpus siliculosus
contig43 0.30-0.48 Mb
Ectocarpus crouanorium
contig1 8.00-8.18 Mb
Ectocarpus fasiculatus
contig191 0.11-0.26 Mb
Scytosiphon promiscuus
chr12 7.09-6.92 Mb
Chordaria linearis
contig44 1.26-1.66 Mb
Cladosiphon ocamuranus
scaf240 0.01-0.19 Mb
Pylaiella littoralis
contig619 0.00-0.15 Mb
Saccharina japonica
LG15 15.54-15.26 Mb
Undaria pinnatifida
chr9 20.47-19.93 Mb
Fucus serratus
chr16 19.87-19.02 Mb
Desmarestia herbacea
chr2 15.08-15.58 Mb
Dictyota dichotoma
chr9 0.00-0.48 Mb
Ec-12 _008170.1 (unknown function)
Ec-12 _008180.1 (alpha/beta hydrolase)
Ec-12 _008190.1 (sulfotransferase, clade12)
Ec-12 _008200.1 (GFPP/FK)
Ec-12 _008230.1 (ankyrin repeat-containing protein)

Figure 4


A.
brown algae
S. ischiensis
other sister groups
0.5
95
echinoderms [6]
100
insects [3]
α-1,3-FucT
82
nematodes [2]
α-1,3-FucT
90
82
Phaeodactylum-tricornutum XP 002180883.1 (diatom)
insects [3]
α-1,3-FucT
α-1,3-FucT
91
99
metazoa [12]
α-1,3/1,4-FucT
100
plants [7]
α-1,3-FucT
Pelagomonas-calceolata CAH0372149.1 (Pelagophyceae)
63
Phaeothamnion-confervicola contig472 1
67
brown algae [8]
100
plants [10]
α-1,3/1,4-FucT
100
Schizocladia-ischiensis contig8 6
Chrysoparadoxa australica Contig 35 2
95
diatoms [4]
100
green algae [5]


B.
β9
α10
Helicobacter pylori 2NZW
Helicobacter pylori 2NZW 231 QYKFNLCFENTQG.YGYVTEKII.DAYF
Bos taurus Q11126 244 QYKFYLAFENSLH.PDYITEKLWKNALQ
Homo sapiens P22083 406 RYKFYLAFENSQH.LDYITEKLWRNALL
Rattus norvegicus Q99JB3 237 TCKFYLSFENSIH.KDYITEKLY.NAFL
Ectocarpus sp7 Ec03_005030.1 611 HYKFLLAFENNNQIRDYVTEKVY.NGLQ
Saccharina latissima contig4.11107 285 RYKFYLAFENNDQLKDYVTEKVY.NGLQ
Fucus serratus contig1665.4342 269 RYKFLLAFENNDELQDYVTEKVY.NGFQ
Ascophyllum nodosum contig945.6 530 RYKFLFAFENDDELQDYVTEKVY.NGFQ


C.
R564
R195
Y627
Y246
K250
K631
E630
E249


D.
Y246
Y627
P460
R115
H112
W33
H458
F76
W337
E249
E630
W420
H112
K225

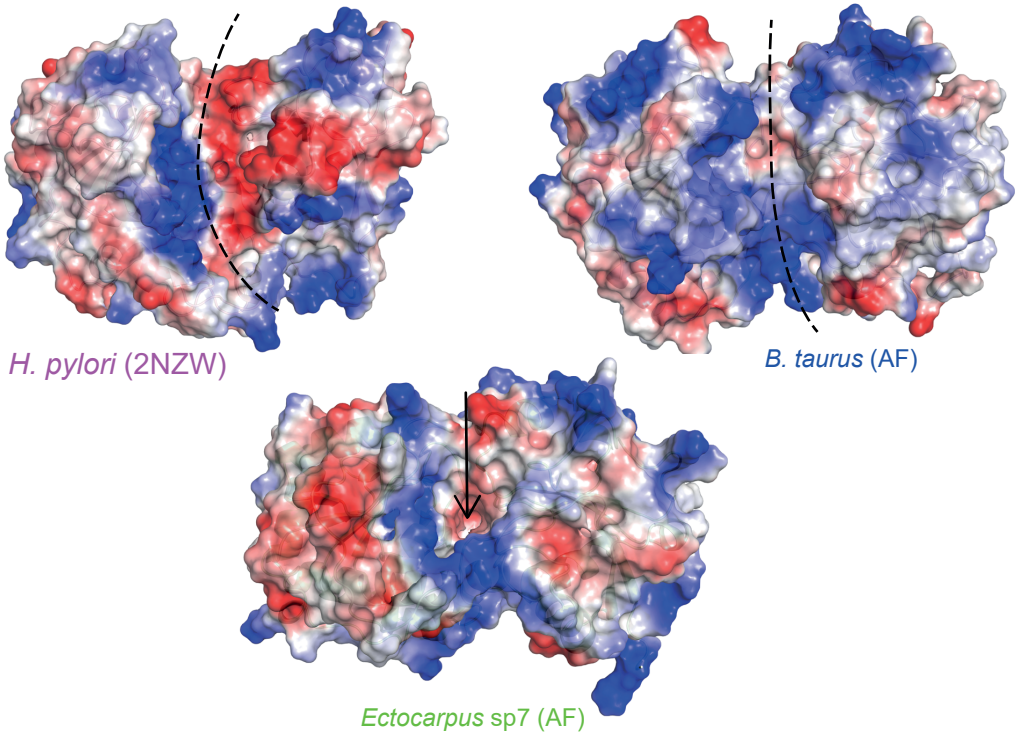

E.
H. pylori (2NZW)
B. taurus (AF)
Ectocarpus sp7 (AF)

Figure 5

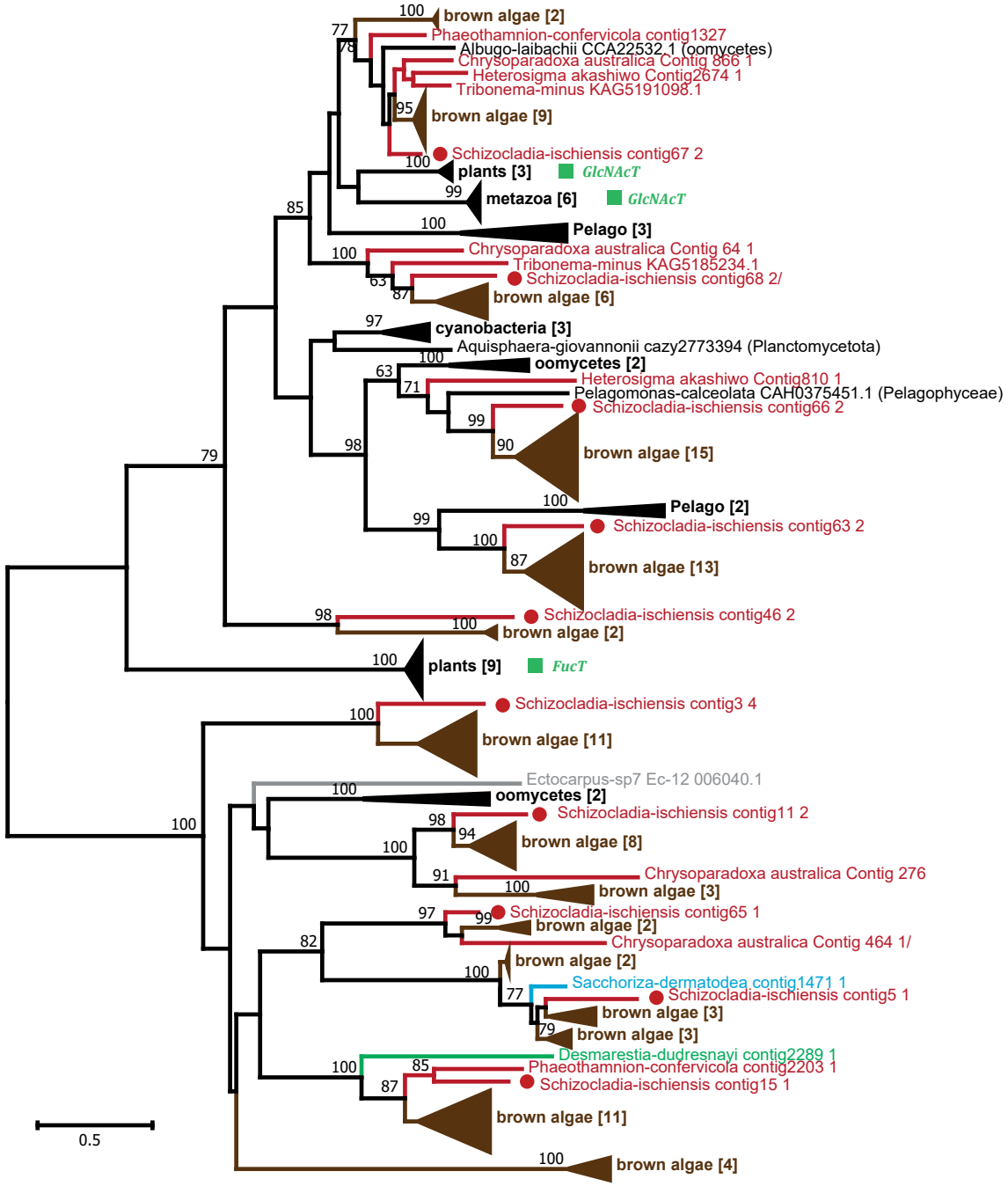

brown algae [2]
Phaeothamnion-confervicola contig1327
Albugo-laibachii CCA22532.1 (oomycetes)
Chrysoparadoxa australica Contig 866 1
Heterosigma akashiwo Contig2674 1
Tribonema-minus KAG5191098.1
brown algae [9]
Schizocladia-ischiensis contig67 2
plants [3]
GlcNAcT
metazoa [6]
GlcNAcT
Pelago [3]
Chrysoparadoxa australica Contig 64 1
Tribonema-minus KAG5185234.1
Schizocladia-ischiensis contig68 2/
brown algae [6]
cyanobacteria [3]
Aquisphaera-giovannonii cazy2773394 (Planctomycetota)
oomycetes [2]
Heterosigma akashiwo Contig810 1
Pelagomonas-calceolata CAH0375451.1 (Pelagophyceae)
Schizocladia-ischiensis contig66 2
brown algae [15]
Pelago [2]
Schizocladia-ischiensis contig63 2
brown algae [13]
Schizocladia-ischiensis contig46 2
brown algae [2]
plants [9]
FucT
Schizocladia-ischiensis contig3 4
brown algae [11]
Ectocarpus-sp7 Ec-12 006040.1
oomycetes [2]
Schizocladia-ischiensis contig11 2
brown algae [8]
Chrysoparadoxa australica Contig 276
brown algae [3]
Schizocladia-ischiensis contig65 1
brown algae [2]
Chrysoparadoxa australica Contig 464 1/
brown algae [2]
Sacchoriza-dermatodea contig1471 1
Schizocladia-ischiensis contig5 1
brown algae [3]
brown algae [3]
Desmarestia-dudresnayi contig2289 1
Phaeothamnion-confervicola contig2203 1
Schizocladia-ischiensis contig15 1
brown algae [11]
brown algae [4]
0.5

Figure 6

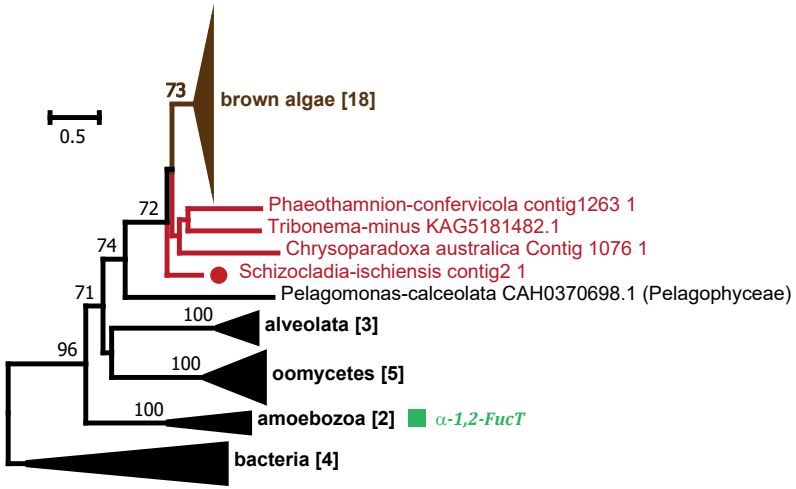

0.5
73
brown algae [18]
72
74
71
96
100
100
100
Phaeothamnion-confervicola contig1263 1
Tribonema-minus KAG5181482.1
Chrysoparadoxa australica Contig 1076 1
Schizocladia-ischiensis contig2 1
Pelagomonas-calceolata CAH0370698.1 (Pelagophyceae)
alveolata [3]
oomycetes [5]
amoebozoa [2]
α-1,2-FucT
bacteria [4]

Figure 7

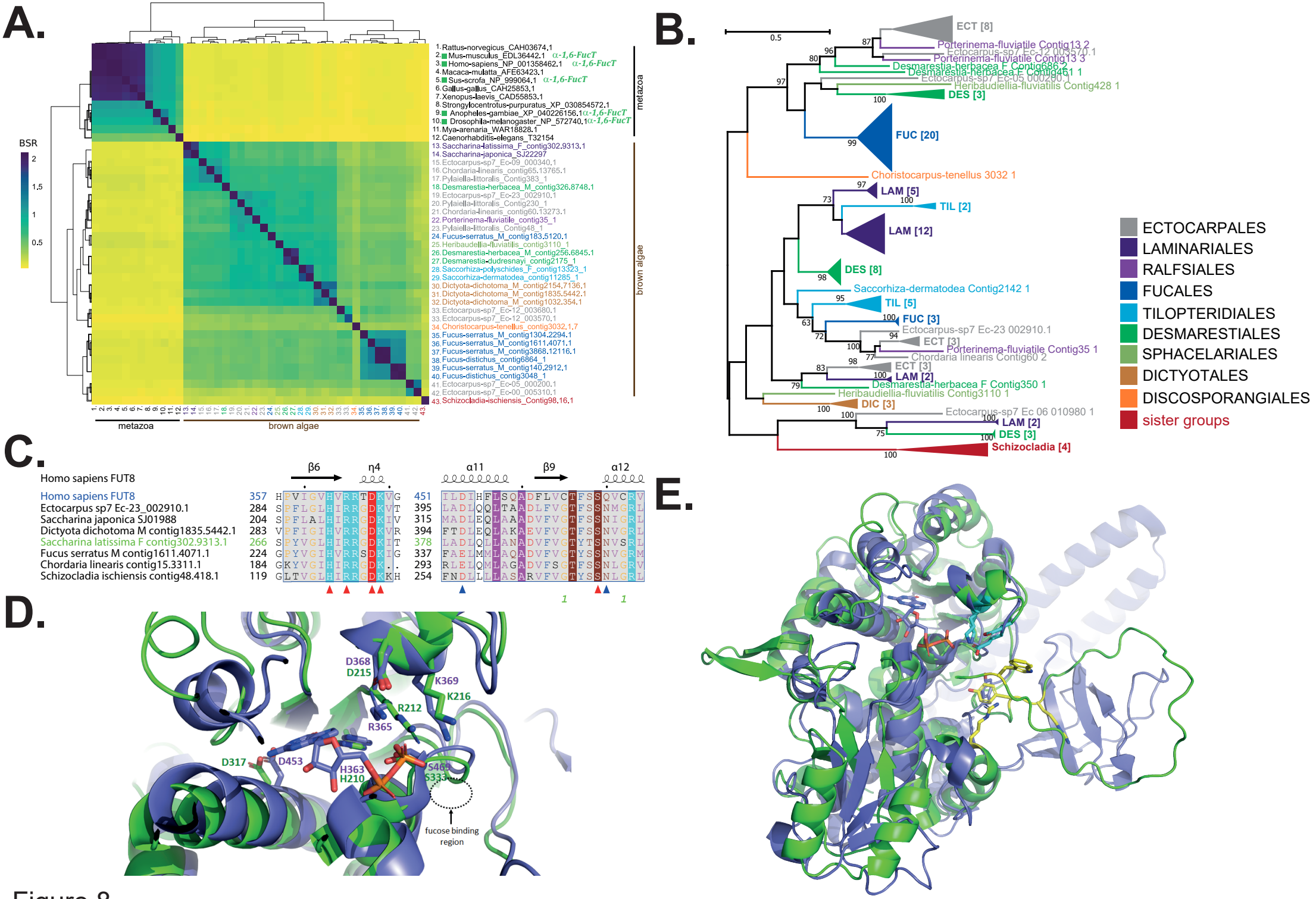

A.
BSR
2
1.5
1
0.5
1.Rattus-norvegicus_CAH03674.1
2.Mus-musculus_EDL36442.1 α-1,6-FucT
3.Homo-sapiens_NP_001358462.1 α-1,6-FucT
4.Macaca-mulatta_AFE63423.1
5.Sus-scrofa_NP_999064.1 α-1,6-FucT
6.Gallus-gallus_CAH25853.1
7.Xenopus-laevis_CAD55853.1
8.Strongylocentrotus-purpuratus_XP_030854572.1
9.Anopheles-gambiae_XP_040226156.1 α-1,6-FucT
10.Drosophila-melanogaster_NP_572740.1 α-1,6-FucT
11.Mya-arenaria_WAR18828.1
12.Caenorhabditis-elegans_T32154
13.Saccharina-latissima_F_contig302.9313.1
14.Saccharina-japonica_SJ22297
15.Ectocarpus-sp7_Ec-09_000340.1
16.Chordaria-linearis_contig65.13765.1
17.Pylaiella-littoralis_Contig383_1
18.Desmarestia-herbacea_M_contig326.8748.1
19.Ectocarpus-sp7_Ec-23_002910.1
20.Pylaiella-littoralis_Contig230_1
21.Chordaria-linearis_contig60.13273.1
22.Porterinema-fluviatile_contig35_1
23.Pylaiella-littoralis_Contig48_1
24.Fucus-serratus_M_contig183.5120.1
25.Heribaudellia-fluviatilis_contig3110_1
26.Desmarestia-herbacea_M_contig256.6845.1
27.Desmarestia-dudresnayi_contig2175_1
28.Saccorhiza-polyschides_F_contig13323_1
29.Saccorhiza-dermatodea_contig11285_1
30.Dictyota-dichotoma_M_contig2154.7136.1
31.Dictyota-dichotoma_M_contig1835.5442.1
32.Dictyota-dichotoma_M_contig1032.354.1
33.Ectocarpus-sp7_Ec-12_003680.1
33.Ectocarpus-sp7_Ec-12_003570.1
34.Choristocarpus-tenellus_contig3032.1.7
35.Fucus-serratus_M_contig1304.2294.1
36.Fucus-serratus_M_contig1611.4071.1
37.Fucus-serratus_M_contig3868.12116.1
38.Fucus-distichus_contig6864_1
39.Fucus-serratus_M_contig140.2912.1
40.Fucus-distichus_contig3048_1
41.Ectocarpus-sp7_Ec-05_000200.1
42.Ectocarpus-sp7_Ec-00_005310.1
43.Schizocladia-ischiensis_Contig98.16.1
metazoa
brown algae
B.
0.5
ECT [8]
Porterinema-fluviatile Contig13 2
Ectocarpus-sp7 Ec-12 003570.1
Porterinema-fluviatile Contig13 3
Desmarestia-herbacea F Contig686 2
Desmarestia-herbacea F Contig461 1
Ectocarpus-sp7 Ec-05 000200.1
Heribaudiellia-fluviatilis Contig428 1
DES [3]
FUC [20]
Choristocarpus-tenellus 3032 1
LAM [5]
TIL [2]
LAM [12]
DES [8]
Saccorhiza-dermatodea Contig2142 1
TIL [5]
FUC [3]
Ectocarpus-sp7 Ec-23 002910.1
ECT [3]
Porterinema-fluviatile Contig35 1
Chordaria linearis Contig60 2
ECT [3]
LAM [2]
Desmarestia-herbacea F Contig350 1
Heribaudiellia-fluviatilis Contig3110 1
DIC [3]
Ectocarpus-sp7 Ec 06 010980 1
LAM [2]
DES [3]
Schizocladia [4]
ECTOCARPALES
LAMINARIALES
RALFSIALES
FUCALES
TILOPTERIDIALES
DESMARESTIALES
SPHACELARIALES
DICTYOTALES
DISCOSPORANGIALES
sister groups
C.
Homo sapiens FUT8
β6
η4
α11
β9
α12
Homo sapiens FUT8 357 HPVIGVHVRRTDKVG 451 ILDIHFLSQADFLVCTFSSQVCRV
Ectocarpus sp7 Ec-23_002910.1 284 SPFLGLHIRRGDKVT 395 LADLQQLTAADLFVGTFSSNMGRL
Saccharina japonica SJ01988 204 SPFLALHIRRGDKIV 315 MADLEQLAAADVFVGTFSSNIGRL
Dictyota dichotoma M contig1835.5442.1 283 VPFIGIHVRRGDKVR 394 FTDLEQLAKAEVFVGTFSSNVGRL
Saccharina latissima F contig302.9313.1 266 SPYVGLHIRRGDKIT 378 LADLQLLANADVFVGTYTSNVSRL
Fucus serratus M contig1611.4071.1 224 GPYVGIHVRRSDKIG 337 FAELMMLAQADVFVGTYSSNVGRM
Chordaria linearis contig15.3311.1 184 GKYVGIHIRRGDK.. 293 RLELQMLAGADVFSGTFSSNLGRV
Schizocladia ischiensis contig48.418.1 119 GLTVGLHIRRGDKKH 254 FNDLLLLASARVFVGTYSSNLGRL
1
1
D.
D368
D215
K369
K216
R212
R365
D317
D453
H363
H210
S468
S333
fucose binding region
E.

Figure 8

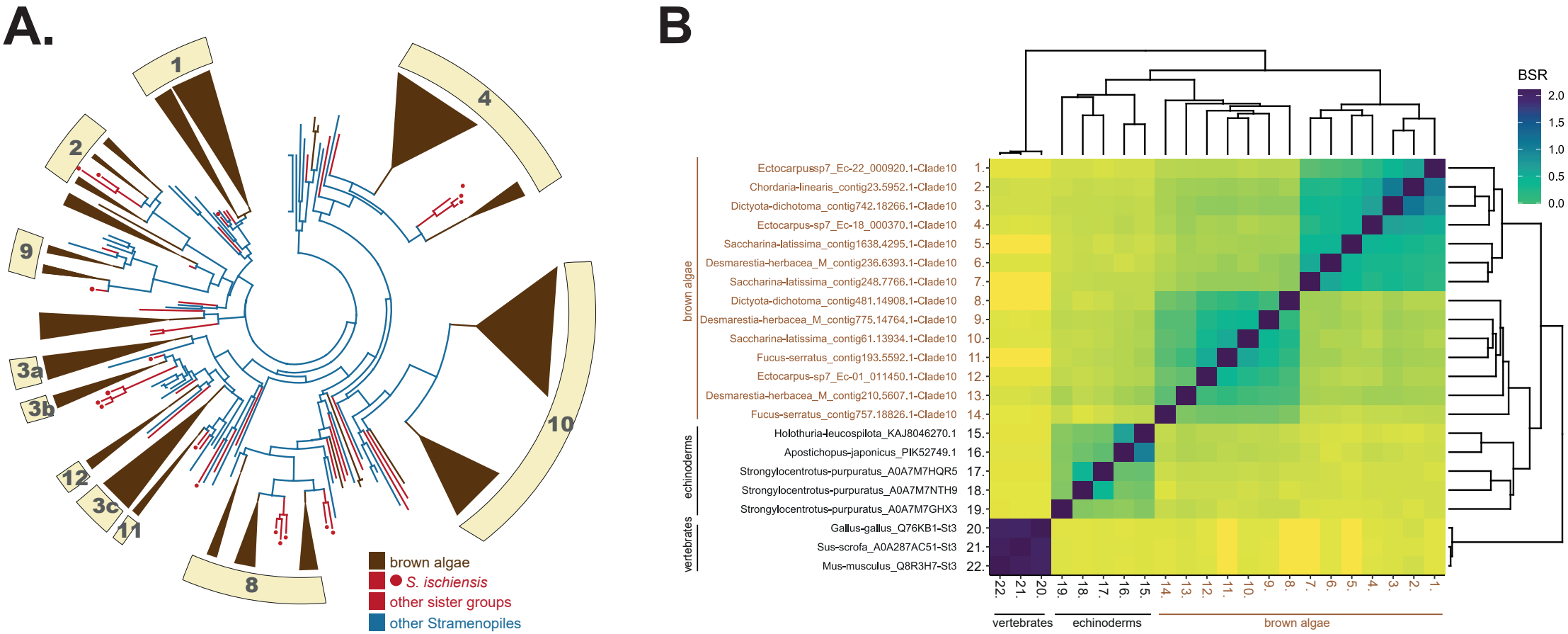

A.
1
2
3a
3b
3c
4
8
9
10
11
12
brown algae
S. ischiensis
other sister groups
other Stramenopiles
B
BSR
2.0
1.5
1.0
0.5
0.0
brown algae
echinoderms
vertebrates
Ectocarpussp7_Ec-22_000920.1-Clade10
Chordaria-linearis_contig23.5952.1-Clade10
Dictyota-dichotoma_contig742.18266.1-Clade10
Ectocarpus-sp7_Ec-18_000370.1-Clade10
Saccharina-latissima_contig1638.4295.1-Clade10
Desmarestia-herbacea_M_contig236.6393.1-Clade10
Saccharina-latissima_contig248.7766.1-Clade10
Dictyota-dichotoma_contig481.14908.1-Clade10
Desmarestia-herbacea_M_contig775.14764.1-Clade10
Saccharina-latissima_contig61.13934.1-Clade10
Fucus-serratus_contig193.5592.1-Clade10
Ectocarpus-sp7_Ec-01_011450.1-Clade10
Desmarestia-herbacea_M_contig210.5607.1-Clade10
Fucus-serratus_contig757.18826.1-Clade10
Holothuria-leucospilota_KAJ8046270.1
Apostichopus-japonicus_PIK52749.1
Strongylocentrotus-purpuratus_A0A7M7HQR5
Strongylocentrotus-purpuratus_A0A7M7NTH9
Strongylocentrotus-purpuratus_A0A7M7GHX3
Gallus-gallus_Q76KB1-St3
Sus-scrofa_A0A287AC51-St3
Mus-musculus_Q8R3H7-St3
1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22.
vertebrates
echinoderms
brown algae

Figure 9