Supplemental Materials: **Insights into the red algae and eukaryotic evolution from the genome of *Porphyra umbilicalis* (Bangiophyceae, Rhodophyta)**

Table of Contents

Supplementary Tables and Figures

Methods

Culture techniques and antibiotic treatment

Full culture techniques for the single genotypic isolate (Pum1) of haploid *Porphyra umbilicalis* are as previously published (1). Pum1 has a generation time of 2.5-6 mo depending upon temperature, nutrient and light levels. Material for the PacBio sequencing came from the 13th generation. Prior to that time, the isolate was treated periodically with an antibiotic cocktail (100 mg/L penicillin, 25 mg/L streptomycin), which resulted in elimination of 75% of the bacterial OTUs present on wild blades (2). A further antibiotic treatment (Table S1) was applied to blade material that was prepared for final DNA extraction and Illumina/PacBio sequencing; subsamples of blade tissue stained with the fluorescent DNA stain DAPI lost the bacterial biofilm present on Day 1 by the end of the antibiotic treatment.

Table S1. Antibiotic treatment schedule for Pum1.

Final concentrations in the culture medium are provided. The blades were transferred to new seawater media at the beginning of Days 1, 4 and Day 6 and flash-frozen on Day 9.

Antibiotic	Days 1-3	Days 4-5	Day 6	Day 7-8	Source
ampicillin	0	200 mg/L	0	0	Sigma
carbenicillin	0	0	0	100 mg/L	Sigma
cefotaxime	0	200 mg/L	100 mg/L	100 mg/L	VWR
chloramphenicol	0	5 mg/L	0	0	Sigma
ciprofloxacin	0	20 mg/L	0	0	Sigma
gentamycin	0	100 mg/L	0	0	VWR
kanamycin	0	600 mg/L	50 mg/L	200 mg/L	Sigma
neomycin	0	250 mg/L	0	0	Sigma
nystatin	0	5 mg/L	0	0	VWR
penicillin G	1 g/L	1 g/L	1 g/L	1 g/L	VWR
polymyxin B	0	0	0	15 mg/L	VWR
streptomycin	0	100 mg/L	0	60 mg/L	Sigma

Nucleic acid preparation

Details of RNA extraction procedures and conditions used for producing ESTs are as previously published (1). To generate a genome assembly with whole genome shotgun sequencing, five DNA samples were prepared. For the first four preparations, we used a Carlson lysis buffer containing 2% CTAB followed by a Qiagen Genomic Tip 500/G kit by the manufacturer's instructions and one round of cesium chloride gradient purification (3). Three of these DNA preparations were used for pilot sequencing, and one was used for Illumina jumping library construction (2 x 3 kb and 2 x 7 kb, Cre-Lox Random Shearing, sequenced at 2 x 300 bp, ~4x) and deep Illumina fragment library construction (2 x 250 bp, 400, 800 bp insert, sequenced at 2 x 251 bp).

For generating DNA for PacBio sequencing, Pum1 material was quickly rinsed in sterile water and flash-frozen in liquid nitrogen. Frozen material (1-2 g) was ground to a fine powder using a mortar and pestle and resuspended in 10-20 ml sterile 0.1 M Tris, pH 8.0, 0.1 M NaCl with 2% cetyl trimethyl ammonium bromide (CTAB, to promote precipitation of polysaccharides), and added proteinase K (200 μg/ml). The solution was incubated at room temperature for about 60 min with gentle shaking. The debris was pelleted by centrifugation, and the supernatant extracted sequentially with phenol, then phenol:chloroform (1:1), and then chloroform. Both nucleic acids and the remaining polysaccharides in the aqueous phase were precipitated by 2 vol of ethanol. The pellet was solubilized overnight (or longer when necessary), and the nucleic acid resuspended in 10 mM Tris, pH 8.0, 1 mM EDTA with the addition of the fluorescent DNA dye bisbenzimide at ~10 μg/ml followed by the addition of 1 g CsCl/ml. The solution was centrifuged (10 min at 10,000 rpm, SS34 rotor) to remove debris, and the clarified solution was centrifuged (12-14 h, 300,000 x *g*, VTi65 rotor). Multiple fluorescent bands were observed in the gradient and collected individually. The high GC bands were screened by PCR to identify bands enriched in *Porphyra* DNA. Sometimes the gradient was repeated for the individual bands to further purify the material.

Genome sequencing and assembly

A pilot assembly (v0.75) was constructed from Illumina reads (Table S2) with ALLPATHS(4) and screened for contaminants. This assembly was fragmented (23,884 contigs, contig L50 7.0 kb). Examination of EST hits on the scaffolds revealed that the larger scaffolds (> 175 kb) lacked EST hits, yet encoded many gene loci. This suggests the sequences come from one or more contaminants that were not present in the sample used for expression sequencing.

Table S2: Sequencing for pilot Illumina assembly.

Fragment libraries were downsampled to the coverages shown for assembly.

Library	Coverage	Insert (bp, mean ± SD)	Sequence reads
UBBC	28.0x	371 +/- 40	2x251
WNSN	78.8x	808 +/- 80	2x251
YCGA	42.2x	410 +/- 40	2x251
A65N2	4.0x	3,901 +/- 406	2x300
A65NA	4.0x	7.234 +/- 514	2x300
IBIB	4.7x	7,212 +/- 495	2x300
IBIC	3.7x	3,954 +/- 395	2x300
Total	**165.4x**		

An improved and highly contiguous (contig L50 = 189.9 kb) draft assembly was generated with PacBio sequencing (57 cells, 26.9 Gbp total sequences, mean read length 7 kb) despite the 66 % G + C-content that had impeded sequencing on the Illumina platform and biased recovery toward enhancement of the number of lower G + C-content bacterial sequences (Table S2). The main assembly was performed with a modified version of the Falcon 2.0 assembler (5), polished with the standard QUIVER polishing pipeline (6), and screened for contaminating sequences from organelles. A pilot Illumina-only assembly built with ALLPATHS (4) (see above) contained contigs that were not present in the FALCON assembly. A total of 1,517 Illumina contigs representing 3.9 Mbp of sequence was added to the V1.0 assembly. Scaffolds with homology to mitochondrial or chloroplast sequence were partitioned out of the main genome into two separate bins. Finally, homozygous SNPs and indels were corrected in the assembly by aligning ~220x of Illumina reads (WNSN library, 2 x 250, 800 bp insert) to the assembly with BWA mem (7) and identifying SNPs and indels with the UnifiedGenotyper tool from the GATK (8). A total of 33 out of the 50 homozygous SNPs and 2,331 out of 2,651 homozygous indels were corrected. The remainder could not be fixed because they were so close to another polymorphism that the GATK (8) UnifiedGenotyper pipeline was unable to resolve multiple close polymorphisms.

Genome partitioning

A draft annotation of genes on the scaffolds and subsequent visualization and inspection of 50 scaffolds chosen randomly among the 693 scaffolds longer than 25 kb showed that, unexpectedly, scaffolds belonged to more than one bin. The three bins that were apparent could be described as: Those with genes with EST support and more than one exon, assumed to be *Porphyra*; those without EST support and more than one exon (an unexpected eukaryote present in the *Porphyra* culture), and those without EST support and only single-exon genes (bacteria present in the *Porphrya* sample, despite antibiotic treatment). Scaffolds shorter than 25 kb could not be partitioned in this simple manner, so heuristic and principal component analysis (PCA) was performed to separate the scaffolds into these three bins. The following parameters were calculated for each scaffold: the number of loci with EST/RNA-Seq support; read depth in at least one representative 454, single-cell, Illumina and PacBio sequencing effort; % GC; fraction of blastx (9) hits to Uniref90 (10) protein that fell into these taxonomic groups: i) Rhodophyta, ii) Bacteria, Actinobacteria, Bacteriodetes/Chlorobi group, Planctomycetes, Proteobacteria, and iii) Opisthokonta, other Eukaryota.

Using these parameters as guides, heuristic rules were iteratively developed by inspecting up to ten scaffolds in each bin and modifying the rules to correct scaffolds partitioned into the wrong bin according to expert review. The scaffolds were partitioned into the following bins: *Porphyra* genome, bacterial metagenomes, other eukaryote genome, unknown. This process continued until it was no longer possible to find clear examples of scaffolds that had been partitioned into the wrong bin. This provided a robust partitioning for the larger scaffolds, but also included scaffolds in the unknown bin explicitly as well as when it was impossible to tell whether they were in the right bin. In addition, scaffolds could be misplaced owing to subjective errors in the manual partitioning. Nonetheless, this provided a good starting point, particularly for scaffolds with BLASTX and/or expression hits, typically > 25kb.

PCA analysis was performed on the parameters calculated for each scaffold listed above using the prcomp() function in the built-in stats package in R (11) with data values scaled and centered. PC1, PC2 and PC3 explained 45.55%, 18.40% and 9.24% of the variance, respectively. Scatter plots of PC1-6 (Fig S3), with points colored by their heuristically-determined bin, showed that in almost all pairwise combinations of principal component, the bins were overlaid on each other and it was impossible to determine simple separators between the bins. However, in one plot, PC2 vs. PC3, (Fig S3), the red, green and cyan points for the other eukaryote, *Porphyra* and prokaryotic contaminant bins were well-separated, approximately along the axes PC2=0 and PC3=0.

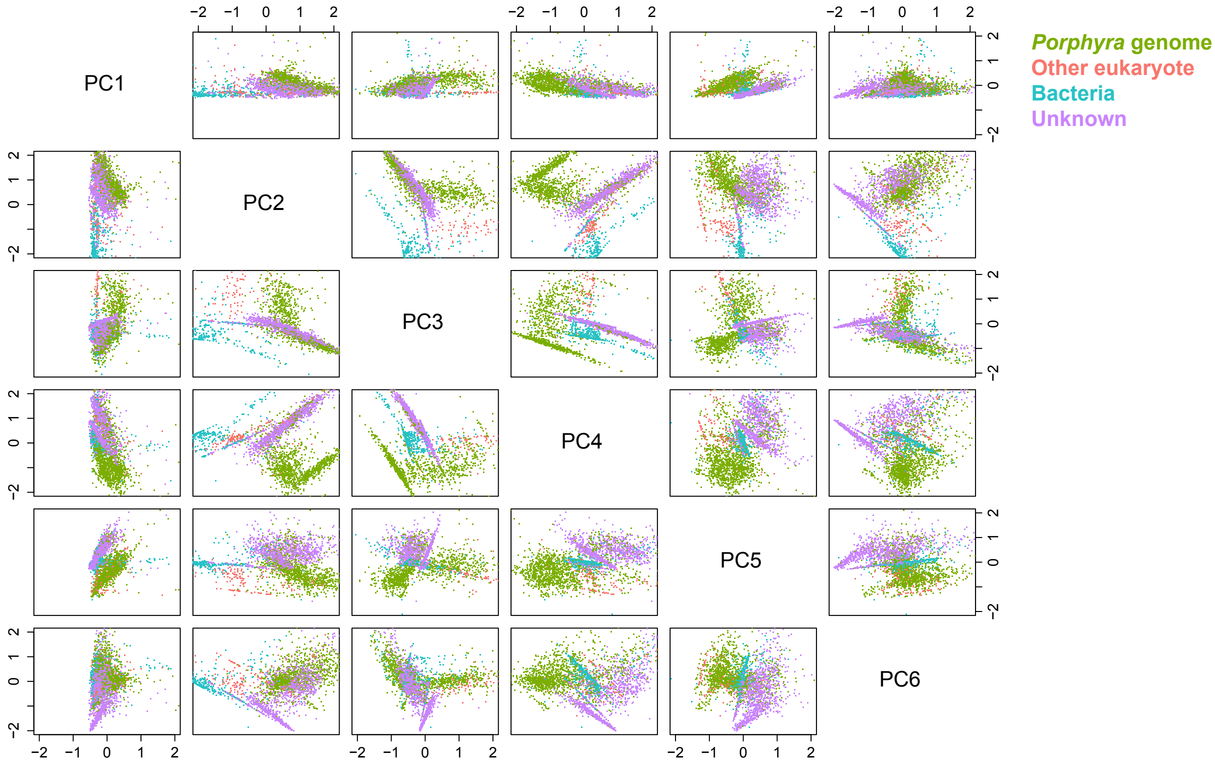


Figure S3: Principal components analysis of scaffolds

In order to find combinations of principal components that separated the scaffolds into correct bins as accurately as possible, we generated a grid of scatter plots of all pairwise combinations of principal components 1-6 plotted against each other where each dot is a scaffold, colored by its heuristic bin membership (see legend on right). Inspection of this plot was used to choose a pair of PCs that give the best separation of the heuristic bins, although these contain errors, particularly away from the centroid of the population.

Inspection of the parameters for the scaffolds on either side of the axes showed that heuristic bin placement was often with low confidence and that the PC2 = 0 axis acts as a good separator between *Porphyra* scaffolds and the other bins, which could in turn be separated by the line PC3 = 0.2 that kept well-characterized, larger scaffolds in their expected bins (Fig S3).

These two partitions enabled scaffolds to be placed in bins accordingly:

$$PC2 > 0 = \textit{Porphyra}$$
$$PC2 \leq 0 \text{ and } PC3 > 0.2 = \text{other eukaryote}$$
$$PC2 \leq 0 \text{ and } PC3 \leq 0.2 = \text{bacteria}$$

Lastly, manual inspection of known *Porphyra* genes revealed that 97 scaffolds had been placed in other eukaryote or bacterial bins by the PCA. At least 0.05 loci with *Porphyra* ESTs/kb were found on these scaffolds, consistent with their belonging to the *Porphyra* genome bin, and this characteristic was used to move them into that bin (Table S4).

Remaining antibiotic-resistant bacteria and the principal eukaryotic contaminant may influence *Porphyra*'s ecological niche or they may be hitch-hikers. The longest associated bacterial scaffolds (~ 5 Mb) could represent near-complete chromosomal assemblies. Representation was predominantly from Proteobacteria, with some sequences from Planctomycetes and the Bacteroidetes/Chlorobi group. Analysis of these scaffolds will be pursued separately.

Table S4: Genome assembly summary

Summary statistics for the three bins into which scaffolds were separated

Statistic	*Porphyra*	Other eukaryote	Bacteria
% GC	65.78	60.21	56.19
Main genome scaffold total	2,125	180	847
Main genome contig total	2,183	180	847
Main genome scaffold sequence total	87.7 Mb	21.2 Mb	64.5 Mb
Main genome contig sequence total	87.6 Mb (0.1% gaps)	21.2 Mb (0.0% gap)	64.5 Mb (0.0% gaps)
Main genome scaffold N/L50	130/202.0 kb	17/473.0kb	14/872.0 kb
Main genome contig N/L50	137/188.0 kb	17/473.0kb	14/872.0 kb
Number of scaffolds > 50 kb	436	58	154
% main genome in scaffolds > 50 kb	88.4%	93.8%	86.2%

De novo repeat finding and repeat masking

RepeatModeler v1.0.8 (12) is a *de novo* repeat sequence finder and was run on the *Porphyra* assembly to generate a library of repetitive sequences.

Repeat Modeler is not able to distinguish between highly-abundant sequences that are members of families of transposable elements or large families of protein-coding genes. Annotation of the initially predicted repeat sequences with Pfam (13) and PANTHER (14) domains was used to categorize the repeat sequences into those with annotations corresponding to known genes (24 sequences were removed from the repeat library); those that were associated with transposable elements (945 sequences; these were kept in the repeat library) and the remaining 50 sequences that needed manual inspection, after which 30 were retained and added to the 945 sequences. These were then filtered to remove sequences 100 bp or shorter, making a library of 943 repeat sequences with median length 727 bp and totaling 1,295,558 bp.

Classification of the repeat sequences with RepeatMasker (15) is part of the Repeat Modeler pipeline; however, this left many repeat sequences unclassified. Recently, TEclass (16) was developed to classify transposable elements in repetitive sequence, and this was used to fill in classifications that RepeatMasker omitted.

The resulting library of repetitive sequences was used to mask 38.5 Mbp (43.9%) in the genome with RepeatMasker (15). Some of these sequences were from unknown families, but 17.0 % of the assembly is represented by various LTRs and 17.7 % by DNA-transposons (Table S5).

Table S5: Repeat content of *Porphyra* genome

Repeat Class and family	Masked (bp)	Percentage of genome
DNA	11,041,644	12.59%
DNA/CMC-Chapaev	259,060	0.30%
DNA/CMC-EnSpm	430,600	0.49%
DNA/MULE-MuDR	418,409	0.48%
DNA/PIF-Harbinger	2,521,640	2.88%
DNA/TcMar-Ant1	129,502	0.15%
DNA/TcMar-Fot1	90,662	0.10%
DNA/TcMar-ISRm11	145,378	0.17%
DNA/TcMar-Tigger	292,070	0.33%
DNA/TcMar-m44	94,756	0.11%
DNA/hAT-Tag1	57,087	0.07%
DNA/hAT-hAT1	21,058	0.02%
LINE	2,990,223	3.41%
LINE/L1-Tx1	46,810	0.05%
LINE/L2	61,812	0.07%
LINE/Penelope	323	0.00%
LTR	5,719,306	6.52%
LTR/Copia	2,827,172	3.22%
LTR/DIRS	64,383	0.07%
LTR/ERVL	14,769	0.02%
LTR/Gypsy	5,934,981	6.77%
LTR/Ngaro	131,465	0.15%
LTR/Pao	35,115	0.04%
Retro	3,049,480	3.48%
SINE	101,880	0.12%
SINE/tRNA-Deu-L2	20,095	0.02%
Unknown	1,729,930	1.97%
nonLTR	26,000	0.03%
rRNA	249,557	0.28%
TOTAL	**38,505,167**	**43.91**

RNA-Seq and expressed sequence tag (EST) expression analysis

Three 2 × 100 bp Illumina Hi-Seq stranded libraries were sequenced for RNA-Seq with a 300 bp insert. Libraries were constructed and sequenced from three regions: blade margin with mature neutral spores, inner blade area with differentiating neutral spores, and central vegetative area. All reads were deposited in the GenBank Sequence Read Archive (accessions SRX115571, SRX115570, SRX115569). They comprise ~ 227 million (226,967,371) read pairs, 45.4 Gbp in total. These reads were aligned to the genome with gmap and assembled into transcripts with the in-house tool pertran generating 30,011 transcript assemblies.

Expressed sequence tag (EST) sequences were previously generated on the 454 platform and filtered (1, 17) from *P. umbilicalis* (2,152,212 sequences) and *P. yezoensis* (a closely-related species, 1,265,156 sequences). These data sets were separately aligned to the genome with gmap and assembled with the in-house tool pertran, generating 4,948 and 3,625 transcript assemblies, respectively.

Gene model prediction

In order to predict gene models, we first determined gene loci in the genome as follows. The transcript assemblies were aligned to the genome with PASA (18) (95% identity and 40% coverage for *P. umbilicalis* and 90% identity and 50% coverage for *P. yezoensis* assemblies). Predicted protein sequences from *Cyanidioschyzon merolae* (19), *Chondrus crispus* (20), *Cyanophora paradoxa* (21), *Porphyridium purpureum* (22), *Galdieria sulphuraria* (23), *Pyropia yezoensis* (24), and Swissprot (25) were aligned to the soft-masked genome sequence with BLASTX (9) and Exonerate (26). Regions with hits from expression data and/or protein homology were extended up to 2 kb at each end unless an adjacent locus on the same strand was reached. These footprints of homology and/or expression evidence in the genome define gene loci and were used as seeds for homology-based gene prediction with FGENESH+ (27), FGENESH_EST (similar to FGENESH+, EST as splice site and intron input instead of protein/translated ORF), and GenomeScan (28). The best predictions (based on support from expression and protein homology, but penalized for overlap with masked repeats) were improved with PASA (18), including addition of terminal untranslated regions (UTRs), correction of splice sites and addition of alternative splice forms, all based on expression evidence. These improved gene models were filtered based on the following statistics of C-score and protein homology coverage. C-score is a BLASTP score ratio to the BLASTP score of the best hit and protein coverage is the highest percentage of protein aligned to the best of homologs. PASA-improved transcripts were selected based on C-score, protein coverage, EST coverage, and its CDS overlapping with repeats. A transcript was selected if its C-score was ≥ 0.5 and protein coverage was ≥ 0.5, or it had EST coverage, but its CDS overlapping with repeats was less than 20%. For gene models whose CDSs overlapped with repeats for more that 20%, C-scores had to be at least 0.7 and homology coverage at least 70% to be selected. The selected gene models were subject to Pfam and Panther analysis and gene models with Pfam/Panther domains associated with transposable elements were removed. This generated annotations at 13,125 loci with 235 alternative splice forms (Table S6). Data were deposited to NCBI within the BioProject PRJNA234409 and as Whole Genome Shotgun project MXAK00000000m version MXAK01000000 to DDBJ/ENA/GenBank. The genome is also available from the JGI plant genomics portal Phytozome (https://phytozome.jgi.doe.gov).

Table S6. Annotation summary statistics

Protein homology was calculated by BLASTP (9) with an E-value cutoff of 1E-5 alignment coverage of the protein by proteins from UniRef90 (10) or other sequenced red algal genomes

Primary transcripts (loci)	13,125
Alternate transcripts	235
Total transcripts	13,360
For primary transcripts	
Average number of exons	1.9
Median exon length	579
Median intron length	203
Gene models with any EST overlap in CDS	7,816
Gene models with EST support over 100% of their lengths	4,413
Gene models with EST support over 90% of their lengths	4,977
Gene models with EST support over 75% of their lengths	5,436
Gene models with EST support over 50% of their lengths	6,145
Gene models with protein homology coverage of 100%	1,147
Gene models with protein homology coverage of over 90%	3,590
Gene models with protein homology coverage of over 50%	6,636
Gene models with any protein homology	8,206
For all gene models	
Number of complete genes (start to stop)	12,287
Number of genes with start but no stop	325
Number of genes with stop but no start	469
Predicted proteins with Pfam annotation	4,790
Predicted proteins with Panther annotation	4,314
Predicted proteins with KOG annotation	2,529
Predicted proteins with KEGG Orthology annotation	2,347
Predicted proteins with E.C. number annotation	2,595

The G + C percentage of the coding regions of the *Porphyra* genome is high, averaging ~72 %, but rising to gene models as high as ~94% (Figs. S7, S8, Table S9)). This made the sequencing process difficult, and it also made gene prediction harder, because the sequence had reduced complexity, complicating alignment and homology searches. Furthermore, gene features like splice sites and stop codons require A and T bases and their rarity in the genome causes gene prediction algorithms to over-predict exon sequence. Over-prediction of repeats and gene loci was reduced by modifying gene prediction parameters to make them more stringent after manual review of a curated set of nearly 100 genes. Where this improvement occasionally led to failure to predict gene models with strong support (e.g., by ESTs), we defined the model by its approximate region of homology to EST contigs, and such regions are listed in the annotation tables that follow by their scaffold positions.

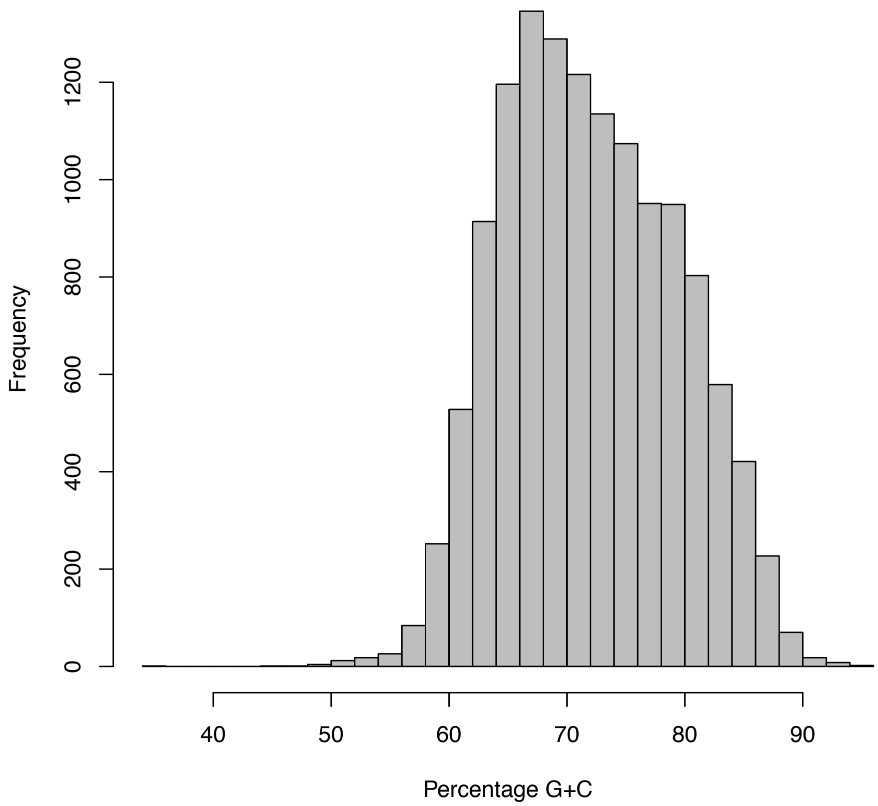


Figure S7: Histogram of G+C-content (%) in primary transcript sequences in *Porphyra* gene models.

Comparative analysis of red algal G + C content

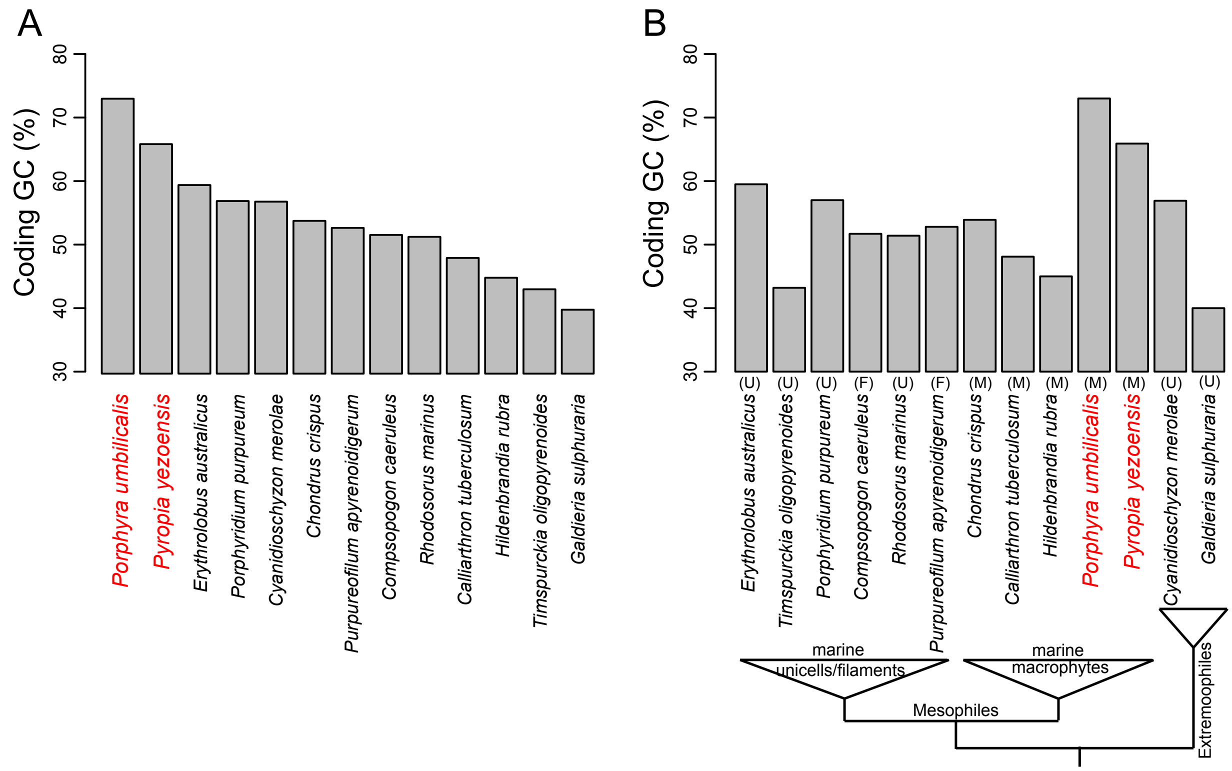


Figure S8: Coding G + C- contents in 13 red algal species.
Porphyra and *Pyropia* are shown in red. (A) Species are sorted by GC values in a descending order. (B) Species are ordered according to red algal phylogeny(29). Marine macrophyte representatives include species in Bangiophyceae and Florideophyceae. Marine unicellular representatives are *Erythrolobus australicus* (Porphyridiophyceae), *Timspurckia oligopyrenoides* (Porphyridiophyceae), *Porphyridium purpureum* (Porphyridiophyceae), and *Rhodosorus marinus* (Stylonematophyceae), and small filaments are *Compsopogon caeruleus* (Compsopogonophyceae) and *Purpureofilum apyrenoidigerum* (Stylonematophyceae). U: unicells; F: filaments; M: macrophytes.

Table S9: Sequence data for G + C studies

Genome and transcriptome data used to study coding region G + C-content and amino acid composition (see Fig. S8).

Taxon	Data type	References
Porphyra umbilicalis	Whole genome	This study
Pyropia yezoensis	Partial genome	(24)
Erythrolobus australicus	Transcriptome	(30)
Porphyridium purpureum	Whole genome	(22)
Purpureofilum apyrenoidigerum	Transcriptome	(31)
Cyanidioschyzon merolae	Whole genome	(19)
Chondrus crispus	Whole genome	(20)
Rhodosorus marinus	Transcriptome	(30)
Compsopogon caeruleus	Transcriptome	(30)
Calliarthron tuberculosum	Partial genome	(32)
Hildenbrandia rubra	Transcriptome	(31)
Timspurckia oligopyrenoides	Transcriptome	(30)
Galdieria sulphuraria	Whole genome	(23)

Analysis of genome completeness

In addition to analysis based on mapping ESTs to the assembly (see main text), a high degree of completeness is supported by careful examination of 41 informational genes and RNA polymerases I, II, and III (Table S10). The assembled *Porphyra* genome contains 40 of the 41 genes examined (Table S10), including a full complement of genes for all subunits of RNA polymerase I, II, and III.

Table S10. Analysis of presence of 41 core informational genes involved in transcription, translation and DNA replication.

Genes included were based initially on informational genes involved in transcription, translation, and DNA synthesis that were present in the CEGMA data set (33), with supplementation with RNA polymerase I, II and III subunits. All sequences were further screened to verify that they were carried as single-copy (no paralogous families) genes across eukaryotic diversity, with only recent, organism-specific duplications allowed. MCM4 is the one gene out of the 41 examined that is missing from the *Porphyra* genome. It is present in most other red algae, but also is missing from the *P. yezoensis* genome, meaning this could be a gene loss from the Bangiales rather than evidence of an incomplete genome.

Gene symbol	Defline	*Porphyra* gene	Length (aa)	Expression support	E-value
RPA1	RNA polymerase (RNAP) I, largest subunit	Pum0220s0021.1	1607	yes	4.0E-75
RPB1	RNAP II, largest subunit	Pum0177s0031.1	1561	yes	0
RPC1	RNAP III, largest subunit	Pum0177s0015.1	1592	yes	0
RPA2	RNAP I, 2nd largest subunit	Pum0107s0041.1	1274	yes	0
RPB2	RNAP II, 2nd largest subunit	Pum0087s0034.1	1235	yes	0
RPC2	RNAP III, 2nd largest subunit	Pum0451s0002.1	1233	yes	0
RPB3	RNAP II, 3rd largest subunit	Pum0112s0005.1	342	yes	9.0E-87
RPAC1	RNAP I, III, common subunit	Pum0711s0005.1	417	yes	2.0E-87
RPB4	RNAP II, subunit 4	Pum0260s0026.1	165	yes	4.0E-86
RPABC1	RNAP I, II, III, common subunit 1	Pum1325s0003.1	212	yes	6.0E-71
RPABC2	RNAP I, II, III, common subunit 2	Pum0292s0015.1	131	yes	1.0E-35
RPB7	RNAP II, subunit 7	Pum1752s0001.1	170	yes	4.0E-122
RPC8	RNAP III, subunit 8	Pum0260s0031.1	199	yes	6.0E-91
RPB11	RNAP II, subunit 11	Pum0108s0009.1	132	yes	2.00E-34
RPAC2	RNAP I, III, common subunit 2	Pum0161s0055.1	119	yes	1.0E-31
RPABC5	RNAP I, II, III, common subunit 5	Pum0183s0030.1	80	yes	2.0E-30
RPABC3	RNAP I, II, III, common subunit 3	Pum0108s0030.1	155	yes	1.0E-34
RPB9	RNAP II, subunit 9	Pum0408s0005.1	125	yes	9.0E-38
RPC11	RNAP III, subunit 11	Pum0082s0074.1	118	yes	6.0E-37
TFIIB	Transcription Factor IIB	Pum0120s0019.1	1017	yes	5.0E-24
TFIIH	Transcription Factor IIH, helicase	Pum0550s0007.1	856	yes	0
ELP3p *	Elongator complex protein 3	Pum0176s0001.1	268	yes	2.0E-140
ELP3p *	Elongator complex protein 3	Pum1773s0001.1	287	yes	2.0E-122
TBP	TATA-binding protein	Pum0161s0003.1	256	yes	2.0E-82
SPT5	Transcription elongation factor SPT5	Pum0196s0021.1	1207	yes	1.0E-148
TIF2	Translation initiation factor eIF4A	Pum0505s0005.1	309	yes	4.0E-157
TIF2 γ	Translation initiation factor eIF2, gamma	Pum0094s0031.1	465	yes	0
EF1α †	Elongation factor 1 alpha	Pum0224s0021.1	450	yes	0
EF1α †	Elongation factor 1 alpha	Pum0259s0019.1	450	yes	0
EF1α †	Elongation factor 1 alpha	Pum0224s0023.1	411	yes	0
EF2	Elongation factor 2	Pum0053s0005.1	842	yes	0
eRF3	Elongation release factor	Pum0507s0002.1	579	yes	0
MCM2	MCM DNA helicase complex, subunit 2	Pum2398s0001.1	424	yes	0
MCM3	MCM DNA helicase complex, subunit 3	Pum0446s0013.1	773	yes	0
MCM4 ‡	MCM DNA helicase complex, subunit 4	No gene model	NA	NO	NA
MCM5	MCM DNA helicase complex, subunit 5	Pum0736s0007.1	1028	yes	0
MCM6	MCM DNA helicase complex, subunit 6	Pum1395s0002.1	818	yes	0
MCM7	MCM DNA helicase complex, subunit 7	Pum2318s0001.1	655	yes	0
MCM8	MCM DNA helicase complex, subunit 8	Pum2064s0001.1	577	yes	6.0E-145
DNA-polε	DNA polymerase epsilon	Pum2553s0001.1	554	yes	5.0E-90
DNA-polδ	DNA polymerase delta	Pum0192s0005.1	932	yes	0
DNA-Topo IIIα	DNA topoisomerase III alpha	Pum0675s0004.1	1182	yes	0
DNA-Topo II	DNA topoisomerase II	Pum0313s0031.1	1427	yes	0
LigI	DNA ligase I	Pum0072s0080.1	750	yes	0

* The ELP3 gene in *Porphyra* is split into two non-overlapping gene models, one for the N-terminal and one for the C-terminal regions of the gene. Both models occur at ends of scaffolds, in orientations that would allow them to form a contiguous reading frame if assembled together. A nearly complete gene appears to be present between the two gene models.

† The Pum0224s0021.1 and Pum0259s0019.1 gene models are nearly identical, but on different scaffolds. Pum0224s0021.1 and Pum0224s0023.1 are adjacent on the same scaffold but interrupted by one gene model, which appears to encode an endonuclease. Pum0224s0023.1 is flanked on the other side by a gene model that appears to be a copy of the intervening endonuclease, but is missing some sequence and has only a small region of EST support.

‡ MCM4 is the one gene out of the 41 examined that is missing from the *Porphyra* genome. It is present in most other red algae, but also is missing from the *P. yezoensis* genome, meaning this could be a gene loss from the Bangiales rather than evidence of an incomplete genome.

Chloroplast genome

The full chloroplast genome (190,173 bp) was assembled (GenBank Accession MF385003), and we predicted an additional gene (orf74) compared to the existing partial plastid genome for *Porphyra umbilicalis* (34). Interestingly, homologs to orf74 by BLAST were limited to plastid genomes of other bangiophyte red algae.

New PacBio sequences (21,380 reads) were mapped to JQ408795.1 using CLC Genomics Workbench v8.5 (https://www.qiagenbioinformatics.com/) with the following parameters: Masking mode = No masking, Update contigs = No, Mismatch cost = 2, Cost of insertions and deletions = Linear gap cost, Insertion cost = 3, Deletion cost = 3, Insertion open cost = 6, Insertion extend cost = 1, Deletion open cost = 6, Deletion extend cost = 1, Length fraction = 0.8, Similarity fraction = 0.8, Global alignment = No, Non-specific match handling = Map randomly. This resulted in 955x coverage of the incomplete plastid sequence. Where ten or more reads disagreed with the partial plastid sequence, the sequence was edited (34 edits). The two gaps in JQ408795.1 (36,278- 36,578bp, 138,849-138,948) were closed, and the Geneious V9.0 ORF tool (translation table 11) predicted the new open reading frame (orf74) in the complete plastid; its start position had been masked by a gap in the partial plastid sequence (36,278 - 36,578 bp in JQ408795.1).

The nuclear genome was scanned for nuclear plastid DNA-like sequences (NUPTs) and nuclear plastid DNA-like sequences (NUMTs) with BLASTn v2.3.0 (9). The mitochondrial (NC_018544.1) and herein completed plastid genomes were used as query sequences with the following parameters: E-value = 10^{-10}; match 2; mismatch −3; gap open = 5; gap extend = 2; all other parameters were set to default values. Regions of nuclear DNA that contained tight clusters of organelle-like DNA interrupted by genomic sequence that did not show sequence identity to organellar DNA were not counted as a single NUPT or NUMT but as separate hits. Spurious hits, such as those where organelle genes matched to homologous genes in the nuclear genome (*e.g.*, organelle ribosomal DNA matching nuclear ribosomal DNA) were omitted. One NUPT and 7 NUMTs were identified in *P. umbilicalis* and their total length is ≤ 2 Kbp. These results agree with earlier investigations where monoplastidic plant and algal taxa exhibited no or few NUPTs and where the number of NUMTs correlated to both numbers of mitochondria and genome size (35, 36).

Relationships of red algal classes

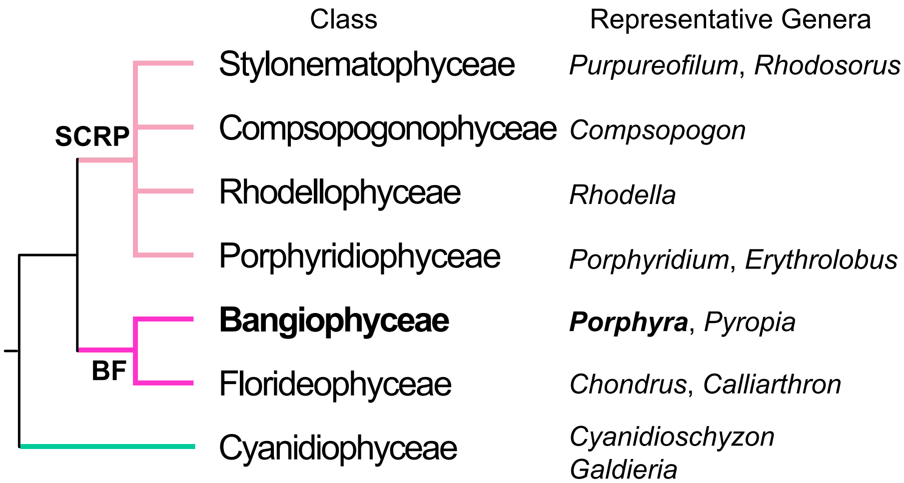


Figure S11. Relatedness of red algal classes

Phylogenomic-based classification of the Rhodophyta reveals a basal clade of extremophilic species (Cyanidiophyceae) and two sister clades (herein: SCRP, BF) of mesophilic species (29, 37). Most described red algal species belong to the BF clade; note position of *Porphyra*. Red shading of SCRP and BF is arbitrary, but blue-green is used for Cyanidiophyceae because members lack phycoerythrin as an accessory photosynthetic pigment and are blue-green.

Evolutionary rate comparison

To compare the evolutionary rates across species, we constructed orthologous gene families for a total of 21 red algal, Viridiplantae, and metazoan taxa (Fig. S12 and Table S13 using OrthoMCL (38) under the default setting with modifications (valueExponentCutoff = −10 and percentMatchCutoff = 40).

Sequences were retrieved for each of the 300 single-copy gene families and aligned using MUSCLE (39) under the default settings. These sequence alignments were trimmed using TrimAl (-automated) (40) and further ‘polished’ using T-Coffee (41). The resulting sequence alignments were then concatenated (101,371 amino acids) and used for tree construction under maximum likelihood using RAxML (42) under the LG + Γ + F model.

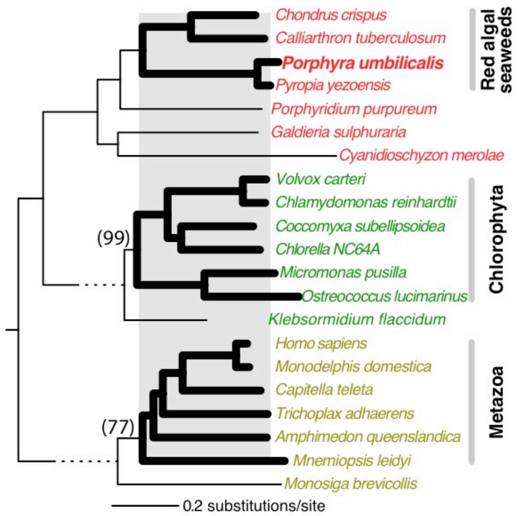


Figure S12. Maximum likelihood tree of red algae, green algae, and metazoans.
All interior nodes are supported with 100 % bootstrap support unless indicated (in parentheses). The clade defining red algal seaweeds is shown with the thick branches, as are the clades of similar evolutionary depth in green algae (Chlorophyta) and metazoans. The gray box indicates the putative split time for these three major groups. The broken lines indicate branches with transformed lengths in order to align the three major groups. Red text: red algae; green text: Viridiplantae; olive text: Opisthokonta.

Table S13. Genome data sets for evolutionary rate comparison

Genome data used for the evolutionary rate comparison (see Figure S12).

Taxon	Data type	References
Porphyra umbilicalis	Whole genome	This study
Pyropia yezoensis	Partial genome	(24)
Chondrus crispus	Whole genome	(20)
Porphyridium purpureum	Whole genome	(22)
Galdieria sulphuraria	Whole genome	(23)
Cyanidioschyzon merolae	Whole genome	(19)
Calliarthron tuberculosum	Partial genome	(32)
Chlamydomonas reinhardtii	Whole genome	(43)
Volvox carteri	Whole genome	(44)
Chlorella variabilis	Whole genome	(45)
Coccomyxa subellipsoidea	Whole genome	(46)
Micromonas pusilla CCMP1545	Whole genome	(47)
Ostreococcus lucimarinus	Whole genome	(48)
Klebsormidium flaccidium	Whole genome	(49)
Homo sapiens	Whole genome	(50)
Monodelphis domestica	Whole genome	(51)
Capitella teleta	Whole genome	(52)
Trichoplax adhaerens	Whole genome	(53)
Amphimedon queenslandica	Whole genome	(54)
Mnemiopsis leidyi	Whole genome	(55)
Monosiga brevicollis	Whole genome	(56)

Meiotic genes

Of the nine most common meiosis-specific genes in eukaryotic sexual organisms (57), SPO11, DMC1, HOP1, HOP2, MND1, MSH4, MSH5, REC8, and MER3, all but three (DMC1, REC8 and MER3) were found in the *Porphyra umbilicalis* genome. In addition, *P. umbilicalis* appears to lack five other genes that function in many other eukaryotes for DNA repair and mitosis in addition to meiosis: MLH2, MLH3, PDS5, RAD21, and SSC3. However, the significance of this requires additional study because some of these genes are not present in clearly sexual red algae (*Chondrus* (20)), and it is possible that other proteins support these functions in *Porphyra*.

Table S14. List of meiotic genes in *Porphyra*

Porphyra gene	Gene symbol	aliases	Defline	Description
Pum0507s0009.1	HOP1.1	HOP	Hsp70-Hsp90 organizing protein	Binds DSBs and oligomerizes early during meiotic prophase. PMID: 18663385
Pum0532s0010.1	HOP1.2	HOP	Hsp70-Hsp90 organizing protein	Binds DSBs and oligomerizes early during meiotic prophase. PMID: 18663385
Pum0207s0034.1	HOP2	AHP2; T6J4.9	Homologous-pairing protein 2 homolog	Ensures accurate and efficient homology searching in meiotic prophase. PMID: 18663385
Pum2135s0001.1	MLH1	COCA2	DNA mismatch repair protein MutL-homolog 1	Heterodimerizes with MLH3 to form MutL gamma which plays a role in meiosis.
Pum0280s0004.1	MND1	F19B15.200; GAJ	Meiotic nuclear division protein 1 homolog	Required for stable heteroduplex DNA formation. PMID: 18663385
Pum1900s0001.1	MRE11		Double-strand break repair protein MRE11	Component of the MRN complex, which plays a central role in double-strand break (DSB) repair, DNA recombination, maintenance of telomere integrity and meiosis.
Pum0982s0002.1	MSH2	MYF24.25	DNA mismatch repair protein Msh2; MutS protein homolog 2	Component of the post-replicative DNA mismatch repair system.
Pum0229s0006.1	MSH4		DNA mismatch repair protein Msh4; MutS protein homolog 4	Directs Holliday junction resolution towards crossover with interference. PMID: 18663385
Pum0953s0002.1	MSH5		DNA mismatch repair protein Msh5; MutS protein homolog 5	Directs Holliday junction resolution towards crossover with interference. PMID: 18663385
Pum1300s0002.1	MSH6	GTBP	DNA mismatch repair protein Msh6; MutS protein homolog 6	Binds base-base mismatches. PMID: 18663385
Pum0200s0021.1	PMS1.1	T14P8.6	DNA mismatch repair protein PMS1	Plays a major role in the repair of heteroduplex sites present in meiotic recombination intermediates.
Pum0200s0022.1	PMS1.2	T14P8.6	DNA mismatch repair protein PMS1	Plays a major role in the repair of heteroduplex sites present in meiotic recombination intermediates.
Pum0125s0033.1	RAD1	REC1	Cell cycle checkpoint protein RAD1	Plays a role in DNA repair. PMID: 15489334
Pum0075s0001.1	RAD50		DNA repair protein RAD50	Dimer holds broken DNA ends together. PMID: 18663385
Pum1410s0001.1	RAD52	RHM52	Recombination DNA repair protein	Initiates DSB repair by homologous recombination. PMID: 18663385
Pum0306s0030.1	SMC1	PSM1	Structural maintenance of chromosomes protein 1	Forms core sister chromatid cohesin subunits. PMID: 18663385
Pum0532s0008.1	SMC2		Structural maintenance of chromosomes protein 2	Essential for chromosome assembly and segregation. PMID: 18663385
Pum0082s0046.1	SMC3	SMC3l1; CSPG6	Structural maintenance of chromosomes protein 3	Forms core sister chromatid cohesin subunits. PMID: 18663385
Pum0256s0001.1	SMC4.1		Structural maintenance of chromosomes protein 4	Essential for chromosome assembly and segregation. PMID: 18663385
Pum0256s0002.1	SMC4.2		Structural maintenance of chromosomes protein 4	Essential for chromosome assembly and segregation. PMID: 18663385
Pum0209s0016.1	SMC5	EMB2782; F1N13.60	Structural maintenance of chromosomes protein 5	Involved in DNA repair and checkpoint responses. PMID: 18663385
Pum0374s0018.1	SMC6	SMC6l1	Structural maintenance of chromosomes protein 6	Binds ssDNA, involved in DNA repair & checkpoint responses. PMID: 18663385
Pum1006s0002.1	SPO11	TOP6A3; TOP6A	Meiotic recombination protein SPO11	Transesterase, creates ds-DNA breaks in homologous chromosomes. PMID: 18663385

Analysis of *Porphyra*'s cytoskeleton

Table S15. Table of cytoskeletal genes found in *Porphyra*

Porphyra gene identifier	Gene Symbol	Aliases or Synonyms	Definition	Description	PubMed ID for key references
Pum0076s0028.1	TubA		Tubulin alpha chain	Alpha subunit of tubulin dimer in microtubules	11441808
Pum0536s0008.1	TubB		Tubulin beta chain	Beta subunit of tubulin dimer in microtubules	11441808
Pum0608s0003.1	TubG		Tubulin gamma chain	MTOC-associated tubulin subunit involved in microtubule nucleation	24075308
scaffold_177:136569-139713	GCP		Gamma complex protein	MTOC-associated protein involved in microtubule nucleation	26316498
Pum0215s0033.1	CofB	cofactor B	tubulin folding cofactor B	Protein involved in tubulin folding	25908846
Pum1971s0002.1	EB1		End binding protein 1	Microtubule plus-end tracking protein	18322465
Pum1280s0002.1	MAP215	MOR1; DIS1;		Microtubule plus-end tracking protein that interacts with EB1 and promotes microtubule	18322465

		CKAP5; ch-TOG		assembly	
Pum0076s0027.1	Centrin	caltractin	Centrin	Ca2+ activated contractile protein, MTOC component; possible location of polar ring	8175926
Pum0144s0010.1	Cmd.1		Calmodulin	Calcium binding protein calmodulin	26528296
Pum0033s0060.1	Cmd.2		Calmodulin	Calcium binding protein calmodulin	26528296
Pum0060s0052.1	Cmd.3		Calmodulin	Calcium binding protein calmodulin	26528296
Pum0031s0074.1	CaN		Cancineurin	Calcium and calmodulin dependent serine/threonine protein phosphatase	25655284
Pum0235s0011.1	CmdR		Calmodulin related protein	EF_hand containing protein related to calmodulin	27462077
Pum0167s0024.1	Kin5.1	BimC	kinesin 5	Kinesin 5 family member, predicted to be a mitotic spindle motor that is homotetrameric and slides antiparallel MTs.	17652157
Pum0032s0002.1	Kin5.2	BimC	kinesin 5	Kinesin 5 family member, predicted to be a mitotic spindle motor that is homotetrameric and slides antiparallel MTs.	17652157
Pum0262s0026.1	Kin7	CENP-E	kinesin 7	Kinesin 7 (CENP-E) family member, predicted to be a mitotic kinesin that localizes to centromeres.	18427114
Pum0517s0010.1	Kin13.1	MCAK, Kif2	kinesin family member	Kinesin 13 family member, predicted to be a mitotic motor that drives microtubule depolymerization	20109574
Pum0535s0005.1	Kin13.2	MCAK, Kif2	kinesin family member	Divergent kinesin 13 family member, predicted to be a mitotic motor that drives microtubule depolymerization	20109574
Pum0229s0039.1	Kin13.3	MCAK, Kif2	kinesin family member	Divergent kinesin 13 family member, predicted to be a mitotic motor that drives microtubule depolymerization	20109574
Pum0401s0008.1	Kin14	NCD, Kar3	kinesin 14	Kinesin 14 (C-terminal) family member; predicted to be mitotic motor that drives microtubule depolymerization, but in plants has been reported to be involved in membrane transport	26322239
Pum0640s0007.1	DyLC8		Dynein light chain, cytoplasmic	Protein that serves as a dynein light chain but is conserved in many organisms (e.g., land plants) lacking cytoplasmic dynein	27964786
Pum0014s0123.1	TCTEX		Dynein, light chain, Tctex-type 1, cytoplasmic	Protein that serves as a dynein light chain but may also have dynein-independent functions at the kinetochore	25928583
Pum1561s0001.1	Act.1		Conventional actin	Cytoskeletal protein that forms microfilaments	8590468
Pum0070s0053.1	Act.2		Conventional actin	Cytoskeletal protein that forms microfilaments	8590468
Pum0231s0007.1	Act.3		Conventional actin	Cytoskeletal protein that forms microfilaments	8590468
Pum0031s0095.1	Act.4		Conventional actin	Cytoskeletal protein that forms microfilaments	8590468
Pum0233s0019.1	ARP4.1	BAF53	Actin-related protein Arp4	Nuclear actin-related protein involved in chromatin remodeling	15743449
Pum1930s0001.1	ARP4.2	BAF53	Actin-related protein Arp4	Nuclear actin-related protein involved in chromatin remodeling	15743449
Pum0387s0002.1	ARP4.3	BAF53	Actin-related protein Arp4	Nuclear actin-related protein involved in chromatin remodeling	15743449
Pum0299s0010.1	Form.1		Formin-related protein	Actin filament nucleator	18977251
Pum0356s0010.1	Form.2		Formin-related protein	Actin filament nucleator	18977251
Pum0499s0004.1	Prof.1		Profilin	Binds to G-actin and speeds microfilament elongation	11732068
Pum0053s0016.1	Prof.2		Profilin	Binds to G-actin and speeds microfilament elongation	11732068
Pum0165s0012.1	Porf.3		Profilin	Binds to G-actin and speeds microfilament elongation	11732068
Pum0070s0054.1	ADF.1		Cofilin	Promotes actin filament disassembly	11732068
Pum0190s0014.1	ADF.2		Cofilin	Promotes actin filament disassembly	11732068
Pum1553s0002.1	Seve.1		Severin	Gelsolin-related protein predicted to depolymerize actin filaments	11732068
Pum2419s0001.1	Seve.2		Severin	Gelsolin-related protein predicted to depolymerize actin filaments	11732068
Pum0427s0014.1	Sept.1		septin	binds GTP; predicted to form filaments and be involved in cytokinesis	25957401
Pum0452s0003.1	Sept.2		septin	binds GTP; predicted to form filaments and be involved in cytokinesis	25957401

Table S16. Presence/absence of selected cytoskeletal genes in the red algae *Porphyra*, *Cyanidioschyzon*, *Galdieria, Porphyridium*, and *Chondrus.*

Gene identifiers are shown in the body of the table. For the *Porphyra* gamma tubulin complex gene, repetitive sequence at the locus prevented automated gene prediction, but the region of homology on scaffold_177 is shown. N.F. = not found.

Protein Class or name	Description	*Porphyra umbilicalis*	*Porphyridium purpureum*	*Galdieria sulphuraria*	*Cyanidioschyzon merolae*	*Chondrus crispus*
Microtubule Structure						
Alpha tubulin	alpha tubulin subunit of alpha-beta tubulin dimer	Pum0076s0028.1	Porphyridium.evm.model.contig_448.6	XP_005704892.1	Phy000UGHU_CYAME	Phy0034NN8_CHOCR_GSCHC2T00004954001, Phy0034QHQ_CHOCR_GSCHC2T00011873001
Beta tubulin	beta tubulin subunit of alpha-beta tubulin dimer	Pum0536s0008.1	Porphyridium.evm.model.contig_2073.5	XP_005707937.1, XP_005707936.1, XP_005703597.1, XP_005708540.1, XP_005703595.1, XP_005707934.1, XP_005705383.1	Phy000UF3H_CYAME	Phy0034OFO_CHOCR_GSCHC2T00006992001, Phy0034PEN_CHOCR_GSCHC2T00009280001
Gamma tubulin	gamma tubulin subunit involved in microtubule nucleation	Pum0608s0003.1	Porphyridium.evm.model.contig_3412.2	XP_005706854.1	Phy000UF4B_CYAME	Phy0034QIB_CHOCR_GSCHC2T00011912001
Microtubule nucleation						
Gamma complex protein	gamma complex protein involved in microtubule nucleation	scaffold_177:136569..139713	Porphyridium.evm.model.contig_2494.36, Porphyridium.evm.model.contig_444.12	XP_005706628.1, XP_005706682.1	Phy000UECH_CYAME CMJ255C, Phy000UEYX_CYAME CMN041C	Phy0034M98_CHOCR_GSCHC2T00001390001, Phy0034NQY_CHOCR_GSCHC2T00005200001
Microtubule folding						
Cofactor A	Tubulin folding cofactor A	**N.F.**	Porphyridium.evm.model.contig_4434.3	XP_005708190.1	**N.F.**	**N.F.**
Cofactor B	tubulin folding cofactor B	Pum0215s0033.1	Porphyridium.evm.model.contig_444.24	XP_005705747.1	Phy000UD48_CYAME	Phy0034M1D_CHOCR_GSCHC2T00000902001
Cofactor D	tubulin folding cofactor D	Pum0374s0020.1	Porphyridium.evm.model.contig_4446.17	XP_005708084.1	Phy000UDMV_CYAME	Phy0034PWU_CHOCR_GSCHC2T00010467001
Cofactor E	tubulin folding cofactor E	**N.F.**	**N.F.**	XP_005705826.1	Phy000UDQF_CYAME CMG030C	Phy0034LPZ_CHOCR_GSCHC2T00000186001
Microtubule plus-end tracking proteins						
EB1	microtubule plus-end tracking protein "end binding protein 1"	Pum1971s0002.1	Porphyridium.evm.model.contig_570.4	XP_005706580.1	Phy000UEA9_CYAME CMJ156C	**N.F.**
XMAP215; Also known as MOR1, DIS1, CKAP5, ch-TOG	microtubule plus-end tracking protein that interacts with EB1 and promote microtubule assembly.	Pum1280s0002.1	Porphyridium.evm.model.contig_3543.1	XP_005708148.1	Phy000UF92_CYAME CMO196C	Phy0034RIB_CHOCR_GSCHC2T00014443001
CLASP	cytoplasmic linker associated protein	**N.F.**	Porphyridium.evm.model.contig_4458.3, Porphyridium.evm.model.contig_4467.12	XP_005704018.1, XP_005705964.1	Phy000UDGY_CYAME CMD154C	Phy0034Q6Z_CHOCR_GSCHC2T00011051001
Microtubule Cross-linking						
MAP65/ASE1/PRC1	microtubule cross-linking protein invovled in mitosis and microtubule organization	**N.F.**	Porphyridium.evm.model.contig_3488.5	XP_005703277.1	Phy000UFF4_CYAME CMP130C	**N.F.**
Calmodulin and its relatives						
Centrin; also known as caltractin	calmodulin-relative involved in microtubule nucleation and other processes	Pum0076s0027.1	Porphyridium.evm.model.contig_510.18	XP_005708752.1	**N.F.**	Phy0034SN7_CHOCR_GSCHC2T00017304001
Calmodulin	calcium binding protein calmodulin	Pum0144s0010.1 Pum0033s0060.1 Pum0060s0052.1	Porphyridium.evm.model.contig_4427.8, Porphyridium.evm.model.contig_439.6, Porphyridium.evm.	XP_005704341.1, XP_005704340.1, XP_005706567.1	Phy000UDIE_CYAME CME034C, Phy000UEI9_CYAME CMK219C	Phy0034QOM_CHOCR_GSCHC2T00012343001

			model.contig_2323.3, Porphyridium.evm.model.contig_4428.4			
Calcineurin B	calcium and calmodulin dependent serine/threonine protein phosphatase	Pum0031s0074.1	Porphyridium.evm.model.contig_2033.20	XP_005706567.1,	Phy000UECQ_CYAME	Phy0034RM6_CHOCR_GSCHC2T00014698001
Calmodulin related proteins	EF_hand containing protein related to calmodulin	Pum0235s0011.1	Porphyridium.evm.model.contig_3451.8, Porphyridium.evm.model.contig_2139.4, Porphyridium.evm.model.contig_4426.5, Porphyridium.evm.model.contig_2211.7, Porphyridium.evm.model.contig_2024.16, Porphyridium.evm.model.contig_2290.1, Porphyridium.evm.model.contig_2025.58, Porphyridium.evm.model.contig_2296.12, Porphyridium.evm.model.contig_4408.9	XP_005705171.1	Phy000UFFY_CYAME CMP171C	Phy0034LWX_CHOCR_GSCHC2T00000640001,
Kinesin motors found in at least one red alga*						
Kinesin 5; also known as BimC	kinesin 5 family member, predicted to be a mitotic spindle motor that is homotetrameric and slides antiparallel MTs.	Pum0167s0024.1 Pum0032s0002.1	PorphyridiumContig_2090.10	GasuXP_005706647.1, GasuXP_005706648.1	Phy000UDD8_CYAME	ChcrPhy0034SO2
Kinesin 7; also known as CENP-E	kinesin 7 (CENP-E) family member, predicted to be a mitotic kinesin that localizes to centromeres.	Pum0262s0026.1	PorphyridiumContig_2065.3	GasuXP_005704229.1	CymePhy000UFSN	ChcrPhy0034R7X
Kinesin 13; also known as MCAK, Kif2	kinesin 13, predicted to be a mitotic motor that drives microtubule depolymerization	Pum0517s0010.1, Pum0535s0005.1, Pum0229s0039.1	**N.F.**	**N.F.**	**N.F.**	**N.F.**
Kinesin 14; also known as NCD and C-terminal kinesin	kinesin 14 (C-terminal) family member; predicted to be mitotic motor that drives microtubule depolymerization	Pum0401s0008.1	PorphyridiumContig_3395.13, PorphyridiumContig_2100.14	GasuXP_005707452.1, gi|545706185|ref|XP_005705410, gi|546313465|ref|XP_005715293	CymePhy000UG3D_CYAME, CymePhy000UGMH	ChcrPhy0034N4O
Kinesin 4; also known as chromo kinesin	kinesin 4 family member; predicted to be involved in mitosis and perhaps vesicle transport	**N.F.**	PorphyridiumContig_4446.19	GasuXP_005705866.1, GaSuXP_005704836.1	**N.F.**	**N.F.**
Kinesin 6	kinesin 6 family member; predicted to be involved in cell division	**N.F.**	PorphyridiumContig_4416.18	GasuXP_005707198.1	**N.F.**	ChcrPhy0034SNV
Kinesin 15	kinesin 15 family member	**N.F.**	**N.F.**	GaSuXP_005707115.1, gi|545703096|ref|XP_005703875	Cyame_CMO070C	**N.F.**
Dynein and associated proteins						
Dynein heavy chain	cytoplasmic dynein HC 1	**N.F.**	**N.F.**	**N.F.**	**N.F.**	XP_005717072.1

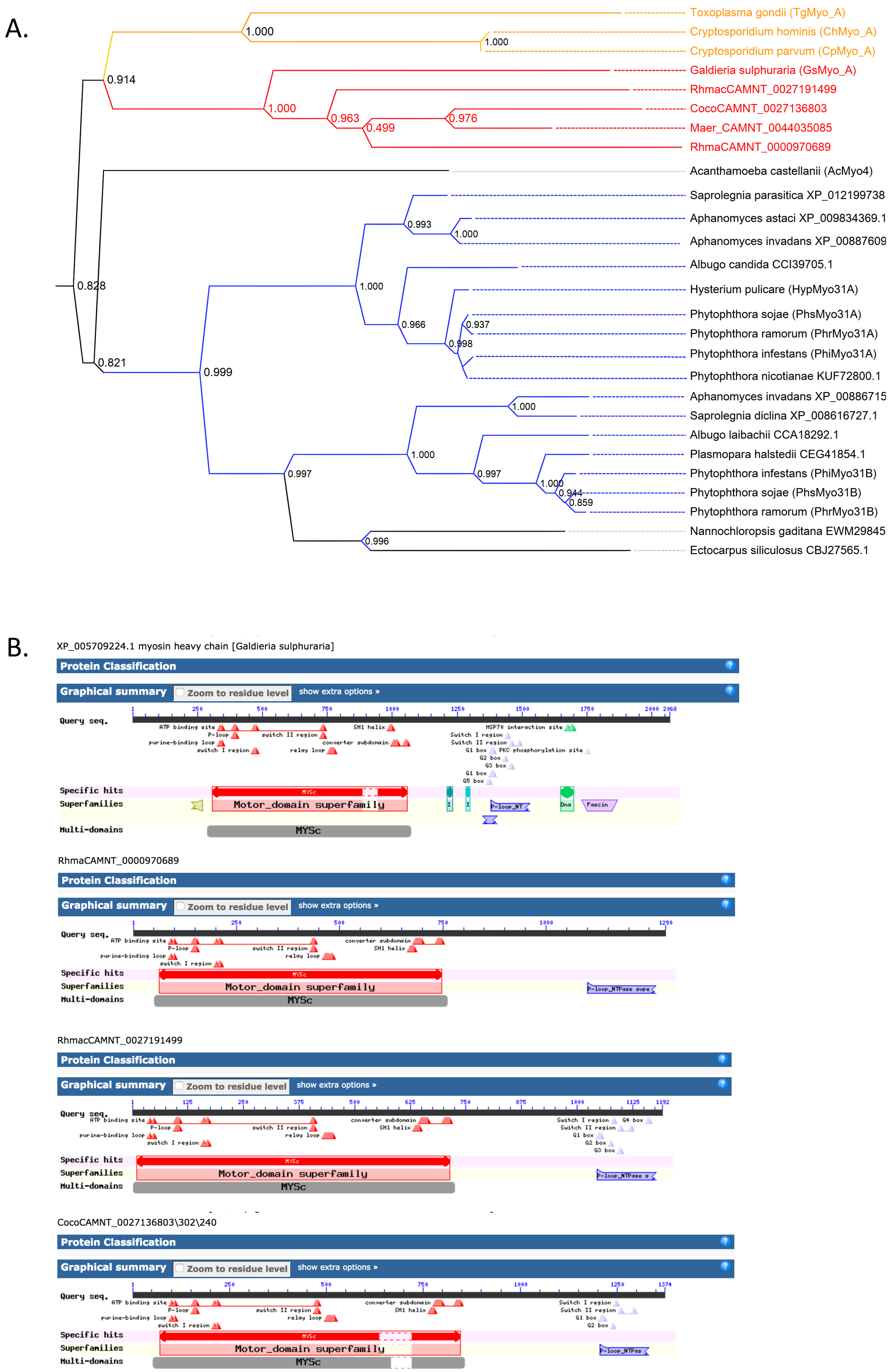
Dynein LC8	dynein light chain, 8kD, also involved non-dynein functions	Pum0640s0007.1	Porphyridium.evm.model.contig_2017.6	XP_005707601.1	Phy000UGEH_CYAME	Phy0034Q64_CHOCR_GSCHC2T00011009001
Actin superfamily						
Actin	Actin filament protein subunit	Pum1561s0001.1, Pum0070s0053.1, Pum0031s0095.1, Pum0231s0007.1	Porphyridium.evm.model.contig_3739.1	XP_005709337.1, XP_005707846.1, XP_005707521.1, XP_005705473.1	Phy000UEW4_CYAME	Phy0034QL5_CHOCR_GSCHC2T00012135001, Phy0034PGJ_CHOCR_GSCHC2T00009415001
Arp4	Nuclear actin-related protein Arp4; predicted to be involved in chromatin remodeling	Pum0233s0019.1, Pum0387s0002.1, Pum1930s0001.1,	Porphyridium.evm.model.contig_3453.2, Porphyridium.evm.model.contig_2122.24, Porphyridium.evm.model.contig_3422.10, Porphyridium.evm.model.contig_2345.4, Porphyridium.evm.model.contig_2156.3	XP_005708436.1, XP_005704222.1, XP_005708774.1, XP_005706259.1	Phy000UGC4_CYAME, Phy000UELG_CYAME, Phy000UG7F_CYAME, Phy000UFYV_CYAME, Phy000UEBS_CYAME	Phy0034RB1_CHOCR_GSCHC2T00013862001, Phy0034MZQ_CHOCR_GSCHC2T00003301001, Phy0034QC7_CHOCR_GSCHC2T00011459001
Actin nucleating proteins						
Formin	protein that nucleates linear arrays of actin filaments	Pum0299s0010.1, Pum0356s0010.1	Porphyridium.evm.model.contig_503.5, Porphyridium.evm.model.contig_565.1, Porphyridium.evm.model.contig_2285.5, Porphyridium.evm.model.contig_4458.10, Porphyridium.evm.model.contig_2495.1	XP_005709381.1, XP_005708214.1, XP_005703173.1, XP_005703511.1	Phy000UEZ2_CYAME, Phy000UF2I_CYAME, Phy000UDN7_CYAME	Phy0034RRW_CHOCR_GSCHC2T00015089001
Regulators of actin polymerization						
Profilin	actin monomer binding and capping protein profilin	Pum0499s0004.1, Pum0053s0016.1, Pum0165s0012.1	Porphyridium.evm.model.contig_4437.11, Porphyridium.evm.model.contig_2058.1	XP_005708857.1, XP_005707058.1	Phy000UFH0_CYAME	Phy0034R4T_CHOCR_GSCHC2T00013528001
Cofilin; also known as actin depolymerizing factor (ADF)	actin filament depolymerizing factor cofilin	Pum0070s0054.1, Pum0190s0014.1	Porphyridium.evm.model.contig_2093.3, Porphyridium.evm.model.contig_3633.1	XP_005702926.1	Phy000UF12_CYAME	Phy0034MBR_CHOCR_GSCHC2T00001585001, Phy0034PHU_CHOCR_GSCHC2T00009495001
Severin	actin-regulatory protein related to gelsolin and villin; contains 3 gelsolin-homology domains	Pum1553s0002.1, Pum2419s0001.1	Porphyridium.evm.model.contig_2167.3, Porphyridium.evm.model.contig_2035b.39	**N.F.**	**N.F.**	Phy0034SG0_CHOCR_GSCHC2T00016843001
Myosin†						
myosin heavy chain	unclassified myosin motor	**N.F.**	**N.F.**	XP_005709224.1	**N.F.**	**N.F.**
Septin						
Septin	septin GTPase; predicted to be a GTPase that forms filaments	Pum0427s0014.1, Pum0452s0003.1	Porphyridium.evm.model.contig_3450.2	**N.F.**	**N.F.**	**N.F.**

Notes:

*Although this table focuses on the set of fully sequenced red algal genomes listed above, we note that kinesin 4 is also found in the transcriptome of *Rhodosorus marinus* (CAMNT_0000974579) and *Compsopogon caeruleus* (CAMNT_0027137013). Kinesin 4 is of particular interest because it has been described as a membrane transport kinesin in plants (58).

† The various transcriptomes of red algae contain several myosins. Some of these seem likely to be contaminants because they are closely related to myosins in other organisms (e.g., Amoebazoa) (not shown). However, myosins from some SCR (SCRP clade) red algae group phylogenetically with the *Galdieria* myosin

(Fig. S17), suggesting that these are *bona fide* red algal myosins. These myosin proteins are: *Compsopogon caeruleus* (CAMNT_0027136803), *Madagascaria erythrocladiodes* (CAMNT_0044035085), *Rhodosorus marinus* (CAMNT_0000970689), and *Rhodella maculata* (CAMNT_0027191499). At present, these motors are classified as orphans, but they show some similarity to the myosin 27 proteins that act as membrane transport motors in apicomplexans (59).



A.
Toxoplasma gondii (TgMyo_A)
Cryptosporidium hominis (ChMyo_A)
Cryptosporidium parvum (CpMyo_A)
Galdieria sulphuraria (GsMyo_A)
RhmacCAMNT_0027191499
CocoCAMNT_0027136803
Maer_CAMNT_0044035085
RhmaCAMNT_0000970689
Acanthamoeba castellanii (AcMyo4)
Saprolegnia parasitica XP_012199738
Aphanomyces astaci XP_009834369.1
Aphanomyces invadans XP_00887609
Albugo candida CCI39705.1
Hysterium pulicare (HypMyo31A)
Phytophthora sojae (PhsMyo31A)
Phytophthora ramorum (PhrMyo31A)
Phytophthora infestans (PhiMyo31A)
Phytophthora nicotianae KUF72800.1
Aphanomyces invadans XP_00886715
Saprolegnia diclina XP_008616727.1
Albugo laibachii CCA18292.1
Plasmopara halstedii CEG41854.1
Phytophthora infestans (PhiMyo31B)
Phytophthora sojae (PhsMyo31B)
Phytophthora ramorum (PhrMyo31B)
Nannochloropsis gaditana EWM29845
Ectocarpus siliculosus CBJ27565.1
1.000
1.000
0.914
1.000
0.963
0.976
0.499
0.993
1.000
0.828
1.000
0.966
0.937
0.998
0.821
0.999
1.000
1.000
0.997
0.997
1.000
0.944
0.859
0.996
B.
XP_005709224.1 myosin heavy chain [Galdieria sulphuraria]
Protein Classification
Graphical summary
Zoom to residue level
show extra options »
Query seq.
Specific hits
Superfamilies
Multi-domains
MYSc
Motor_domain superfamily
RhmaCAMNT_0000970689
Protein Classification
Graphical summary
Zoom to residue level
show extra options »
Query seq.
Specific hits
Superfamilies
Multi-domains
MYSc
Motor_domain superfamily
RhmacCAMNT_0027191499
Protein Classification
Graphical summary
Zoom to residue level
show extra options »
Query seq.
Specific hits
Superfamilies
Multi-domains
MYSc
Motor_domain superfamily
CocoCAMNT_0027136803\302\240
Protein Classification
Graphical summary
Zoom to residue level
show extra options »
Query seq.
Specific hits
Superfamilies
Multi-domains
MYSc
Motor_domain superfamily

Figure S17. Phylogenetic analysis of red algal myosins.

A: Branch of a much larger phylogenetic tree (> 1900 myosin motor domain sequences, not shown) that indicates with high confidence that the red algal myosins (red) form a clade that excludes myosins from other organisms; this supports the idea that these are legitimate red algal myosins. A group of apicomplexan myosins (orange) branch with the red algal myosins with lesser but still high confidence. In this tree, the red algal myosins branch near a class of oomycete myosins (blue), but this relationship varies with the tree-building parameters and so is less robust (not shown). **B:** The red algal myosins share a common tail element, a nucleotide-binding P-loop. The function for this domain in the context of a myosin tail domain is not known, but similar P-loops are found in the tails of some myosin V proteins (e.g., human myosin Vc), proteins that are involved in membrane transport.

Methods: For (A), we used as a published core dataset the myosin motor domain alignment (60), which included the *Galdieria sulphuraria* myosin (GsMyoA, XP_005709224.1). The names used by Odronitz and Kollmar (60) are given in parentheses. We added to this dataset the red algal myosins identified by TBLASTN (9) search of the transcriptomes of *Compsopogon caeruleus* (Coco0027136803), *Madagascaria erythrocladiodes* (Maer_0044035085*), Rhodella maculata* (Rhmac0027191499), and *Rhodosorous marinus* (Rhma0000970689), as well as some other myosins indicated by accession numbers. The remainder of the sequences shown here and in the rest of the tree were from the supplement (60). The new sequences were trimmed to the core motor domain as previously defined (60), then all gaps were removed from the aligned sequences and subjected to alignment and tree-building using the software package PASTA (61) under default settings. The resulting tree was visualized by Archaeopteryx (62), and it agreed well with the myosin classifications as defined by Odronitz and Kollmar (60). The subtree shown in (A) was edited to add color and remove branches from identical sequences. The images shown in (B) were produced by the NCBI Conserved Domain Database server (63) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) using default settings.

Figure S18. Phylogenetic tree of kinesin motor domain sequences

Porphyra umbilicalis (red boxes) and other red algae (blue boxes) in evolutionary context to other organisms. Subfamilies with low numbers of red algal sequences are in cooler colors (e.g., kinesin-1 is dark blue while kinesin 9 is yellow). For the red algal sequences, the abbreviations used are *Porphyra umbilicalis* (Pum), *Porphyridium purpureum (Porphyridium), Chondrus crispus* (Chcr), *Galdieria sulphuraria* (Gasu), and *Cyanidioschyzon merolae* (Cyme). Kinesins in *Porphyra umbilicalis* and several other fully-sequenced red algae (i.e. *Chondrus crispus*, *Galdieria sulphuraria*, and *Cyanidioschyzon merolae*) were identified by BLAST (9) against the predicted proteomes of the target red algal species (see Table S9) using previously identified red algal kinesins as probes. Independently, we also searched the *Porphyra umbilicalis* predicted proteome by PSI-BLAST using PSSMs built from two iterations of searches of the full NCBI refseq database; this PSI-BLAST search was sufficient to identify all previously identified red algal kinesins. We combined these sequences with the set of kinesin sequences provided on the kinesin home page (https://labs.cellbio.duke.edu/kinesin/) and also provided in the supplementary material of the kinesin phylogenetic analysis performed by Wickstead and Gull (64).The tree was produced by FASTREE as implemented by PASTA (61), using default parameters. The confidence values are are local support estimates provided by FASTREE (65). Sequences were assigned to groups by identifying the deepest moderately-supported node (defined here as a support value of 0.80 or above, indicated by grey dots that did not encompass another previously defined group). Well-supported nodes (black dots) were defined as having support values of 0.93 or better. For this work, we used the group names as assigned by Wickstead and colleagues (64), who followed the pattern of Lawrence et al. (66) Kinesins 17-20 as assigned by Wickstead et al. (64) were somewhat variable in our trees (e.g., sometimes grouping within other well-supported groups), and so we have colored those groups grey. Similar overall groupings with variation in deep and/or weak nodes were seen with parallel PASTA runs performed using alternative alignment tools and evolutionary models. No variation was seen in the assignment of the *Porphyra* sequences to major groups except in the case of the divergent kinesins at the base of the kinesin 13 group; these sometimes grouped strongly with kinesin 13 and other times ended up in other parts of the tree (not shown).

Photosynthesis, Photoprotection, Stress Genes

Eleven One Helix Proteins (OHPs) were found in *P. umbilicalis*. Fewer OHP proteins are predicted in *Chondrus crispus* (Four NCBI: CO650034, CO650196, CO650444, XM_005717496.1_2), *Cyanidioschyzon merolae* (One NCBI: XP_005537856), *Porphyridium purpureum* (Seven Genome Scaffold: Porphyridium.evm.model.contig_2050.7; Porphyridium.evm.model.contig_3409.2; Porphyridium.evm.model.contig_2120.3; Porphyridium.evm.model.contig_4593.1; Porphyridium.evm.model.contig_983.1; Porphyridium.evm.model.contig_2120.4; Porphyridium.evm.model.contig_504.4), *Porphyra yezoensis* (Six Genome Scaffold: PyezoenV1.contig_14470_g3479; PyezoenV1.contig_13292_g3169; PyezoenV1.contig_42319_g9580; PyezoenV1.contig_21247_g5217; PyezoenV1.contig_10091_g2413; PyezoenV1.contig_606_g32), and the glaucophyte *Cyanophora paradoxa* (Five NCBI: DAA33879; DAA33880; DAA33881; DAA33882; DAA33883). The large *OHP* gene family in *P. umbilicalis*, and potentially an associated increase in OHP protein expression, may contribute to the organism's capacity to resist intense light exposure during low tide.

Table S19. Table of all photosynthetic genes and photoprotection genes

Porphyra gene identifier	Gene symbol	Aliases	Defline	Description
Pum0463s0006.1	OHP1	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly. Chondrus homolog induced by high light and high temperature
Pum0463s0011.1	OHP2	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly. Chondrus homolog induced by high light and high temperature
Pum1825s0001.1	OHP3	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum0014s0082.1	OHP4	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum0014s0085.1	OHP5	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum0505s0019.1	OHP6	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum0066s0011.1	OHP7	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum0617s0014.1	OHP8	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum0280s0002.1	OHP9	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum1810s0002.1	OHP10	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
scaffold_230:11830-12195	OHP11	Hlip, Scp	one-helix member of LHC Superfamily	Proposed roles in stress response, chlorophyll sequestration, and PSII assembly
Pum1594s0001.1	CpeT		Chromophore lyase CpcT/CpeT family	Covalently attaches bilin pigment to phycobilisome protein
Pum0077s0046.1	CpcS		Chromophore lyase CpeS/CpcS family	Covalently attaches bilin pigment to phycobilisome protein
Pum0082s0051.1	CpeZ		Phycoerythrin lyase	Covalently attaches bilin pigment to phycobilisome protein
Pum0135s0033.1	CpcF		Phycocyanobilin lyase, beta subunit	Covalently attaches bilin pigment to phycobilisome protein
Pum0192s0003.1	Lhcr1		Chlorophyll a light-harvesting protein, isoform 1	Photosystem I-associated light-harvesting antenna protein
Pum1672s0002.1	Lhcr2		Chlorophyll a light-harvesting protein, isoform 2	Photosystem I-associated light-harvesting antenna protein
Pum0031s0122.1	Lhcr3		Chlorophyll a light-harvesting protein, isoform 3	Photosystem I-associated light-harvesting antenna protein
Pum0483s0007.1	Lhcr4		Chlorophyll a light-harvesting protein, isoform 4	Photosystem I-associated light-harvesting antenna protein
Pum0127s0012.1	Lhcr5		Chlorophyll a light-harvesting protein, isoform 5	Photosystem I-associated light-harvesting antenna protein
Pum2012s0001.1	Lhcr6		Chlorophyll a light-harvesting protein, isoform 6	Photosystem I-associated light-harvesting antenna protein
Pum0422s0022.1	Lhcr7		Chlorophyll a light-harvesting protein, isoform 7	Photosystem I-associated light-harvesting antenna protein
Pum0577s0001.2	Lhcr8		Chlorophyll a light-harvesting protein, isoform 8	Photosystem I-associated light-harvesting antenna protein
Pum2112s0001.1	Lhcr9		Chlorophyll a light-harvesting protein, isoform 9	Photosystem I-associated light-harvesting antenna protein
Pum0359s0001.1	Lhcr10		Chlorophyll a light-harvesting protein, isoform 10	Photosystem I-associated light-harvesting antenna protein
Pum0093s0001.1	Lhcr11		Chlorophyll a light-harvesting protein, isoform 11, partial	Photosystem I-associated light-harvesting antenna protein
Pum0289s0024.1	RedCap1		LHC-like protein unique to red lineage	
Pum1564s0002.1	RedCap2		LHC-like protein unique to red lineage	
Pum0575s0008.1	PSAO		photosystem I subunit O	
Pum1462s0001.1	PSBO1		Oxygen-evolving enhancer protein 1 of photosystem II	
Pum1288s0002.1	PSBO2		Oxygen-evolving enhancer protein 1 of photosystem II	
Pum0818s0007.1	PSBP1		photosystem II oxygen-evolving enhancer protein 2	
Pum0098s0016.1	PSBP2		PsbP-like protein	
Pum0086s0010.1	PSBQ2		oxygen-evolving enhancer protein 3-2	
Pum0203s0001.1	psbU1		photosystem II complex extrinsic protein precursor U	
Pum1541s0001.1	psbU2		photosystem II complex extrinsic protein precursor U	
Pum1049s0001.1	PSB27	CGLD45	photosystem II Psb27 protein	
Pum0084s0022.1	PETC1		Rieske iron-sulfur subunit of the cytochrome b6f complex, chloroplast precursor	
Pum0084s0021.1	PETC2		Rieske iron-sulfur subunit of the cytochrome b6f complex, chloroplast precursor	
Pum0234s0022.1	PETC3		Rieske iron-sulfur subunit of the cytochrome b6f complex, chloroplast precursor	
Pum0924s0003.1	PETF1		ferredoxin	
Pum2905s0001.1	PETF2		ferredoxin 2	
Pum1510s0002.1	PETF3		ferredoxin 3	
Pum0513s0003.1	PETF4		ferredoxin 4	
Pum0537s0012.1	FNR		ferredoxin--NADP+ reductase	
Pum0055s0023.1	PETJ		cytochrome c6	
Pum0709s0005.1	ATPC1		chloroplast ATP synthase gamma chain	

Pum2300s0002.1	ATPC2		chloroplast ATP synthase gamma chain 2	
Pum0332s0024.1	PGR5		PROTON GRADIENT REGULATION 5	
Pum0493s0004.1	PGRL1		PGR5-like protein 1	
Pum0523s0007.1	apcC		phycobilisome core linker protein	
Pum0023s0023.1		gamma 31A	Phycoerythrin gamma subunit, 31 kDa homolog, isoform 1	Caps phycoerythrin rods in phycobilisome
Pum2278s0001.1		gamma 31B	Phycoerythrin gamma subunit, 31 kDa homolog, isoform 2	Caps phycoerythrin rods in phycobilisome
Pum0077s0041.1		gamma 31C	Phycoerythrin gamma subunit, 31 kDa homolog, isoform 3	Caps phycoerythrin rods in phycobilisome
Pum0027s0069.1		gamma 31D	Phycoerythrin gamma subunit, 31 kDa homolog, isoform 4	Caps phycoerythrin rods in phycobilisome
Pum0652s0007.1		gamma 33A	Phycoerythrin gamma subunit, 33 kDa homolog, isoform 1	Caps phycoerythrin rods in phycobilisome
Pum1573s0002.1		gamma 33B	Phycoerythrin gamma subunit, 33 kDa homolog, isoform 2	Caps phycoerythrin rods in phycobilisome

Table S20: Table of antioxidant and heat-shock proteins.

Porphyra gene identifier	Gene Symbol	Aliases or Synonyms	Defline	Description
Pum0810s0005.1	HSP10a	cpn10; GroES	Hsp10	Heat shock 10 kDa protein
Pum0587s0005.1	HSP20a		Hsp20	Heat shock 20 kDa protein
Pum0127s0002.1	HSP20b		Hsp20	Heat shock 20 kDa protein
Pum0698s0002.1	HSP20c		Hsp20	Heat shock 20 kDa protein
Pum0587s0008.1	HSP20d		Hsp20	Heat shock 20 kDa protein
Pum0284s0002.1	HSP20e		Hsp20	Heat shock 20 kDa protein
Pum0415s0001.1	HSP40a	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum0445s0009.1	HSP40b	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum0021s0013.1	HSP40c	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum0342s0011.1	HSP40d	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum1226s0002.1	HSP40e	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum0577s0009.1	HSP40f	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum1753s0001.1	HSP40h	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum0422s0019.1	HSP40i	DnaJ	Hsp40	Heat shock 40 kDa protein
scaffold_325:113547-112309	HSP40j	DnaJ	Hsp40	Heat shock 40 kDa protein
scaffold_686:17489-14504	HSP40k	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum0098s0012.1	HSP40l	DnaJ	Hsp40	Heat shock 40 kDa protein
Pum1766s0002.1	HSP60a		Hsp60	Heat shock 60 kDa protein
Pum0256s0045.1	HSP60b		Hsp60	Heat shock 60 kDa protein
Pum0212s0028.1	HSP70a	DnaK	Hsp70	Heat shock 70 kDa protein
Pum0550s0009.1	HSP70b	DnaK	Hsp70	Heat shock 70 kDa protein
Pum0311s0011.1	HSP70c	DnaK	Hsp70	Heat shock 70 kDa protein
Pum0267s0009.1	HSP70d	DnaK	Hsp70	Heat shock 70 kDa protein
Pum0381s0017.1	HSP70e	DnaK	Hsp70	Heat shock 70 kDa protein
Pum0306s0026.1	HSP70f	DnaK	Hsp70	Heat shock 70 kDa protein
Pum0087s0060.1	HSP90a		Hsp90	Heat shock 90 kDa protein
Pum0877s0006.1	HSP90b		Hsp90	Heat shock 90 kDa protein
Pum0167s0022.1	HSP90c		Hsp90	Heat shock 90 kDa protein
Pum0032s0076.1	CLP1		Clp protease ATP binding subunit	Clp protease ATP binding subunit
Pum0433s0006.1	TCP1		T-complex protein 1 subunit alpha	A molecular chaperone member of the TCP1 complex
Pum0202s0002.1	TCP1		T-complex protein 1 subunit alpha	A molecular chaperone member of the TCP1 complex
Pum0082s0053.1	TCPB1		T-complex protein 1 subunit beta	A molecular chaperone member of the TCP1 complex
Pum0432s0001.1	TCPD		T-complex protein 1 subunit delta	A molecular chaperone member of the TCP1 complex
Pum0107s0025.1	TCPE		T-complex protein 1 subunit epsilon	A molecular chaperone member of the TCP1 complex
Pum1502s0002.1	TCPZ		T-complex protein 1 subunit zeta	A molecular chaperone member of the TCP1 complex
Pum1958s0001.1	TCPQ1		T-complex protein 1 subunit theta	A molecular chaperone member of the TCP1 complex
Pum1752s0002.1	APX1		ascorbate peroxidase	
Pum0153s0010.1	APX2		ascorbate peroxidase	
Pum0519s0021.1	SOD1		Cu-Zn superoxide dismutase	
Pum0419s0001.1	SOD2		Cu-Zn superoxide dismutase	
Pum0208s0031.1	MnSOD1		manganese superoxide dismutase	
Pum0377s0012.1	CAT1		catalase	
Pum0377s0010.1	CAT2		catalase	
Pum0152s0048.1	CAT3		catalase	
Pum0192s0009.1	CAT4		catalase	
Pum0335s0005.1	CAT5		catalase	
Pum0098s0036.1	GR1		glutathione reductase	
Pum0032s0044.1	GR2		glutathione reductase	
scaffold_686:17489-14504	GR3		glutathione reductase	
Pum0414s0008.1	MDAR1		monodehydroascorbate reductase	
Pum2909s0001.1	MDAR2		monodehydroascorbate reductase	
Pum1972s0001.1	DHAR1		dehydrascorbate reductase	
Pum0290s0005.1	GPX1		glutathione peroxidase	

Pum0212s0008.1	V-HPO	V-BPO	vanadium-dependent haloperoxidase	Catalyzes a reaction that removes hydrogen peroxide and produces halogenated compounds
Pum0408s0001.1	GRX1		glutaredoxin	glutathione dependent reductase
Pum0445s0024.1	GRX2		glutaredoxin	glutathione dependent reductase
Pum0220s0005.1	APR1		adenylylsulfate (APS) reductase	
Pum0982s0005.1	MSR1		peptide methionine sulfoxide reductase	
Pum0082s0007.1	PDI1		Protein disulfide-isomerase	Catalyzes the formation and breakage of disulfide bonds between cysteine residues.
Pum0189s0001.1	PDI2		Protein disulfide-isomerase	Catalyzes the formation and breakage of disulfide bonds between cysteine residues.

Table S21. Antioxidant proteins in *Porphyra*, *Cyanidioschyzon*, *Galdieria*, *Chondrus* and *Porphyridium*.

The table shows the number of genes encoding antioxidant proteins on the genomes of *Porphyra* and selected red algae

Gene	*P. umbilicalis*	*G. sulphuraria*	*C. crispus*	*C. merolae*	*P. purpureum*
ascorbate peroxidase (APX)	2	2	2	1	2
Cu/Zn superoxide dismutase (SOD)	2	0	2	0	0
Fe/Mn superoxide dismutase (SOD)	1	2	2	3	3
catalase (CAT)	5	1	2	1	1
glutathione reductase (GR)	2	1	1	1	2
monodehydroascorbate reductase (MDHAR)	2	2	1	1	1
dehydroascorbate reductase (DHAR)	2	0	2	0	1
glutathione peroxidase (GPX)	1	1	2	0	3
peroxiredoxin (PRX)	1	3	3	3	2
glutaredoxin (GRX)	3	2	4	2	2
thioredoxin reductase (TRxR)	1	1	1	3	1

Vitamin E

Tocopherols, known collectively as vitamin E, are lipid-soluble antioxidants synthesized exclusively by photosynthetic organisms (67). In plants this group of compounds play a critical photoprotective role (68, 69), preventing photooxidative damage to polyunsaturated fatty acids from reactive oxygen species (70). The *P. umbilicalis* genome encodes several genes of the tocopherol biosynthesis pathway as characterized in *Arabidopsis thaliana*, including: p-hydroxyphenylpyruvate dioxygenase (HPPD; *PDS1*), homogentisate phytyltransferase (HPT; *VTE2*), tocopherol cyclase (*VTE1*), 2-methyl-6-phytyl-1,4-benzoquinone methyltransferase (MPBQ MT; *VTE3*), and γ-tocopherol methyltransferase (γ-TMT; *VTE4*) (Table S22). Curiously a second, virtually identical hit for *VTE4* was also found on another scaffold (scaffold 164). However, the identification of frameshifts alongside two in-frame stop codons within this gene copy suggest that it is a pseudogene. γ-TMT catalyses the synthesis of α and β-tocopherol from γ and δ-tocopherol respectively, and α-tocopherol is considered to have the greatest vitamin E bioactivity as the most efficient antioxidant (71). That there are two copies of the *VTE4* gene in *Porphyra*, one of which was subsequently pseudogenised raises interesting questions regarding the selective pressures governing the composition of tocopherol forms present in *Porphyra*. While biochemical evidence suggests *P. umbilicalis* contains α-tocopherol (72), the relative abundance of the different tocopherol isomers is, to our knowledge, currently unknown.

Table S22. Tocopherol biosynthetic enzymes

Genes encoding tocopherol biosynthetic enzymes in *Porphyra* and other sequenced red algae. The gene identifier is shown for red algal sequences. N.F. not found.

Gene	*P. umbilicalis*	*C. crispus*	*G. sulphuraria*	*C. merolae*	*P. purpureum*	*P. yezoensis*
HPPD (PDS1)	Pum0014s0114.1	Phy0034PO2	XP_005704388.1	CMI063C	contig_3399.12	contig_10349_g2486
HPT (VTE2)	Pum0106s0040.1	Phy0034MHL	XP_005705799.1	CMN202C	contig_2502.20	contig_11812_g2810
TCY (VTE1)	Pum0277s0012.1	N.F.	XP_005702972.1	CML326C	contig_2120.2	N.F.
MPBQ MT (VTE3)	Pum0438s0002.1	Phy0034SI0	XP_005704060.1	CMD011C	contig_3419.7	contig_15138_g3632
γ-tocopherol methyltransferase* (γTMT; VTE4)	Pum0445s0014.1	Phy0034MEZ	XP_005706819.1	CMT560C	contig_728.1	contig_24223_g5970

*A second gene encoding γ-tocopherol methyltransferase was also found (Pum0164s0002.1), but frameshifts in the coding sequence lead to multiple stop codons. It is, therefore, likely that this is a pseudogene.

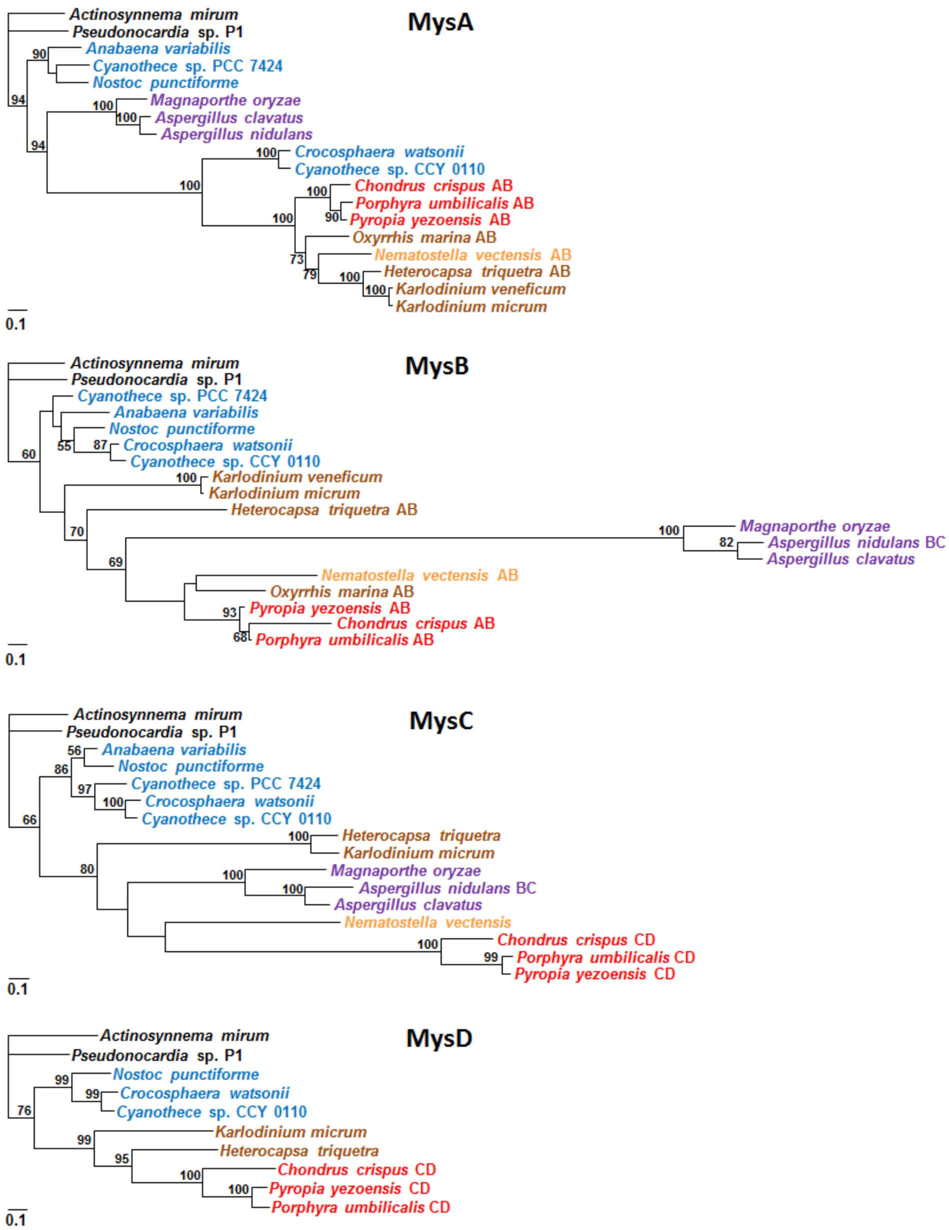


Figure S23. Phylogenetic affiliation of MAA biosynthetic genes

Phylogenetic affiliation of MAA biosynthetic genes from selected Actinobacteria (black), Cyanobacteria (blue), red algae (red), dinoflagellate (brown), fungi (purple) and a cnidarian (orange). Maximum likelihood trees were constructed with RAxML (42) from protein alignments of 391 (MysA), 149 (MysB), 305 (MysC) and 346 (MysD) amino acid positions respectively; numbers at nodes denote > 50% bootstrap support (100 replicates). Letters following species names denote fusion of the following genes: AB, *MysA* with *MysB*; BC, *MysB* with *MysC*; CD, *MysC* with *MysD*. See Figure S24 for database accessions of genes/proteins.

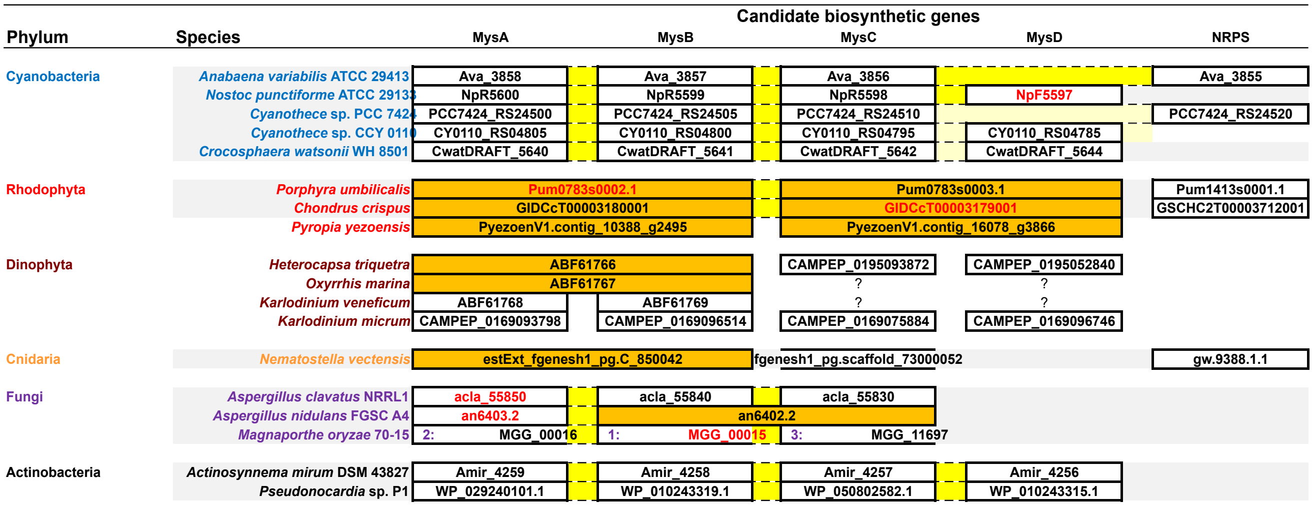


Figure S24. Gene clusters and fusions in the MAA biosynthetic pathway

Genomic arrangements of MysA/B and MysC/D fusions are shown for various phyla and species. Gene fusions are boxed in orange; unfused, but adjacent genes are yellow boxed with dashes; and close genes, but with other genes in between, are pale yellow boxed with dashes. Organisms with sequenced genomes are shown with gray background; red gene identifiers indicate genes on the opposite strand; gene order in purple deviates from the typical genomic arrangement. Cyanobacteria (blue), Rhodophyta (red), Dinophyta (brown), Cnidaria (orange), fungi (purple), Actinobacteria (black). Data for other species are from the following publications and databases: *A. variabilis*, *N. punctiforme*, *Cyanothese* spp, *C. watsonii* (73) and NCBI (https://www.ncbi.nlm.nih.gov/); *C. crispus* (20) and MMO-EMBRC (http://mmo.embrc-france.fr/en/algae/chondrus); *P.yezoensis* (24) and NRIFS (http://nrifs.fra.affrc.go.jp/ResearchCenter/5_BB/genomes/nori/index.html); *H. triqueta* (74) and MMETSP (30) (http://data.imicrobe.us/project/view/104); *O. marina*, *K. veneficum* (74) and NCBI (https://www.ncbi.nlm.nih.gov/); *K. micrum* from MMETSP (30); *N. vectensis* (75) and JGI (http://genome.jgi.doe.gov/); *A. clavatus*, *A. nidulans*, *M. oryzae*, *A. mirum*, *Pseudonocardia* (73, 76) from NCBI (https://www.ncbi.nlm.nih.gov/).

Biosynthesis of isoprenoids

Carotenoids

As was the case for the transcriptome data from *P. umbilicalis* (1) the current genome version is missing candidates for 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MDS), zeta-carotene desaturase (ZDS) and carotenoid isomerase (CRTISO), while candidate genes for all other carotenoid biosynthetic steps could be identified (Table S25). For the cytochrome 450-protein CYP97B29 (PuCHY1), its predicted function as a carotenoid hydroxylase (1) was recently confirmed experimentally (77). The genomes of the prasinophyte genera *Ostreococcus* and *Micromonas* also lack orthologs of ZDS and CRTISO, indicating that their function may have been taken over by yet uncharacterized enzymes in *P. umbilicalis* and in prasinophytes. The lack of a canonical MDS, however, would be unprecedented. The genomes of *C. merolae*, *G. sulphuraria* and *C. crispus* contain MDS genes, and the *P. yezoensis* genome data contain a partial MDS sequence (PyezoenV1.contig_45452_g10060). Moreover, RNAseq data from

P. umbilicalis contain fragments of a MDS candidate gene (comp35301_c2_seq1, comp31598_c0_seq1), suggesting the gene is present in the genome, but not represented in the current assembly.

Sterols

Despite extensive mining of the genome data we found no evidence for genes or gene remnants of the cytosolic mevalonic acid (MVA) pathway in *P. umbilicalis* (with the exception of a hydroxymethylglutaryl-CoA synthase homolog that may be involved in a different biosynthetic pathway) thus corroborating previous results from analyses of transcriptomic data (1, 29). Therefore, it appears very likely that in *P. umbilicalis* the isoprene units for sterol biosynthesis are supplied by the plastidic MEP pathway, as has been shown for other red algae and green algae in general (29, 78). Regarding sterol biosynthesis, we found candidates for all genes necessary for the biosynthesis of desmosterol and cholesterol (Table S25) as already inferred from our transcriptome data of *P. umbilicalis* (1). The three candidate genes for sterol methyl transferases that had been identified in the transcriptome data, however, are not present in the current genome version and probably were contaminants. Therefore, there no longer is *in-silico* evidence for *P. umbilicalis* being able to synthesize C-29 sterols.

Biosynthesis of tetrapyrroles

Regarding tetrapyrrole biosynthesis up to protoporphyrin IX, only a candidate for uroporphyrinogen III synthase (UROS) is missing in the current genome version (Table S25). In the RNA-Seq data from *P. umbilicalis*, however, two fragments of an UROS candidate gene (comp47334_c0_seq1, comp30123_c0_seq1) could be identified. Similarly, candidates for all genes necessary for biosynthesis of chl *a* are present in the nuclear genome with the exceptions of divinyl protochlorophyllide reductase (DVR) that could only be identified in the RNAseq data (comp98546_c0_seq1; comp32085_c0_seq1). Mg-chelatase subunit I (ChlI) and subunit CHL27 (ycf59) of the aerobic Mg-protoporphyrin IX monomethylester cyclase are encoded on the plastid genome of *P. umbilicalis* (34). Notably, *Porphyra* and other red algae contain two candidate genes for heme oxygenase (HO) involved in the biosynthesis of phycobilins and scavenging of free heme. One potential HO (HO1) is encoded in the plastid genome and is closely related to the cyanobacterial HO encoded by *pbsA*, while HO2 is nuclear-encoded and most similar to HO (HY1) from streptophytes.

Table S25. Isoprenoid (sterols and carotenoids), tocopherol, tetrapyrrole and MAA biosynthetic proteins in *Porphyra*

Porphyra gene identifier	Gene	Aliases & synonyms	Defline	Description
Isoprenoid Biosynthesis: plastidic MEP pathway				
Pum0237s0005.1	DXS1	DXPS, CLA1, DEF		first enzyme of non-mevalonate MEP pathway of isopentenyl diphosphate biosynthesis in plastids; catalyzes: pyruvate + D-glyceraldehyde 3-phosphate = 1-deoxy-D-xylulose 5-phosphate + CO_2
Pum0237s0013.1	DXS2	DXPS, CLA1, DEF		first enzyme of non-mevalonate MEP pathway of isopentenyl diphosphate biosynthesis in plastids; catalyzes: pyruvate + D-glyceraldehyde 3-phosphate = 1-deoxy-D-xylulose 5-phosphate + CO_2
Pum0033s0055.1	DXR	ispC, yaeM	1-deoxy-D-xylulose 5-phosphate reductoisomerase, plastid precursor; EC 1.1.1.267	second enzyme of non-mevalonate MEP pathway of isopentenyl diphosphate biosynthesis in plastids; catalyzes: 1-deoxy-D-xylulose 5-phosphate + NADPH + H+ = 2-C-methyl-D-erythritol 4-phosphate + NADP+
Pum0098s0022.1	MCT	CMS, ispD, ygbP	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase, plastid precursor; EC 2.7.7.60	third enzyme of non-mevalonate MEP pathway of isopentenyl diphosphate biosynthesis in plastids; catalyzes: CTP + 2-C-methyl-D-erythritol 4-phosphate = diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol
scaffold_23:704958-706033	CMK	CDP-ME, CDPMEK, ispE, ychB	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, plastid precursor; EC 2.7.1.148	fourth enzyme of non-mevalonate MEP pathway of isopentenyl diphosphate biosynthesis in plastids; catalyzes: ATP + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol = ADP + 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol
Pum0459s0015.1	HDS	ispG, GcpE	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase, plastid precursor; EC 1.17.7.1	sixth enzyme of non-mevalonate MEP pathway of isopentenyl diphosphate biosynthesis in plastids; catalyzes: (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + H2O + 2 oxidized ferredoxin = 2-C-methyl-D-erythritol 2,4-cyclodiphosphate + 2 reduced ferredoxin; PMID: 15849308
Pum0582s0002.1	HDR	IDS, ispH, LytB	4-hydroxy-3-methylbut-2-enyl diphosphate reductase, plastid precursor; EC 1.17.1.2	seventh enzyme of non-mevalonate MEP pathway of isopentenyl diphosphate biosynthesis in plastids; catalyzes: dimethylallyl diphosphate + NAD(P)+ + H2O = (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H + H+ ; PMID:11004185
Carotenoid Biosynthesis				
Pum0377s0009.1	GGPPS	GGPS; CrtE	geranylgeranyl diphosphate synthase, plastid precursor; EC 2.5.1.29	involved in carotenoid biosynthesis in plastids; catalyzes: (2E,6E)-farnesyl-diphosphate:isopentenyl-diphosphate farnesyltranstransferase; the plant/algal enzyme also uses geranyl diphosphate and dimethylallyl diphosphate as donors, enabling synthesis of geranylgeranyl diphosphate from one dimethylallyl diphosphate and three isopentenyl diphosphate molecules
Pum0392s0019.1	PSY	crtB	phytoene synthase, plastid precursor; EC 2.5.1.32	involved in carotenoid biosynthesis in plastids; catalyzes: head-to-head condensation of two geranylgeranyl-diphosphate molecules yielding 15-cis-phytoene; committed step in carotenoid biosynthesis
Pum1802s0002.1	PDS	crtP	phytoene desaturase, plastid precursor; EC 1.3.5.5	involved in carotenoid biosynthesis in plastids; catalyzes: (1) phytoene + acceptor <=> phytofluene + reduced acceptor, (2) phytofluene + acceptor <=> zeta-carotene + reduced acceptor; uses plastoquinone as electron aceptor
Pum0025s0060.1	Z-ISO		zeta-carotene isomerase, plastid precursor; EC 5.2.1.12	involved in carotenoid biosynthesis in plastids; catalyzes: 9,15,9'-tricis-zeta-carotene = 9,9'-dicis-zeta-carotene; PMID: 20335404
Pum0242s0024.1†	LCY1	crtL	lycopene cyclase, plastid precursor; EC 5.5.1.18 or 5.5.1.19	involved in carotenoid biosynthesis in plastids; catalyzes cyclization of linear psi-end group to cyclic end group; may catalyze formation of epsilon- (EC 5.5.1.18) or beta- (EC 5.5.1.19) end group, exact function awaits experimental confirmation; DOI: 10.1111/j.1529-8817.2012.01229.x
Pum0111s0004.1	CYP97B29	PuCHY1	cyt P450 carotene hydroxylase, plastid precursor; most similar to CYP97E and CYP97F from heterokont algae; EC 1.14.-.-	involved in carotenoid biosynthesis in plastids; cytochrome P450 monoxygenase; catalyzes hydroxylation at C3 of carotenoid beta- (EC 1.14.13.129) and epsilon- (EC 1.14.99.45) end groups; DOI: 10.1111/j.1529-8817.2012.01229.x; PMID: 24942088
Isoprenoid Biosynthesis: cytoplasmic MVA pathway				
Pum0353s0009.1	HMGS		hydroxymethylglutaryl-CoA synthase; EC 2.3.3.10	involved in mevalonic acid (MVA) pathway for biosynthesis of isopentenyl diphosphate in the cytosol; catalyzes: acetyl-CoA + H2O + acetoacetyl-CoA = (S)-3-hydroxy-3-methylglutaryl-CoA + CoA; PMID: 21645203, 24847714
Sterol Biosynthesis				
Pum0076s0049.1	FPPS		farnesyl-diphosphate synthase; EC 2.5.1.10	involved in biosynthesis of terpenoid backbone of isoprenoids such as sterols; catalyzes: geranyl diphosphate + isopentenyl diphosphate = diphosphate +

				(2E,6E)-farnesyl diphosphate
Pum0184s0009.1	SQS	ERG9	squalene synthase; EC 2.5.1.21	involved in sterol biosynthesis; catalyzes: 2 (2E,6E)-farnesyl diphosphate + NAD(P)H + H+ = squalene + 2 diphosphate + NAD(P)+ (2) presqualene diphosphate + NAD(P)H + H+ = squalene + diphosphate + NAD(P)+
Pum0203s0004.1	SQE	SQP, CYP17, ERG1	squalene monoxygenase; EC 1.14.13.132	involved in sterol biosynthesis; catalyzes: squalene + NADPH + H+ + O2 = (3S)-2,3-epoxy-2,3-dihydrosqualene + NADP+ + H2O
Pum2748s0001.1	LAS	ERG7; OSC	lanosterol synthase; EC 5.4.99.7	involved in sterol biosynthesis; catalyzes: (3S)-2,3-epoxy-2,3-dihydrosqualene = lanosterol; highly similar to CAS = cycloartenol synthase (EC 5.4.99.8); putatively identified as lanosterol synthase based on cholesterol and desmosterol as major sterols synthesized in Porphyra
Pum0269s0017.1	CYP51	14DM; ERG11	sterol 14alpha-demethylase; EC 1.14.13.70	involved in sterol biosynthesis; catalyzes: a 14alpha-methylsteroid + 3 O2 + 3 NADPH + 3 H+ = a Delta14-steroid + formate + 3 NADP+ + 4 H2O
Pum0130s0027.1	FACKEL	ERG24	DELTA14-sterol reductase; EC 1.3.1.70	involved in sterol biosynthesis; catalyzes reduction of 14-double bond of conjugated DELTA8,14-sterols and DELTA7,14-sterols; acts on a range of steroids with a 14(15)-double bond
Pum0463s0003.1	SMO1	ERG25	methylsterol monoxygenase; EC 1.14.13.72	involved in sterol biosynthesis; catalyzes: 3beta-hydroxy-4beta-methyl-5alpha-cholest-7-ene-4alpha-carbaldehyde + NAD(P)H + H+ + O2 = 3beta-hydroxy-4beta-methyl-5alpha-cholest-7-ene-4alpha-carboxylate + NAD(P)+ + H2O
Pum0197s0018.1	SMO2	ERG25	methylsterol monoxygenase; EC 1.14.13.72	involved in sterol biosynthesis; catalyzes: 3beta-hydroxy-4beta-methyl-5alpha-cholest-7-ene-4alpha-carbaldehyde + NAD(P)H + H+ + O2 = 3beta-hydroxy-4beta-methyl-5alpha-cholest-7-ene-4alpha-carboxylate + NAD(P)+ + H2O
Pum0109s0008.1	HSD-D	ERG26	3beta-hydroxysteroid-4alpha-carboxylate 3-dehydrogenase (decarboxylating); EC 1.1.1.170	involved in sterol biosynthesis; catalyzes: a 3beta-hydroxysteroid-4alpha-carboxylate + NAD(P)+ = a 3-oxosteroid + CO2 + NAD(P)H
Pum0513s0015.1	ERG2		cholestenol DELTA-isomerase, yeast-type; EC 5.3.3.5	involved in sterol biosynthesis; catalyzes: 5alpha-Cholest-7-en-3beta-ol = 5alpha-cholest-8-en-3beta-ol
Pum0965s0003.1	ERG3		DELTA7-sterol 5(6)-desaturase, yeast-type; EC 1.14.19.20	involved in sterol biosynthesis; catalyzes: a DELTA7-sterol + 2 ferrocytochrome b5 + O2 + 2 H+ = a DELTA5,7-sterol + 2 ferricytochrome b5 + 2 H2O
Pum0521s0006.1	DWF5	7-DHCR	7-dehydrocholesterol reductase; EC 1.3.1.21	involved in sterol biosynthesis; catalyzes: cholesterol + NADP+ = cholesta-5,7-dien-3beta-ol + NADPH + H+
Pum0093s0006.1	DWF1	ERG4	DELTA24(24(1))-sterol reductase; EC 1.3.1.71	involved in sterol biosynthesis; catalyzes: ergosta-5,7,22,24(24(1))-tetraen-3beta-ol + NADPH + H+ = ergosterol + NADP+; acts on a range of steroids with a 24(24(1))-double bond
Tocopherol Biosynthesis				
Pum0014s0114.1	HPPD	4-HPPD	4-hydroxyphenylpyruvate dioxygenase, cytsolic enzyme; EC 1.13.11.27	involved in tocopherol biosynthesis; catalyzes: 4-hydroxyphenylpyruvate + O2 = homogentisate + CO2
Pum0106s0040.1	HPT	VTE2	homogentisate phytyltransferase, plastid precursor; EC 2.5.1.115	involved in tocopherol biosynthesis; catalyzes: phytyl diphosphate + homogentisate = diphosphate + 2-methyl-6-phytylbenzene-1,4-diol + CO2; PMID: 12586887
Pum1830s0002.1	HST		homogentisate solanesyltransferase, plastid precursor; EC 2.5.1.117	involved in plastoquinone biosynthesis; catalyzes: all-trans-nonaprenyl diphosphate + homogentisate = diphosphate + 2-methyl-6-all-trans-nonaprenylbenzene-1,4-diol + CO2
Pum0438s0002.1	MPBQ-MT	VTE3	2-methyl-6-phytyl-1,4-hydroquinone methyltransferase, plastid precursor; EC 2.1.1.295	bifunctional enzyme involved in biosynthesis of tocopherols by catalyzing (1): S-adenosyl-L-methionine + 2-methyl-6-phytylbenzene-1,4-diol = S-adenosyl-L-homocysteine + 2,3-dimethyl-6-phytylbenzene-1,4-diol; and in biosynthesis of plastoquinol by catalyzing (2): S-adenosyl-L-methionine + 2-methyl-6-all-trans-nonaprenylbenzene-1,4-diol = S-adenosyl-L-homocysteine + plastoquinol
Pum0277s0012.1	TCY	VTE1	tocopherol cyclase, plastid precursor; EC 5.5.1.24	involved in tocopherol biosynthesis; can accept multiple substrates potentially catalyzing: (1) delta-tocopherol = 2-methyl-6-phytylbenzene-1,4-diol; (2) delta-tocotrienol = 6-geranylgeranyl-2-methylbenzene-1,4-diol; (3)gamma-tocopherol = 2,3-dimethyl-6-phytylbenzene-1,4-diol; (4) gamma-tocotrienol = 6-geranylgeranyl-2,3-dimethylbenzene-1,4-diol
Pum0445s0014.1 *	TMT	VTE4	tocopherol O-methyltransferase, plastid precursor; EC 2.1.1.95	involved in tocopherol biosynthesis; catalyzes: S-adenosyl-L-methionine + gamma-tocopherol = S-adenosyl-L-homocysteine + alpha-tocopherol; The enzymes from plants and photosynthetic bacteria have similar efficiency with the gamma and delta isomers of tocopherols and tocotrienols.
Tetrapyrrole Biosynthesis				
Pum0608s0008.1	cpGTS	GluRS	glutamyl-tRNA synthetase, plastid precursor; EC 6.1.1.17	involved in tetrapyrrole biosynthesis in plastids; catalyzes: ATP + L-glutamate + tRNAGlu = AMP + diphosphate + L-glutamyl-tRNAGlu; PMID: 17227226
Pum0192s0006.1	HEMA	GluTR	glutamyl-tRNA reductase, plastid precursor; EC 1.2.1.70	involved in tetrapyrrole biosynthesis in plastids; catalyzes: L-glutamate 1-semialdehyde + NADP+ + tRNAGlu = L-

				glutamyl-tRNAGlu + NADPH + H+; PMID: 7957167, 17227226
Pum0014s0112.1	GSA	GSAT, GSA-AT, hemL	glutamate-1-semialdehyde 2,1-aminomutase, plastid precursor; EC 5.4.3.8	involved in tetrapyrrole biosynthesis in plastids; catalyzes: (S)-4-amino-5-oxopentanoate = 5-aminolevulinate; cofactor: pyridoxal 5'-phosphate; PMID: 9144156, 17227226
Pum0111s0021.1	ALAD	PBS; HEM2; hemB	porphobilinogen synthase, plastid precursor; EC 4.2.1.24	involved in tetrapyrrole biosynthesis in plastids; catalyzes: 2 x 5-aminolevulinate = porphobilinogen + 2 H2O; cofactor: zink; enzyme homo-octameric; PMID: 9144156, 17227226
scaffold_530:26000-27050	PBGD	HEM3; hemC	hydroxymethylbilane synthase, plastid precursor; EC 2.5.1.61	involved in tetrapyrrole biosynthesis in plastids; catalyzes: 4 x porphobilinogen + H2O <=> hydroxymethylbilane + 4 NH3; In the presence of UROS (EC 4.2.1.75), which is often called cosynthase, the product is cyclized to form uroporphyrinogen III; if UROS is absent, the hydroxymethylbilane cyclizes spontaneously to form uroporphyrinogen I; PMID: 17227226
Pum0482s0015.1	UROD	URO-D; UPD; hemE	uroporphyrinogen decarboxylase, plastid precursor; EC 4.1.1.37	involved in tetrapyrrole biosynthesis in plastids; catalyzes: uroporphyrinogen III = coproporphyrinogen + 4 CO2; PMID: 17227226
Pum0288s0012.1	CPX	CPO; CPOX; hemF	coproporphyrinogen oxidase (aerobic), plastid precursor; EC 1.3.3.3	involved in tetrapyrrole biosynthesis in plastids; catalyzes: coproporphyrinogen-III + O2 + 2 H+ = protoporphyrinogen-IX + 2 CO_2 + 2 H_2O; PMID: 17227226
Pum0215s0022.1	PPX	PPO; PPOX Protox; HemY	protoporphyrinogen oxidase, plastid precursor; EC 1.3.3.4	involved in tetrapyrrole biosynthesis in plastids; catalyzes: protoporphyrinogen IX + 3 O2 = protoporphyrin IX + 3 H2O2; PMID: 17227226
Pum0027s0127.1	FC	FeChel	ferrochelatase, plastid precursor; EC 4.99.1.1	involved in tetrapyrrole / heme biosynthesis in plastids; catalyzes: protoheme + 2 H+ = protoporphyrin + Fe2+; PMID: 21565502, 21565502, 17227226
Pum0200s0032.1	UPM	SirA	uroporphyrinogen III methyltransferase; EC 2.1.1.107	Involved in siroheme biosynthesis in plastids; catalyzes: 2 S-adenosyl-L-methionine + uroporphyrinogen III = 2 S-adenosyl-L-homocysteine + precorrin-2; DOI: 10.4161/psb.5.1.10173; gene model probably erroneous!
Pum0033s0026.1	ChlD		Magnesium chelatase subunit D; EC 6.6.1.1	Involved in chlorophyll biosynthesis in plastids; Inserts Mg(II) into protoporphyrin IX
Pum0809s0010.1	ChlH		Magnesium chelatase subunit H; EC 6.6.1.1	Involved in chlorophyll biosynthesis in plastids; Inserts Mg(II) into protoporphyrin IX
Pum0393s0013.1	ChlM	PPMT	Mg protoporphyrin IX S-adenosyl methionine O-methyl transferase; EC 2.1.1.11	Involved in chlorophyll biosynthesis in plastids; Adds methyl group to Mg protoporphyrin IX
Pum2734s0001.1	POR		NADPH:protochlorophyllide oxidoreductase, light-dependent; EC 1.3.1.33	Involved in chlorophyll biosynthesis in plastids; Reduces divinyl protochlorophyllide a to divinyl chlorophyllide a
Pum0141s0036.1	ChlP	GGR	geranylgeranyl diphosphate reductase, plastid precursor; EC 1.3.1.83	involved in biosynthesis of phytol for formation of chlorophylls and tocopherols; catalyzes: phytyl diphosphate + 3 NADP+ = geranylgeranyl diphosphate + 3 NADPH + 3 H+
Pum0098s0032.1	ChlG		Chlorophyll a synthase; EC 2.5.1.62	Involved in chlorophyll biosynthesis in plastids; Adds phytol tail to chlorophylls
Pum0031s0107.1	PPH		Pheophytinase	Involved in chlorophyll catabolism; Alpha/beta hydrolase, dephytylates pheophytin yielding pheophorbide
Pum0176s0043.1	HO2	HemO, HY1, psbA	Heme oxygenase (biliverdin-producing, ferredoxin); EC 1.14.15.20	Involved in phycobilin biosynthesis and heme scavenging in plastids; Oxidizes the cyclic tetrapyrrole heme to the linear tetrapyrrole biliverdin IXa
Pum0348s0001.1	PebA		15,16-dihydrobiliverdin: ferredoxin oxidoreductase; EC 1.3.7.2	Involved in phycobilin biosynthesis in plastids; Converts biliverdin IXa to 15,16-dihydrobiliverdin
Pum0518s0019.1	PebB		Phycoerythrobilin: ferredoxin oxidoreductase; EC 1.3.7.3	Involved in phycobilin biosynthesis in plastids; Converts 15,16-dihydrobiliverdin to phycoerythrobilin
Mycosporine-like Amino Acid (MAA) Biosynthesis				
Pum0783s0002.1	MysA/MysB	DDGS/OMT fusion, DHQS-like/OMT fusion	demethyl 4-deoxy-gadusol synthase/O-methyltransferase fusion protein	involved in the biosynthesis of mycosporine-like amiono acids (MAA); putatively catalyzes the formation of 4-deoxygadusol from sedoheptulose 7-phosphate via the intermediate demethyl 4-deoxygadusol
Pum0783s0003.1	MysC/MysD		C-N ligase/D-ala-D-ala ligase-like fusion protein	involved in the biosynthesis of mycosporine-like amino acids (MAA); putatively catalyzes the formation of shinorine from 4-deoxygadusol via the intermediate mycosporine-glycine

*There is a second copy of TMT (Pum0164s0002.1) likely a pseudogene.

[†]There is a region of partial homology to LCY on scaffold_2881, with expression support between coordinates 1-1295 that probably encodes a second paralog LCY2.

Circadian clock-related genes

Table S26. *Porphyra* circadian clock-related genes

Porphyra gene identifier	Gene Symbol	Aliases or Synonyms	Defline	Description
Pum1359s0001.1	SKP1		Subunit of E3 ubiquitin ligase	Protein with some homology to a component of the circadian clock of *Chlamydomonas reinhardtii*
Pum0470s0014.1	XRN1		N-terminus of single-stranded RNA 5'-3' exoribonuclease	Protein with some homology to a component of the circadian clock of *Chlamydomonas reinhardtii*
Pum0401s0002.1	aCRY		Animal like cryptochrome	Protein with some homology to a component of the circadian clock of *Chlamydomonas reinhardtii*
Pum0223s0025.1	pCRY		Plant Cryptochrome	Protein with some homology to a component of the circadian clock of *Chlamydomonas reinhardtii*, probably not acting as plant cryptochrome blue light receptor, because it is closest to the DNA photolyase PHR2 from *Chlamydomonas*
Pum0341s0009.2	SPY		Spindly	Protein with some homology to a component of the circadian clock of *Arabidopsis thaliana*
Pum0031s0117.1	XCT		XAP5 circadian Timekeeper	Protein with some homology to a component of the circadian clock of *Arabidopsis thaliana*
Pum0501s0009.1	CRB		Chloroplast RNA binding protein	Protein with some homology to a component of the circadian clock of *Arabidopsis thaliana*
Pum2644s0001.1	CRY3		Cryptochrome 3 (L. esculentum)	Protein with some homology to a component of the circadian clock of *Lycopersicon esculentum*
Pum0070s0024.1	CUL1		Cullin 1	Protein with some homology to a component of the circadian clock of *Arabidopsis thaliana*
Pum0031s0079.1	LWD2		Light- regulated WD 2	Protein with some homology to a component of the circadian clock of *Arabidopsis thaliana*
Pum0205s0019.1	CK2-alpha1		Casein kinase 2 alpha chain 1	Protein with some homology to a component of the circadian clock of *Arabidopsis thaliana*
Pum0205s0018.1	CK2-alpha2		Casein kinase 2 alpha chain 2	Protein with some homology to a component of the circadian of *Arabidopsis thaliana*
Pum0608s0002.1	CK2-beta4		Casein kinase 2 beta chain 4	Protein with some homology to a component of the circadian of *Arabidopsis thaliana*
Pum0244s0013.1	CSN2		COP9 signalosome complex subunit 2	Protein with some homology to a component of the circadian clock of *Neurospora crassa*
Pum0392s0007.1	PP1		Serine/Threonine protein phosphatase PP1	Protein with some homology to a component of the circadian clock of *Neurospora crassa*
Pum0924s0004.1	LARK		LARK-RNA binding protein	Protein with some homology to a component of the circadian clock of *Drosophila melanogaster*
Pum0093s0012.1	PP2A-B	Twins isoform A	Protein phosphatase 2A subunit B(= twins)	Protein with some homology to a component of the circadian clock of *Drosophila melanogaster*
Pum0392s0002.1	PP2A-C	Widerborst isoform C	Protein phosphatase 2A subunit C (=Widerbrost)	Protein with some homology to a component of the circadian clock of *Drosophila melanogaster*
Pum0403s0022.1	ATM		Ataxia telangiectasia mutated	Protein with some homology to a component of the circadian clock of *Homo sapiens*
Pum0183s0021.1	CHK1		Check point Kinase 1	Protein with some homology to a component of the circadian clock of *Homo sapiens*
Pum1377s0003.1	CK1epsilon		Casein Kinase 1 epsilon	Protein with some homology to a component of the circadian clock of *Homo sapiens*
Pum0172s0029.1	GSK3 beta		Glycogen synthase kinase 3 beta (*M.musculus*)	Protein with some homology to a component of the circadian clock of *Mus musculus*
Pum0218s0017.1	HSF1		Heat-shock factor 1	Protein with some homology to a component of the circadian clock of *Mus musculus*
Pum0072s0088.1	NAT5		N-acetyl-transferase 5	Protein with some homology to a component of the circadian clock of *Homo sapiens*
Pum0314s0016.1	PP5		Protein phosophatase 5 (*M. musculus*)	Protein with some homology to a component of the circadian clock of *Mus musculus*
Pum0608s0009.1	RBM4		RNA binding protein 4 (M.musculus)	Protein with some homology to a component of the circadian clock of *Mus musculus*
Pum0303s0002.1	SIRT1		Sirtuin 1 (*M. musculus*)	Protein with some homology to a component of the circadian clock of *Mus musculus*

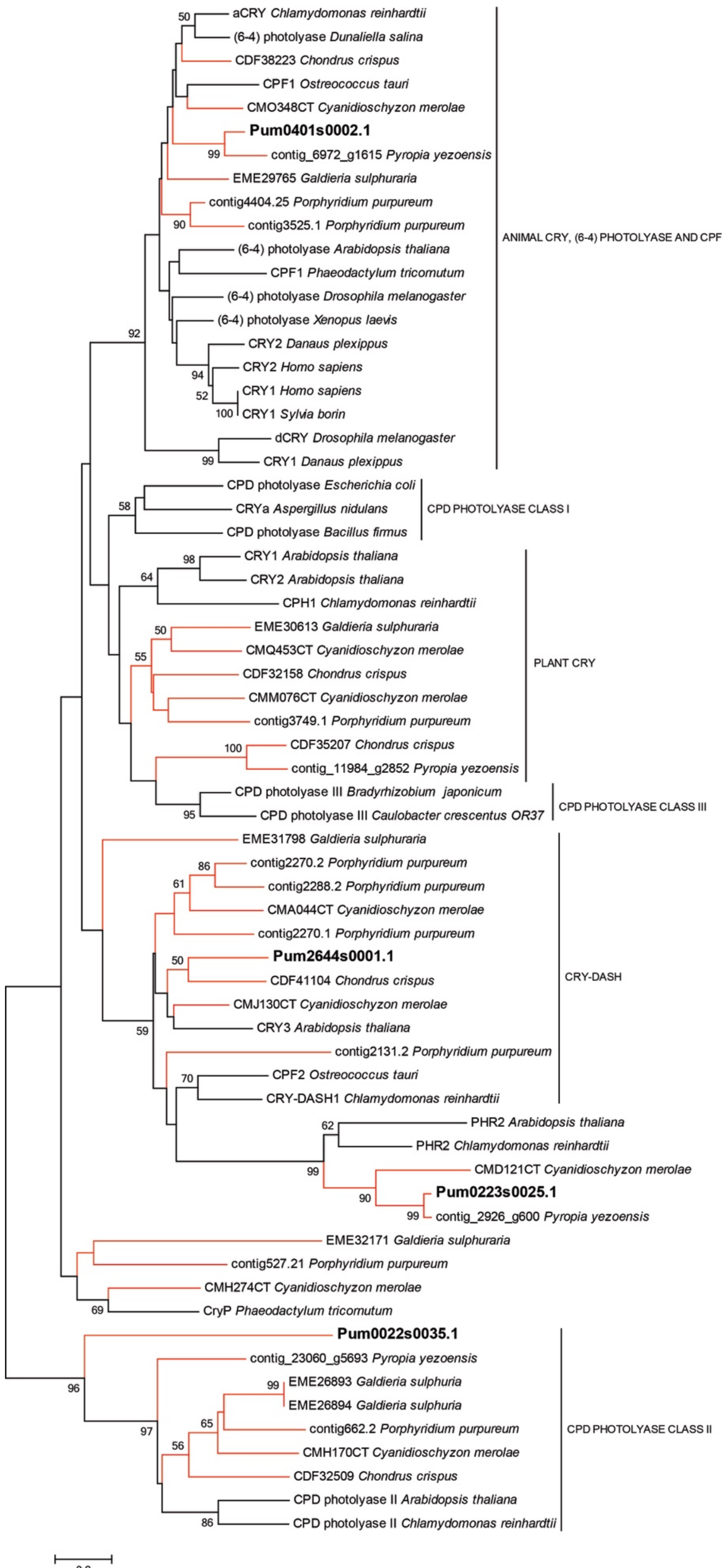

aCRY Chlamydomonas reinhardtii
(6-4) photolyase Dunaliella salina
CDF38223 Chondrus crispus
CPF1 Ostreococcus tauri
CMO348CT Cyanidioschyzon merolae
Pum0401s0002.1
contig_6972_g1615 Pyropia yezoensis
EME29765 Galdieria sulphuraria
contig4404.25 Porphyridium purpureum
contig3525.1 Porphyridium purpureum
(6-4) photolyase Arabidopsis thaliana
CPF1 Phaeodactylum tricornutum
(6-4) photolyase Drosophila melanogaster
(6-4) photolyase Xenopus laevis
CRY2 Danaus plexippus
CRY2 Homo sapiens
CRY1 Homo sapiens
CRY1 Sylvia borin
dCRY Drosophila melanogaster
CRY1 Danaus plexippus
ANIMAL CRY, (6-4) PHOTOLYASE AND CPF
CPD photolyase Escherichia coli
CRYa Aspergillus nidulans
CPD photolyase Bacillus firmus
CPD PHOTOLYASE CLASS I
CRY1 Arabidopsis thaliana
CRY2 Arabidopsis thaliana
CPH1 Chlamydomonas reinhardtii
EME30613 Galdieria sulphuraria
CMQ453CT Cyanidioschyzon merolae
CDF32158 Chondrus crispus
CMM076CT Cyanidioschyzon merolae
contig3749.1 Porphyridium purpureum
CDF35207 Chondrus crispus
contig_11984_g2852 Pyropia yezoensis
PLANT CRY
CPD photolyase III Bradyrhizobium japonicum
CPD photolyase III Caulobacter crescentus OR37
CPD PHOTOLYASE CLASS III
EME31798 Galdieria sulphuraria
contig2270.2 Porphyridium purpureum
contig2288.2 Porphyridium purpureum
CMA044CT Cyanidioschyzon merolae
contig2270.1 Porphyridium purpureum
Pum2644s0001.1
CDF41104 Chondrus crispus
CMJ130CT Cyanidioschyzon merolae
CRY3 Arabidopsis thaliana
contig2131.2 Porphyridium purpureum
CPF2 Ostreococcus tauri
CRY-DASH1 Chlamydomonas reinhardtii
CRY-DASH
PHR2 Arabidopsis thaliana
PHR2 Chlamydomonas reinhardtii
CMD121CT Cyanidioschyzon merolae
Pum0223s0025.1
contig_2926_g600 Pyropia yezoensis
EME32171 Galdieria sulphuraria
contig527.21 Porphyridium purpureum
CMH274CT Cyanidioschyzon merolae
CryP Phaeodactylum tricornutum
Pum0022s0035.1
contig_23060_g5693 Pyropia yezoensis
EME26893 Galdieria sulphuria
EME26894 Galdieria sulphuria
contig662.2 Porphyridium purpureum
CMH170CT Cyanidioschyzon merolae
CDF32509 Chondrus crispus
CPD photolyase II Arabidopsis thaliana
CPD photolyase II Chlamydomonas reinhardtii
CPD PHOTOLYASE CLASS II
0.2

Figure S27. Phytolyase/cryptochrome phylogeny

Phylogenetic analysis of the photolyases/cryptochromes family members identified from *Porphyra umbilicalis* using the neighbor-joining method in MEGA6 (79). The CPD photolyases class II were used as outgroup. Bootstraps of 1000 replicates were performed and expressed as a percentage. The *P. umbilicalis* photolyase/cryptochrome family members were set in bold (**Pum**) at larger size. The scale bar indicates amino acid substitutions per site. Sequences from red algae are indicated by red branches. Only bootstrap values above 50 are shown. Accession numbers: aCRY *Chlamydomonas reinhardtii*, Cre06.g278251.t1.1 (sequences of *C. reinhardtii* are from *Chlamydomonas reinhardtii* v5.5, https://phytozome.jgi.doe.gov/pz/portal.html); dCRY *Drosophila melanogaster*, BAA35000.1; CRY1 *Danaus plexippus*, AAX58599.1; CRY1 *Homo sapiens*, NP_004066.1; CRY2 *Homo sapiens*, Q49AN0.2; CRY1 *Sylvia borin*, Q6ZZY0.1; CRY2 *Danaus plexippus*, ABA62409.1; (6-4) photolyase *Drosophila melanogaster,* NP_477188.1; (6-4) photolyase *Arabidopsis thaliana*, 3FY4; (6-4) photolyase *Xenopus laevis*, NP_001081421.1; (6-4) photolyase *Dunaliella salina*, AAX56342.1; CPF1 *Ostreococcus tauri,* AAU14280.1; CRY-DASH1 *Chlamydomonas reinhardtii*, Cre02.g078939.t1.1; CPH1 *Chlamydomonas reinhardtii* , Cre06.g295200.t1.2; CryP *Phaeodactylum tricornutum*, 54342(sequences of *P. tricornutum* are from JGI, http://genome.jgi.doe.gov/Phatr2/Phatr2.home.html); CPF1 *Phaeodactylum tricornutum*, 27429; CRY1 *Arabidopsis thaliana*, NP_567341.1; CRY2 *Arabidopsis thaliana*, AEE27693.1; CRY3 *Arabidopsis thaliana*, 2IJG; CPF2 *Ostreococcus tauri*, Q5IFN2.1; CPD photolyase *Escherichia coli*, WP_032176081.1; CPD photolyase *Bacillus firmus*, 2017201A; CRYa *Aspergillus nidulans*, XP_657991.1; CPD photolyase II *Arabidopsis thaliana*, NP_849651.1; CPD photolyase II *Chlamydomonas reinhardtii*, Cre12.g533000.t1.2; PHR2 *Arabidopsis thaliana*, AAM64933.1; PHR2 *Chlamydomonas reinhardtii*, Cre12.g534550.t1.1; CPD photolyase III *Bradyrhizobium japonicum*, WP_038967081.1; CPD photolyase III *Caulobacter crescentus* OR37, ENZ83876.1. The accession numbers not specified are from NCBI. Sequences of *Chondrus crispus*, *Cyanidioschyzon merolae* and *Galdieria sulphuria* are from http://plants.ensembl.org/index.html. Sequences of *Porphyridium purpureum* are from http://cyanophora.rutgers.edu/porphyridium/. Sequences of *Pyropia yezoensis* are from http://nrifs.fra.affrc.go.jp/ResearchCenter/5_BB/genomes/nori/index.html.

Central carbon metabolism

Table S28. Carbon metabolic genes in *Porphyra*

Inventory of annotated genes related to carbon concentrating mechanisms and central carbon metabolism. Subcellular localizations were predicted using TargetP 1.1 and characterized as “probable” with confidence values of 1 or 2 and “possible” with scores of 3-5. The following proteins were not found in the *Porphyra* assembly: delta- and epsilon- carbonic anhydrase, pyruvate phosphate dikinase, pyruvate carboxylase, pyruvate decarboxylase, phosphoenolpuyruvate synthase, aldose-1-epimerase, hexokinase, L-arabinokinase, NADP-dependent glyceraldehyde-3-phosphate dehydrogenase, 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase, glyceraldehyde-3-phosphate:ferredoxin oxidoreductase, glyceraldehyde-3-phosphate dehydrogenase (NAD(P)($^+$)), multiple inositol-polyphosphate phosphatase / 2,3-bisphosphoglycerate 3-phosphatase.

	Porphyra gene identifier	Gene Symbol	Defline	Notes
Inorganic CCMs	Pum0065s0011.1	bCA1	beta-carbonic anhydrase	Possibly chloroplastic
	Pum0106s0035.1	bCA2	beta-carbonic anhydrase	Possibly chloroplastic
	Pum0072s0091.1	aCA1	alpha-carbonic anhydrase	Probably secreted
	Pum0087s0047.1	aCA2		Probably secreted
	Pum0869s0006.1	aCA3		Possibly cytoplasmic
	Pum0335s0021.1	aCA4		Probably mitochondrial
	Pum0328s0034.1	aCA5		Possibly secreted
	Pum1457s0001.1	aCA6		Possibly cytoplasmic
	Pum0933s0007.1	aCA7		Probably cytoplasmic
	Pum0184s0012.1	aCA8		Probably mitochondrial
	Pum2057s0001.1	gCA1	gamma-carbonic anhydrase	Probably mitochondrial
	Pum0353s0003.1	gCA2		Probably mitochondrial
	Pum0290s0001.1	gCA3		Possibly cytoplasmic
	Pum0504s0017.1	SLC4a1	Solute carrier family 4 protein / Anion exchanger	Probably cytoplasmic
	Pum0536s0006.1	SLC4a2		Probably chloroplastic
Organic CCMs	Pum0262s0021.1	*SSADH*	succinate-semialdehyde dehydrogenase	Probably cytoplasmic
	Pum0205s0026.1	*PEPC*	Phosphoenolpyruvate carboxylase	Probably cytoplasmic
	Pum0294s0013.1	*PepCK*	Phosphoenolpyruvate carboxykinase	Possibly mitochondrial
	Pum2527s0001.1	*MDH1*	Malate dehydrogenase	Possibly cytoplasmic, see also TCA cycle
	Pum0393s0028.1	*MDH2*		Probably mitochondrial
	Pum0523s0015.1	*MDH3*		Possibly chloroplastic
	Pum0161s0014.1	*NAD-ME1*		Possibly cytoplasmic
	Pum0058s0083.1	*NADP-ME1*	NADP-dependent malic enzyme	Possibly mitochondrial
	Pum1804s0001.1	*PK1*	Pyruvate kinase	Possibly cytoplasmic, see also Glycolysis
	Pum1688s0001.1	*PK3*		Possibly cytoplasmic
	Pum0760s0007.1	*PK2*		Possibly cytoplasmic
Pentose phosphate pathway	Pum0399s0018.1	*PGI1*	glucose-6-phosphate isomerase	see also glycolysis
	Pum0254s0024.1	*PGI2*		
	Pum0237s0011.1	*G6PD1*	glucose-6-phosphate 1-dehydrogenase	
	Pum0197s0010.1	*G6PD2*	glucose-6-phosphate 1-dehydrogenase	
	Pum0625s0004.1	*6PGL*	6-phosphogluconolactonase	
	Pum1630s0002.1	*6PGD1*	6-phosphogluconate dehydrogenase	
	Pum0518s0009.1	*6PGD2*	6-phosphogluconate dehydrogenase	
	Pum2821s0001.1	*RPI*	ribose 5-phosphate isomerase	
	Pum0459s0007.1	*TKL*	transketolase	
	Pum0237s0025.1	*TAL*	transaldolase	
	Pum2042s0001.1	*FBA1*	fructose-bisphosphate aldolase	see also glycolysis
	Pum3097s0001.1	*FBA2*		
	Pum0096s0001.1	*FBA3*		
	Pum0482s0016.1	*FBA4*		
	Pum0256s0023.1	*FBA5*		
	Pum0333s0011.1	*PFK1*	6-phosphofructokinase	see also glycolysis
	Pum0835s0007.1	*PFK2*		
	Pum0518s0007.1	*FBPase1*	fructose-1,6-bisphosphatase	see also glycolysis
	Pum0214s0001.1	*FBPase2*		
	Pum0513s0004.1	*FBPase3*		
Glycolysis and fermentation	Pum0225s0014.1	PGM1	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	
	Pum0289s0016.1	PGM2	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	
	Pum0333s0028.1	PGM3	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	
	Pum0399s0018.1	*PGI1*	glucose-6-phosphate isomerase	see also pentose phosphate pathway

	Pum0254s0024.1	*PGI2*	glucose-6-phosphate isomerase	
	Pum0224s0012.1	*G5PE*	glucose-6-phosphate 1-epimerase	
	Pum0333s0011.1	*PFK1*	6-phosphofructokinase	see also pentose phosphate pathway
	Pum0835s0007.1	*PFK2*	6-phosphofructokinase	
	Pum0518s0007.1	*FBPase1*	fructose-1,6-bisphosphatase	see also pentose phosphate pathway
	Pum0214s0001.1	*FBPase2*	fructose-1,6-bisphosphatase	
	Pum0513s0004.1	*FBPase3*	fructose-1,6-bisphosphatase	
	Pum2042s0001.1	*FBA1*	fructose-bisphosphate aldolase	see also pentose phosphate pathway
	Pum3097s0001.1	*FBA2*	fructose-bisphosphate aldolase	
	Pum0096s0001.1	*FBA3*	fructose-bisphosphate aldolase	
	Pum0482s0016.1	*FBA4*	fructose-bisphosphate aldolase	
	Pum0256s0023.1	*FBA5*	fructose-bisphosphate aldolase	
	Pum0340s0003.1	*TIM1*	triosephosphate isomerase	
	Pum0709s0010.1	*TIM2*	triosephosphate isomerase	
	Pum0438s0011.1	*PGK1*	phosphoglycerate kinase	
	Pum1817s0001.1	*PGK2*	phosphoglycerate kinase	
	Pum0161s0062.1	*GAPDH1*	glyceraldehyde 3-phosphate dehydrogenase	
	Pum0071s0045.1	*GAPDH2*	glyceraldehyde 3-phosphate dehydrogenase	
	Pum0519s0001.1	*ENO1*	enolase	
	Pum2805s0001.1	*ENO2*	enolase	
	Pum0164s0005.1	*ENO3*	enolase	
	Pum1804s0001.1	*PK1*	pyruvate kinase	see also organic carbon concentrating mechanisms
	Pum1688s0001.1	*PK3*	pyruvate kinase	
	Pum0760s0007.1	*PK2*	pyruvate kinase	
	Pum0294s0013.1	*PepCK*	phosphoenolpyruvate carboxykinase	see also organic carbon concentrating mechanisms
	Pum0108s0010.1	*PDC-E1b*	pyruvate dehydrogenase complex E1 component	
	Pum0463s0001.1	*DLDH*	dihydrolipoamide dehydrogenase	see also TCA cycle
	Pum0101s0002.1	*PDC-E2*	dihydrolipoyllysine-residue acetyltransferase	
	Pum0153s0012.1	*ACS1*	acetyl-CoA synthetase	
	Pum0111s0054.1	*ACS2*	acetyl-CoA synthetase	
	Pum1467s0001.1	*LDH1*	L-lactate dehydrogenase	
	Pum1620s0001.1	*LDH2*	L-lactate dehydrogenase	
TCA cycle	Pum0580s0012.1	*CISY1*	Citrate synthase	
	Pum0176s0002.1	*ACO1*	Aconitase	
	Pum0254s0015.1	*IDH1*	Isocitrate dehydrogenase	
	Pum1933s0002.1	*IDH2*	Isocitrate dehydrogenase	
	Pum0399s0005.1	*OGDH1*	2-oxoglutarate dehydrogenase, E1 component	
	Pum0523s0014.1	*DLST*	2-oxoglutarate dehydrogenase E2 component	
	Pum0463s0001.1	*DLDH*	dihydrolipoamide dehydrogenase	see also glycolysis
	Pum0031s0088.1	*SUCA*	Succinyl-CoA ligase	
	Pum1533s0001.1	*SDH*	Succinate dehydrogenase	
	Pum0225s0009.1	*FUM*	Fumarate hydratase	
	Pum2527s0001.1	*MDH1*	Malate dehydrogenase	see also organic carbon concentrating mechanisms
	Pum0393s0028.1	*MDH2*	Malate dehydrogenase	
	Pum0523s0015.1	*MDH3*	Malate dehydrogenase	

Carbonic anhydrases (CAs) are an important part of inorganic carbon concentration mechanisms in algae (80). The *Porphyra* genome encodes 13 CAs, 8 of the alpha type (5 of which were supported by expression data), two beta CAs (both expressed) and three gamma CAs, all of which were expressed (see Table S28). While red algal alpha and gamma carbonic anhydrases form a sister group of CAs from the green lineage and were thus probably ancestral in Plantae, there are two or possibly three groups of red algal beta CAs (Fig. S29). They may either have been acquired several times independently in the recent evolution of red algae or constitute ancestral proteins with several losses having led to the current distribution. In the latter case, *Porphyra*-type beta-CAs were probably already present in the common ancestor of eukaryotes and later lost in the green lineage while a second copy may have been acquired by the common ancestor of Archaeplastida, possibly in the course of primary endosymbiosis.

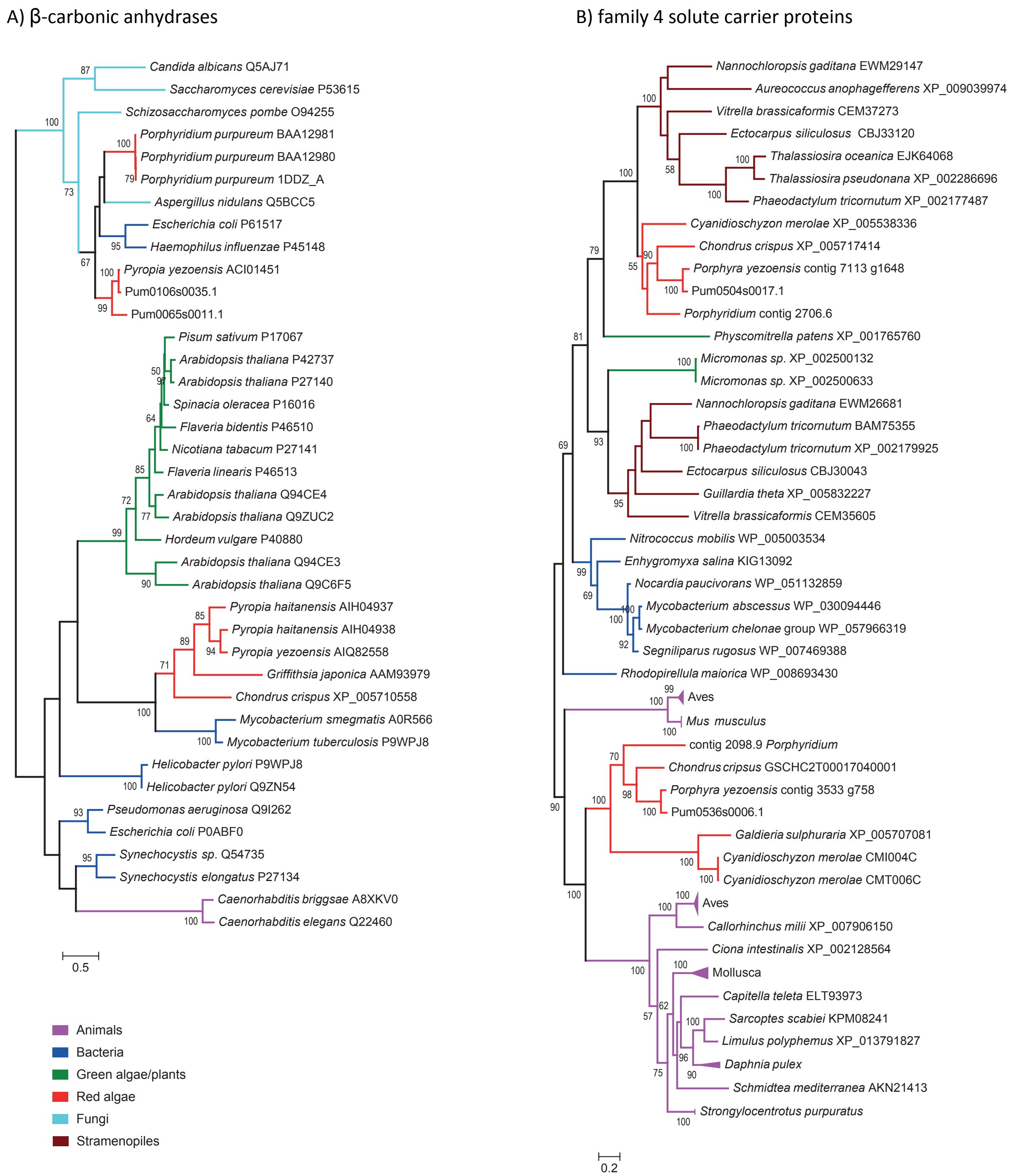


Figure S29. Carbonic anhydrase and family 4 solute carrier proteins
Unrooted maximum likelihood tree of *Porphyra* beta-CAs (A) and SLC4as (B) as well as the most similar sequences retrieved from NCBI nr and published red algal genomes. Alignments were generated with MAFFT and manually curated. Trees were constructed using phyml 3.0 using the LG substitution model and 100 bootstrap replicates.

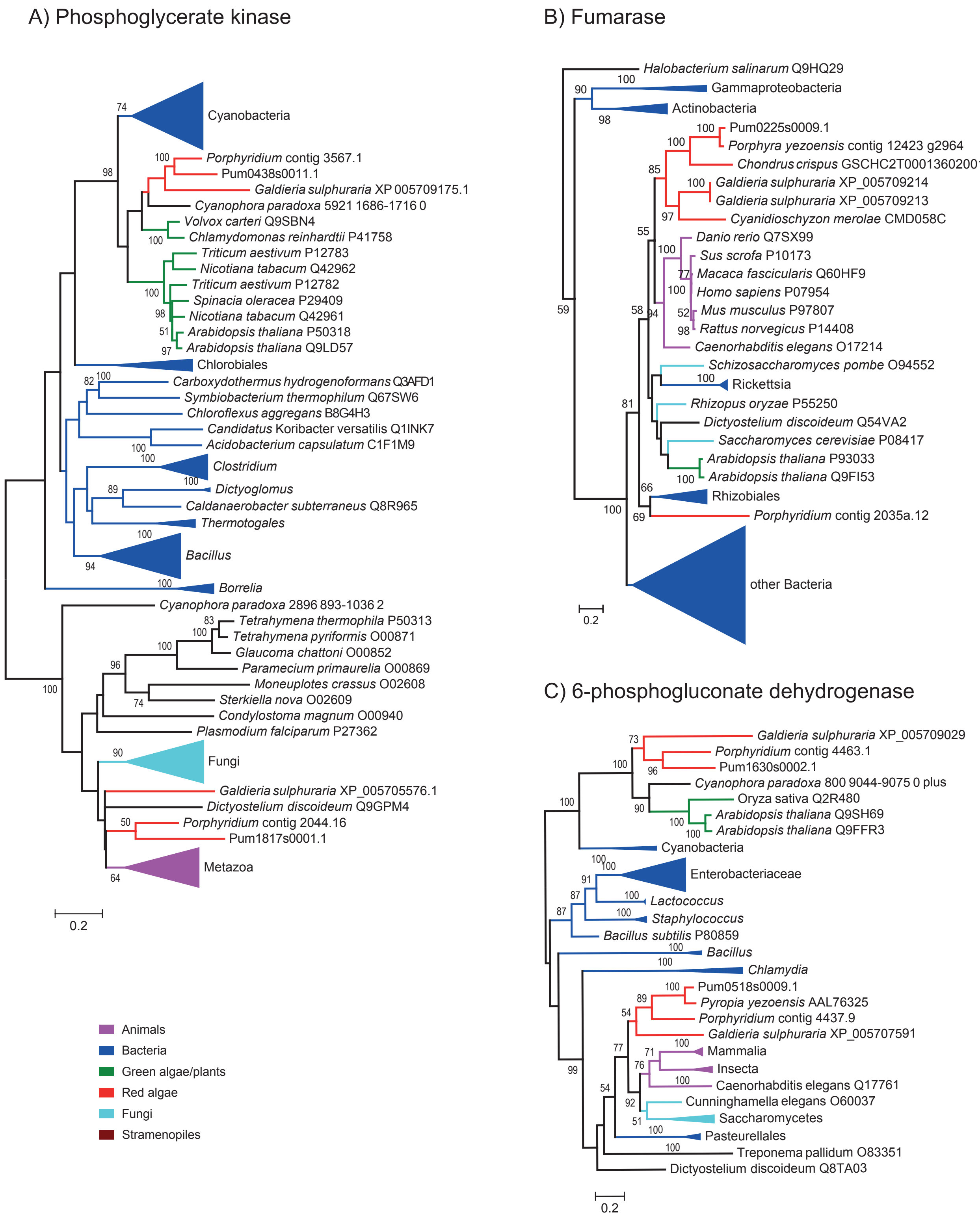

A) Phosphoglycerate kinase
Cyanobacteria
Porphyridium contig 3567.1
Pum0438s0011.1
Galdieria sulphuraria XP 005709175.1
Cyanophora paradoxa 5921 1686-1716 0
Volvox carteri Q9SBN4
Chlamydomonas reinhardtii P41758
Triticum aestivum P12783
Nicotiana tabacum Q42962
Triticum aestivum P12782
Spinacia oleracea P29409
Nicotiana tabacum Q42961
Arabidopsis thaliana P50318
Arabidopsis thaliana Q9LD57
Chlorobiales
Carboxydothermus hydrogenoformans Q3AFD1
Symbiobacterium thermophilum Q67SW6
Chloroflexus aggregans B8G4H3
Candidatus Koribacter versatilis Q1INK7
Acidobacterium capsulatum C1F1M9
Clostridium
Dictyoglomus
Caldanaerobacter subterraneus Q8R965
Thermotogales
Bacillus
Borrelia
Cyanophora paradoxa 2896 893-1036 2
Tetrahymena thermophila P50313
Tetrahymena pyriformis O00871
Glaucoma chattoni O00852
Paramecium primaurelia O00869
Moneuplotes crassus O02608
Sterkiella nova O02609
Condylostoma magnum O00940
Plasmodium falciparum P27362
Fungi
Galdieria sulphuraria XP_005705576.1
Dictyostelium discoideum Q9GPM4
Porphyridium contig 2044.16
Pum1817s0001.1
Metazoa
0.2
Animals
Bacteria
Green algae/plants
Red algae
Fungi
Stramenopiles
B) Fumarase
Halobacterium salinarum Q9HQ29
Gammaproteobacteria
Actinobacteria
Pum0225s0009.1
Porphyra yezoensis contig 12423 g2964
Chondrus crispus GSCHC2T00013602001
Galdieria sulphuraria XP_005709214
Galdieria sulphuraria XP_005709213
Cyanidioschyzon merolae CMD058C
Danio rerio Q7SX99
Sus scrofa P10173
Macaca fascicularis Q60HF9
Homo sapiens P07954
Mus musculus P97807
Rattus norvegicus P14408
Caenorhabditis elegans O17214
Schizosaccharomyces pombe O94552
Rickettsia
Rhizopus oryzae P55250
Dictyostelium discoideum Q54VA2
Saccharomyces cerevisiae P08417
Arabidopsis thaliana P93033
Arabidopsis thaliana Q9FI53
Rhizobiales
Porphyridium contig 2035a.12
other Bacteria
0.2
C) 6-phosphogluconate dehydrogenase
Galdieria sulphuraria XP_005709029
Porphyridium contig 4463.1
Pum1630s0002.1
Cyanophora paradoxa 800 9044-9075 0 plus
Oryza sativa Q2R480
Arabidopsis thaliana Q9SH69
Arabidopsis thaliana Q9FFR3
Cyanobacteria
Enterobacteriaceae
Lactococcus
Staphylococcus
Bacillus subtilis P80859
Bacillus
Chlamydia
Pum0518s0009.1
Pyropia yezoensis AAL76325
Porphyridium contig 4437.9
Galdieria sulphuraria XP_005707591
Mammalia
Insecta
Caenorhabditis elegans Q17761
Cunninghamella elegans O60037
Saccharomycetes
Pasteurellales
Treponema pallidum O83351
Dictyostelium discoideum Q8TA03
0.2

Figure S30. Phosphoglycerate kinases, fumarases, and 6-phosphogluconate dehydrogenses

Unrooted maximum likelihood tree of *Porphyra* phosphoglycerate kinases (A), fumarases (B) and 6-phosphogluconate dehydrogenases (C) as well as the most similar sequences retrieved from NCBI nr and published red algal genomes. Alignments were generated with MAFFT and manually curated. Trees were constructed using phyml 3.0 using the LG substitution model and 100 bootstrap replicates.

CAs are complemented by family 4 solute carrier proteins, which may serve to pump bicarbonate into the cell (81). The corresponding *Porphyra* proteins are predicted to be targeted to the cell (SLC4a1 = Pum0504s0017.1) and chloroplast membrane (SLC4a2 = Pum0536s0006.1), respectively. As shown in Figure S30, SLC4a1 is a typical plant protein; it is present in several red algae and was probably transferred to the stramenopile lineage during secondary endosymbiosis. SLC4a2 was also found in a range of red algae but clearly branches in a group of otherwise metazoan proteins. It is possible that SLC4a2-like proteins are ancestral in eukaryotes, as proposed for the *Porphyra* beta CAs, but in this scenario, additional duplication events and losses would be necessary to explain why red algal SLC4a2s branch within a group of metazoan proteins. Alternatively, red algae may have acquired this protein horizontally. In any case, the combination of CAs and SLC4as likely allow *Porphyra* to capture external CO_2 as carbonate and to import it into the chloroplasts, where it can be converted to CO_2 and fixed by the RUBISCO.

Regarding potential organic carbon concentrating mechanisms, the *Porphyra* genome encodes two malic enzymes: one requiring NAD (NAD-ME1 = Pum0161s0014.1) and one dependent on NADP (NADP-ME1 = Pum0058s0083.1), three malate dehydrogenases (MDH1-3 = Pum2527s0001.1, Pum0393s0028.1, Pum0523s0015.1), and one phosphoenolpyruvate carboxylase (PepC = Pum0205s0026.1). However, like in *Chondrus* (20), no phosphoenolpyruvate synthase (PPS), pyruvate phosphate dikinase (PPDK), or pyruvate carboxylase (PYC) were found. Unlike *Chondrus*, but like *Cyanidioschyzon* and *Galdieria*, *Porphyra* also encodes a phosphoenolpyruvate carboxykinase (PEPCK = Pum0294s0013.1). Assuming that one of the two pyruvate kinases (PK1-2 = Pum1804s0001.1, Pum0760s0007.1) available in *Porphyra* may serve for PEP synthesis, *Porphyra* thus, like most plants, possesses all necessary genes to perform C_4 photosynthesis. However, currently, no evidence exists that these enzymes exhibit different activities in different parts of the thallus, or during the light and dark period, respectively. Furthermore, a single-cell C_4-type carbon concentrating mechanism, as proposed for instance in diatoms (82), in which CO_2 is released from a C_4 compound only in the chloroplast seems unlikely, since none of the MEs nor the PEPCK are predicted to be localized in the latter.

Most remaining genes supporting central carbon metabolism are present and conserved in land plants (see Table S28) but in a few cases red algal proteins branched within or as a sister group to proteins from the animal lineage as also observed for SLC4a2. This was notably the case for one of the two phosphoglycerate kinases (PGK2 = Pum1817s0001.1), the fumarase (FUM = Pum0225s0009.1), and one of the two 6-phosphogluconate dehydrogenases (6PGD1 = Pum0518s0009.1) (Figure S30). In all of these cases, a contamination of the sequencing data is highly improbable, as similar enzymes can be found in several red algal genomes. In the case of phosphoglycerate kinases and 6-phosphogluconate dehydrogenases, one of the two copies each was found in a group with green plants and rooted by cyanobacteria (Fig. S30), so that we can assume this copy was acquired by the ancestor of Archaeplastida during primary endosymbiosis. The second copy of each of these genes may correspond to the ancestral eukaryotic copy that was already present in the host of the primary endosymbiosis. The reason that several red algae have, despite the evolutionary bottlenecks they have passed (20, 31), retained both copies remains to be explored. With respect to the fumarase gene, only a single copy was retained by most red algae, and it differs from the copy found in the green lineage. Neither of these copies branches with cyanobacteria, so the most plausible explanation is that both proteins are ancestral eukaryotic proteins and different copies were conserved.

Cell signaling

Table S31. Organic and ionic osmoregulation

Biological process	*Porphyra* gene identifier	Gene Symbol	Aliases or Synonyms	Defline	Description
Compatible solute, trehalose	Pum0106s0007.1	TRE1	GH37	trehalase	hydrolase enzyme that catalyzes the conversion of trehalose to glucose
	Pum0106s0006.1	TRE2	GH37	trehalase	hydrolase enzyme that catalyzes the conversion of trehalose to glucose
	Pum0082s0055.1	TPS1		Alpha,alpha-trehalose-phosphate synthase	synthesize trehalose (PMID: 9681010); Required for vegetative growt(PMID: 15181208); starch and sucrose degradation (PMID: 15667325); Involved in the regulation of glucose sensing and signaling genes (PMID: 15667325); synthesize floridoside or isofloridoside (PMID: 25323590)
	Pum0021s0026.1	TPS2		Alpha,alpha-trehalose-phosphate synthase	synthesize trehalose (PMID: 9681010); Required for vegetative growt(PMID: 15181208); starch and sucrose degradation (PMID: 15667325); Involved in the regulation of glucose sensing and signaling genes (PMID: 15667325); synthesize floridoside or isofloridoside (PMID: 25323590)
	Pum0615s0010.1	TPP1		Probable trehalose-phosphate phosphatase	Catalyzes the hydrolysis of trehalose 6-phosphate to trehalose and phosphate; prevents the accumulation of toxic levels of trehalose 6-phosphate. Trehalose accumulation may improve abiotic stress tolerance
	Pum0615s0014.1	TPP2		Probable trehalose-phosphate phosphatase	Catalyzes the hydrolysis of trehalose 6-phosphate to trehalose and phosphate; prevents the accumulation of toxic levels of trehalose 6-phosphate. Trehalose accumulation may improve abiotic stress tolerance
	Pum0165s0020.1	TRES		Trehalose synthase/amylase TreS	(bacteria) Catalyzes the reversible interconversion of maltose and trehalose by transglucosylation. TreS plays a key role in the utilization of trehalose for the production of glycogen, and might also function as a sensor and/or regulator of trehalose levels within the cell.Thus, when trehalose levels in the cell become dangerously low, TreS can expedite the conversion of glycogen to maltose via its amylase activity and then convert the maltose to trehalose; but this enzyme also can expedite or promote the conversion of trehalose to glycogen when cytoplasmic trehalose levels become too high (PMID: 15511231)
Compatible solute, floridoside	Pum1445s0001.1	AGAL1.1		Alpha-galactosidase 1	may function to degrade floridoside; may function in cell wall loosening and cell wall expansion.
	Pum0127s0045.1	AGAL1.2		Alpha-galactosidase 1	may function to degrade floridoside; may function in cell wall loosening and cell wall expansion.
	Pum0352s0011.1	AGAL1.3		Alpha-galactosidase 1	may function to degrade floridoside; may function in cell wall loosening and cell wall expansion.
Compatible solute, betaine	Pum0215s0031.1	ADH1		alcohol dehydrogenase	synthesizes betaine aldehyde
	Pum0023s0007.1	CHDH1		Choline dehydrogenase	This protein is involved in step 1 of the subpathway that synthesizes betaine aldehyde from choline; Choline + acceptor = betaine aldehyde + reduced acceptor.
	Pum0313s0014.1	BETA1		Oxygen-dependent choline dehydrogenase	This protein is involved in step 1 of the subpathway that synthesizes betaine aldehyde from choline; Choline + acceptor = betaine aldehyde + reduced acceptor.
	Pum1084s0002.1	BETA2		Oxygen-dependent choline dehydrogenase	This protein is involved in step 1 of the subpathway that synthesizes betaine aldehyde from choline; Choline + acceptor = betaine aldehyde + reduced acceptor.
Compatible solute, proline					
	Pum0721s0004.1	PROB		Glutamate 5-kinase	Catalyzes the transfer of a phosphate group to glutamate to form L-glutamate 5-phosphate.
	Pum0234s0025.1	PYCRL		Pyrroline-5-carboxylate reductase	synthesizes L-proline from L-glutamate 5-semialdehyde
	Pum0103s0061.1	POX	PROD	Proline dehydrogenase	Converts proline to delta-1-pyrroline-5-carboxylate
Ionic solutes					
	Pum0199s0002.1	GLR1		Ionotropic glutamate receptor	Ionotropic glutamate receptor, involved in cellular signalling
	Pum0218s0032.1	GLR3		Ionotropic glutamate receptor	Ionotropic glutamate receptor, involved in cellular signalling
	Pum0413s0014.1	TPC1		Two-pore calcium channel	Two-pore calcium channel, contains Ca2+-binding EF-hand between two two voltage-gated ion channel domains
	Pum0151s0018.1	HCN1		HCN channel	Hyperpolarisation activated cyclic nucleotide gated channel (K+channel)
	Pum0170s0003.1	OSCA1		hyperosmolarity induced calcium channel	Similar to hyperosmolarity induced calcium channel (OSCA)

	Pum0066s0025.1	OSCA2		hyperosmolarity induced calcium channel	Similar to hyperosmolarity induced calcium channel (OSCA)
	Pum0897s0002.1	MSL1		MscS-like channel	Similar to mechanosensitive MSL channels
	Pum0721s0003.1	CAX1		Ca2+/H+ exchanger	Ca2+/H+ exchanger, involved in Ca2+ and pH homeostasis
	Pum0352s0014.1	NCKX1		K+ dependent Na+/Ca2+ exchanger	K+ dependent Na+/Ca2+ exchanger, involved in Ca2+ homeostasis
	Pum0335s0015.1	NCKX2		similar to K+ dependent Na+/Ca2+ exchanger	similar to K+ dependent Na+/Ca2+ exchanger
	Pum0286s0023.1	NCKX3		similar to K+ dependent Na+/Ca2+ exchanger	similar to K+ dependent Na+/Ca2+ exchanger, but possesses a single sodium/potassium/calcium exchanger domain (IPR004481) whereas NCKX has tandem domains
	Pum0286s0024.1	NCKX4		similar to K+ dependent Na+/Ca2+ exchanger	similar to K+ dependent Na+/Ca2+ exchanger, but possesses a single sodium/potassium/calcium exchanger domain (IPR004481) whereas NCKX has tandem domains
	Pum0235s0011.1	EFCAX1		Ca2+/H+ exchanger protein with EF-Hand	Ca2+/H+ exchanger protein with Ca2+-binding EF-Hand
	Pum1522s0001.1	ECA1		P2A-type Ca2+-ATPase	P2A-type Ca2+-ATPase, ER-type Ca2+ pump, involved in Ca2+ homeostasis
	Pum0171s0015.1	ACA1		P2B-type Ca2+-ATPase	P2A-type Ca2+-ATPase, autoinhibited Ca2+ pump, involved in CA2+ homeostasis
	Pum1746s0001.1	NAK1		P2C-type Na+/K+-ATPase	P2C-type Na+/K+-ATPase, involved in Na+ and K+ transport
	Pum0108s0015.1	NAK2		P2C-type Na+/K+-ATPase	P2C-type Na+/K+-ATPase, involved in Na+ and K+ transport
	Pum0322s0022.1	NAK3		P2C-type Na+/K+-ATPase	P2C-type Na+/K+-ATPase, involved in Na+ and K+ transport
	Pum0127s0014.1	AHA1		P3A-type H+-ATPase	P3A-type H+-ATPase, autoinhibited H+-ATPase, involved in pH homeostasis
	Pum0027s0055.1	SPF1		P5A-type ATPase	P5A-type ATPase, substrate not known
	Pum0076s0036.1	VTC3		protein kinase: type 2C protein phosphatase	hybrid protein kinase::type 2C protein phosphatase, putative role in the regulation of ascorbate biosynthesis
	Pum1621s0001.1	GMEa		GDP-D-mannose-3,5-epimerase	GDP-D-mannose-3,5-epimerase, catalyses reversible epimerisation of GDP-D-mannose to GDP-L-galactose
	Pum0268s0015.1	IMP1		Myo-inositol-1-phosphatase	Myo-inositol-1-phosphatase, catalyses the converson of inositol monophosphate to inositol and inorganic phosphate
	Pum2015s0001.1	LgalDH		L-galactose dehydrogenase	Catalyses the NAD-dependent oxidation of L-galactose at C1 to produce L-galactono-1,4-lactone
	Pum0077s0055.1	GLDH		L-galactono-1,4-lactone dehydrogenase	Catalyses the C1 oxidation of L-galactono-1,4-lactone, the terminal step in the biosynthesis of ascorbic acid

Table S32. Calcium-dependent kinases

Distribution of calcium dependent protein kinases (CDPK) and the new class of calcium dependent tyrosine kinase like proteins (CDTKL) in red algal genomes and transcriptomes.

Genome	Transcriptome MMET ID			CDPK	CDTKL
y		Florideophyceae	*Porphyra umbilicalis*	0	4
y		Bangiophyceae	*Chondrus crispus*	0	2
y		Porphyridiophyceae	*Porphyridium purpureum*	4	0
y		Cyanidiophyceae	*Galdieria sulphuraria*	0	1
y		Cyanidiophyceae	*Cyanidioschyzon merolae*	0	1
n	MMETSP0312	Compsopogonophyceae	*Compsopogon coeruleus*	2	2
n	MMETSP0011_2	Stylonematophyceae	*Rhodosorus marinus*	5	1
n	MMETSP0315	Stylonematophyceae	*Rhodosorus marinus*	5	1
n	MMETSP0167	Rhodellophyceae	*Rhodella maculata*	5	1
n	MMETSP0314	Rhodellophyceae	*Rhodella maculata*	5	2
n	MMETSP1172	Porphyridiophyceae	*Timspurckia oligopyrenoides*	2	2
n	MMETSP1353	Porphyridiophyceae	*Erythrolobus australicus*	4	1
n	MMETSP0313	Porphyridiophyceae	*Porphyridium aerugineum*	2	3
y		Glaucophyta	*Cyanophora paradoxa*	0	0

Table S33. Calcium sensor proteins

	Gene ID v1.4	**Gene name**	**Description**
Ca^{2+}-binding proteins	Pum0031s0074.1	CNB1	Calcineurin B
	Pum0033s0060.1	CAM1	Calmodulin
	Pum0144s0010.1	CAM2	Calmodulin
	Pum0060s0052.1	CAML1	Calmodulin-like
CDTKLs	Pum0027s0071.1	CDTKL1	Calcium dependent tyrosine kinase like
	Pum0151s0026.1	CDTKL2	Calcium dependent tyrosine kinase like
	Pum1603s0001.1	CDTKL3	Calcium dependent tyrosine kinase like
	Pum1872s0001.1	CDTKL4	Calcium dependent tyrosine kinase like
Proteins exhibiting similarity to Ca^{2+} sensor kinases but lacking known Ca^{2+} regulatory domains	Pum0306s0027.1	STT1	serine/threonine protein kinase
	Pum0098s0013.1	STT2	serine/threonine protein kinase
	Pum0483s0009.1	STT3	serine/threonine protein kinase
	Pum0672s0009.1	STT4	serine/threonine protein kinase
	Pum0422s0024.1	STT5	serine/threonine protein kinase
	Pum0368s0022.1	STT6	serine/threonine protein kinase
	Pum0081s0005.1	STT8	serine/threonine protein kinase
	Pum0466s0004.1	STT9	serine/threonine protein kinase

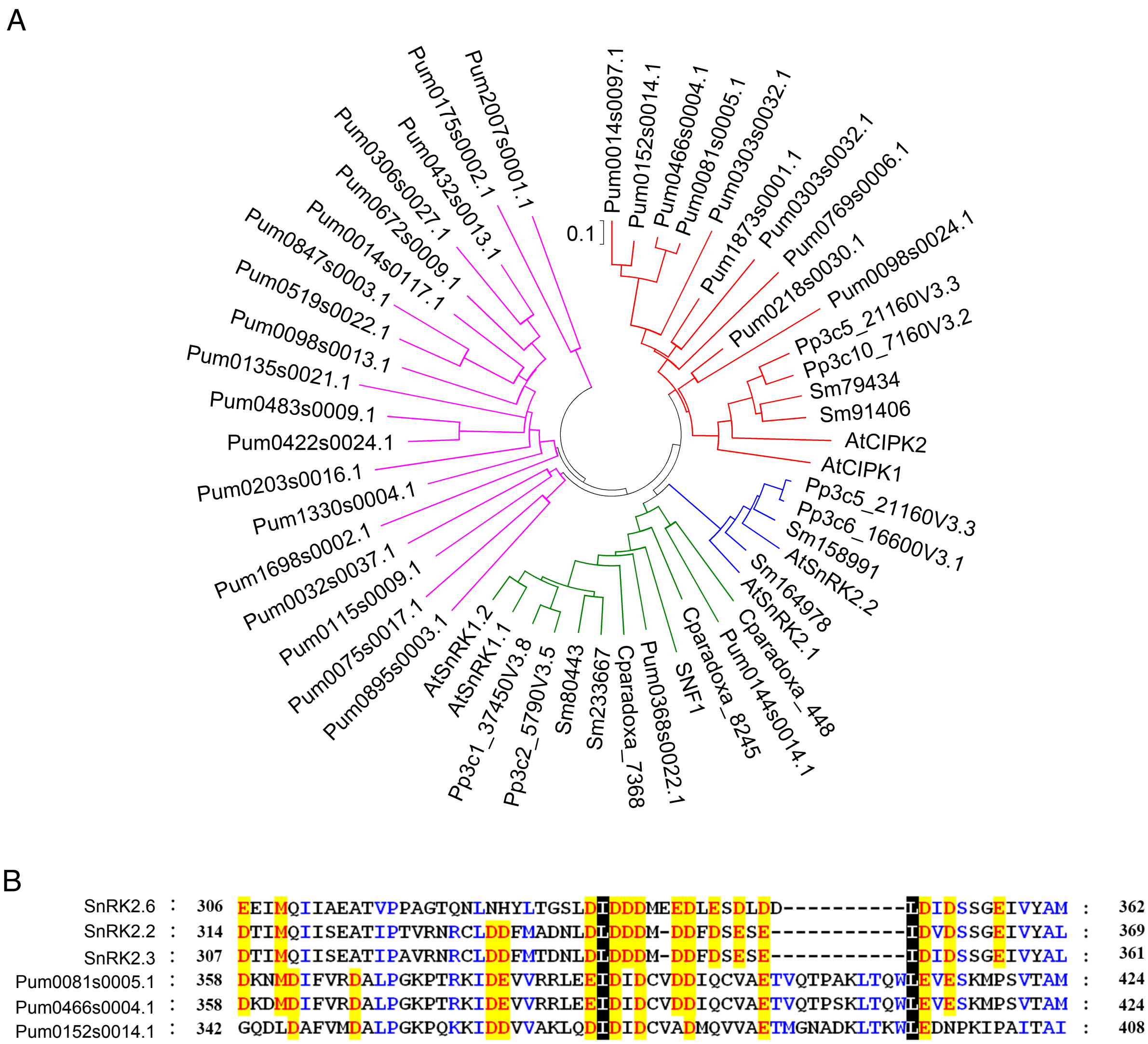


Figure S34. Phylogenetic analysis of SnRK proteins in *Porphyra umbilicalis*.

(A) Neighbor-joining tree was generated based on full-length protein sequences of SnRK proteins from *Porphyra umbilicalis* (Pumb), *Arabidopsis* (At), *Physcomitrella patens* (Pp), *Selaginella moellendorfii* (Sm), *Cyanophora paradoxa* (Cparadoxa). The green branches show the SnRK1 clade, red branches show the SnRK3/CIPK subclade, and the purple branches show *Porphyra*-specific SnRK-like kinases. (B) Sequence alignment of the C-terminal domain of *Arabidopsis* SnRK2s and three *Porphyra* SnRK proteins of the SnRK3/CIPK clade. Conserved acidic residues are highlighted in red.

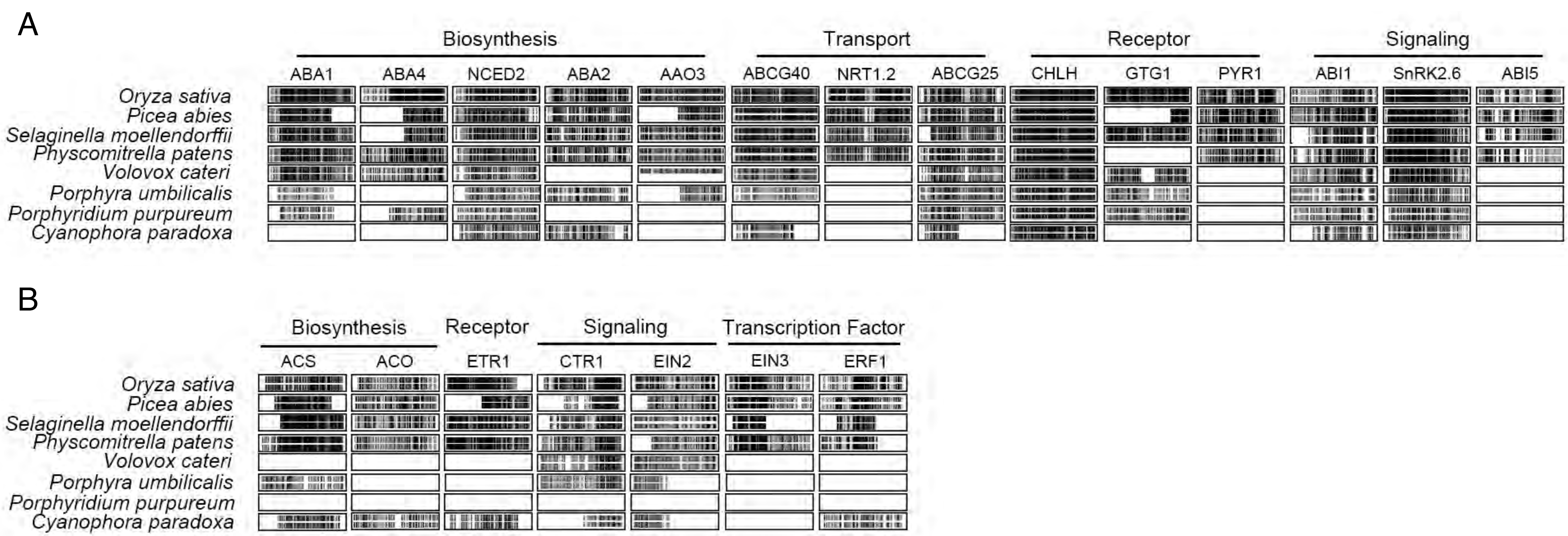

A
Biosynthesis
Transport
Receptor
Signaling
ABA1
ABA4
NCED2
ABA2
AAO3
ABCG40
NRT1.2
ABCG25
CHLH
GTG1
PYR1
ABI1
SnRK2.6
ABI5
Oryza sativa
Picea abies
Selaginella moellendorffii
Physcomitrella patens
Volovox cateri
Porphyra umbilicalis
Porphyridium purpureum
Cyanophora paradoxa
B
Biosynthesis
Receptor
Signaling
Transcription Factor
ACS
ACO
ETR1
CTR1
EIN2
EIN3
ERF1
Oryza sativa
Picea abies
Selaginella moellendorffii
Physcomitrella patens
Volovox cateri
Porphyra umbilicalis
Porphyridium purpureum
Cyanophora paradoxa

Figure S35. Sequence conservation of proteins involved in biosynthesis and signal transduction of abscisic acid (ABA) and ethylene.

Displays of pairwise sequence alignment of known *Arabidopsis* protein to their closest counterpart in *Oryza sativa* (monocot angiosperm), *Picea abies* (gymnosperm), *Selaginella moellendorfii* (lycophyte), *Physcomitrella patens* (moss), *Volvox carteri* (chlorophyte), *Porphyra umbilicalis*, *Porphyridium purpureum* (red alga), and *Cyanophora paradoxa* (glaucophyte). Black and grey vertical lines represent identical and similar amino acids, respectively.

Table S36 ABA and ethylene related gene identifiers

Gene	*Arabidopsis thaliana*	*Oryza sativa*	*Picea abies*	*Selaginella moellendorffii*	*Physcomitrella patens*	*Volvox carteri*	*Porphyra umbilicalis*	*Porphyridium purpureum*	*Cyanophora paradoxa*
ABA									
ABA1	AT5G67030	LOC_Os04g37619.1	MA_10435618g0010	172667	Pp3c12_10330V3.3	Vocar.0004s0501.1	Pum0264s0015.1	Porphyridium_2132.7	
ABA4	AT1G67080	LOC_Os01g03750.1	MA_102175g0010	68744	Pp3c12_21540V3.1	Vocar.0011s0240.1		Porphyridium_3458.4	
NCED2	AT4G18350	LOC_Os03g44380.1	MA_10428505g0020	233638	Pp3c16_17210V3.1	Vocar.0057s0024.1	Pum0231s0028.1	Porphyridium_3463.3	Cparadoxa_11135_140-245_0
ABA2	AT1G52340	LOC_Os03g59610.1	MA_2492g0010	88869	Pp3c8_9180V3.7		Pum0076s0014.1		Cparadoxa_7779_2762-3149_0
AAO3	AT2G27150	LOC_Os07g18120.1	MA_10437199g0010	444239	Pp3c5_19220V3.2	Vocar.0007s0135.1	Pum0533s0004.1		
ABCG40	AT1G15520	LOC_Os09g16380.1	MA_105476g0010	96758	Pp3c22_9340V3.1	Vocar.0049s0028.1	Pum0422s0015.1		Cparadoxa_664_1045-1184_0
NRT1.2	AT1G69850	LOC_Os06g38294.1	MA_10431946g0010	103849	Pp3c17_16320V3.1				
ABCG25	AT1G71960	LOC_Os11g07600.1	MA_18507g0010	90169	Pp3c6_8850V3.3	Vocar.0028s0079.1	Pum0229s0030.1	Porphyridium_2187.6	Cparadoxa_7055_14-155_0
CHLH	AT5G13630	LOC_Os03g20700.1	MA_108278g0010	144086	Pp3c10_13960V3.2	Vocar.0003s0179.1	Pum0809s0010.1	Porphyridium_431.16	Cparadoxa_9833_1088-1404_1
GTG1	AT1G64990	LOC_Os04g51180.1	MA_8491653g0010	235006		Vocar.0074s0009.1	Pum0177s0012.1	Porphyridium_522.10	
PYR1	AT4G17870	LOC_Os10g42280.1	MA_27987g0010	80077	Pp3c26_15240V3.1				
ABI1	AT4G26080	LOC_Os01g40094.1	MA_120757g0010	113714	Pp3c11_18330V3.1	Vocar.0013s0168.1	Pum0197s0026.1	Porphyridium_510.16	Cparadoxa_7581_9474-9635_2
SnRK2.6	AT4G33950	LOC_Os03g41460.1	MA_10430788g0010	158991	Pp3c5_21160V3.3	Vocar.0029s0126.1	Pum0368s0022.1	Porphyridium_2056.5	Cparadoxa_8245_2813-2940_0
ABI5	AT2G36270	LOC_Os01g64000.1	MA_61649g0010	266892	Pp3c20_7230V3.2				
Ethylene									
ACO	AT2G19590.1	Os06g37590.1	MA_54476g0010	101250	Pp3c12_18540V3.2				Cparadoxa_714_3289-3380_1
ACS	AT2G43750.1	Os01g74650.3	MA_10434700g0020	232271	Pp3c22_15210V3.2		Pum0076s0061.1		Cparadoxa_11763_1543-1591_0
CTR1	AT5G03730.2	Os02g32610.3	MA_35694g0010	10499	Pp3c12_3550V3.3	Vocar.0035s0070.1	Pum0138s0034.1		Cparadoxa_7087_4628-4660_2
EIN2	AT5G03280.1	Os07g06130.2	MA_17425g0010	447098	Pp3c27_2050V3.2	Vocar.0072s0016.1	Pum2079s0001.1		Cparadoxa_545_3473-3705_0
EIN3	AT3G20770.1	Os03g20780.1	MA_53004g0010	112063	Pp3c11_15260V3.1				
ERF1	AT3G23240.1	Os03g64260.1	MA_9905g0010	68470	Pp3c2_33310V3.1				Cparadoxa_9552_40-760_0
ETR1	AT1G66340.1	Os03g49500.2	MA_67531g0010	110685	Pp3c27_3830V3.3				Cparadoxa_3562_1930-1945_0

Cell cycle regulation

In plants, the major drivers of the cell cycle are the A- and B- type cyclin-dependent kinases (CDKs), where CDKA controls the transition from G_1/S and CDKB controls the transition from G_2/M (83). CDKC genes are thought to play a role in transcriptional control of the cell cycle, while CDKD activates A-type CDKs (84). *Porphyra* has 6 cyclin-dependent kinase (CDK) genes, 4 cyclin genes, single orthologs of the retinoblastoma tumor suppressor gene RB (=RBR in Table S37), transcription factors E2F1 and DP1, as well as a cyclin-dependent kinase regulatory subunit 1, CKS1 (Table S37). Single orthologs of CDKA, CDKB, CDKC, CDKD, CDKE, and CDKG were identified in *Porphyra.* Recently, CDKG1 has been implicated in regulating the number of mitotic divisions in *Chlamydomonas* (85).

Of the four cyclin genes identified in *Porphyra*, three of the four (*CYCC1*, *CYCT1, CYCU3*) are thought to play a role in transcriptional regulation (86). Cyclin D genes are thought to act both during the transition from G_1/S and also G_2/M (83); however, we did not find a cyclin D gene in *Porphyra* or in any other red alga following extensive search of red algal transcriptomes (*Compsopogon caeruleus, Erythrolobus australicus, Madagascaria erythrocaldiodes, Porphyra purpurea, Porphyra umbilicalis, Rhodella maculata, Rhodosorus marinus, Timspurckia oligopyrenoides* (1, 29, 30)) and red algal genomes (*Cyanidioschyzon merolae* (19)*, Galdieria sulphuraria* (87), *Chondrus crispus* (20), *Porphyra umbilicalis* (this work)). All hits had E-values $>10^{-6}$ and lacked a conserved LxCxE motif.

Cyclin D genes are widely conserved in plants and animals where they play a key role in regulating the G_1/S transition by partnering with cyclin-dependent kinases to phosphorylate and regulate retinoblastoma-related proteins (also known as pocket proteins). Most cyclin D proteins have a conserved LxCxE motif in the N-terminal half of the protein that is important for their interaction with their CDK partner (88, 89). Cyclin D genes are found in land plants (Embryophyta) (88) and green algae (Chlorophyta) (44, 90). Lacking cyclin D suggests that the G1/S checkpoint is regulated by a different cyclin and/or CDK or both. Alternatively, cyclin D could have an alternative motif to the CDK binding LxCxE motif, however, no blast hit had a degenerate LxCxE motif (e.g. LxCxx, LxxxE, LxCxx, xCxE). CYCA might be performing the duties of the D-type cyclins in red algae. Interestingly, no cyclin E2 gene nor a cyclin H gene was identified in *Porphyra,* yet both are present in *C. merolae* (91, 92).

Table S37. Cell cycle genes in *Porphyra*

Porphyra gene identifiers are shown in the first column. Where homology exists on a scaffold, but the automated gene prediction pipeline was unable to generate a gene model, the scaffold and coordinate range are shown.

Porphyra gene locus	Gene Symbol	Defline
Pum0177s0017.1	CDKA1	cyclin-dependent kinaseA-1
Pum1681s0001.1	CDKB1	cyclin-dependent kinase B1-1
Pum0234s0030.1	CDKC1	cyclin-dependent kinase C-1
Pum0452s0004.1	CDKD1	cyclin-dependent kinase D-1
Pum2102s0001.1	CDKE1	cyclin-dependent kinase E-1
scaffold_2658:345-1481	CDKG1	cyclin-dependent kinase G-1
Pum0328s0032.1	CYCC1	cyclin C1-1
Pum0340s0018.1	CYCA	G2/mitotic-specific cyclin-A
Pum0280s0025.1	CYCT1	cyclin T1
Pum2183s0001.1	RDP1	dual specificity phosphatase

Pum0082s0070.1	E2F1	transcription factor E2F
Pum0393s0018.1	CYCU3	cyclin U3-1
Pum0881s0003.1	DP1	transcription factor Dp-1
Pum0066s0041.1	CKS1	cyclin-dependent kinase regulatory subunit 1
Pum0014s0124.1	RB	retinoblastoma-related protein 1

Detecting microbial pathogens: Putative receptors

Little is known about the molecular mechanisms underpinning the algal immune system, but specific recognition of pathogens and grazers is reported from *Chondrus* (93) and *Gracilaria* (94). This suggests that red algae have pathogen recognition receptors.

In land plants, pathogen recognition is mediated by extracellular and intracellular receptors, some of which directly recognize microbe-derived elicitors. Schematically, recognition of saccharidic elicitors typically involves lectin domains (e.g symbiotic signals detected by the receptor SYMRK(95)), while peptide elicitors are recognised by various repeat domains (e.g. LRR domains in the flagellin receptor FLS2). Those microbe sensing receptors typically have associated ligand binding and signal transduction domains (NLRs, RLKs and LecRKs in plants (96)) to trigger defense signalling pathways. Many resistance genes characterized in higher plants belong to large gene families of up to several hundred members, which contain a Leucine Rich Repeat (LRR) domain, coupled to NB-ARC, Coiled-coil (CC) and Toll-Interleukin 1R (TIR), the latter being shared with some microbial pattern recognition receptors of vertebrates.

Because of their involvement in pathogen recognition in animals and plants (96), we searched for *Porphyra* genes containing TIR, NB-ARC and LRR domains, including genes encoding proteins containing at least one of these domains or other ligand binding domains. Because many immune receptors form multigenic families, we systematically screened all gene families of more than 5 members for protein domains relevant to signal-transduction and ligand binding. Families of interest identified in *Porphyra* are summarized in Fig. S38, together with their occurrence in selected other red algae.

	Porphyra umbilicalis	*Pyropia yezoensis*	*Chondrus crispus*	*Galdieria sulphuraria*	*Porphyridium purpureum*	*Cyanidioschyzon merolae*
TIR	✓	✓	✓	✓	✓	✓
NB-ARC	✓	✓	✓	✗ tbc	✗ tbc	✗ tbc
TIR NB-ARC WD40	✗	✗	✓	✗	✗	✗
NB-ARC WD40	✓	✓	✓	✗	✗	✗
NB-ARC TPR	✓	✓	✓	✗	✗	✗
IPT IPT Malectin Malectin	✓	✓	✓	✗	✗	✗
Malectin	✓	✓	✓	✗	✓	✗
C-type vWFA vWFA	✓	✓	✗	✗	✗	✗
C-type	✓	✓	✗	✗	✗	✗

Figure S38. Putative components of microbial sensing pathways in red algae.
Expanded gene families were inferred using OrthoMCL ((38), Version 2.0, cluster granularity 1.5). PFAM (13) domain predictions were assigned to each protein using Blast2Go ((97), v4.0.7). Signal peptides and transmembrane domains were predicted by using SignalP ((98), Version 4.1) and TMHMM2.0 (99), both with default settings. Manual blast search was performed to query *P. yezoensis* (taxid:2788), *C. crispus* (taxid:2769), *P. purpureum* (taxid:35688), *G. sulphuraria* (taxid:130081) and *C. merolae* (taxid:45157) with selected protein families and individual protein domains. *P. yezoensis* representative accessions: MALECTIN-IPT Contig_8376_q1961; NBARC-TPR contig_16619_g4046; NBARC-WD40 contig_16239_g3929; vWFA-C-lectin contig_14935_g3587. *C. crispus* representative accessions: MALECTIN-IPT : XP_005712575; TIR-NBARC-WD40: XP_005719051; TIR-NBARC-WD40: XP_005719051; NBARC-TPR XP_005712496.1. tbc, to be confirmed

Potential intracellular receptors

In contrast to higher plants, but similar to *Chondrus*, *Porphyra* only contains 9 LRR genes; no TIR-NBS-LRR or CC-NB-ARC-LRR homologue of plant resistance genes was identified. However, we identified (Table S39) at least 51 genes containing NB-ARC signalling domains associated with TPR/SLR repeats (33 genes) or WD40 repeats (18 genes). Surprisingly, no TIR domain were detected in the NB-ARC – WD40 family (nor in the corresponding *Pyropia* family), contrasting with *Chondrus*. WD40 repeats are typically involved in specific and highly variable ligand-binding, including non-self and immune receptors in the fungus *Podospora* (100). Therefore, the *Porphyra* NB-ARC-WD40 genes represent likely candidates to fulfil this function, as we suggested in a similar analysis conducted on the *Chondrus* genome. Additionally, the NB-ARC-TPR family appears equally interesting due to its relatively large size (33 genes), and its potential implication in immunity inferred in *Ectocarpus* (101).

Potential extracellular receptors

Lectins play major roles in pathogen recognition in plants and animals (102) thanks to their highly specific binding to –mostly - sugar moieties. Lectin interactions have been implicated in cell-cell recognition in red algae, including in the specific targeting of *Heterosiphonia* apical cells by the oomycete pathogen *Olpidiopsis pyropiae* (103). On one hand, as observed in *Chondrus,* no orthologs of extracellular plant immune receptors (including CERK1, CeBIP, WAK1, LYM1, LYK3, DORN1, LysM proteins) were detected in *Porphyra umbilicalis*. However, a screen for general lectin domains identified two legume-like lectins and at least three proteins having vWFA domain and C-type lectin domain. The latter domain was not detected in *Chondrus* proteins (Fig. S38), and has been extensively associated with recognition of bacterial and fungal polysaccharides in animals (102), and C-lectins displaying similar association with vWFA domains were found induced in *C. elegans* upon pathogen infection (104), and were considered likely pathogen recognition proteins in this organism.

Malectin-domain family

Finally we detected a family of at least 18 proteins containing up to 4 N-terminal malectin domains (Pfam 11721), associated with one or two IPT Immunoglobulin-like fold domains. In land plants, the malectin domain is found in tandem on the extracellular part of some plant RLKs and has been associated with the control of plant cell wall integrity (105), and its role in mediating interaction with beneficial and pathogenic microorganisms is an emerging area of research (95, 106). Interestingly, species distribution of this domain on PFAM indicates the malectin and malectin-like domain expansion is restricted to higher plants (roughly 20 to 180 proteins with, on average, one or two domains per protein). The IPT domain is often associated with animal cell surface receptors, where it also specifically binds peptides (107). This malectin/IPT family detected in *Porphyra* is conserved in *Pyropia* and *Chondrus* (Fig. S38) with similar domain organization, the closest domain organization detected in PFAM concerns a cyanobacterial protein. Detection of signal peptides and transmembrane helix highlighted these genes as likely cell surface receptors involved in cell-wall/cell adhesion or immunity in *Porphyra.*

Table S39. Candidate microbe-sensing genes

***Porphyra* gene identifier**	**Defline**
Pum0027s0083.1	C-type lectins - vWFA domain containing protein
Pum0692s0013.1	C-type lectins - vWFA domain containing protein
Pum0711s0009.1	C-type lectins - vWFA domain containing protein
Pum1585s0002.1	Legume-like Lectin domain containing protein
Pum1659s0001.1	Legume-like Lectin domain containing protein
Pum0023s0017.1	Malectin/IPT domain containing protein
Pum0053s0004.1	Malectin/IPT domain containing protein
Pum0081s0011.1	Malectin/IPT domain containing protein
Pum0098s0040.1	Malectin/IPT domain containing protein
Pum0116s0042.1	Malectin/IPT domain containing protein
Pum0176s0034.1	Malectin/IPT domain containing protein
Pum0178s0013.1	Malectin/IPT domain containing protein
Pum0224s0017.1	Malectin/IPT domain containing protein
Pum0229s0001.1	Malectin/IPT domain containing protein
Pum0256s0024.1	Malectin/IPT domain containing protein
Pum0257s0008.1	Malectin/IPT domain containing protein
Pum0284s0007.1	Malectin/IPT domain containing protein
Pum0284s0017.1	Malectin/IPT domain containing protein

Pum0285s0029.1	Malectin/IPT domain containing protein
Pum0300s0001.1	Malectin/IPT domain containing protein
Pum0319s0024.1	Malectin/IPT domain containing protein
Pum0319s0026.1	Malectin/IPT domain containing protein
Pum0322s0002.1	Malectin/IPT domain containing protein
Pum0438s0019.1	Malectin/IPT domain containing protein
Pum0512s0003.1	Malectin/IPT domain containing protein
Pum0512s0004.1	Malectin/IPT domain containing protein
Pum0536s0003.1	Malectin/IPT domain containing protein
Pum0954s0002.1	Malectin/IPT domain containing protein
Pum1283s0003.1	Malectin/IPT domain containing protein
Pum1490s0001.1	Malectin/IPT domain containing protein
Pum0022s0001.1	NBARC TPR /SLR protein
Pum0022s0008.1	NBARC TPR /SLR protein
Pum0022s0015.1	NBARC TPR /SLR protein
Pum0022s0016.1	NBARC TPR /SLR protein
Pum0023s0002.1	NBARC TPR /SLR protein
Pum0048s0021.1	NBARC TPR /SLR protein
Pum0052s0039.1	NBARC TPR /SLR protein
Pum0066s0014.1	NBARC TPR /SLR protein
Pum0066s0015.1	NBARC TPR /SLR protein
Pum0081s0014.1	NBARC TPR /SLR protein
Pum0130s0022.1	NBARC TPR /SLR protein
Pum0135s0001.1	NBARC TPR /SLR protein
Pum0135s0051.1	NBARC TPR /SLR protein
Pum0150s0006.1	NBARC TPR /SLR protein
Pum0150s0009.1	NBARC TPR /SLR protein
Pum0164s0033.1	NBARC TPR /SLR protein
Pum0167s0002.1	NBARC TPR /SLR protein
Pum0167s0003.1	NBARC TPR /SLR protein
Pum0184s0031.1	NBARC TPR /SLR protein
Pum0190s0009.1	NBARC TPR /SLR protein
Pum0218s0019.1	NBARC TPR /SLR protein
Pum0238s0005.1	NBARC TPR /SLR protein
Pum0240s0016.1	NBARC TPR /SLR protein
Pum0283s0002.1	NBARC TPR /SLR protein
Pum0283s0002.2	NBARC TPR /SLR protein
Pum0287s0003.1	NBARC TPR /SLR protein
Pum0288s0018.1	NBARC TPR /SLR protein
Pum0305s0008.1	NBARC TPR /SLR protein
Pum0317s0027.1	NBARC TPR /SLR protein
Pum0526s0002.1	NBARC TPR /SLR protein
Pum0565s0011.1	NBARC TPR /SLR protein
Pum0603s0007.1	NBARC TPR /SLR protein
Pum0651s0001.1	NBARC TPR /SLR protein
Pum0014s0030.1	NBARC WD40 domain containing protein
Pum0023s0011.1	NBARC WD40 domain containing protein
Pum0048s0005.1	NBARC WD40 domain containing protein
Pum0048s0031.1	NBARC WD40 domain containing protein
Pum0051s0003.1	NBARC WD40 domain containing protein
Pum0051s0004.1	NBARC WD40 domain containing protein
Pum0065s0005.1	NBARC WD40 domain containing protein
Pum0093s0020.1	NBARC WD40 domain containing protein
Pum0099s0030.1	NBARC WD40 domain containing protein
Pum0135s0014.1	NBARC WD40 domain containing protein
Pum0135s0018.1	NBARC WD40 domain containing protein
Pum0169s0026.1	NBARC WD40 domain containing protein
Pum0184s0016.1	NBARC WD40 domain containing protein
Pum0188s0024.1	NBARC WD40 domain containing protein
Pum0295s0009.1	NBARC WD40 domain containing protein
Pum0362s0006.1	NBARC WD40 domain containing protein
Pum0494s0008.1	NBARC WD40 domain containing protein
Pum1045s0002.1	NBARC WD40 domain containing protein

Carbohydrate-active enzymes (CAZymes)

Porphyra encodes many CAZymes (Table S40) and many are related to cell wall synthesis and remodeling (Table S40, Figure S41, Table S42) and to carbohydrate transport and storage (Table S43). We identified a complete enzymatic pathway for breakdown of floridean starch (Table S44), the storage carbohydrate of red algae (108).

Table S40. Table of carbohydrate active and cell wall-related enzymes

Porphyra gene identifiers are shown, except where scaffold identifier and start and stop coordinates in the form (scaffold:start-stop), supported by both homology and expression.

***Porphyra* gene identifier**	**Gene Symbol**	**Aliases or Synonyms**	**Defline**	**Description**
Pum1985s0002.1			epimerase	an enzyme that catalyzes the stereochemical inversion of the configuration about an asymmetric carbon atom
Pum0143s0006.1	DPE2	GH77	4-alpha-glucanotransferase	Transfers a segment of a (1->4)-alpha-D-glucan to a new position in an acceptor, which may be glucose or a (1->4)-alpha-D-glucan. Glycoside hydrolase family 77. Carbohydrate-binding module family 20.
Pum0288s0022.1		GT2	N-acetylglucosaminyltransferase	Catalysis of the transfer of an N-acetylglucosaminyl residue from UDP-N-acetyl-glucosamine to a sugar.Glycosyltransferase family 2.
Pum0093s0047.1	PMTY	GT39	Glycosyltransferase, family 39	Transfers mannose from Dol-P-mannose to Ser or Thr residues on proteins. Dolichyl D-mannosyl phosphate + protein = dolichyl phosphate + O-D-mannosylprotein. Family 39 glycosyltransferase
Pum0066s0039.1		GH16	beta-agarase or beta-porphyranase	Putative beta-agarase or beta-porphyranase, Cleaves the beta-1,4-linkages between beta-D-galactose and alpha-L-3,6-anhydro-galactose residues in agarose. Or cleaves the beta-1,4-linkages between beta-D-galactose and alpha-L-galactose-6-sulfate residues in porphyran. Family 16 glycoside hydrolase.
Pum0523s0003.1		GH85	Glycoside hydrolase, family 85	Putative endo-beta-N-acetylglucosaminidase. Family 85 glycoside hydrolase
Pum1556s0001.1		GT	putative UDP-xylosyltransferase	Catalysis of the transfer of a xylosyl group from UDP-xylose to an acceptor molecule. Non-classified glycosyltransferase, related to the family 77 glycosyltransferases.
Pum0127s0045.1		GH36	alpha-galactosidase	Catalysis of the hydrolysis of terminal, non-reducing alpha-D-galactose residues in alpha-D-galactosides, including galactose oligosaccharides, galactomannans and galactohydrolase. Family 36 of the glycoside hydrolases
Pum0076s0057.1		GT4	glycosyltransferase	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 4 of the glycosyltransferases.
Pum0530s0012.1		GT25	Family 25 glycosyltransferase	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 25 of the glycosyltransferases
Pum0266s0019.1	GPSI	GT35	glycogen phosphorylase	((1->4)-alpha-D-glucosyl)(n) + phosphate = ((1->4)-alpha-D-glucosyl)(n-1) + alpha-D-glucose 1-phosphate.Family 35 of the glycosyltransferases.
Pum0564s0005.1	CesA	GT2	cellulose synthase	UDP-glucose + (1,4-beta-D-glucosyl)(n) = UDP + (1,4-beta-D-glucosyl)(n+1). Family 2 of the glycosyltransferases.
Pum0359s0008.1		GH16	beta-agarase or beta-porphyranase	Putative beta-agarase or beta-porphyranase, Cleaves the beta-1,4-linkages between beta-D-galactose and alpha-L-3,6-anhydro-galactose residues in agarose. Or cleaves the beta-1,4-linkages between beta-D-galactose and alpha-L-galactose-6-sulfate residues in porphyran. Family 16 glycoside hydrolase.
Pum0517s0012.1		GH113	Glycoside hydrolase, family 113	Hydrolase activity, acting on glycosyl bonds. Putative endo-beta-1,4-mannanase. Family 113 glycoside hydrolase.
Pum0014s0039.1		GT7	N-acetylgalactosaminyltransferase	Catalysis of the transfer of an N-acetylgalactosaminyl residue from UDP-N-acetyl-galactosamine to an oligosaccharide. Family 7 glycosyltransferase.
Pum0372s0002.1		GT2	beta-1,3-glucosyltransferase	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 2 of the glycosyltransferases.
Pum0202s0004.1		GT14	Glycosyltransferase, family 14	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 14 of the glycosyltranferases
Pum0087s0027.1	Alg5	GT2	Glycosyltransferase, family 2	Alg5-like. Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 2 glycosyltransferase.
Pum0070s0047.1		GH35	beta-galactosidase	Hydrolysis of terminal non-reducing beta-D-galactose residues in beta-D-galactosides. Family 35 glycoside hydrolase.
Pum0153s0034.1		GH35	beta-galactosidase	Hydrolysis of terminal non-reducing beta-D-galactose residues in beta-D-galactosides. Family 35 glycoside hydrolase.
Pum0587s0004.1	SBE2	GH13	1,4-alpha-glucan-branching enzyme	Catalyzes the formation of the alpha-1,6-glucosidic linkages in starch by scission of a 1,4-alpha-linked oligosaccharide from growing alpha-1,4-glucan chains and the subsequent attachment of the oligosaccharide to the alpha-1,6 position. Family 13 glycoside hydrolase.
Pum3137s0001.1		GT34	Glycosyltransferase family 34	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 34 of the glycosyltransferases.
Pum0877s0005.1	BAM	GH14	beta-amylase	Catalysis of the reaction: (1,4-alpha-D-glucosyl)(n+1) + H2O = (1,4-alpha-D-glucosyl)(n-1) + alpha-maltose. This reaction is the hydrolysis of 1,4-alpha-glucosidic linkages in polysaccharides so as to remove successive maltose units from the non-reducing ends of the chains.Family 14 glycoside hydrolase.
Pum1271s0004.1		GH	putative alpha-N-acetylgalactosaminidase, non-classified GH	Glycosidase that has specific alpha-N-acetylgalactosaminidase activity. Non-classified glycoside hydrolase.
Pum0178s0001.1	Man1	GH47	Mannosyl-oligosaccharide 1,2-alpha-mannosidase, GH47	Man1. Catalysis of the hydrolysis of the terminal (1->2)-linked alpha-D-mannose residues in the oligo-mannose oligosaccharide Man(9)(GlcNAc)(2). Family 47 glycoside hydrolase.
Pum2732s0001.1		GT	Glycosyltransferase, non-classified	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Non-classified glycosyltransferase, related to the family 90 glycosyltransferases.
Pum0330s0004.1		GH16	beta-agarase or beta-porphyranase, GH16	Putative beta-agarase or beta-porphyranase, Cleaves the beta-1,4-linkages between beta-D-galactose and alpha-L-3,6-anhydro-galactose residues in agarose. Or cleaves the beta-1,4-linkages between beta-D-galactose and alpha-L-galactose-6-sulfate residues in porphyran. Family 16 glycoside hydrolase.
Pum2090s0001.1		GT2	Glycosyltransferase, GT2	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 2 glycosyltransferase
Pum0306s0017.1		GT4	glycosyltransferase, GT4	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 4 glycosyltransferase

Pum0546s0006.1		GT4	glycosyl transferase, GT4	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 4glycosyltransferase.
Pum0686s0007.1	GANAB	GH31	alpha-glucosidase, GH31	GANAB-like. Catalysis of the hydrolysis of a glycosyl bond. Family 31 glycoside hydrolase.
Pum1932s0001.1		GH38	alpha-mannosidase, GH38	Catalysis of the hydrolysis of terminal, non-reducing alpha-D-mannose residues in alpha-D-mannosides. Family 38 glycoside hydrolase.
Pum0579s0002.1		GT5	starch synthase, GT5	Catalysis of the reaction: ADP-glucose + (1,4)-alpha-D-glucosyl(n) = ADP + (1,4)-alpha-D-glucosyl(n+1). Family 5 glycosyltransferase.
Pum0678s0001.1	CESA	GT2	cellulose synthase, GT2	UDP-glucose + (1,4-beta-D-glucosyl)(n) = UDP + (1,4-beta-D-glucosyl)(n+1). Family 2 of the glycosyltransferases.
Pum0231s0012.1	CSL	GT2	Glycosyltransferase, GT2	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 2 of the glycosyltransferases.
Pum0082s0055.1		GT20	Putative isofloridoside and floridoside phosphate synthase/phosphatase, GT20	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Putative isofloridoside and floridoside phosphate synthase/phosphatase fusion. Family 20 of the glycosyltransferases
Pum0014s0081.1		GT	glycosyltransferase, non-classified GT	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Non-classified glycosyl transferase, related to the family 90 glycosyltransferases.
Pum0587s0001.1		GT90	Glycosyltransferase, GT90	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 90 glycosyl transferase.
Pum1806s0001.1		GT14	putative xylosyltransferase, GT14	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 14 glycosyltransferase.
Pum2856s0001.1		GT14	putative xylosyltransferase, GT14	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 14 glycosyltransferase.
Pum0103s0048.1	PMT	GT39	O-mannosyltransferase, GT39	Catalysis of the transfer of a mannosyl group to an acceptor molecule, typically another carbohydrate or a lipid. Family 39 glycosyltransferase.
Pum0505s0024.1		GT77	glycosyltransferase, GT77	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 77 glycosyltransferase.
Pum0126s0016.1		GT4	glycosyltransferase, GT4	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 4 of the glycosyltransferases.
Pum1445s0001.1		GH36	alpha-galactosidase, GH36	Catalysis of the hydrolysis of terminal, non-reducing alpha-D-galactose residues in alpha-D-galactosides, including galactose oligosaccharides, galactomannans and galactohydrolase. Family 36 of the glycoside hydrolases.
Pum2295s0001.1	Alg11	GT4	GDP-Man:Man3GlcNAc2-PP-Dol alpha-1,2-mannosyltransferase, GT4	Mannosyltransferase involved in the last steps of the synthesis of Man5GlcNAc(2)-PP-dolichol core oligosaccharide on the cytoplasmic face of the endoplasmic reticulum. Catalyzes the addition of the 4th and 5th mannose residues to the dolichol-linked oligosaccharide chain.Family 4 glycosyltransferase.
Pum1499s0001.1		GT11	glycosyltransferase, GT11	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Non-classified glycosyltransferase, related to the family 11 glycosyltransferases.
Pum0505s0008.1		GT1	glycosyltransferase, GT1	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 1 glycosyltransferase
Pum0021s0026.1		GT20	Putative isofloridoside and floridoside phosphate synthase/phosphatase, GT20	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Putative isofloridoside and floridoside phosphate synthase/phosphatase fusion. Family 20 of the glycosyltransferases
Pum0266s0010.1	GPSI	GT35	glycogen phosphorylase, GT35	Catalysis of the reaction: glycogen + phosphate = maltodextrin + alpha-D-glucose 1-phosphate. Family 35 glycosyltransferase.
Pum0071s0023.1		GH5	glycoside hydrolase, GH5	Catalysis of the hydrolysis of a glycosyl bond. Putative beta-glucosidase. Beta-glucanases participate in the metabolism of beta-glucan, a structural component of the cell wall. Family 5 glycosyl hydrolase.
Pum0053s0020.1	CSL	GT2	glycosyl transferase, GT2	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 2 glycosyltransferase.
Pum0488s0015.1		GT24	UDP-glucose:glycoprotein glucosyltransferase, GT24	Catalysis of the addition of UDP-glucose on to asparagine-linked (N-linked) oligosaccharides of the form Man7-9GlcNAc2 on incorrectly folded glycoproteins. Family 24 glycosyltransferase.
Pum0692s0001.1	Alg9	GT22	Alpha-1,2-mannosyltransferase, GT22	Putative alpha-1,2-mannosyltransferase. Catalyzes the transfer of mannose from Dol-P-Man to lipid-linked oligosaccharides. Family 22 glycosyltransferase.
Pum0200s0016.1	PMT	GT39	O-mannosyltransferase, GT39	Transfers mannose from Dol-P-mannose to Ser or Thr residues on proteins. Dolichyl D-mannosyl phosphate + protein = dolichyl phosphate + O-D-mannosylprotein. Family 39 glycosyltransferase.
Pum0121s0002.1	STT3A	GT66	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit STT3A, GT66	Catalytic subunit of the N-oligosaccharyl transferase (OST) complex which catalyzes the transfer of a high mannose oligosaccharide from a lipid-linked oligosaccharide donor to an asparagine residue within an Asn-X-Ser/Thr consensus motif in nascent polypeptide chains. N-glycosylation occurs cotranslationally and the complex associates with the Sec61 complex at the channel-forming translocon complex that mediates protein translocation across the endoplasmic reticulum (ER). SST3A seems to be involved in complex substrate specificity. STT3A is present in the majority of OST complexes and mediates cotranslational N-glycosylation of most sites on target proteins, while STT3B-containing complexes are required for efficient cotranslational glycosylation and mediate glycosylation of sites that have been skipped by STT3A. Family 66 glycosyltransferase.
Pum0380s0001.1		GH97	Alpha galactosidase, GH97	Hydrolysis of terminal, non-reducing alpha-D-galactose residues in alpha-D-galactosides, including galactose oligosaccharides, galactomannans and galactolipids. Related to the family 97 glycoside hydrolases.
Pum0125s0011.1		GT14	putative N-acetylglucosaminyltrans	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Family 14 glycosyltransferase.

			ferase, GT14	
Pum0014s0111.1	PMT	GT39	mannosyltransferase, GT39	Putative dolichyl-phosphate-mannose-protein mannosyltransferase. Catalysis of the transfer of a mannosyl group to an acceptor molecule, typically another carbohydrate or a lipid. Family 39 glycosyltransferase
Pum0014s0057.1		GH1	glycosyl hydrolase, GH1	Catalysis of the hydrolysis of a glycosyl bond. Family 1 glycosyl hydrolase.
Pum2512s0001.1		GT	Glycosyltransferase, non-classified	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Non-classified glycosyltransferase
Pum0212s0025.1		GT	Glycosyltransferase, non-classified	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Non-classified glycosyltransferase
Pum0196s0023.1		GH5	endo-beta-glucanase, GH5	Catalysis of the hydrolysis of a glycosyl bond. Family 5 glycoside hydrolase
Pum0033s0021.1	Alg8	GT57	Probable dolichyl pyrophosphate Glc1Man9GlcNAc2 alpha-1,3-glucosyltransferase, GT57	Alg8-like. Adds the second glucose residue to the lipid-linked oligosaccharide precursor for N-linked glycosylation. Transfers glucose from dolichyl phosphate glucose (Dol-P-Glc) onto the lipid-linked oligosaccharide Glc1Man9GlcNAc(2)-PP-Dol. Family 57 glycosyltransferase.
Pum0256s0042.1		GT39	O-mannosyltransferase, GT39	Catalysis of the transfer of a mannosyl group to an acceptor molecule, typically another carbohydrate or a lipid. Family 39 glycosyltransferase.
Pum0108s0018.1		GT	glycosyltransferase, non-classified	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Non-classified glycosyltransferase
Pum0014s0049.1		GT	putative beta-N-acetylglucosaminyltransferase, non-classified glycosyltransferase	Catalysis of the transfer of an N-acetylglucosaminyl residue from UDP-N-acetyl-glucosamine to a sugar. Non-classified glycosyltransferase, related to the family 14 glycosyltransferases.
Pum0231s0012.1	CSL	GT2	Glycosyltransferase, GT2, putative mannan synthase	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Putative mannan synthase, GDP-mannose + (mannan)(n) = GDP + (mannan)(n+1). Family 2 of the glycosyltransferases.
Pum0372s0002.1	CSL	GT2	Glycosyl transferase, GT2, putative mannan synthase	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor). Putative mannan synthase, GDP-mannose + (mannan)(n) = GDP + (mannan)(n+1). Family 2 of the glycosyltransferases.
Pum0229s0037.1	iam	GH13	isoamylase, GH13	Hydrolysis of (1->6)-alpha-D-glucosidic branch linkages in starch. Family 13 glycoside hydrolase
Pum0127s0022.1	pulA	GH13	putative pullulanase, GH13	Hydrolysis of (1->6)-alpha-D-glucosidic linkages in pullulan, starch and glycogen. Family 13 glycoside hydrolase
Pum0587s0004.1	AmyAc	GH13	1,4-alpha-glucan-branching enzyme, GH13	Catalyzes the formation of the alpha-1,6-glucosidic linkages in starch by scission of a 1,4-alpha-linked oligosaccharide from growing alpha-1,4-glucan chains and the subsequent attachment of the oligosaccharide to the alpha-1,6 position. Family 13 glycoside hydrolase
Pum0428s0003.1	MGD		monogalactosyldiacylglycerol synthase, GT28	Involved in the synthesis of the major structural component of photosynthetic membranes. Required for proper thylakoid membrane biogenesis. Family 28 glycosyl transferase
Pum1634s0001.1		GT61	glycosyl transferase, GT61	Catalysis of the transfer of a glycosyl group from one compound (donor) to another (acceptor).Family 61 of the glycosyl transferases
Pum1598s0001.1	GWD		alpha glucan water dikinase	Mediates the incorporation of phosphate into starch-like alpha-glucan, mostly at the C-6 position of glucose units. Acts as an overall regulator of starch mobilization. Required for starch degradation, suggesting that the phosphate content of starch regulates its degradability
scaffold_2401:1713-604	CHST	ST	carbohydrate sulfotransferase, scaffold_2401:1713-604	Catalysis of the transfer of a sulfate group from 3'-phosphoadenosine 5'-phosphosulfate to the hydroxyl group of a carbohydrate acceptor, producing the sulfated derivative and 3'-phosphoadenosine 5'-phosphate.
scaffold_317:58759-57639	MGD1	MGDG	monogalactosyldiacylglycerol synthase, GT28	Involved in the synthesis of the major structural component of photosynthetic membranes. Required for proper thylakoid membrane biogenesis. Family 28 glycosyl transferase

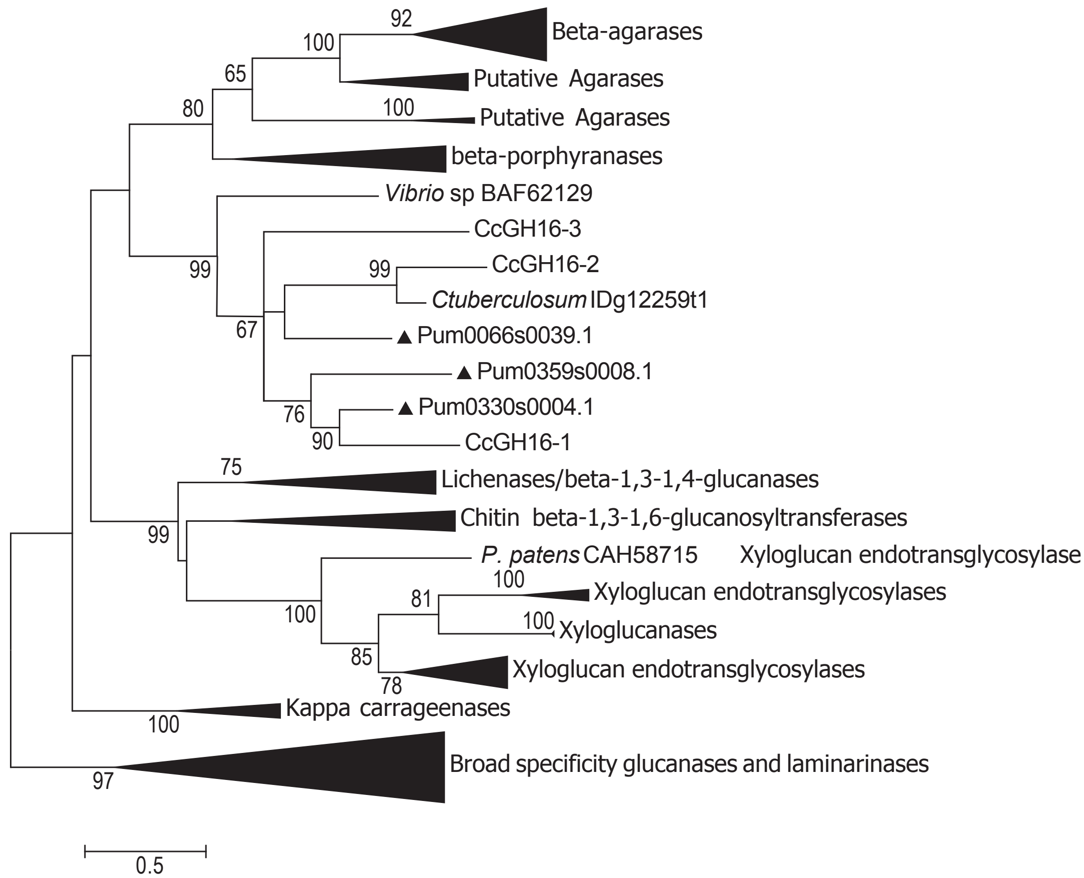


Figure S41. GH16 phylogeny.

Figure showing glycosyl hydrolase (GH16) enzymes in a clade with bacterial agarases and porphyranases. *Porphyra* enzymes are marked with a solid triangle. Full organism names and GenBank accession numbers are found in Table S42. The phylogenetic tree was constructed using the maximum likelihood approach and the program MEGA 6.06 (79). Bootstrap values > 65 are shown.

Table S42. Table of accessions used in GT2 and GT16 phylogenetic trees

List of sequences used for phylogenetic analyses and their corresponding Genbank accession number

Gene Identifier	Genbank accession number	Species
GT2		
SELMODRAFT 75715	EFJ37830	*Selaginella moellendorffii*
SELMODRAFT_86720	EFJ32022	*Selaginella moellendorffii*
SELMODRAFT_119123	EFJ14964	*Selaginella moellendorffii*
SELMODRAFT_73698	EFJ38258	*Selaginella moellendorffii*
SELMODRAFT_141535	EFJ35421	*Selaginella moellendorffii*
SELMODRAFT_156278	EFJ14922	*Selaginella moellendorffii*
SELMODRAFT_163575	EFJ38300	*Selaginella moellendorffii*
OsCesA2	Q84M43	*Oryza sativa Japonica Group*
AtCesA1	O48946	*Arabidopsis thaliana*
OsCesA1	Q6AT26	*Oryza sativa Japonica Group*
PHYPADRAFT_202222	EDQ81939	*Physcomitrella patens subsp. patens*
PHYPADRAFT_228825	EDQ65967	*Physcomitrella patens subsp. patens*
PHYPADRAFT_105213	EDQ77129	*Physcomitrella patens subsp. patens*
PHYPADRAFT_119553	EDQ77472	*Physcomitrella patens subsp. patens*
PHYPADRAFT_213586	EDQ68054	*Physcomitrella patens subsp. patens*
PHYPADRAFT_169568	EDQ58223	*Physcomitrella patens subsp. patens*
AtCesA2	O48947	*Arabidopsis thaliana*
SELMODRAFT_231191	EFJ29751	*Selaginella moellendorffii*
SELMODRAFT_235890	EFJ08899	*Selaginella moellendorffii*
OsCslF1	Q6ZF89	*Oryza sativa Japonica Group*
OsCslF2	Q84S11	*Oryza sativa Japonica Group*
AtCslD1	O49323	*Arabidopsis thaliana*
OsCslD2	Q9LHZ7	*Oryza sativa Japonica Group*
OsCslD1	Q8W3F9	*Oryza sativa Japonica Group*
AtCslD2	Q9LFL0	*Arabidopsis thaliana*
SELMODRAFT_427507	EFJ10084	*Selaginella moellendorffii*
SELMODRAFT_230382	EFJ34908	*Selaginella moellendorffii*
SELMODRAFT_84069	EFJ33203	*Selaginella moellendorffii*
SELMODRAFT_85326	EFJ32798	*Selaginella moellendorffii*
SELMODRAFT_417829	EFJ20635	*Selaginella moellendorffii*
PHYPADRAFT_150074	EDQ53447	*Physcomitrella patens subsp. patens*
PHYPADRAFT_228818	EDQ56470	*Physcomitrella patens subsp. patens*
PHYPADRAFT_228820	EDQ56471	*Physcomitrella patens subsp. patens*
PHYPADRAFT_228816	EDQ66012	*Physcomitrella patens subsp. patens*
PHYPADRAFT_228817	EDQ66047	*Physcomitrella patens subsp. patens*
PHYPADRAFT_228822	EDQ55208	*Physcomitrella patens subsp. patens*
PHYPADRAFT_211198	EDQ70765	*Physcomitrella patens subsp. patens*
AtCslB1	O80898	*Arabidopsis thaliana*
AtCslB2	O80899	*Arabidopsis thaliana*
OsCslH1	Q339N5	*Oryza sativa Japonica Group*
OsCslH2	Q7XUT9	*Oryza sativa Japonica Group*
AtCslG1	Q570S7	*Arabidopsis thaliana*
AtCslG2	Q8VYR4	*Arabidopsis thaliana*
AtCslE1	Q8VZK9	*Arabidopsis thaliana*
OsCslE1	Q651X7	*Arabidopsis thaliana*
OsCslE2	Q0DXZ1	*Arabidopsis thaliana*
Phyinf Ces1	ABP96902	*Phytophthora infestans*
Phyinf Ces2	ABP96903	*Phytophthora infestans*
Phyinf Ces4	ABP96905	*Phytophthora infestans*
CESA	AEZ51039	*Pythium aphanidermatum*
Phyinf Ces3	ABP96904	*Phytophthora infestans*

CESA	AFB77612	*Albugo candida*
Pum0678s0001.1		*Porphyra umbilicalis*
esContig8639		*Porphyra purpurea*
CesA	ABX71734	*Pyropia yezoensis*
Pum0564s0005.1		*Porphyra umbilicalis*
CHC_T00009301001	CDF32287	*Chondrus crispus*
CESA	ABX71734	*Griffithsia monilis*
CHC T00008424001	CDF35770	*Chondrus crispus*
Ava_1568	YP_322086	*Anabaena variabilis ATCC 29413*
Npun F1469	YP_001865112	*Nostoc punctiforme PCC 73102*
MC7420_3817	ZP_05027470	*Microcoleus chthonoplastes PCC 7420*
CESA_Polysphondylium_pallidum_PN500	EFA83461	*Polysphondylium pallidum PN500*
CESA_Acytostelium_subglobosum	BAK54004	*Acytostelium subglobosum*
DDB_0191159	XP_646256	*Dictyostelium discoideum AX4*
contig 34751 g8353		*Pyropia yezoensis*
Pum0053s0020.1		*Porphyra umbilicalis*
R_sphaer_ACM01522	ACM01522	*Rhodobacter sphaeroides KD131*
bcsA	ZP_04185545	*Bacillus cereus AH1271*
TaBcsA	AIF67965	*Terribacillus aidingensis*
LlBcsA	CAA04445	*Lactococcus lactis subsp. lactis*
BsBcsA	WP 046312903	*Bacillus safensis*
AxBcsA	P19449	*Acetobacter xylinum*
KhBcsA	Q59167	*Komagataeibacter hansenii*
PfBcsA	P58931	*Pseudomonas fluorescens*
BCSA ECOLI K12	NP_417990	*Escherichia coli K-12 MG1655*
Sent_Q93IN2	Q93IN2	*Salmonella enterica subsp. enterica serovar Typhimurium str. LT2*
EsCBJ28249	CBJ28249	*Ectocarpus siliculosus*
EsCBJ28245	CBJ28245	*Ectocarpus siliculosus*
EsCBJ30572	CBJ30572	*Ectocarpus siliculosus*
EsCBJ30317	CBJ30317	*Ectocarpus siliculosus*
EsCBJ31065	CBJ31065	*Ectocarpus siliculosus*
EsCBJ31064	CBJ31064	*Ectocarpus siliculosus*
EsCBJ28779	CBJ28779	*Ectocarpus siliculosus*
EsCBN77533	CBN77533	*Ectocarpus siliculosus*
NGA Contig35416	JU981242	*Nannochloropsis gaditana CCMP526*
NGA Contig13416	JU963368	*Nannochloropsis gaditana CCMP527*
contig 8821 g2086		*Pyropia yezoensis*
Pum0231s0012.1		*Porphyra umbilicalis*
contig 3060 g650		*Pyropia yezoensis*
Pum0372s0002.1		*Porphyra umbilicalis*
MICPUN 97997	ACO69823	*Micromonas sp. RCC299*
Ot17g00840	CAL57811	*Ostreococcus tauri*
PHYPADRAFT_183385	EDQ71200	*Physcomitrella patens subsp. patens*
PHYPADRAFT_179490	EDQ75948	*Physcomitrella patens subsp. patens*
PHYPADRAFT_207003	EDQ75893	*Physcomitrella patens subsp. patens*
SELMODRAFT_178995	EFJ17366	*Selaginella moellendorffii*
SELMODRAFT_230176	EFJ35401	*Selaginella moellendorffii*
AtCslA2	Q9FNI7	*Arabidopsis thaliana*
OsCslA1	Q7PC76	*Oryza sativa Japonica Group*
AtCslA1	Q84W54	*Arabidopsis thaliana*
OsCslA2	Q7PC67	*Oryza sativa Japonica Group*
AtCslC5	AtCslC5	*Arabidopsis thaliana*
OsCslC3	Q7PC69	*Oryza sativa Japonica Group*
AtCslC4	Q9LJP4	*Arabidopsis thaliana*
OsCslC1	Q8LIY0	*Oryza sativa Japonica Group*
SELMODRAFT_128946	EFJ09973	*Selaginella moellendorffii*
SELMODRAFT_442658	EFJ23711	*Selaginella moellendorffii*
SELMODRAFT_140198	EFJ38196	*Selaginella moellendorffii*

SELMODRAFT_118058	EFJ15722	*Selaginella moellendorffii*
PHYPADRAFT_67528	EDQ80343	*Physcomitrella patens subsp. patens*
PHYPADRAFT_168024	EDQ61574	*Physcomitrella patens subsp. patens*
PHYPADRAFT_222021	EDQ57558	*Physcomitrella patens subsp. patens*
PHYPADRAFT_131551	EDQ68168	*Physcomitrella patens subsp. patens*
PHYPADRAFT_152670	EDQ51504	*Physcomitrella patens subsp. patens*
PHYPADRAFT_176780	EDQ79502	*Physcomitrella patens subsp. patens*
PHYPADRAFT_191476	EDQ61813	*Physcomitrella patens subsp. patens*
GH16		
Aoff_AF223419	AF223419	*Asparagus officinalis*
Hvulg_BAJ89557	BAJ89557	*Hordeum vulgare subsp. vulgare*
At_AAB18368	AAB18368	*Arabidopsis thaliana*
At_AAD45127	AAD45127	*Arabidopsis thaliana*
At_AAK59660	AAK59660	*Arabidopsis thaliana*
Hvulg_ABY79073	ABY79073	*Hordeum vulgare*
Ghirs_AAO92743	AAO92743	*Gossypium hirsutum*
Ntab_BAA13163	BAA13163	*Nicotiana tabacum*
At_AAC98464	AAC98464	*Arabidopsis thaliana*
1UMZ		*Populus tremula*
Ptrem_AAN87142	AAN87142	*Populus tremula x Populus tremuloides*
At_AAB18366	AAB18366	*Arabidopsis thaliana*
At_AAD21783	AAD21783	*Arabidopsis thaliana*
Dcarota_AAK30204	AAK30204	*Daucus carota*
Hvulg_ACS13756	ACS13756	*Hordeum vulgare*
Tmaj_CAA48325	CAA48325	*Tropaeolum majus*
2UWB		*Tropaeolum majus*
Phypatens_CAH58715	CAH58715	*Physcomitrella patens*
Sc_AAB65002	AAB65002	*Saccharomyces cerevisiae*
Calb_EAL02960	EAL02960	*Candida albicans SC5314*
Ylip_CAG83185	CAG83185	*Yarrowia lipolytica CLIB122*
Sc_AAB67443	AAB67443	*Saccharomyces cerevisiae*
Calb_EAL03096	EAL03096	*Candida albicans SC5314*
Sc_CAA67525	CAA67525	*Saccharomyces cerevisiae*
Calb_EAL01616	EAL01616	*Candida albicans SC5314*
1ZM1		*Fibrobacter succinogenes*
Fsucc_AAA24896	AAA24896	*Fibrobacter succinogenes*
1U0A		*Paenibacillus macerans*
Pmacerans_CAA39426	CAA39426	*Paenibacillus macerans*
1GBG		*Bacillus licheniformis*
Blichen_CAA40547	CAA40547	*Bacillus licheniformis*
3O5S		*Bacillus subtilis subsp. subtilis str. 168*
Bsub_CAA86922	CAA86922	*Bacillus subtilis subsp. subtilis str. 168*
CcGH16-2	XP_005711574	*Chondrus crispus*
Ctuberculosum_IDg12259t1		*Calliarthron tuberculosum IDg12259t1*
Pum0066s0039.1		*Porphyra umbilicalis*
CcGH16-3	XP_005713054	*Chondrus crispus*
Pum0359s0008.1		*Porphyra umbilicalis*
Pum0330s0004.1		*Porphyra umbilicalis*
CcGH16-1	XP_005710570	*Chondrus crispus*
Vibrio_sp_BAF62129	BAF62129	*Vibrio sp. PO-303*
4AWD		*Bacteroides plebeius*
Bplebeius_EDY95423	EDY95423	*Bacteroides plebeius DSM 17135*
Zgal_CBM41185	CBM41185	*Zobellia galactanivorans*
3JUU		*Zobellia galactanivorans*
Zgal_CAZ95074	CAZ95074	*Zobellia galactanivorans*
Zgal_CBM41184	CBM41184	*Zobellia galactanivorans*
Zgal_CBM41183	CBM41183	*Zobellia galactanivorans*
3ILF		*Zobellia galactanivorans*

Zgal_CBM41182	CBM41182	*Zobellia galactanivorans*
Uncultured_bacterium_AAP49324	AAP49324	*uncultured bacterium*
Alteromonas_sp_AGW43026	AGW43026	*Alteromonas sp. GNUM-1*
Agarivorans_sp_ADY17921	ADY17921	*Agarivorans sp. HZ105*
Vibrio_sp_BAF34350	BAF34350	*Vibrio sp. PO-303*
Aalbus_ABW77762	ABW77762	*Agarivorans albus*
Janthinobacterium_sp_ABO93616	ABO93616	*Janthinobacterium sp. SY12*
Pseudoalteromonas_sp_ACF40223	ACF40223	*Pseudoalteromonas sp. JT-6*
Strep_coelicolor_CAB61795	CAB61795	*Streptomyces coelicolor A3(2)*
Fyaey_AEK80424	AEK80424	*Flammeovirga yaeyamensis*
Patlantica_AAA91888	AAA91888	*Pseudoalteromonas atlantica*
Aeromonas_sp_AAF03246	AAF03246	*Aeromonas sp.*
Pseudoalteromonas_sp_ADA72015	ADA72015	*Pseudoalteromonas sp. AG52*
uncultured_bacterium_AAP49346	AAP49346	*uncultured bacterium*
Cagarivorans_AGU13985	AGU13985	*Catenovulum agarivorans YM01*
1O4Y		*Zobellia galactanivorans*
Zgal_CAZ98338	CAZ98338	*Zobellia galactanivorans*
1O4Z		*Zobellia galactanivorans*
Zgal_AAF21821	AAF21821	*Zobellia galactanivorans*
Bplebeius_EDY95404	EDY95404	*Bacteroides plebeius DSM 17135*
4ASM		*Zobellia galactanivorans*
Zgal_CAZ98378	CAZ98378	*Zobellia galactanivorans*
3WZ1		*Microbulbifer thermotolerans*
Mthermo_BAD29947	BAD29947	*Microbulbifer thermotolerans*
Pseudomonas_sp_BAD88713	BAD88713	*Pseudomonas sp. ND137*
Melongatus_BAC99022	BAC99022	*Microbulbifer elongatus*
Sdegradans_ABD80437	ABD80437	*Saccharophagus degradans 2-40*
Zgal_CAZ94309	CAZ94309	*Zobellia galactanivorans*
Zobellia_sp_AGS43006	AGS43006	*Zobellia sp. M-2*
Pseudoalteromonas_sp_ADD92366	ADD92366	*Pseudoalteromonas sp. LL1*
1DYP		*Pseudoalteromonas carrageenovora*
Pcarrageenovora_CAA50624	CAA50624	*Pseudoalteromonas carrageenovora*
2W39		*Phanerochaete chrysosporium*
Pchrysosporium_BAC67687	BAC67687	*Phanerochaete chrysosporium*
Cdestructor_AIX48711	AIX48711	*Cherax destructor*
3DGT		*Streptomyces sioyaensis*
Ssioyaensis_AAF31438	AAF31438	*Streptomyces sioyaensis*
Sequence 4 from patent US 6140096	AAE53061	Sequence 4 from patent US 6140096
4BOW		*Zobellia galactanivorans*
Zgal_CAZ96583	CAZ96583	*Zobellia galactanivorans*
4CTE		*Zobellia galactanivorans*
Zgal_CAZ95067	CAZ95067	*Zobellia galactanivorans*
3ILN		*Rhodothermus marinus*
Rmarinus_AAC69707	AAC69707	*Rhodothermus marinus*
3ATG		*Cellulosimicrobium cellulans*
Ccellulans_AAC38290	AAC38290	*Cellulosimicrobium cellulans*
2HYK		*Nocardiopsis sp. F96*
Nocardiopsis_BAE54302	BAE54302	*Nocardiopsis sp. F96*
2VY0		*Pyrococcus furiosus*
Pfuriosus_AAC25554	AAC25554	*Pyrococcus furiosus DSM 3638*
Tneapolitana_CAA88008	CAA88008	*Thermotoga neapolitana*
4DFS		*Thermotoga petrophila*
Tpetrophila_ABQ46917	ABQ46917	*Thermotoga petrophila RKU-1*
3AZX		*Thermotoga maritima MSB8*
Tmaritima_AAD35118	AAD35118	*Thermotoga maritima MSB8*
Paecilomyces_ADZ46179	ADZ46179	*Paecilomyces sp. FLH30*

Sugars synthesized in the chloroplast must be transported across the two chloroplast membranes. *Porphyra* has 4 nucleotide-sugar transporters, 4 sugar phosphate transporters, 3 triose phosphate/phosphoenolpyruvate transporters, 2 hexose phosphate transporters, and one gene for a glucose-6-phosphate transporter (Table S43). These PO_4^{3-} antiporters transport 3-C and 6-C photosynthetic products out to the cytoplasm, in exchange for inorganic phosphate, and most of these have homologs in *Chondrus* (20). Floridean starch is synthesized and stored in the cytoplasm (109) (Table S44).

Table S43. Carbohydrate transport and storage proteins

***Porphyra* gene identifier**	**Gene symbol**	**Aliases**	**Defline**	**Description**
Pum0709s0008.1	GONST1	VRG4	GDP-mannose transporter	Imports cytoplasmic GDP-mannose into Golgi lumen.
Pum1790s0001.1	GPT1		Glucose-6-phosphate/phosphate translocator 1	Belongs to the TPT transporter family, chloroplastic.
Pum0231s0008.1	PPT2.1		Phosphoenolpyruvate/phosphate translocator 2	Chloroplastic; Also transports dihydroxyacetone phosphate. PMID: 17981875
Pum0159s0013.1	PPT2.2		Phosphoenolpyruvate/phosphate translocator 2	Chloroplastic; Also transports dihydroxyacetone phosphate. PMID: 17981875
Pum0025s0049.1	SLC35C1	NSTP-10; FUCT1	GDP-fucose transporter 1	Cytoplasm GDP-fucose import into Golgi lumen. PMID: 17615350
Pum0450s0013.1	SLC35C2	F12K11.18;	Probable sugar phosphate/phosphate translocator	Belongs to the TPT transporter family. PMID: 14593172
Pum0505s0015.1	SLC35D2.1	SQV-7.1	UDP-sugar transporter	Transports cytosolic nucleotide sugars into Golgi lumen for monophosphates. PMID: 16978188
Pum2965s0001.1	SLC35D2.2	SQV-7.1	UDP-sugar transporter	Transports cytosolic nucleotide sugars into Golgi lumen for monophosphates. PMID: 16978188
Pum0164s0036.1	SLC35E3.1	UDP-GALT1	Probable sugar phosphate/phosphate translocator	Belongs to the TPT transporter family. PMID: 14593172
Pum0242s0002.1	SLC35E3.2		Probable sugar phosphate/phosphate translocator	Belongs to the TPT transporter family. PMID: 14593172
Pum1548s0001.1	SLC35E3.3		Probable sugar phosphate/phosphate translocator	Belongs to the TPT transporter family. PMID: 14593172
Pum0060s0047.1	TPT		Triose phosphate/phosphate translocator	Nucleotide-sugar transporter.
Pum0307s0014.1	UHPC.1	CPB0691; CP_0082	Hexose phosphate transport protein	Transport protein for sugar phosphate uptake. PMID: 10871362
Pum0286s0012.1	UHPC.2	CPB0691; CP_0082	Hexose phosphate transport protein	Transport protein for sugar phosphate uptake. PMID: 10871362

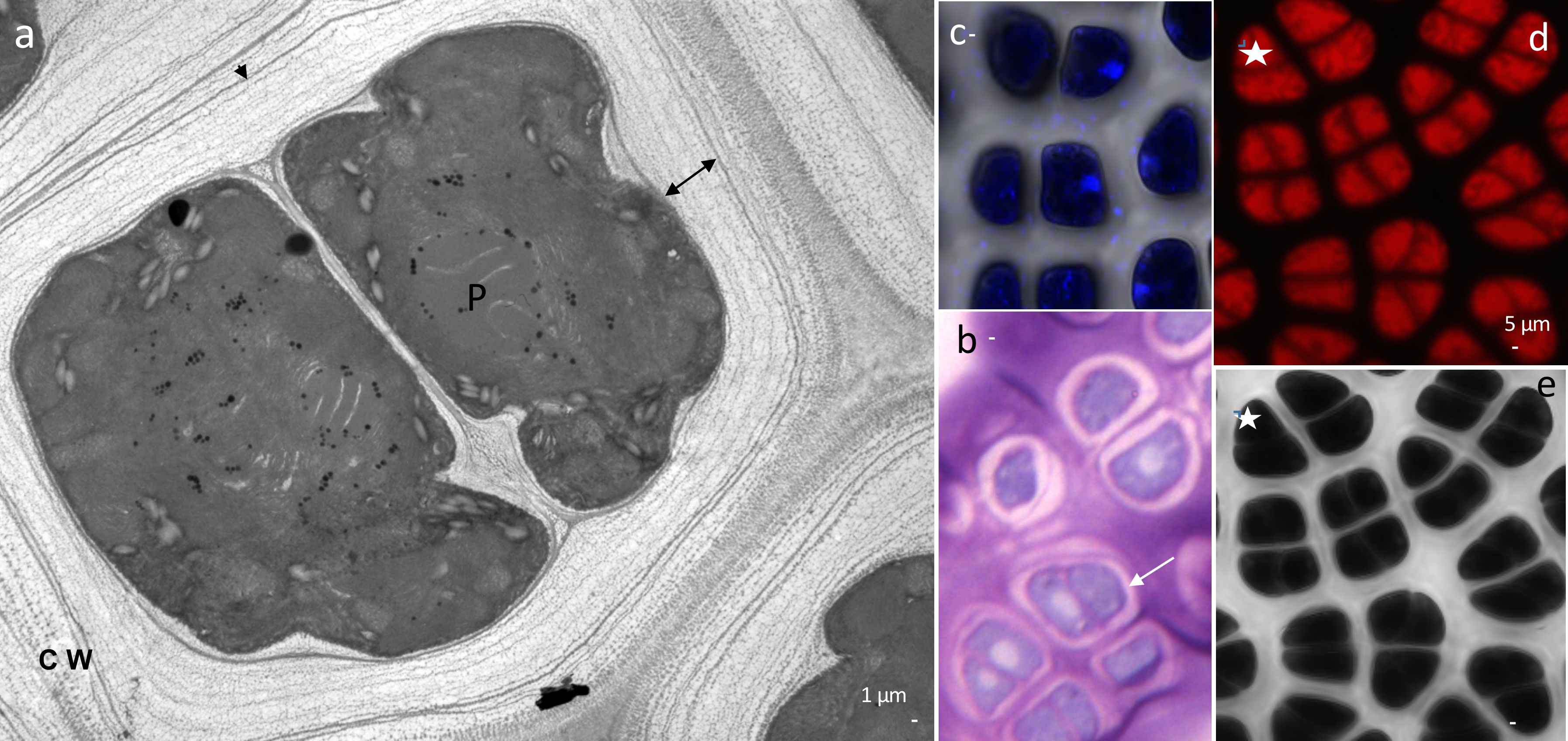
Table S44. Enzymes involved in starch metabolism in selected sequenced red and green species.

Enzyme	GH/GT	*P. umbilicalis*	*C. crispus*	*C. merolae*	*G. sulphuraria*	*A. thaliana*	*O. tauri*
ADP-glucose pyrophosphorylase		0	0	0	0	6	3
Starch synthase (ADPG)	GT5	0	0	0	0	5	5
Starch synthase (UDPG)	GT5	1	1	1	1	0	0
GBSS I	GT5	0	0	0	0	1	1
branching enzymes	GH13	2	1	1	1	3	2
isoamylases	GH13	1	2	2	2	3	3
pullanase	GH13	1	1	1	0	1	1
archael DBE	GH133	0	1	0	0	0	0
phosphorylases	GT35	2	1	1	2	2	3
glucanotransferase (dpe 1)	GH77	0	0	0	0	1	1
trans-glucosidase (dpe 2)	GH77	1	1	2	0	1	1
beta-amylase	GH14	1	1	1	3	3	2
alpha-amylase	GH13	0	0	1	3	3	3
glucan water dikinase		1	1	1	0	2	3
phosphoglucan water dikinase		1	1	0	0	1	2
laforin (sex4)		0	1	0	0	1	1

Microscopic analysis of the blade of *Porphyra umbilicalis*

Figure S45. The vegetative blade with a focus on the cell wall.

A. Ultrastructure of recently divided daughter cells showing the thick cell wall and large central chloroplast with pyrenoid (P). A barbed line spans the inner cell wall (CW); note its different substructure compared to the more compact cell wall between cells. The inner cell wall includes numerous layered periodicities (arrowhead) that encircle and encapsulate the cells. Pieces of blade were fixed in 1% paraformaldehyde + 2% glutaraldehyde in seawater containing 0.2 mM phosphate with 1% osmium tetroxide post-fixation, followed by dehydration, and embedment in Epon. b) Toluidine blue O (pH 4.4) stained-semithin section showing strong metachromasia of sulfated cell wall; the white arrow points to the different character of the innermost cell wall encapsulating cells compared to the cell wall between cells (compare with panel a), c) DAPI-stained nuclei (blue) showing their typical eccentric position (composite confocal images); d-e) companion pair of confocal images (star = same location) showing that the chloroplast fills most of the cell, as demonstrated by the red autofluorescence of chlorophyll. (Micrographs courtesy of Elisabeth Gantt, an original Co-PI). Scale bars: A, 1 μm; B-E, 5 μm.


a
P
c w
1 µm
c
b
d
5 µm
e

Metal metabolism and *Porphyra*

Table S46. Predicted transporter proteins for metals

Porphyra gene identifiers	Gene Symbol	Defline	Description
Pum0445s0027.1	ZIP1	divalent metal transporter	member of subfamily II of the ZIP divalent metal transporter family
Pum0406s0007.1	ZIP2	divalent metal transporter	member of subfamily II of the ZIP divalent metal transporter family
Pum2683s0001.1	ZIP3	divalent metal transporter	member of subfamily II of the ZIP divalent metal transporter family
scaffold_142:17368-16403	ZIP4	divalent metal transporter	member of LZT subfamily of the ZIP divalent metal transporter family
Pum1767s0001.1	ZIP5	divalent metal transporter	member of LZT subfamily of the ZIP divalent metal transporter family
Pum0072s0096.1	PCS1	putative phytochelatin synthase	Phytochelatin synthase is the enzyme responsible for the synthesis of heavy-metal-binding peptides (phytochelatins) from glutathione and related thiols; this protein is orthologous to PCS in C. merolae [PMID: 19060382]
Pum0976s0004.1	VIT1	putative vacuolar iron transporter	Orthologous to the vacuolar iron transporter VIT1 from *Arabidopsis*
Pum0764s0001.1	FTR1	putative iron permease	Orthologous to FTR1 from *Chlamydomonas*, the permease component of a high-affinity iron transporter
Pum0032s0084.1	FOX1	putative multicopper ferroxidase	Orthologous to FOX1 from *Chlamydomonas*, the ferroxidase component of the high-affinity iron transporter
Pum3055s0001.1	FIF1	putative iron-assimilation protein	This protein belongs to a family of algal low-iron-inducible periplasmic proteins (FEA) and membrane-bound, low-iron-inducible proteins (ISIP2) potentially involved in Fe uptake
Pum0052s0005.1	FIF2	putative iron-assimilation protein	This protein belongs to a family of algal low-iron-inducible periplasmic proteins (FEA) and membrane-bound low-iron inducible proteins potentially involved in Fe uptake
Pum0144s0027.1	FIF3	putative iron-assimilation protein	This protein belongs to a family of algal low-iron-inducible periplasmic proteins (FEA) and membrane-bound low-iron inducible proteins potentially involved in Fe uptake
Pum1770s0003.1	HMA1	P1B-type ATPase	Member of the HMA1-like subgroup of divalent metal transporting ATPases. This subgroup seems to be specific to Archaeplastida. Members have been linked to Zn, Cd and Cu homeostasis.
Pum0134s0026.1	MTA1	P1B-type ATPase	Member of Cu-transporting subgroup of P1B-type ATPases; lacks HMA domain
Pum0319s0022.1	MTA2	P1B-type ATPase	Member of Cu-transporting subgroup of P1B-type ATPases
Pum0387s0001.1	MTA3	P1B-type ATPase	Member of Cu-transporting subgroup of P1B-type ATPases; lacks HMA domain
Pum0150s0032.1	MTA4	P1B-type ATPase	Member of Cu-transporting subgroup of P1B-type ATPases
Pum0499s0016.1	MTA5	P1B-type ATPase	Member of Cu-transporting subgroup of P1B-type ATPases; lacks HMA domain
Pum0272s0013.1	MTA6	P1B-type ATPase	Member of Cu-transporting subgroup of P1B-type ATPases
Pum0058s0069.1	CTR1	CTR-type copper ion transporter	Member of the widely conserved family of high-affinity copper transport proteins (Ctr proteins); CTR1 and CTR2 are more similar to homologs from land plants than green algae.
Pum0225s0008.1	CTR2	CTR-type copper ion transporter	Member of the widely conserved family of high-affinity copper transport proteins (Ctr proteins); CTR2 and CTR1 are more similar to homologs from land plants than green algae.
Pum0222s0004.1	MTP1	divalent metal transporter, CDF family	The cation diffusion facilitator (CDF) proteins are typically found to be involved in either Zn, Mn or Fe efflux. This protein is a member of the Zn-transporting subgroup.
Pum0333s0033.1	MTP2	divalent metal transporter, CDF family	The cation diffusion facilitator (CDF) proteins are typically found to be involved in either Zn, Mn or Fe efflux. This protein is a member of the Zn-transporting subgroup.
Pum0152s0012.1	MTP3	divalent metal transporter, CDF family	The cation diffusion facilitator (CDF) proteins are typically found to be involved in either Zn, Mn or Fe efflux. This protein is a member of the Zn-transporting subgroup.
Pum0056s0045.1	MTP4	zinc transporter	The cation diffusion facilitator (CDF) proteins are typically found to be involved in either Zn, Mn or Fe efflux. This protein is a member of the Zn-transporting subgroup.
Pum2079s0001.1	NRM1	Mn2+/Fe2+ transporter, NRAMP family	The natural resistance-associated macrophage protein (NRAMP) family contains divalent cation transporters. Most are involved in Fe^{2+} or Mn^{2+} homeostasis. This gene model encodes a protein similar to NRAMP1 and NRAMP2 from *Chlamydomonas*, which are putatively involved in Mn transport.
Pum0300s0007.1	NRM2	Mn2+/Fe2+ transporter, NRAMP family	The natural resistance-associated macrophage protein (NRAMP) family contains divalent cation transporters. Most are involved in Fe^{2+} or Mn^{2+} homeostasis. This gene model encodes a protein similar to NRAMPs from bacteria with yet to be characterized function.

Copper (Cu) is an essential trace inorganic nutrient that is required by several enzymes in plants and algae including cytochrome *c* oxidase in the mitochondrial electron transport chain and plastocyanin in the photosynthetic electron transport chain. *Porphyra* has a cytochrome *c* oxidase, but unlike plants and green algae, *Porphyra* (and all red algae with sequenced genomes)*,* does not have plastocyanin, the soluble, Cu-dependent, chloroplast-localized, electron transfer protein. Instead, red algae encode the Fe-dependent cytochrome c_6 on the chloroplast genome, which transfers electrons between plastoquinol and cytochrome *f* instead of plastocaynin. Given the presence of plastocyanin in plants, green algae and cyanobacteria, plastocyanin may have existed in the common ancestor to Archaeplastida, but was lost under selective pressure, such as Cu limitation, in the red algal ancestor. However, many extant red algae, including *Porphyra,* have other Cu-dependent proteins (e.g., Cu-Zn superoxide dismutase, multi- copper oxidase, amine oxidase). Additionally, supporting a requirement for Cu in *Porphyra*, we found two genes encoding assimilatory CTR-like Cu transporters and six genes encoding Cu-transporting P_{1B}-type ATPases.

We were surprised to find such a large family of genes encoding Cu-transporting ATPases in the *Porphyra* genome. While land plants commonly have large families of Cu-ATPases, it is not common for sequenced red algal genomes. While there are 6 in *Porphyra* as determined by BLASTP (9) search and subfamily membership based on protein similarity network (MTA1-6 in Table S46, Figure S47), BLASTP (9) search against red algal proteomes and family sub-structure analysis by protein similarity network (Figure S47) suggests the presence of 3 in *Chondus crispus* (Phy0034QQM_CHOCR_GSCHC2T00012475001, Phy0034PX3_CHOCR_GSCHC2T00010478001, Phy0034MDT_CHOCR_GSCHC2T00001736001)*, Pyropia yezoensis* (PyezoenV1_contig_28701_g7060, PyezoenV1_contig_25883_g6380, PyezoenV1_contig_12758_g3046) and *Galdieria sulphuraria* (XP_005707370_1, XP_005703073_1, XP_005702399_1), and 1 in *Porphyridium purpureum* (evm.model.contig_2242.1) and *Cyanidioschyzon merolae* (Phy000UFGW_CYAME). These transporters are likely involved in Cu-detoxification through the endomembrane pathway and/or Cu-loading of secreted Cu-dependent proteins.

Coincident with the loss of plastocyanin, *Porphyra* and other red algal genomes are also missing orthologs of the two Cu transporters that supply Cu to plastocyanin (PAA1 and PAA2 in *A. thaliana*) (110, 111) (Figure S47). HMA1, which is another P_{1B}-type ATPase found in the chloroplast and linked to Cu homeostasis (112), is encoded in the *Porphyra* genome further supporting evolution of this transporter in a common ancestor of Archaeplastida (113). Unlike the chloroplast-localized PAA1/PAA2, which evolved from the cyanobacterial symbiont after endosymbiosis and was lost in red algae, the chloroplast- localized HMA1 was acquired through horizontal gene transfer from a chlamydial ancestor (114) and has been retained in all sequenced genomes from Archaeplastida. Interestingly, although given the selective pressure to retain this transporter in Archaeplastida, it is not found in stramenopiles or haptophytes (Figure S47).

The *Porphyra* genome also encodes members of several broad-specificity divalent metal transporters (Table S46). The assimilatory transporters of the NRAMP and ZIP families may transport Fe^{2+}, Mn^{2+}, and Zn^{2+} at the plasma membrane or across intracellular membranes. We also found members of the CDF family, which could serve as detoxification pumps for excess Zn^{2+} (Table S46).

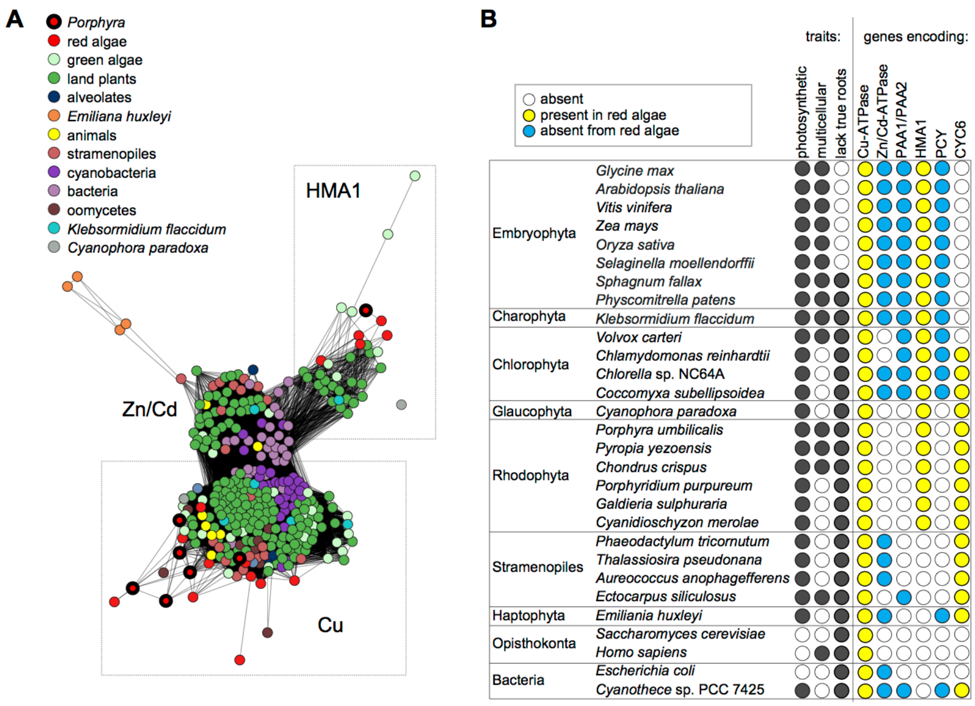


Figure S47. Protein similarity network of P1B-type ATPases and gene occurence profiles.

(A) Protein similarity network of P1B-type ATPases. Each node (circle) represents an individual protein sequence. Edges (lines) connecting nodes represent similarity between two proteins using a local threshold Evalue (from BLASTP (9)) of 1E-75; the *Cyanophora paradoxa* HMA1 does not connect to the network given this threshold, but identity as an HMA1 ortholog was determied by reciprocal best blast to *Arabidopsis* HMA1. The P_{1B}-type ATPase family is divided into three main groups: The Zn/Cd pumps (absent from red algae), Cu pumps (only the endomembrane-type found in red algae) and the Archaeplastida-specific HMA1 group. (B) Identification of P_{1B}-type ATPase subfamilies and presence of genes encoding either cytochrome c6 (CYC6) or plastocyanin (PCY). A filled circle signifies the presence of at least one homolog of the indicated protein family (Cu- ATPase, Cu-transporting subfamily of the P_{1B}-type ATPases; Zn/Cd-ATPase, Zn/Cd- transporting subfamily of the P_{1B}-type ATPases; PAA1/PAA2, chloroplast-localized Cu-specific P_{1B}-type ATPases; HMA1, similar to Zn/Cd-subfamily of P_{1B}-type ATPases but with unknown function). The coloring denotes that protein family is (yellow) or is not (blue) found in red algal genomes.

Table S48. Metal transporter proteins used in protein network analysis.

Sequence identifiers of metal transporters depicted in Figure S47.

Sub-family	Sequence identifier	node color	Organism
Cu	60775	plant	*Selaginella moellendorffii*
Cu	84115	plant	*Selaginella moellendorffii*
Cu	92276	plant	*Selaginella moellendorffii*
Cu	122320	plant	*Selaginella moellendorffii*
Cu	129166	plant	*Selaginella moellendorffii*
Cu	231359	plant	*Selaginella moellendorffii*
Cu	233397	plant	*Selaginella moellendorffii*
Cu	10377_0_0	glaucophyte	*Cyanophora paradoxa*
Cu	Aco002803.1	plant	*Ananas comosus*
Cu	Aco003959.1	plant	*Ananas comosus*
Cu	Aco006600.1	plant	*Ananas comosus*
Cu	Aco011744.1	plant	*Ananas comosus*
Cu	Aco015634.1	plant	*Ananas comosus*
Cu	AT1G63440.1	plant	*Arabidopsis thaliana*
Cu	AT4G33520.2	plant	*Arabidopsis thaliana*
Cu	AT4G33520.3	plant	*Arabidopsis thaliana*
Cu	AT5G21930.1	plant	*Arabidopsis thaliana*
Cu	AT5G21930.2	plant	*Arabidopsis thaliana*
Cu	AT5G21930.3	plant	*Arabidopsis thaliana*
Cu	AT5G44790.1	plant	*Arabidopsis thaliana*
Cu	Cparadoxa_9099_1430-1791_0_plus	glaucophyte	*Cyanophora paradoxa*
Cu	Cre09.g406400.t1.1	green algae	*Chlamydomonas reinhardtii*
Cu	Cre10.g422201.t1.1	green algae	*Chlamydomonas reinhardtii*
Cu	Cre10.g424775.t1.1	green algae	*Chlamydomonas reinhardtii*
Cu	Cre10.g424775.t2.1	green algae	*Chlamydomonas reinhardtii*
Cu	Cre16.g682369.t1.1	green algae	*Chlamydomonas reinhardtii*
Cu	EJK63495.1	stramenopiles	*Thalassiosira oceanica*
Cu	EJK74991.1	stramenopiles	*Thalassiosira oceanica*
Cu	Esi0009_0142	stramenopiles	*Ectocarpus siliculosus*
Cu	Esi0023_0054	stramenopiles	*Ectocarpus siliculosus*
Cu	Esi0125_0069	stramenopiles	*Ectocarpus siliculosus*
Cu	estExt_fgenesh1_pg.C_90003	green algae	*Coccomyxa subellipsoidea C-169*
Cu	estExt_fgenesh1_pm.C_140140	green algae	*Coccomyxa subellipsoidea C-169*
Cu	estExt_Genewise1Plus.C_40563	green algae	*Coccomyxa subellipsoidea C-169*
Cu	fgenesh1_pm.6_#_228	green algae	*Micromonas pusilla CCMP1545*
Cu	fgenesh1_pm.C_Chr_11000053	green algae	*Ostreococcus lucimarinus*
Cu	Genemark1.1168_g	green algae	*Coccomyxa subellipsoidea C-169*
Cu	gi|113475254|ref|YP_721315_1|	cyanobacteria	*Trichodesmium erythraeum IMS101*
Cu	gi|119486994|ref|ZP_01620866_1|	cyanobacteria	*Lyngbya sp_PCC8106*
Cu	gi|119509118|ref|ZP_01628269_1|	cyanobacteria	*Nodularia spumigena CCY9414*
Cu	gi|126657891|ref|ZP_01729044_1|	cyanobacteria	*Cyanothece sp_CCY0110*
Cu	gi|126657983|ref|ZP_01729135_1|	cyanobacteria	*Cyanothece sp_CCY0110*
Cu	gi|126658952|ref|ZP_01730094_1|	cyanobacteria	*Cyanothece sp_CCY0110*
Cu	gi|157412487|ref|YP_001483353_1|	cyanobacteria	*Prochlorococcus marinus str. MIT 9215*
Cu	gi|158334993|ref|YP_001516165_1|	cyanobacteria	*Acaryochloris marina MBIC11017*
Cu	gi|161519848|ref|YP_001583275_1|	bacteria	*Burkholderia multivorans ATCC 17616*
Cu	gi|16329860|ref|NP_440588_1|	cyanobacteria	*Synechocystis sp. PCC 6803*
Cu	gi|16331210|ref|NP_441938_1|	cyanobacteria	*Synechocystis sp. PCC 6803*
Cu	gi|166367885|ref|YP_001660158_1|	cyanobacteria	*Microcystis aeruginosa NIES-843*
Cu	gi|17231274|ref|NP_487822_1|	cyanobacteria	*Nostoc sp. PCC 7120*
Cu	gi|186681031|ref|YP_001864227_1|	cyanobacteria	*Nostoc punctiforme PCC73102*
Cu	gi|209522816|ref|ZP_03271374_1|	cyanobacteria	*Arthrospira maxima CS-328*
Cu	gi|2493016|sp|Q59385_4|COPA_ECOLI	bacteria	*Escherichia coli*
Cu	gi|254415441|ref|ZP_05029202_1|	cyanobacteria	*Coleofasciculus chthonoplastes PCC 7420*
Cu	gi|254423193|ref|ZP_05036911_1|	cyanobacteria	*Synechococcus sp. PCC 7335*
Cu	gi|254423286|ref|ZP_05037004_1|	cyanobacteria	*Synechococcus sp. PCC 7335*
Cu	gi|254423596|ref|ZP_05037314_1|	cyanobacteria	*Synechococcus sp. PCC 7335*
Cu	gi|282898413|ref|ZP_06306404_1|	cyanobacteria	*Raphidiopsis brookii D9*
Cu	gi|282899569|ref|ZP_06307533_1|	cyanobacteria	*Cylindrospermopsis raciborskii CS-505*

Cu	gi|291566996|dbj|BAI89268_1|	cyanobacteria	*Arthrospira platensis NIES-3*
Cu	gi|294924541|ref|XP_002778831_1|	alveolates	*Perkinsus marinus*
Cu	gi|294924545|ref|XP_002778832_1|	alveolates	*Perkinsus marinus*
Cu	gi|294941437|ref|XP_002783102_1|	oomycetes	*Perkinsus marinus*
Cu	gi|325180876|emb|CCA15286_1|	oomycetes	*Albugo laibachii Nc14*
Cu	gi|325181010|emb|CCA15420_1|	oomycetes	*Albugo laibachii Nc14*
Cu	gi|332710730|ref|ZP_08430671_1|	cyanobacteria	*Moorea producens 3L*
Cu	gi|340370150|ref|XP_003383609_1|	animals	*Amphimedon queenslandica*
Cu	gi|354567903|ref|ZP_08987070_1|	cyanobacteria	*Fischerella sp. JSC-11*
Cu	gi|357263630|gb|EHJ12612_1|	cyanobacteria	*Crocosphaera watsonii WH 0003*
Cu	gi|359457420|ref|ZP_09245983_1|	cyanobacteria	*Acaryochloris sp. CCMEE 5410*
Cu	gi|75907770|ref|YP_322066_1|	cyanobacteria	*Anabaena variabilis ATCC 29413*
Cu	Gorai.010G096400.1	plant	*Gossypium raimondii*
Cu	Gorai.010G096400.2	plant	*Gossypium raimondii*
Cu	Gorai.010G096400.3	plant	*Gossypium raimondii*
Cu	GRMZM2G010152_P01	plant	*Zea mays*
Cu	GRMZM2G029951_P01	plant	*Zea mays*
Cu	GRMZM2G143512_P01	plant	*Zea mays*
Cu	GRMZM2G144083_P01	plant	*Zea mays*
Cu	GRMZM2G151406_P01	plant	*Zea mays*
Cu	GRMZM2G315931_P01	plant	*Zea mays*
Cu	GRMZM5G855347_P01	plant	*Zea mays*
Cu	GRMZM5G855347_P02	plant	*Zea mays*
Cu	GSVIVT01012131001	plant	*Vitis vinifera*
Cu	GSVIVT01019769001	plant	*Vitis vinifera*
Cu	GSVIVT01025299001	plant	*Vitis vinifera*
Cu	GSVIVT01035474001	plant	*Vitis vinifera*
Cu	GSVIVT01038740001	plant	*Vitis vinifera*
Cu	gwEuk.5.7.1	green algae	*Ostreococcus lucimarinus*
Cu	gwEuk.7.33.1	green algae	*Ostreococcus lucimarinus*
Cu	jgi|Fracy1|136121|jgi|Fracy1|136121|gw1_21_2_1:2-937	stramenopiles	*Fragilariopsis_cylindrus*
Cu	jgi|Fracy1|196037|jgi|Fracy1|196037|e_gw1_26_2_1:14-901	stramenopiles	*Fragilariopsis_cylindrus*
Cu	jgi|Phatr2|10677|jgi|Phatr2|10677|e_gw1_4_11_1:1-864	stramenopiles	*Phaeodactylum tricornutum*
Cu	jgi|Phatr2|52468|jgi|Phatr2|52468|phatr1_ua_kg_chr_19000019:1-812	stramenopiles	*Phaeodactylum tricornutum*
Cu	jgi|Phyca11|506025|jgi|Phyca11|506025|fgenesh2_kg_PHYCAscaffold_17_#_110_#_Contig3297_1:565-1309	oomycetes	*Phytophthora capsici LT1534*
Cu	jgi|Phyca11|535284|jgi|Phyca11|535284|estExt2_fgenesh1_pg_C_PHYCAscaffold_330117:54-657	oomycetes	*Phytophthora capsici LT1534*
Cu	jgi|Phyca11|535285|jgi|Phyca11|535285|estExt2_fgenesh1_pg_C_PHYCAscaffold_330119:57-946	oomycetes	*Phytophthora capsici LT1534*
Cu	jgi|Phyca11|546403|jgi|Phyca11|546403|estExt2_Genewise1Plus_C_PHYCAscaffold_210077:20-821	oomycetes	*Phytophthora capsici LT1534*
Cu	jgi|Phyca11|548155|jgi|Phyca11|548155|estExt2_Genewise1Plus_C_PHYCAscaffold_270385:1-819	oomycetes	*Phytophthora capsici LT1534*
Cu	jgi|Thaps3|263051|jgi|Thaps3|263051|thaps1_ua_kg_chr_7000008:1-921	stramenopiles	*Thalassiosira pseudonana*
Cu	jgi|Thaps3|264357|jgi|Thaps3|264357|thaps1_ua_kg_chr_16a000004:10-657	stramenopiles	*Thalassiosira pseudonana*
Cu	kfl00006_0120	Kebsomidium flaccidum	*Kebsomidium flaccidum*
Cu	kfl00186_0040	Kebsomidium flaccidum	*Kebsomidium flaccidum*
Cu	kfl00342_0110	Kebsomidium flaccidum	*Kebsomidium flaccidum*
Cu	kfl00630_0010	Kebsomidium	*Kebsomidium flaccidum*

		flaccidum	
Cu	lcl|Pum0134s0026_1	Porphyra	*Porphyra*
Cu	lcl|Pum0150s0032_1	Porphyra	*Porphyra*
Cu	lcl|Pum0272s0013_1	Porphyra	*Porphyra*
Cu	lcl|Pum0319s0022_1	Porphyra	*Porphyra*
Cu	lcl|Pum0387s0001_1	Porphyra	*Porphyra*
Cu	lcl|Pum0499s0016_1	Porphyra	*Porphyra*
Cu	LOC_Os02g07630.1	plant	*Oryza sativa*
Cu	LOC_Os02g10290.1	plant	*Oryza sativa*
Cu	LOC_Os03g08070.1	plant	*Oryza sativa*
Cu	LOC_Os03g08070.2	plant	*Oryza sativa*
Cu	LOC_Os03g08070.3	plant	*Oryza sativa*
Cu	LOC_Os04g46940.1	plant	*Oryza sativa*
Cu	LOC_Os06g45500.1	plant	*Oryza sativa*
Cu	LOC_Os08g37950.1	plant	*Oryza sativa*
Cu	Medtr2g035840.1	plant	*Medicago truncatula*
Cu	Medtr3g105190.1	plant	*Medicago truncatula*
Cu	Medtr3g105190.2	plant	*Medicago truncatula*
Cu	Medtr4g094232.1	plant	*Medicago truncatula*
Cu	Medtr4g094695.1	plant	*Medicago truncatula*
Cu	Medtr5g010420.1	plant	*Medicago truncatula*
Cu	Medtr8g012390.1	plant	*Medicago truncatula*
Cu	Medtr8g012390.2	plant	*Medicago truncatula*
Cu	Medtr8g079250.1	plant	*Medicago truncatula*
Cu	Medtr8g079250.2	plant	*Medicago truncatula*
Cu	Medtr8g079250.3	plant	*Medicago truncatula*
Cu	Medtr8g079250.4	plant	*Medicago truncatula*
Cu	MicpuC2.est_orfs.1_395_4276783:18	green algae	*Micromonas pusilla CCMP1545*
Cu	MicpuC2.EuGene.0000030510	green algae	*Micromonas pusilla CCMP1545*
Cu	MicpuC2.EuGene.0000100405	green algae	*Micromonas pusilla CCMP1545*
Cu	Phy0034PUF_CHOCR_GSCHC2T00010269001	red algae	*Chondrus crispus*
Cu	Phy0034PX3_CHOCR_GSCHC2T00010478001	red algae	*Chondrus crispus*
Cu	Phy0034QQM_CHOCR_GSCHC2T00012475001	red algae	*Chondrus crispus*
Cu	Potri.001G019100.1	plant	*Populus trichocarpa*
Cu	Potri.001G019100.2	plant	*Populus trichocarpa*
Cu	Potri.001G105800.1	plant	*Populus trichocarpa*
Cu	Potri.001G105800.2	plant	*Populus trichocarpa*
Cu	Potri.001G158900.1	plant	*Populus trichocarpa*
Cu	Potri.001G158900.2	plant	*Populus trichocarpa*
Cu	Potri.001G205400.1	plant	*Populus trichocarpa*
Cu	Potri.003G024000.1	plant	*Populus trichocarpa*
Cu	Potri.003G075700.1	plant	*Populus trichocarpa*
Cu	Potri.003G125600.1	plant	*Populus trichocarpa*
Cu	Potri.003G125600.2	plant	*Populus trichocarpa*
Cu	Potri.003G125700.1	plant	*Populus trichocarpa*
Cu	Potri.003G125700.2	plant	*Populus trichocarpa*
Cu	Potri.003G204800.1	plant	*Populus trichocarpa*
Cu	Potri.006G220200.1	plant	*Populus trichocarpa*
Cu	Potri.018G047800.1	plant	*Populus trichocarpa*
Cu	Potri.018G047800.2	plant	*Populus trichocarpa*
Cu	Pp3c1_2740V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2740V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2740V3.3.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2770V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2770V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2880V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2880V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2880V3.3.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2920V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2920V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2960V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2960V3.2.p	plant	*Physcomitrella patens*

Cu	Pp3c1_2960V3.3.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2960V3.4.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2980V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c1_2980V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c15_4170V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c15_4170V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c16_16360V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c16_16360V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c18_10710V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c18_10710V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c18_10710V3.3.p	plant	*Physcomitrella patens*
Cu	Pp3c18_10710V3.4.p	plant	*Physcomitrella patens*
Cu	Pp3c20_13520V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c20_13520V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c20_13520V3.3.p	plant	*Physcomitrella patens*
Cu	Pp3c22_10860V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c22_10860V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3c8_13750V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3c8_13750V3.2.p	plant	*Physcomitrella patens*
Cu	Pp3s34_20V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3s37_40V3.1.p	plant	*Physcomitrella patens*
Cu	Pp3s37_40V3.2.p	plant	*Physcomitrella patens*
Cu	PyezoenV1_contig_12758_g3046	red algae	*Pyropia yezoensis*
Cu	PyezoenV1_contig_25883_g6380	red algae	*Pyropia yezoensis*
Cu	PyezoenV1_contig_28701_g7060	red algae	*Pyropia yezoensis*
Cu	Solyc01g105160.2.1	plant	*Solanum lycopersicum*
Cu	Solyc02g068490.2.1	plant	*Solanum lycopersicum*
Cu	Solyc08g061610.2.1	plant	*Solanum lycopersicum*
Cu	Solyc08g080870.2.1	plant	*Solanum lycopersicum*
Cu	Solyc08g080890.2.1	plant	*Solanum lycopersicum*
Cu	Solyc11g062100.1.1	plant	*Solanum lycopersicum*
Cu	sp|A7X6S1|COPA_STAA1	bacteria	*Staphylococcus aureus*
Cu	sp|O32220|COPA_BACSU	bacteria	*Bacillus subtilis*
Cu	sp|P35670|ATP7B_HUMAN	animals	*Homo sapiens*
Cu	sp|P38360|ATU1_YEAST	fungi	*Saccharomyces cerevisiae*
Cu	sp|P38995|ATU2_YEAST	fungi	*Saccharomyces cerevisiae*
Cu	sp|P58341|ATCU1_RHIME	bacteria	*Rhizobium meliloti*
Cu	sp|P58342|ATCU2_RHIME	bacteria	*Rhizobium meliloti*
Cu	sp|Q04656|ATP7A_HUMAN	animals	*Homo sapiens*
Cu	sp|Q64430|ATP7A_MOUSE	animals	*Mus musculus*
Cu	sp|Q64446|ATP7B_MOUSE	animals	*Mus musculus*
Cu	sp|Q9X5X3|ATCU_SINMW	bacteria	*Sinorhizobium medicae*
Cu	Vocar.0005s0334.1.p	green algae	*Volvox carteri*
Cu	Vocar.0009s0109.1.p	green algae	*Volvox carteri*
Cu	Vocar.0023s0021.1.p	green algae	*Volvox carteri*
Cu	Vocar.0062s0026.1.p	green algae	*Volvox carteri*
Cu	XP_005702399_1	red algae	*Galderia sulphuraria*
Cu	XP_005703073_1	red algae	*Galderia sulphuraria*
Cu	XP_005707370_1	red algae	*Galderia sulphuraria*
HMA1	62740	green algae	*Micromonas sp. RCC299*
HMA1	AT4G37270.1	plant	*Arabidopsis thaliana*
HMA1	Cre17.g720400.t1.1	green algae	*Chlamydomonas reinhardtii*
HMA1	e_gwEuk.14.49.1	green algae	*Chlamydomonas reinhardtii*
HMA1	estExt_Genemark1.C_120086	green algae	*Coccomyxa subellipsoidea C-169*
HMA1	eugene.1400010141	green algae	*Ostreococcus lucimarinus*
HMA1	Gorai.007G218200.1	plant	*Gossypium raimondii*
HMA1	Gorai.007G218200.2	plant	*Gossypium raimondii*
HMA1	GRMZM2G067853_P01	plant	*Zea mays*
HMA1	GRMZM2G067853_P02	plant	*Zea mays*
HMA1	kfl00442_0010p	Kebsomidium flaccidum	*Kebsomidium flaccidum*
HMA1	LOC_Os06g47550.1	plant	*Oryza sativa*
HMA1	LOC_Os06g47550.2	plant	*Oryza sativa*
HMA1	LOC_Os06g47550.4	plant	*Oryza sativa*
HMA1	Medtr4g094985.1	plant	*Medicago truncatula*
HMA1	MicpuC2.estExt_fgenesh1_pg.C_80	green algae	*Micromonas pusilla CCMP1545*

	068		
HMA1	Phy0034R1M_CHOCR_GSCHC2T 00013295001	red algae	*Chondrus crispus*
HMA1	Potri.007G049000.1	plant	*Populus trichocarpa*
HMA1	Potri.007G049000.2	plant	*Populus trichocarpa*
HMA1	Potri.007G049000.3	plant	*Populus trichocarpa*
HMA1	Potri.007G049000.4	plant	*Populus trichocarpa*
HMA1	Potri.007G049000.5	plant	*Populus trichocarpa*
HMA1	Potri.007G049000.6	plant	*Populus trichocarpa*
HMA1	Pp3440_14	red algae	*Porphyridium purpureum*
HMA1	Pp3c3_31850V3.1.p	plant	*Physcomitrella patens*
HMA1	Pp3c3_31850V3.2.p	plant	*Physcomitrella patens*
HMA1	Pp3c4_3060V3.1.p	plant	*Physcomitrella patens*
HMA1	Pp3c4_3060V3.2.p	plant	*Physcomitrella patens*
HMA1	Pum1770s0003_1	Porphyra	*Porphyra*
HMA1	PyezoenV1_contig_15565_g3723	red algae	*Pyropia yezoensis*
HMA1	Solyc02g092920.2.1	plant	*Solanum lycopersicum*
HMA1	Vocar.0021s0128.1.p	green algae	*Volvox carteri*
HMA1	XP_005705513_1	red algae	*Galderia sulphuraria*
Zn/Cd	60690	plant	*Selaginella moellendorffii*
Zn/Cd	89397	plant	*Selaginella moellendorffii*
Zn/Cd	118425	plant	*Selaginella moellendorffii*
Zn/Cd	149298	plant	*Selaginella moellendorffii*
Zn/Cd	AC205008.4_FGP002	plant	*Zea mays*
Zn/Cd	Aco011454.1	plant	*Ananas comosus*
Zn/Cd	AT2G19110.1	plant	*Arabidopsis thaliana*
Zn/Cd	AT4G30110.1	plant	*Arabidopsis thaliana*
Zn/Cd	AT4G30120.1	plant	*Arabidopsis thaliana*
Zn/Cd	EJK66294.1	stramenopiles	*Thalassiosira oceanica*
Zn/Cd	gi|156390845|ref|XP_001635480_1|	animals	*Nematostella vectensis*
Zn/Cd	gi|16331905|ref|NP_442633_1|	cyanobacteria	*Synechocystis sp. PCC 6803*
Zn/Cd	gi|16331908|ref|NP_442636_1|	cyanobacteria	*Synechocystis sp. PCC 6803*
Zn/Cd	gi|16752171|ref|NP_445538_1|	bacteria	*Chlamydophila pneumoniae AR39*
Zn/Cd	gi|212639169|ref|YP_002315689_1|	bacteria	*Anoxybacillus flavithermus*
Zn/Cd	gi|221124502|ref|XP_002166500_1|	animals	*Hydra magnipapillata*
Zn/Cd	gi|23097731|ref|NP_691197_1|	bacteria	*Oceanobacillusi heyensis HTE831*
Zn/Cd	gi|24981380|gb|AAN65675_1|AE01 6194_5	bacteria	*Pseudomonas putida*
Zn/Cd	gi|254421578|ref|ZP_05035296_1|	cyanobacteria	*Synechococcus sp. PCC 7335*
Zn/Cd	gi|254425877|ref|ZP_05039594_1|	cyanobacteria	*Synechococcus sp. PCC 7335*
Zn/Cd	gi|257089794|ref|ZP_05584155_1|	bacteria	*Enterococcus faecalis CH188*
Zn/Cd	gi|261418985|ref|YP_003252667_1|	bacteria	*Geobacillus sp_Y412MC61*
Zn/Cd	gi|269838297|ref|YP_003320525_1|	bacteria	*Sphaerobacter thermophilus DSM 20745*
Zn/Cd	gi|27314522|gb|AAO03659_1|AE01 6744_62	bacteria	*Staphylococcus epidermidis*
Zn/Cd	gi|294887453|ref|XP_002772117_1|	alveolates	*Perkinsus marinus*
Zn/Cd	gi|296131980|ref|YP_003639227_1|	bacteria	*Thermincola potens JR*
Zn/Cd	gi|308070777|ref|YP_003872382_1|	bacteria	*Paenibacillus polymyxa E681*
Zn/Cd	gi|312905523|ref|ZP_07764637_1|	bacteria	*Enterococcus faecalis TX0635 = WH245*
Zn/Cd	gi|323450224|gb|EGB06106_1|	stramenopiles	*Aureococcus anophagefferens*
Zn/Cd	gi|335038349|ref|ZP_08531614_1|	bacteria	*Caldalkalibacillus thermarum TA2.A1*
Zn/Cd	gi|357005254|ref|ZP_09070253_1|	bacteria	*Paenibacillus larvae subsp. larvae B-3650*
Zn/Cd	gi|357199463|gb|AET57360_1|	bacteria	*Paenibacillus terrae HPL-003*
Zn/Cd	gi|54036914|sp|P63686_1|CTPD_M YCBO	bacteria	*Mycobacterium bovis AF2122/97*
Zn/Cd	gi|81323283|sp|Q8RNN5|Q8RNN5_ LEGPN	bacteria	*Legionella pneumophila*
Zn/Cd	gi|81852879|sp|Q8YSC8|Q8YSC8_ ANASP	bacteria	*Nostoc sp. PCC 7120*
Zn/Cd	gi|81857901|sp|Q9KEV5|Q9KEV5_ BACHD	bacteria	*Bacillus halodurans*
Zn/Cd	Gorai.004G098400.1	plant	*Gossypium raimondii*
Zn/Cd	Gorai.009G003000.1	plant	*Gossypium raimondii*
Zn/Cd	Gorai.009G003000.3	plant	*Gossypium raimondii*
Zn/Cd	GRMZM2G099191_P01	plant	*Zea mays*
Zn/Cd	GRMZM2G099191_P02	plant	*Zea mays*
Zn/Cd	GRMZM2G175576_P01	plant	*Zea mays*

Zn/Cd	GRMZM2G175576_P02	plant	*Zea mays*
Zn/Cd	GRMZM2G455491_P01	plant	*Zea mays*
Zn/Cd	GSVIVT01010896001	plant	*Vitis vinifera*
Zn/Cd	gw1.5.79.1	green algae	*Coccomyxa subellipsoidea C-169*
Zn/Cd	gwEuk.10.60.1	green algae	*Ostreococcus lucimarinus*
Zn/Cd	jgi|Emihu1|229251|jgi|Emihu1|229251|gm1_15400047:240-654	Emiliana huxleyi	*Emiliana huxleyi*
Zn/Cd	jgi|Emihu1|252377|jgi|Emihu1|252377|gm1_132400001:9-477	Emiliana huxleyi	*Emiliana huxleyi*
Zn/Cd	jgi|Emihu1|452140|jgi|Emihu1|452140|estExtDG_Genemark1_C_570041:207-710	Emiliana huxleyi	*Emiliana huxleyi*
Zn/Cd	jgi|Emihu1|452141|estExtDG_Genemark1_C_570046	Emiliana huxleyi	*Emiliana huxleyi*
Zn/Cd	jgi|Fracy1|168079|jgi|Fracy1|168079|estExt_Genewise1_C_40176:119-583	stramenopiles	*Fragilariopsis_cylindrus*
Zn/Cd	jgi|Fracy1|188713|jgi|Fracy1|188713|e_gw1_9_39_1:105-780	stramenopiles	*Fragilariopsis_cylindrus*
Zn/Cd	jgi|Fracy1|206868|jgi|Fracy1|206868|estExt_Genewise1Plus_C_40033:229-706	stramenopiles	*Fragilariopsis_cylindrus*
Zn/Cd	jgi|Phatr2_bd|387|jgi|Phatr2_bd|387|e_gw1_17x31_1_1:27-705	stramenopiles	*Phaeodactylum tricornutum*
Zn/Cd	jgi|Phatr2_bd|391|jgi|Phatr2_bd|391|e_gw1_17x31_2_1:27-705	stramenopiles	*Phaeodactylum tricornutum*
Zn/Cd	jgi|Phatr2|52367|jgi|Phatr2|52367|phatr1_ua_kg_chr_16000037:12-711	stramenopiles	*Phaeodactylum tricornutum*
Zn/Cd	jgi|Phatr2|676|jgi|Phatr2|676|gw1_16_21_1:5-666	stramenopiles	*Phaeodactylum tricornutum*
Zn/Cd	jgi|Thaps3|261657|jgi|Thaps3|261657|thaps1_ua_kg_chr_3000035:1-666	stramenopiles	*Thalassiosira pseudonana*
Zn/Cd	kfl00136_0120	Kebsomidium flaccidum	*Kebsomidium flaccidum*
Zn/Cd	LOC_Os07g12900.1	plant	*Oryza sativa*
Zn/Cd	Medtr2g036380.1	plant	*Medicago truncatula*
Zn/Cd	Medtr2g036380.2	plant	*Medicago truncatula*
Zn/Cd	Medtr2g036380.3	plant	*Medicago truncatula*
Zn/Cd	Medtr2g036380.4	plant	*Medicago truncatula*
Zn/Cd	Medtr2g036380.5	plant	*Medicago truncatula*
Zn/Cd	Medtr4g127580.1	plant	*Medicago truncatula*
Zn/Cd	MicpuC2.EuGene.0000010855	green algae	*Micromonas pusilla CCMP1545*
Zn/Cd	Potri.006G076900.1	plant	*Populus trichocarpa*
Zn/Cd	Pp3c2_8660V3.1.p	plant	*Physcomitrella patens*
Zn/Cd	Pp3c2_8660V3.2.p	plant	*Physcomitrella patens*
Zn/Cd	Pp3c2_8660V3.3.p	plant	*Physcomitrella patens*
Zn/Cd	Pp3c2_8660V3.4.p	plant	*Physcomitrella patens*
Zn/Cd	Pp3c6_24750V3.1.p	plant	*Physcomitrella patens*
Zn/Cd	Pp3c6_24750V3.2.p	plant	*Physcomitrella patens*
Zn/Cd	Solyc07g009130.2.1	plant	*Solanum lycopersicum*
Zn/Cd	sp|O32219|CADA_BACSU	bacteria	*Bacillus subtilis*
Zn/Cd	sp|Q59465|HMCT_HELPY	bacteria	*Helicobacter pylori*
Zn/Cd	sp|Q59998|ATZN_SYNY3	bacteria	*Synechocystis sp. (strain PCC 6803 / Kazusa)*
Zn/Cd	ZntA	bacteria	*Escherichia coli str. K-12 substr. MG1655*

Macronutrients

Nitrogen

Table S49. Ammonium transporters

Our analysis of ammonium transporters from *Porphyra* and other sequenced red algae shows more AMTs in *Porphyra* (7) compared to *Pyropia* (probably 5 AMTs), *Chondrus* (1), *Porphyridium* (3), *Galdieria* (2) and *Cyanidioschyzon* (2).

Taxon	Gene ID	Gene symbol
Porphyra umbilicalis	Pum0165s0019.1	AMT1.1
	Pum0463s0020.1	AMT1.2
	Pum1775s0001.1	AMT1.3
	Pum0126s0003.1	AMT1.4
	Pum1656s0001.1	AMT2.1
	Pum0027s0002.1	AMT2.2
	Pum0022s0083.1	AMT3
Pyropia yezoensis	PyezoenV1.contig_16335_g3953; identical to BAV55984.1	AMT1.4
	PyezoenV1.contig_15829_g3790	AMT2.1
	PyezoenV1._15136_g3631	AMT2.2
	PyezoenV1.contig_36752_g8709	AMT3
	Gene fragments related to AMT: PyezoenV1.contig_28057_g6901 PyezoenV1.contig_16336_g3954 PyezoenV1_16334_g3952	AMT
Porphyridium	Porphyridium.evm.model.contig_2333.1	AMT1
	Porphyridium.evm.model.contig_3836.1	AMT1
	Porphyridium.evm.model.contig_2277.5	AMT1
Chondrus crispus	Phy0034M7Z_CHOCR_GSCHC2T00001276001	AMT1
Galdieria sulphuraria	XP_005708658.1	AMT
	XP_005706620.1	AMT
Cyanidioschyzon merolae	Phy000UGI8_CYAME	AMT
	Phy000UEGA_CYAME	AMT

Table S50. Nitrogen metabolism genes

Porphyra gene identifiers	Gene Symbol	Aliases or Synonyms	Defline	Description
Pum0165s0019.1	*AMT1.1*		ammonia transporter	ammonium transmembrane transporter channel; PMID 8062823; plasma membrane (WoLFP);
Pum0463s0020.1	*AMT1.2*		ammonia transporter	ammonium transmembrane transporter channel; PMID 8062823; chloroplast or mitochondrial(WoLFP & Target P)
Pum1775s0001.1	*AMT1.3*		ammonia transporter	ammonium transmembrane transporter channel; PMID 8062823; signal peptide (Target P), cytosol or vacuole membrane (WoLFP)
Pum0126s0003.1	*AMT1.4*		ammonia transporter	ammonium transmembrane transporter channel; PMID 8062823; not mitochondrial or chloroplast; no N terminal signal protein
Pum1656s0001.1	*AMT2.1*		ammonia transporter	ammonium transmembrane transporter channel; PMID 8062823; possibly chloroplast (pTarget & WoLFRPSort)
Pum0027s0002.1	*AMT2.2*		ammonia transporter	ammonium transmembrane transporter channel; PMID 8062823; integral membrane protein (TargtP, WoLFP)
Pum0022s0083.1	*AMT3*		ammonia transporter	ammonium transmembrane transporter channel; PMID 8062823l; not chloroplast or mitochondrial, no Nterminal signal protein (WoLFP) but has some putative transmembrane domains.
Pum0612s0003.1	*DUR3.1*	UT-2	urea active transporter protein	urea active transporter protein DOI 10.1007/s12562-015-0947-7; possible plasma membrane (WoLFP)
Pum0148s0030.1	*DUR3.2*	UT-3	urea active transporter protein	urea active transporter protein DOI 10.1007/s12562-015-0947-7; integral membrane protein (WoLFP, TargetP) possibly vacuolar
scaffold_169:36146-40685	*DUR1,2*		urea amidolyase-like protein; UAL	Urea carboxylase and allophanate amidolyase; pubmed 18550080; 23503846
Pum1246s0003.1	*GLN1*	GLNA-like; GS1	glutamine synthetase;	Glutamine synthetase; cytosolic isoform;PMID:10482686
Pum0499s0012.1	*GLN2*	GLN1, GS1	glutamine synthetase	Glutamine synthetase; mitochondrial isoform, PMID:10482687
Pum0058s0079.1	*SIR*		Assimilatory sulfite reductase (ferrodoxin)	Hydrogen sulfide + 6 oxidized ferredoxin [iron-sulfur] cluster + 3 H_2O = sulfite + 6 reduced ferredoxin [iron-sulfur] cluster + 6 H^+; PMID:8695637
Pum2215s0001.1	*NRT2*	ACH1, LIN1,	high affinity nitrate	PMID: 10066586; plasma membrane (WoLFP)

		NRT2;1AT	transporter; NRT2;	
Pum0014s0024.1	*NIA*		nitrate reductase[NADH]; NRT	cytosolic,(EC:1.7.1.1); PMID:3393528
scaffold_1826:2943-1543	*GPR1*	GSA	glutamate -5-semialdehyde dehydrogenase	
Pum0280s0026.1	*NIR1*		Ferrodoxin nitrite reductase	NH_3 + 2 H_2O + 6 oxidized ferredoxin = nitrite + 6 reduced ferredoxin + 7 H^+.;chloroplastic; E1.7.7.1; PMID:7894060
Pum0415s0006.1	*GDH*		glutamate dehyrdogenase1-like	mitochondrial; PUBMED9734815
Pum0244s0012.1	*OTC*		Ornithine carbamoyltransferase	Octase; Ornithine carbamoyltransferase; plastid
Pum0126s0008.1	*ASSYR*		Argininosuccinate synthase-related	chloroplastic; Citrulline--aspartate ligase; E.C. 6.3.4.5; PMID:10617198
Pum0060s0016.1	*ASL1*		argininosuccinate lyase	E.C. 4.3.2.1
Pum0171s0026.1	*EMB1027*		arginine-tRNA ligase	E.C. 6.1.1.19; PMID:16251277; chloroplastic (WoLFP, not targetP);Arginyl-tRNA synthetase
Pum0262s0019.1	*delta-OAT*		ornithine transaminase	E.C. 2.6.1.13; not chloroplast or mitochondrial
Pum0111s0007.1	*delta-OAT2*		acetylornithine aminotransferase	acetylornithine aminotransferase
Pum0223s0017.1	*ODC*		ornithine decarboxylase	E.C. 4.1.1.17; OMID: 9733552l putrescine biosynthetic process from ornithine L-ornithine --> putrescine + CO_2
Pum2451s0001.1	*SOX*	MCP	sulfite oxidase	Sulfite + O_2 + H2O = sulfate + H_2O_2; PMID 11598126
Pum0148s0020.1	*CBRR1*		NADH-cytochrome 5 reductase-related	NADH + 2 ferricytochrome b5 = NAD^+ + H^+ + 2 ferrocytochrome b5.; mitochondrial; PMID 17227547
Pum0537s0012.1	*FNR2*	PETH2	Ferre doxin NADP reductase	chloroplast?, PMID 10948265
Pum0288s0034.1	*NADE*		glutamine-dependent NAD(+) synthetase	$NAD(^+)$ synthase [glutamine-hydrolyzing] or $NAD(^+)$ synthetase; UNIPROT
Pum0234s0010.1	*ECM4-omega-R*	GTO,1, 2 or 3	Glutathione S-transferase omega-related	glutathione reductase (omega);PMID:16709151; all three yeast homologs are active as dimethylarsinic acid reductases but have additional functions
Pum0254s0035.1	*CICDH*		isocitrate dehydrogenase	(EC:1.1.1.42); NADPH regeneration, citric acid cycle; Isocitrate + NADP+ = 2-oxoglutarate + CO_2 + NADPH.
Pum0075s0018.1	NBLA		Phycobilisome degradation protein NblA like	Belongs to the ycf18/nblA family

In addition to its AMTs (Table S49), *Porphyra* has a single high affinity nitrate transporter protein (*NRT*2*)* and a single nitrate reductase (*NIA*, Table S50; Figure S51), as is typical of other red algae (23). *Porphyra* also encodes a single nitrite reductase (*NIR*1) with homology (95%) to an unnamed protein in *Pyropia. Porphy*ra, like *Chondrus* (20) and *Pyropia*, has two urea-active transporters (*DUR 3.1,* Pum0612s0003.1; *DUR3.2* Pum0148s0030.1) and a bifunctional urea amidolyase, also encoded by the DUR1,2 gene from which ammonium is produced from urea:

Urea + ATP + HCO_3^- --urea carboxylase→ allophonate + ADP + Pi

Allophanate + 3 H_2O + H^+ ---allophanate lyase→2 NH_4^+ + 2 HCO_3^-

Porphyra has both components of the glutamine synthetase [EC 6.3.1.2] and glutamate synthase (GOGAT; EC 1.4.1.13) pathway to assimilate NH_4^+ (110). However, like *Chondrus crispus* (23)*, Porphyra,* and *Pyropia yezoensis* lack a nuclear-encoded glutamate synthase (GOGAT). However, *Chondrus* (YP_007627474.1)*, Porphyra* (AFC40043.1) and *P. yezoensis* (YP_537056) have plastid encoded glutamate synthase *gltB* (or GOGAT). *Porphyra* has two nuclear-encoded *GLN* genes for glutamine synthetases; the GLN1protein, Pum0499s0012.1, appears to be targeted to the mitochondria while the GLN2 protein, Pum1246s0003.1, appears to be localized to the cytosol (Figure S51). Therefore, the GS/GOGGAT ammonia assimilation pathway in these red algae is coordinated over two or three cellular compartments (mitochondria, cytoplasm and chloroplast) and different organelle and nuclear genomes. Amino acid transporters are needed to shuttle amino acids (glutamate, glutamine) and dicarboxylic acids between these cellular compartments. The amino acid transporters/carriers have not yet been annotated for *Porphyra*. *Pyropia, Chondrus, Cyanidioschyzon*, *Galderia,* and *Porphyridium* all have nuclear encoded *GLN* gene(s). Glutamate dehydrogenase, encoded by a single *GDH* gene (Pum0415s0006.1), is a mitochondrial-localized protein; GDH may salvage NH_4^+ released in amino acid recycling pathways.

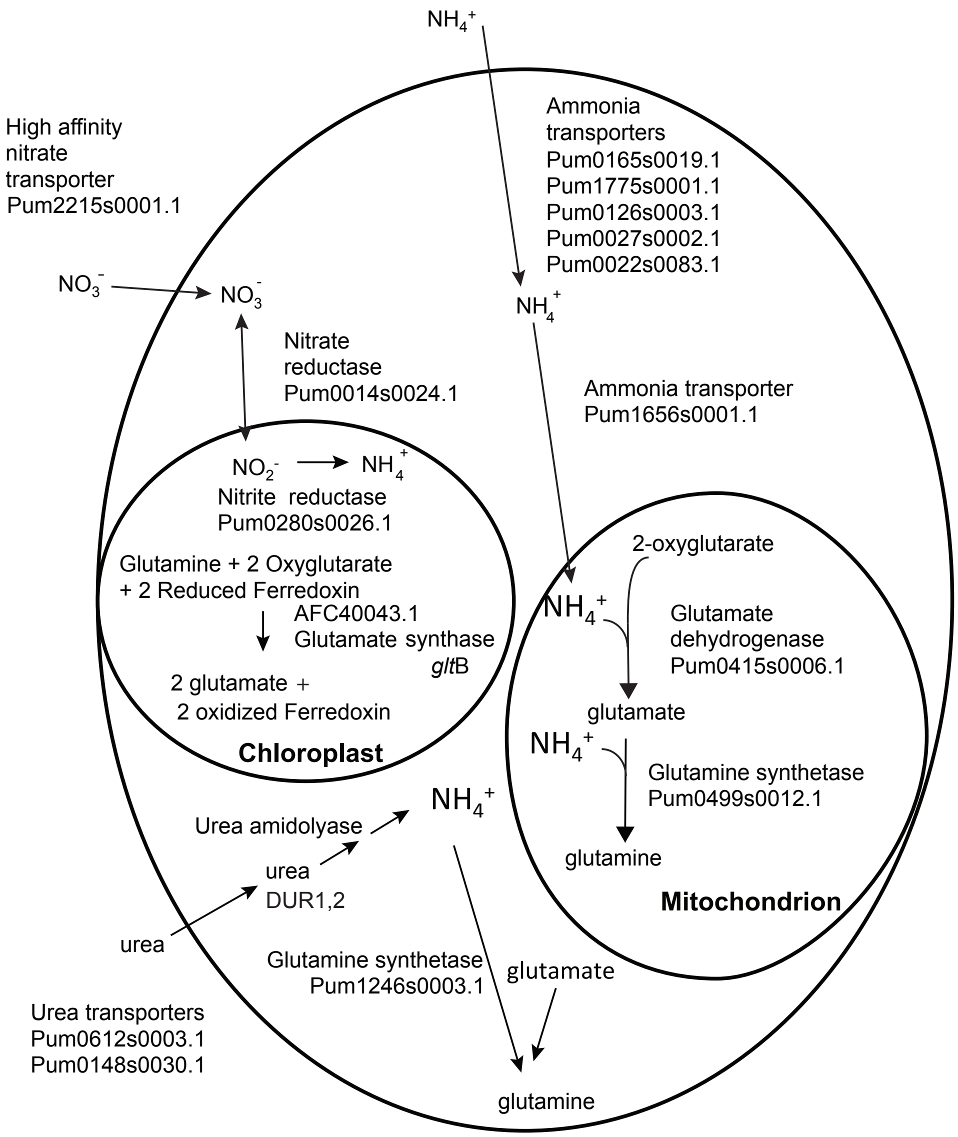


Figure S51. Summary of nitrogen uptake and conversions in *Porphyra umbilicalis*

Porphyra and *Pyropia*'s significant uptake capacity for dissolved inorganic nitrogen (DIN) and dissolved organic nitrogen (DON) support the bioremediation capability of these crops in aquaculture. Wu et al. (111) estimated that 3688.15 metric tons of tissue N and 105.6 mt of tissue P were removed in 2012-2013 by *P. yezoensis* in China's largest cultivation area in the southern Yellow Sea.

The phycobilisomes of red algae breakdown under nitrogen starvation, but when nutrient stress is relieved, they reassemble (112). Thus, the phycobilisome has long been considered to serve a dual role in photosynthesis and macronutrient storage to help the alga survive acute nutrient stress in the field. Similarly to *Pyropia* (24), we found one nuclear-encoded *NblA* gene in *Porphyra* (Table S50); this gene supports breakdown of phycobilisomes under nutrient stress. This finding provides additional data for phycobilisome breakdown under nutrient limitation in *Porphyra*, which becomes green on the shore in summer when stratification of the seawater results in starvation of DIN and other nutrients.

Phosphorus.

Table S52. List of genes involved in acquisition of phosphorus in *Porphyra* (and ATP synthase units)

Porphyra gene identifiers	Gene Symbol	Aliases or synonyms	Defline	Description
Phosphate metabolism				
Pum0154s0031.1	ACP2	LAP	Lysosomal acid phosphatase	Encodes the beta subunit of lysosomal acid phosphatase. PMID: 24275569
Pum0076s0064.1	ACP2.2		Acid Phosphatase	Putative lysomal acid phosphatase.
Pum0217s0007.1	ACPP	PPAP	Similar to Prostatic acid phosphatase	A non-specific tyrosine phosphatase.
Pum0400s0006.1	ALKPPC .1	PHOA	Alkaline phosphatase	Hydrolyzes phosphate esters at alkaline pH. PMID: 24865297
Pum0204s0023.1	ALKPPC .2	PHOA	Alkaline phosphatase	Hydrolyzes phosphate esters at alkaline pH. PMID: 24865297
Pum0397s0012.1	LPP3		Phosphatidic acid phosphatase	Putative lipid phosphate phosphatase 3. PMID: 15875012; 11130713
Pum0191s0021.1	MET14	APSK	Adenylyl-sulfate kinase	Phosphorylates adenosine 5'-phosphsulfate.
Pum0381s0009.1	PPK		Polyphosphate kinase	Transfers ATP terminal phosphate to a long-chain polyphosphate. PMID: 17940044
Phosphate transport				
Pum0507s0005.1	PHO4		High-affinity phosphate transporter	High-affinity (phosphate-repressible) phosphate transporter. PMID: 2531109
Pum0153s0028.1	PHT4;4		Phosphate transporter 4;4	Chloroplastic; Probable anion transporter. PMID: 19513231
Pum0690s0007.1	PIT		Putative phosphate permease	Inorganic phosphate transporter. PMID: 7542800.
Pum0113s0020.1	SLC20A 1.2	PIT-1	Sodium-dependent phosphate transporter	Sodium-phosphate symporter. PMID: 9878855
ATP Synthesis				
Pum0070s0075.1	ATPA	H28O16.1	ATP synthase subunit alpha, mitochondrial	Forms catalytic core in F(1). PMID: 9851916
Pum0738s0007.1	ATPB	ATP2-1	ATP synthase subunit beta, mitochondrial	Forms catalytic core in F(1). PMID: 9851916
Pum0328s0038.1	ATPD		ATP synthase subunit delta, mitochondrial	Part of the catalytic core CF(1). PMID: 14671022
Pum0223s0027.1	ATP5E	F19C24.25	ATP synthase subunit epsilon, mitochandria	Part of the catalytic core CF(1). PMID: 14671022
Pum2300s0002.1	ATPG.1	ATPC	ATP synthase subunit gamma	Part of the catalytic core CF(1). PMID: 14671022
Pum0709s0005.1	ATPG.2	ATPC	ATP synthase subunit gamma, chloroplastic	Part of the catalytic core CF(1). PMID: 14671022
Pum0166s0011.1	ATPG.3	ATPC	ATP synthase subunit gamma, mitochondrial	Part of the catalytic core CF(1). PMID: 14671022
Pum0021s0048.1	VAP.1		V-type ATP synthase proteolipid subunit	Membrane integral V0 complex proton-conducting pore.
Pum2119s0001.1	VAP.2		V-type ATP synthase proteolipid subunit	Membrane integral V0 complex proton-conducting pore.
Pum0154s0030.1	ATP6V1 A		V-type proton ATPase catalytic subunit A	V-type proton ATPase catalytic subunit A. PMID: 7619836
Pum0025s0064.1	VHA-A.1	F3K23.17	V-type proton ATPase subunit A	Required for assembly and activity of the V-ATPase. PMID: 22223895
Pum1670s0001.1	VHA-A.2	F3K23.17	V-type proton ATPase subunit A	Required for assembly and activity of the V-ATPase. PMID: 22223895
Pum0049s0015.1	VHA-A.3	F19H22.18 0	V-type proton ATPase subunit A	Required for assembly and activity of the V-ATPase. PMID: 22223895
Pum1686s0002.1	VHA-B		V-type proton ATPase subunit B	Non-catalytic subunit of the peripheral V1 complex. PMID: 7619836
Pum0290s0011.1	VATC	ATP6C; ATP6C1	V-type proton ATPase subunit C 1	Necessary catalytic sector. PMID: 21183079
Pum2407s0002.1	VHA-D	VHA36; VHA36-1; VATD; VATPD	V-type proton ATPase subunit D 1	Subunit of the peripheral V1 complex. PMID: 19376835; 12537569
Pum0148s0017.1	VHA-E2	VMA9	V-type proton ATPase subunit E 2	Subunit of the integral membrane V0 complex. PMID: 11950611
Pum1027s0001.1	VATG	VMA10	V-type proton ATPase subunit G	Catalytic subunit of the peripheral V1 complex.
Pum1393s0003.1	VHA-H		V-type proton ATPase subunit H	Activates ATPase activity and couples it with proton flow. PMID: 22223895

Pum1541s0002.1	ACLA-3		ATP-citrate synthase alpha chain protein 3	Synthesis of cytosolic acetyl-CoA. PMID: 12869764
Pum0335s0012.1	ACLB-1		ATP-citrate synthase beta chain protein 1	Synthesis of cytosolic acetyl-CoA. PMID: 12869764
Pum0178s0009.1	VATF	VHA14	V-type proton ATPase subunit F	Membrane integral V0 complex proton-conducting pore.
Pum0111s0026.1	VHA-C		V-type proton ATPase subunit c''	Membrane integral V0 complex proton-conducting pore.
Pum2221s0001.1	VATD-1		V-type proton ATPase subunit d	Subunit of the integral membrane V0 complex of vacuolar ATPase.

Phosphate transporters (Tables S43, S52) are similar to those described in *Chondrus crispus* (20) and *Galdieria sulphuria* (23). The *Porphyra* genome encodes a polyphosphate kinase (Pum0381s0009.1) that synthesizes high-energy polymers of phosphate. Polyphosphatase, the enzyme responsible for catalyzing the release from polyphosphate granules was not found in the *Porphyra* genome assembly; however, because polyphosphate kinase is a reversible enzyme, the absence of the polyphosphatase gene suggests that *Porphyra* utilizes one enzyme to catalyze both reactions. The two alkaline phosphatases in *Porphyra* suggest that the alga can utilize the dissolved organic phosphorus (DOP) in the nearshore waters; alkaline phosphate is a phosphomonoesterase that utilizes phosphorous from many ester-linked substrates (113).

ATP synthases are responsible for ATP synthesis, each consisting of F_0 and F_1 sectors. F_0 is composed of subunits a and c while F_1 is composed of subunits alpha, beta, gamma, delta, and epsilon. In *P. umbilicalis* we were able to account for the mitochondrial F_1 subunits. (Table S52). We also analyzed vacuolar ATPase; V-ATPase is essential for several biological processes, because it is responsible for controlling the pH, through the transport of protons, for a variety of intracellular compartments in eukaryotic cells. V-ATPase is a heteromultimeric enzyme composed of a peripheral catalytic V1 complex (components A to H) attached to an integral membrane V0 proton pore complex (components: a, c, c', c'', d and e). All but the component V1 E were identified in the *P. umbilicalis* genome (Table S52).

Sulfur.

Porphyra has the enzymes necessary to assimilate sulfate including 2 Na^+/sulfate symporters and one H^+/sulfate transporter (Table S53). Sulfatases including arylsufatases are absent from red algal genomes including *Porphyra*, but the alga has other types of enzymes that allow it to remodel its cell wall (see Carbohydrate-active enzymes [CAZymes]).

Table S53. Table of proteins involved in sulfur metabolism in *Porphyra*

Porphyra gene identifiers	Gene Symbol	Aliases or Synonyms	Defline	Description
Pum0914s0001.1	ACK		acetate kinase	
scaffold_306:99067-100209	PAT		phosphate acetyltransferase	
Pum0647s0006.1	SUMF1		Sulfatase modifying enzyme	modifies the sulfatase - activation. There do not appear to be any arylsulfatases in the red algae (*Porphyra, Chondrus, Pyropea, Porphyridium*)
Pum2231s0001.1	SLT1		Sodium sulfate transporter	
Pum0877s0003.1	SLTL1		Sodium sulfate transporter-type	Putative sodium sulfate transport or for other anion (e.g. chromate etc)
Pum0191s0021.1	APK1	MET14, PAPSS	Adenosine 5'-phosphosulfate kinase	
Pum0220s0005.1	APR1		Adenosine 5'-phosphosulfate reductase	
Pum0835s0008.1	ATS1	APS, MET3	ATP sulfurylase	
Pum0058s0079.1	SIR1	MET5, cycJ	Sulfite reductase	
Pum0622s0006.1	GSS2		glutathione synthase	
Pum1549s0001.1	GSS1	GSH, GCLM, GCLC	gamma-glutamylcysteine synthetase	glutamate cysteine ligase, GCL, catalytic subunit
Pum0121s0018.1	SAM	METK, MAT	S-adenosylmethionine synthase	
Pum0096s0025.1	SQD1		UDP-sulfoquinovose synthase	
Pum0220s0013.1	UXS1		UDP-glucuronic acid decarboxylase	
Pum0565s0001.1	UXS2		UDP-glucuronic acid decarboxylase	
Pum0574s0012.1	UGE1	GAL10, GALE, RHD	UDP-glucose 4-epimerase -	single gene in reds
Pum2784s0001.1	SQD2		sulfolipid synthase -	single gene, best blast hits are reds and plants
Pum0428s0003.1	MGD1		monogalactosyldiacylglycerol synthase (MGDGS)	
Pum0923s0003.1	DGD1		digalactosyldiacylglycerol synthase	
Pum0835s0008.1	SAT	ATPS	Sulfate adenylyltransferase	This enzyme catalyzes the reaction of ATP and sulfate to produce diphosphate and adenylyl sulfate. This enzyme participates in 3 metabolic pathways: purine metabolism, selenoamino acid metabolism, and sulfur metabolism.
Pum0946s0005.1			Protein Disulfide Isomerase A family	Protein Disulfide Isomerase (PDIa) family, redox active TRX domains; composed of eukaryotic proteins involved in oxidative protein folding in the endoplasmic reticulum (ER) by acting as catalysts and folding assistants.
Pum0167s0012.1			Putative protein disulfide-isomerase	Fuchs S, De Lorenzo F, Anfinsen CB. Studies on the mechanism of the enzymic catalysis of disulfide interchange in proteins. J. Biol. Chem. 242 (1967) 398-402. [PMID:6022836]
Pum0427s0017.1	txn	Trx	Thioredoxin	Participates in various redox reactions through the reversible oxidation of its active center dithiol to a disulfide and catalyzes dithiol-disulfide exchange reactions (By similarity). Plays a role in the reversible S-nitrosylation of cysteine residues in target proteins, and thereby contributes to the response to intracellular nitric oxide. Nitrosylates the active site Cys of CASP3 in response to nitric oxide (NO), and thereby inhibits caspase-3 activity. Induces the FOS/JUN AP-1 DNA binding activity in ionizing radiation (IR) cells through its oxidation/reduction status and stimulates AP-1 transcriptional activity (By similarity)
Pum0371s0016.1	PDIA6		Probable protein disulfide-isomerase A6	Catalyzes the rearrangement of -S-S- bonds in proteins.

Arsenic uptake and metabolism

Inorganic arsenic is taken up by all marine organisms (reviewed in (114)). Detoxification genes in *Porphyra* include transporters and methyl transferases (Table S54).

Table S54. Arsenic related genes

Porphyra gene identifier	Gene Symbol	Aliases or Synonyms	Defline	Description
Pum0021s0045.1	ARSA2		Arsenite-translocating ATPase	ArsA ATPase functionas as an efflux pump located on the inner membrane of the cell that catalyzes the extrusion of arsenite, antimonite and arsenate. PMID: 10970874
Pum0218s0039.1	ARSA1		Arsenical pump-driving ATPase; Arsenite-stimulated ATPase	Arsenical pump-driving ATPase; Arsenite-stimulated ATPase; Golgi to ER traffic protein 3; Guided entry of tail-anchored proteins 3
Pum0461s0015.1	ARSC1		Arsenate reductase	Arsenate reductase; Arsenical pump modifier
Pum0176s0045.1	AS3MT	CYT19	Arsenite methyltransferase	Catalyzes the transfer of a methyl group from AdoMet to trivalent arsenicals producing methylated and dimethylated arsenicals.

Lipids

When plants and algae are grown at lower temperatures, the amount of fatty acid unsaturation within a membrane system usually increases, which enhances temperature and stress tolerance (115).) The unsaturated, ω3-long-chain polyunsaturated fatty acid (LCPUFA) eicosapentaenoic acid (EPA, 20:5) is found in *P. umbilicalis* (116), with most of the fatty acid being palmitic acid (16:0).

Most lipid-related genes in *Porphyra* are present as single copy genes (Table S55), in agreement with our analysis of *Porphyra* ESTs (1) and other red algal genomes (19, 20, 22). In photosynthetically active organisms, fatty acid synthesis takes place in the plastid and involves two multimeric enzyme complexes, acetyl-CoA carboxylase and fatty acid synthase. As in other red algae, but in contrast to members of the green lineage, the subunits of the acetyl-CoA carboxylase - with exception of the biotin carboxylase - are encoded in the plastidial genome, not in the nucleus. No candidate gene resembling the plant type thioesterase (FATA/B in *Arabidopsis*) was found. However, we identified a candidate with strong sequence similarity to the thioesterase from *Phaeodactylum tricornutum* (117), in agreement with data from *Porphyridium purpureum* (22), enabling export of C_{16} and C_{18} fatty acids to be released from the plastid for further assembly into lipids in the endoplasmic reticulum (ER).
Inositol phospholipids (phosphoinositides)

Inositol phospholipids (phosphoinositides)

A phosphatidylinositol-4- kinase (P14K) candidate (Pum1010s0001.1) was found, which would lead to synthesis of PI(4)P (Table S55). We found candidates for synthesis of PI(4,5)P_2 from either PI(4)P or PI(5)P by a PIPkinase type I (Pum0438s0007.1) (Table S55). Septins (Table S15) bind preferentially to PI(4,5)P_2 and can assemble into filaments as well as playing a role in membrane domains (118). We did not, however, find evidence for synthesis of PI(3,5)P2 in

Porphyra; this phosphoinositide binds formins as part of phragmoplast expansion and proper establishment of the orientation of the cell plate in mosses and angiosperms (119, 120). We also discovered a PI-bisphophatase candidate (Pum0152s0049.1), and a candidate for phospholipase C (Pum1738s0001.1) that synthesizes the signaling molecules DAG and IP_3 from $PI(4,5)P_2$ (Table S55)(118).

Table S55. List of red algal genes related to lipid-synthesis including phosphoinositides

Porphyra gene identifiers	Gene Symbol	Aliases or Synonyms	Defline	Description
Lipids and fatty acids				
Pum0153s0015.1	OLE1		fatty acid delta-9 desaturase	stearoyl delta-9 desaturase (membrane-bound enzyme, not FAD5), notFAB2
Pum0051s0008.1	FAD2		fatty acid delta-12/omega-6 desaturase	delta-12/omega-6 fatty acid desaturase, involved in linoleic acid biosynthesis from oleic acid
Pum0721s0007.1	FAD3		fatty acid delta-15/omega-3 desaturase	delta-15/omega-3 fatty acid desaturase, catalyzes the formation of eicosapentaenoic acid (20:5) from arachidonic acid (20:4), also involved in alpha-linolenic acid biosynthesis
Pum0199s0006.1	FAD6		fatty acid delta-6 desaturase	delta-6 desaturase, involved in synthesis of gamma-linolenic acid from linoleic acid
Pum0197s0017.1	ELO1		fatty acid delta-6 elongase	delta-6 elongase, catalyzes the elongation of gamma-linolenic acid
Pum1976s0001.1	d5DES		fatty acid delta-5 desaturase	delta5 desaturase, involved biosynthesis of arachidonic acid
Pum0205s0013.1	FAD4		PG delta-3 desaturase	introduces a delta-3 double bond into a C16 fatty acid at the sn2-position of phosphatidylglycerol
Pum0103s0015.1	accC.1	BCR1	biotin carboxylase	acetyl-CoA carboxylase, biotin carboxylase subunit
Pum0014s0045.1	accC.2	BCR1	biotin carboxylase	acetyl-CoA carboxylase, biotin carboxylase subunit
Pum0290s0029.1	accC.3	BCR1	biotin carboxylase	acetyl-CoA carboxylase, biotin carboxylase subunit
Pum0215s0018.1	accC.4	BCR1	biotin carboxylase	acetyl-CoA carboxylase, biotin carboxylase subunit
Pum0152s0008.1	ACP1	acpP	acyl carrier protein 1	acyl carrier protein 1
Pum1510s0001.1	KASI-II.1	FAB1	3-Ketoacyl-ACP synthase	3-Ketoacyl-[acyl carrier protein] synthase (KASI_KASII superfamily protein)
Pum0076s0014.1	KAR1		3-Ketoacyl-ACP reductase	3-Ketoacyl-[acyl carrier protein] reductase
Pum0144s0012.1	HAD		3-Hydroxyacyl-ACP dehydratase	3-Hydroxyacyl-[acyl carrier protein] dehydratase
Pum1413s0003.1	ENR	fabI	Enoate reductase	Enoyl-[acyl carrier protein] reductase
Pum0111s0016.1	ATS1	GPAT1	Glycerol-3-phosphate:acyl-ACP acyltransferase	Glycerol-3-phosphate:acyl-[acyl carrier protein] acyltransferase
Pum0322s0008.1	ATS2	LPAAT	Lyso-phosphatidate:acyl-ACP acyltransferase	Lyso-phosphatidate:acyl-[acyl carrier protein] acyltransferase
Pum0517s0009.1	CDS1		CDP-DAG synthase	cytidinediphosphate diacylglycerol synthase
Pum0098s0007.1	CDS2		CDP-DAG synthase	cytidinediphosphate diacylglycerol synthase
Pum0096s0025.1	SQD1		UDP-sulfoquinovose synthase	UDP-sulfoquinovose synthase
Pum2784s0001.1	SQD2	SLS	UDP-sulfoquinovose: alpha-diacylglycerol-sulfoquinovosyltransferase	UDP-sulfoquinovose: alpha-diacylglycerol-sulfoquinovosyltransferase
Pum0428s0003.1	MGD1	mgdA	MGD synthase 1	Monogalactosyldiacylglycerol synthase 1
Pum0081s0018.1	MGD2	mgdA	MGD synthase 2	Monogalactosyldiacylglycerol synthase 2
Pum0923s0003.1	DGD1		DGD synthase 1	digalactosyldiacylglycerol synthase 1
Pum0121s0018.1	METM	SAS1	S-Adenosylmethionine synthetase	S-Adenosylmethionine synthetase
Pum0082s0003.1	ECT1		PtdEtn synthase	cytidinediphosphate ethanolamine synthase
Pum1463s0001.1	INO1		inositol-3-phosphate synthase	inositol-3-phosphate synthase
Pum0175s0009.1	PGPS		PtdGro synthase	phosphatidylglycerol synthase
Pum0810s0006.1	LACS1	ACSL1	long chain acyl-CoA synthetase	long chain acyl-CoA synthetase 1 (alias acyl-CoA synthetase long-chain family member 1)
Pum0099s0037.1	LACS2		long chain acyl-CoA synthetase	long chain acyl-CoA synthetase 3 (alias acyl-CoA synthetase long-chain family member 3)
Pum0254s0010.1	TGD3		trigalactosyldiacylglycerol 3	possibly involved in ER to chloroplast lipid transfer, ATPase domain of a multimeric complex
Pum0233s0024.1	GPAT2		glycerol-3-phosphate acyltransferase	glycerol-3-phosphate acyltransferase
Pum0233s0023.1	GPAT3		glycerol-3-phosphate acyltransferase	glycerol-3-phosphate acyltransferase
Pum1773s0002.1	LPCAT1		lysophosphatidylcholine acyltransferase	lysophosphatidylcholine acyltransferase
Pum0116s0034.1	LPCAT2		lysophosphatidylcholine acyltransferase	lysophosphatidylcholine acyltransferase
Pum0668s0001.1	ACPThE		thioesterase	acyl-[acyl carrier protein] thioesterase candidate, not FATA/FAT1
Pum0357s0001.1	MBOAT1		membrane bound O-acyl transferase-like protein	membrane bound O-acyl transferase (MBOAT) family protein isoform 1
Pum0510s0005.1	MBOAT2		membrane bound O-acyl transferase-like protein	membrane bound O-acyl transferase (MBOAT) family protein isoform 2
Pum0082s0057.1	LCB1	SPT	serine palmitoyltransferase subunit 1	serine palmitoyltransferase, long chain base subunit 1, multimeric enzyme with LCB2, catalyzes first step in sphingolipid biosynthesis
Pum0161s0033.1	LCB2	SPT	serine palmitoyltransferase	serine palmitoyltransferase, long chain base subunit 1, multimeric

			subunit 2	enzyme with LCB2, catalyzes first step in sphingolipid biosynthesis; note that the catalytical center is located in LCB2
Pum0065s0003.1	TSC10		3-dehydrosphinganine reductase	3-dehydrosphinganine reductase, ketosphinganine reductase, second step in sphingolipid biosynthesis
Pum0847s0006.1	D4DES1	DES1; DEGS	sphingolipid delta-4 desaturase	sphingolipid desaturase, introduces a delta-4-trans double bond into sphinganine
Pum0207s0017.1	D4DES2	DES1; DEGS	sphingolipid delta-4 desaturase	sphingolipid desaturase, introduces a delta-4-trans double bond into sphinganine
Pum0106s0025.1	LAG1	LAC1; LASS1	ceramide synthase	similar to S. cerevisiae Longevity-assurance proteins LAG1 and LAC1 involved in ceramide biosynthesis
Pum0218s0026.1	FAH1	SCS7	fatty acid alpha-hydroxylase	ceramide hydroxylase, introduces an alpha-hydroxyl group on similar to SCS7 (S. cerevisiae) and FAH1 (A. thaliana)
Pum0827s0007.1	LCBK	LCB4	long chain base kinase	long chain base (LCB) kinase, similar to S. cerevisiae LCB4
Signaling/ Cell Membrane				
Pum0033s0033.1	PLD.1		Phospholipase D	Hydrolyzes glycerol-phospholipids at terminal phosphodiesteric bond. PMID: 7551587; PMID: 12183179
Pum0501s0015.1	PLD.2		Phospholipase D	Hydrolyzes glycerol-phospholipids at terminal phosphodiesteric bond. PMID: 7551587; PMID: 12183179
Pum1896s0001.1	PLD.3		Phospholipase D	Hydrolyzes glycerol-phospholipids at terminal phosphodiesteric bond. PMID: 7551587; PMID: 12183179
PI kinases				
Pum1010s0001.1	PI4K-α		Phosphatidylinositol-4-kinase-α	involved in PI(4)P-synthesis from PI
Pum0438s0007.1	PIP5K		Phosphatidylinositol-4-phosphate 5-kinase family protein	enzyme involved in synthesis of PI(4,5)P2 from PI(4)P
Pum0152s0049.1	PI5PII		Phosphoinositide 5-Phosphatase Type II	phosphadidylinositol-bisphosphatase candidate
Pum1738s0001.1	PIPLC		PREDICTED: 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase delta-1	enzyme involved in synthesis of IP3 and DAG from PI(4,5)P2

Red algae lack a plastid desaturation pathway with the exception of the phosphatidylglycerol-specific Δ3-desaturase FAD4, which differs from other membrane-bound desaturases by its unusual protein domain structure (121). This suggests that palmitic acid (16:0) and stearic acid (18:0) are exported from the plastid, rather than oleic acid (18:1) as occurs in plants and other algae. In the ER, the fatty acids are most likely incorporated into phosphatidylcholine and further elongated and desaturated to form the LCPUFAs arachidonic acid (20:4) and eicosapentaenoic acid (20:5, EPA). Besides 16:0, these VLCPUFA are the predominant fatty acids in all membrane lipids including the plastid galactolipids monogalactosyldiacylglycerol and digalactosyldiacylglycerol (22, 122-124). This suggests an important role for ER to chloroplast lipid trafficking in *Porphyra* (and other red algae) since the C_{20} fatty acids found in plastid lipids have to be ER-derived. In A*rabidopsi*s four proteins play a role in chloroplast lipid import: TGD1, 2, 3 and 4 (125). In *Arabidopsis* TGD1, 2 and 3 resemble a bacterial type ABC transporter, located at the inner chloroplast envelope. TGD4 is part of an outer envelope complex. No candidate was found for TGD4 in *Porphyra*, but as described previously (1)), orthologues for TGD1, 2 and 3 were identified. TGD1 and 2 are encoded in the chloroplast genome; only TGD3 (Pum1407s0003.1) which contains the ATPase domain seems to be nuclear-encoded (Table S55).

Currently no experimental data are available regarding sphingolipid metabolism in *Porphyra* (or other red algae). However, we could identify all genes necessary for the synthesis of sphing-4-enine (sphingosine), sphingosine-1-phosphate and ceramide (Table S56). Possible candidates for a glycosylceramide synthase as well as inositol phosphorylceramide synthase were identified, suggesting that *Porphyra* may be capable of the synthesis of complex structural sphingolipids. An unusual feature of the the fatty acid α-hydroxylase (Pum0218s0026.1) involved in sphingolipid metabolism is its N-terminal cytochrome $b_{5\,;}$ the ER fatty acid Δ9-desaturase (Pum0153s0015.1) has a C-terminal cytochrome b_5-domain, similar to that of fungal orthologues.

Table S56. List of genes involved in sphingolipid metabolism of *Porphyra umbilicalis*.

Porphyra gene identifier	Gene symbol	Aliases or Synonyms	Definition	Description
Pum0082s0057.1	LCB1	SPT	serine palmitoyltransferase subunit 1	serine palmitoyltransferase, long chain base subunit 1, multimeric enzyme with LCB2, catalyzes first step in sphingolipid biosynthesis
Pum0161s0033.1	LCB2	SPT	serine palmitoyltransferase subunit 2	serine palmitoyltransferase, long chain base subunit 1, multimeric enzyme with LCB2, catalyzes first step in sphingolipid biosynthesis; note that the catalytical center is located in LCB2
Pum0065s0003.1	TSC10		3-dehydrosphinganine reductase	3-dehydrosphinganine reductase, ketosphinganine reductase, second step in sphingolipid biosynthesis
Pum0847s0006.1	D4DES1	DES1; DEGS	sphingolipid delta-4 desaturase	sphingolipid desaturase, introduces a delta-4-trans double bond into sphinganine
Pum0207s0017.1	D4DES2	DES1; DEGS	sphingolipid delta-4 desaturase	sphingolipid desaturase, introduces a delta-4-trans double bond into sphinganine
Pum0106s0025.1	LAG1	LAC1; LASS1	ceramide synthase	similar to S. cerevisiae Longevity-assurance proteins LAG1 and LAC1 involved in ceramide biosynthesis
Pum0218s0026.1	FAH1	SCS7	fatty acid alpha-hydroxylase	ceramide hydroxylase, introduces an alpha-hydroxyl group on similar to SCS7 (S. cerevisiae) and FAH1 (A. thaliana)

Amino acids

All twenty amino acids present in proteins are synthesized in *Porphyra*, using mostly plant-like pathways (Table S57). During translation, amino acids are attached to tRNAs by ligases (Table 58). The flavor of nori (*P. yezonensis*) is attributed to the amino acids alanine, glutamate, and asparagine (114), which are also abundant in the similar tasting laver (*Porphyra*). Proline, an osmolyte as well as an important constituent of proteins, is also abundant in *Porphyr*a, which is a good source of most required amino acids in the human diet, except for the sulfated amino acids cysteine and methionine (Figure S59). Selenocysteine is not produced.

Table S57. Amino acid metabolic enzymes in *Porphyra*

Porphyra gene identifier	Gene symbol	Defline	Enzyme commission (E.C.) number
Pum0832s0002.1	PAAH	3-hydroxybutyryl-CoA dehydrogenase	1.1.1.157
Pum1467s0001.1	LDH	L-lactate dehydrogenase	1.1.1.27
Pum0237s0028.1	HOM	homoserine dehydrogenase	1.1.1.3
Pum0127s0007.1	HIBAD	3-hydroxyisobutyrate dehydrogenase	1.1.1.31
Pum0229s0028.1	COMC	(2R)-3-sulfolactate dehydrogenase (NADP+)	1.1.1.338
Pum0254s0035.1	IDH	isocitrate dehydrogenase	1.1.1.42
Pum0518s0009.1	PGD	6-phosphogluconate dehydrogenase	1.1.1.44
Pum0197s0010.1	G6PD1	glucose-6-phosphate 1-dehydrogenase	1.1.1.49
Pum0237s0011.1	G6PD2	glucose-6-phosphate 1-dehydrogenase	1.1.1.49
Pum0167s0014.1	GRHPR1	glyoxylate/hydroxypyruvate reductase	1.1.1.81
Pum0283s0001.1	IMD1	3-isopropylmalate dehydrogenase	1.1.1.85
Pum0316s0012.1	IMD2	3-isopropylmalate dehydrogenase	1.1.1.85
Pum0353s0024.1	ILVC	ketol-acid reductoisomerase	1.1.1.86
Pum0327s0003.1	PHGDH	D-3-phosphoglycerate dehydrogenase	1.1.1.95
Pum0103s0009.1	GPX1	glutathione peroxidase	1.11.1.-
Pum0027s0004.1	GPX2	glutathione peroxidase	1.11.1.-
Pum0153s0010.1	APX1	L-ascorbate peroxidase	1.11.1.11
Pum1752s0002.1	APX2	L-ascorbate peroxidase	1.11.1.11
Pum0152s0048.1	CAT	catalase	1.11.1.6
Pum0720s0003.1	TDO2	tryptophan 2,3-dioxygenase	1.13.11.11
Pum0229s0012.1	ADO	cysteamine dioxygenase	1.13.11.19
Pum0268s0009.1	CDO1	cysteine dioxygenase	1.13.11.20
Pum0647s0007.1	ADI1	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	1.13.11.53
Pum0400s0004.1	HAAO	3-hydroxyanthranilate 3,4-dioxygenase	1.13.11.6
Pum0111s0004.1		cytochrome P450	1.14.-.-
Pum0720s0005.1	PAH	phenylalanine-4-hydroxylase	1.14.16.1
Pum2435s0001.1	DAPB	4-hydroxy-tetrahydrodipicolinate reductase	1.17.1.8
Pum0445s0005.1	RRM1	ribonucleoside-diphosphate reductase subunit M1	1.17.4.1
Pum0070s0052.1	ASD	aspartate-semialdehyde dehydrogenase	1.2.1.11
Pum0060s0049.1	ARGC1	N-acetyl-gamma-glutamyl-phosphate reductase	1.2.1.38
Pum0262s0021.1	ALDH3A1	aldehyde dehydrogenase (NAD(P)+)	1.2.1.5
Pum0533s0004.1	AOX	aldehyde oxidase	1.2.3.1
Pum0399s0005.1	OGDH	2-oxoglutarate dehydrogenase E1 component	1.2.4.2
Pum0021s0027.1	TRYA	arogenate dehydrogenase (NADP+)	1.3.1.78
Pum0104s0018.1	IVD	isovaleryl-CoA dehydrogenase	1.3.8.4
Pum0106s0028.1	ACADM	acyl-CoA dehydrogenase	1.3.8.7
Pum1518s0001.1	AROB	3-dehydroquinate synthase	1.4.1.24
Pum0407s0005.1	GLUD1	glutamate dehydrogenase (NAD(P)+)	1.4.1.3
Pum0415s0006.1	GLUD2	glutamate dehydrogenase (NADP+)	1.4.1.4
Pum0072s0060.1	AOC2	primary-amine oxidase	1.4.3.21
Pum0072s0085.1	DAO	D-amino-acid oxidase	1.4.3.3
Pum0342s0010.1	GLDC	glycine dehydrogenase	1.4.4.2
Pum0234s0025.1	PYCR	pyrroline-5-carboxylate reductase	1.5.1.2
Pum0183s0040.1	PAO	polyamine oxidase	1.5.3.14
Pum0178s0003.1	DLD1	dihydrolipoamide dehydrogenase	1.8.1.4
Pum0463s0001.1	DLD2	dihydrolipoamide dehydrogenase	1.8.1.4
Pum0032s0044.1	GSR1	glutathione reductase (NADPH)	1.8.1.7

Pum0927s0005.1	BHMT2	homocysteine S-methyltransferase	2.1.1.10
Pum0414s0003.1	METE	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	2.1.1.14
Pum0451s0005.1	BSMB	dimethylglycine N-methyltransferase	2.1.1.161
Pum0506s0003.1	ASH1L	histone-lysine N-methyltransferase ASH1L	2.1.1.43
Pum0323s0008.1	EGTD	L-histidine N-alpha-methyltransferase	2.1.1.44
Pum0622s0003.1	AMT	aminomethyltransferase	2.1.2.10
Pum0244s0012.1	OTC	ornithine carbamoyltransferase	2.1.3.3
Pum1617s0002.1	ILV2	acetolactate synthase I/II/III large subunit	2.2.1.6
Pum0592s0012.1		acetyltransferase	2.3.1.-
Pum0205s0017.1	DBT	2-oxoisovalerate dehydrogenase E2 component (dihydrolipoyl transacylase)	2.3.1.168
Pum0368s0015.1	SAT1	serine acetyltransferase	2.3.1.30
Pum0076s0020.1	ARG7	glutamate N-acetyltransferase / amino-acid N-acetyltransferase	2.3.1.35
Pum0523s0014.1	DLST	2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide succinyltransferase)	2.3.1.61
Pum0365s0008.1	GGT6	gamma-glutamyltranspeptidase / glutathione hydrolase	2.3.2.2
Pum0353s0009.1	HMGCS1	hydroxymethylglutaryl-CoA synthase	2.3.3.10
Pum0184s0006.1	LEU2	2-isopropylmalate synthase	2.3.3.13
Pum0615s0013.1	PPAT	amidophosphoribosyltransferase	2.4.2.14
Pum0307s0015.1	HISG	ATP phosphoribosyltransferase	2.4.2.17
Pum3080s0001.1	TRPD	anthranilate phosphoribosyltransferase	2.4.2.18
Pum0234s0020.1	SRM	spermidine synthase	2.5.1.16
Pum0116s0043.1	GST	glutathione S-transferase	2.5.1.18
Pum0076s0061.1	CYSD1	cysteine synthase	2.5.1.47
Pum0375s0007.1	CGS1	cystathionine gamma-synthase	2.5.1.48
Pum0330s0018.1	DHS1	3-deoxy-7-phosphoheptulonate synthase	2.5.1.54
Pum0785s0004.1	DHS2	3-deoxy-7-phosphoheptulonate synthase	2.5.1.54
Pum0121s0018.1	MAT1	S-adenosylmethionine synthetase	2.5.1.6
Pum0615s0007.1	ASP1	aspartate aminotransferase	2.6.1.-
Pum0519s0015.1	ASP2	aspartate aminotransferase	2.6.1.1
Pum0111s0007.1	WIN1	acetylornithine aminotransferase	2.6.1.11
Pum0262s0019.1	OAT	ornithine--oxo-acid transaminase	2.6.1.13
Pum0296s0004.1	GFPT	glucosamine--fructose-6-phosphate aminotransferase (isomerizing)	2.6.1.16
Pum0491s0008.1	BCAT	branched-chain amino acid aminotransferase	2.6.1.42
Pum1678s0002.1	AGXT	alanine--glyoxylate aminotransferase	2.6.1.45
Pum2925s0001.1	TAT	tyrosine aminotransferase	2.6.1.5
Pum0262s0020.1	PSAT1	phosphoserine aminotransferase	2.6.1.52
Pum1777s0002.1	ARO9	aromatic amino acid aminotransferase II	2.6.1.58
Pum0335s0013.1	KYAT1	kynurenine---oxoglutarate transaminase / cysteine-S-conjugate beta-lyase / glutamine---phenylpyruvate transaminase	2.6.1.7
Pum0135s0036.1	ALD1	LL-diaminopimelate aminotransferase	2.6.1.83
Pum0668s0010.1	HPA1	histidinol-phosphate aminotransferase	2.6.1.9
Pum0541s0004.1	GLYK	D-glycerate 3-kinase	2.7.1.31
Pum0027s0080.1	SK1	shikimate kinase	2.7.1.71
Pum0914s0001.1	ACKA	acetate kinase	2.7.2.1
Pum0721s0004.1	PRO1	glutamate 5-kinase	2.7.2.11
Pum1952s0001.1	ASK1	aspartate kinase	2.7.2.4
Pum0766s0002.1	PCYT2	ethanolamine-phosphate cytidylyltransferase	2.7.7.14
Pum0191s0021.1	PAPSS	3'-phosphoadenosine 5'-phosphosulfate synthase	2.7.7.4
Pum0082s0003.1	EPT1	ethanolaminephosphotransferase	2.7.8.1
Pum0268s0005.1	HIBCH	3-hydroxyisobutyryl-CoA hydrolase	3.1.2.4
Pum0191s0007.1	HISB	imidazoleglycerol-phosphate dehydratase / histidinol-phosphatase	3.1.3.15
Pum0116s0016.1	PSPH	phosphoserine phosphatase	3.1.3.3
Pum1610s0001.1	ENOPH1	enolase-phosphatase E1	3.1.3.77
Pum0014s0057.1	BGLU1	beta-glucosidase	3.2.1.21
Pum0104s0022.1	AHCY	adenosylhomocysteinase	3.3.1.1
Pum1573s0001.1	LAP1	leucyl aminopeptidase	3.4.11.1
Pum1589s0001.1	CNDP1	beta-Ala-His dipeptidase	3.4.13.20
Pum0465s0007.1	ANSA	L-asparaginase	3.5.1.1
Pum2174s0001.1	ACY1	aminoacylase	3.5.1.14
Pum0047s0009.1	NIT2	omega-amidase	3.5.1.3
Pum0071s0011.1	AMI1	formamidase	3.5.1.49

Pum0033s0031.1	NLP1	N-carbamoylputrescine amidase	3.5.1.53
Pum1715s0001.1	PYRB	aspartate carbamoyltransferase	3.5.2.3
Pum0031s0024.1	AGUA	agmatine deiminase	3.5.3.12
Pum0130s0036.1	NIT1	nitrilase	3.5.5.1
Pum0230s0005.1	HISI	phosphoribosyl-ATP pyrophosphohydrolase / phosphoribosyl-AMP cyclohydrolase	3.6.1.31
Pum2444s0001.1	ASDA	aspartate decarboxylase	4.1.1.12
Pum0223s0017.1	ODC	ornithine decarboxylase	4.1.1.17
Pum0288s0024.1	DPD1	diaminopimelate decarboxylase	4.1.1.20
Pum0066s0006.1	AMD1	S-adenosylmethionine decarboxylase	4.1.1.50
Pum0328s0008.1	TRP1	anthranilate synthase	4.1.3.27
Pum0673s0005.1	MEE32	3-dehydroquinate dehydratase / shikimate dehydrogenase	4.2.1.10
Pum0215s0028.1	APIP	methylthioribulose-1-phosphate dehydratase	4.2.1.109
Pum0381s0007.1	ECHS1	enoyl-CoA hydratase	4.2.1.17
Pum0645s0001.1	LEU1	3-isopropylmalate dehydratase	4.2.1.33
Pum0510s0016.1	ILVD	dihydroxy-acid dehydratase	4.2.1.9
Pum0443s0008.1	ADT1	arogenate/prephenate dehydratase	4.2.1.91
Pum0125s0027.1	DHQS	3-dehydroquinate synthase	4.2.3.4
Pum0033s0027.1	AROC	chorismate synthase	4.2.3.5
Pum2549s0001.1	ILVA	threonine dehydratase	4.3.1.19
Pum0060s0016.1	ASL1	argininosuccinate lyase	4.3.2.1
Pum0317s0016.1	ASL2	adenylosuccinate lyase	4.3.2.2
Pum0614s0001.1	DHDPS1	4-hydroxy-tetrahydrodipicolinate synthase	4.3.3.7
Pum0608s0010.1	ACS	1-aminocyclopropane-1-carboxylate synthase	4.4.1.14
Pum0413s0011.1	HISA	phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase	5.3.1.16
Pum0022s0045.1	MRI1	methylthioribose-1-phosphate isomerase	5.3.1.23
Pum0953s0005.1	TRPC	indole-3-glycerol phosphate synthase	5.3.1.24
Pum0470s0016.1	PGAM1	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	5.4.2.11
Pum0289s0016.1	PGAM2	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	5.4.2.12
Pum0277s0010.1	MUT	methylmalonyl-CoA mutase	5.4.99.2
Pum2195s0001.1	CM1	chorismate mutase	5.4.99.5
Pum0053s0022.1	PANC	pantoate--beta-alanine ligase	6.3.2.1
Pum0424s0010.1	CARNS1	carnosine synthase	6.3.2.11
Pum1549s0001.1	GCLC	glutamate--cysteine ligase catalytic subunit	6.3.2.2
Pum0622s0006.1	GSS	glutathione synthase	6.3.2.3
Pum0086s0024.1	CPS1	carbamoyl-phosphate synthase (ammonia)	6.3.4.16
Pum0014s0116.1	ADSS	adenylosuccinate synthetase	6.3.4.4
Pum0126s0008.1	ASS1	argininosuccinate synthase	6.3.4.5
Pum0290s0029.1	DUR1	urea carboxylase	6.3.4.6
Pum0014s0045.1	PCCA	propionyl-CoA carboxylase alpha chain	6.4.1.3
Pum0293s0018.1	GDCH	glycine cleavage system H protein	
Pum0103s0061.1	PRODH	proline dehydrogenase	
Pum0288s0034.1	NADSYN1	Glutamine-dependent NAD(+) synthetase	
Pum0285s0010.1	GUA1	GMP synthase (glutamine-hydrolyzing)	

Table S58. Amino acid tRNA-ligases

Porphyra gene identifier	Gene symbol	Aliases	Defline	Description
Pum0014s0059.1	TYRS	FAC31	Tyrosine-tRNA ligase	Catalyzes the attachment of tyrosine to tRNA(Tyr) PMID: 22714903
Pum0171s0026.1	ARGS		Arginine-tRNA ligase	ATP, L-arginine, and tRNA(Arg) combine to form AMP and L-arginyl-tRNA(Arg). PMID: 17211581
Pum0058s0046.1	ASNS1	NARS	Asparagine-tRNA ligase, cytoplasmic	ATP, L-asparagine, and tRNA(Asn) are added together to generate AMP and L-asparaginyl-tRNA(Asn). PMID: 16926386
Pum0098s0020.1	ASPS.1		Aspartate-tRNA(Asp/Asn) ligase	Catalyzes the attachment of aspartate to tRNA (Asp) in a two step reaction. PMID: 16251277
Pum0692s0007.1	ASPS.2	DARS; APS	Aspartate-tRNA(Asp/Asn) ligase, cytoplasmic	Catalyzes the attachment of aspartate to tRNA (Asp) in a two step reaction. PMID: 16251277
Pum0416s0016.1	CYSS		Cysteine--tRNA ligase	ATP, L-cysteine, and tRNA(Cys) combine to form AMP

				and L-cysteinyl-tRNA(Cys) with a Zn(2+) cofactor in the Cytoplasm. PMID: 18487258
Pum0608s0008.1	GLTX	OVA3	Glutamate--tRNA ligase	Catalyzes the attachment of glutamate to tRNA(Glu) in a two-step reaction. PMID: 18431481
Pum2283s0001.1	GRS1.1	F1N18.8; GARS	Glycine--tRNA ligase	Catalyzes the attachment of glycine to tRNA(Gly). Produces diadenosine tetraphosphate, a molecule needed for cell regulation pathways. PMID: 22223895
Pum2191s0001.1	HISS		Histidine-tRNA ligase, cytoplasmic	ATP, L-histidine and tRNA(His) combine to form AMP and L-histidyl-tRNA(His) in the cytoplasm. PMID: 15875012
Pum0516s0002.1	ILES.1		Isoleucine-tRNA ligase	Catalyzes the attachment of isoleucine to tRNA(Ile). PMID: 12917641
Pum0239s0026.1	ILES.2	IARS	Isoleucine-tRNA ligase, cytoplasmic	ATP, L-isoleucine, and tRNA(Ile) combine to form AMP and L-isoleucyl-tRNA(Ile). PMID: 16251277
Pum0510s0013.1	LEUS.1	LARS	Leucine-tRNA ligase	Catalyzes the attachment of leucine to tRNA (Leu) in a two step reaction. PMID: 16251277
Pum0482s0003.1	LEUS.2		Leucine-tRNA ligase, cytoplasmic	Catalyzes the attachment of leucine to tRNA (Leu) in a two step reaction. PMID: 16251277
Pum0745s0005.1	LYSS		Lysine-tRNA ligase	Involved in the formation of lysol-transfer RNA which is then ready to insert lysine into proteins. PMID: 18159947; 7576245
Pum0924s0002.1	FARSB.1	F28P22.26; FRS1	Phenylalanine--tRNA ligase beta subunit	ATP, L-phenylalanine, and tRNA(Phe) combine to form AMP and L-phenylalanyl-tRNA(Phe). PMID: 16251277
Pum0924s0001.1	FARSB.2	F28P22.26; FRS1	Phenylalanine--tRNA ligase beta subunit, cytoplasmic	ATP, L-phenylalanine, and tRNA(Phe) combine to form AMP and L-phenylalanyl-tRNA(Phe). PMID: 16251277
Pum0107s0029.1	FARSA.1	MPHES; FRS2.1; PHES	Phenylalanine-tRNA ligase alpha subunit	ATP, L-phenylalanine, and tRNA(Phe) combine to form AMP and L-phenylalanyl-tRNA(Phe). PMID: 16251277
Pum0116s0033.1	FARSA.2	MPHES; FRS2.1; PHES	Phenylalanine-tRNA ligase alpha subunit, cytoplasmic	ATP, L-phenylalanine, and tRNA(Phe) combine to form AMP and L-phenylalanyl-tRNA(Phe). PMID: 16251277
Pum0234s0045.1	FARS2	FARS1	Phenylalanine-tRNA ligase,chloroplastic/mitochondrial	Is responsible for the charging of tRNA(Phe) with phenylalanine in mitochondrial translation. PMID: 22223895
Pum1865s0001.1	PRORS.1	PROS	Proline--tRNA ligase	Catalyzes the attachment of proline to tRNA(Pro) in a two-step reaction. PMID: 12949112
Pum1980s0001.1	PRORS.2	PROS	Proline--tRNA ligase	Catalyzes the attachment of proline to tRNA(Pro) in a two-step reaction. PMID: 16251277
Pum0365s0006.1	WARS		Tryptophan--tRNA ligase, cytoplasmic	Belongs to the class-I aminoacyl-tRNA synthetase family. PMID: 15875012
Pum0255s0015.1	VALS.1	VALS1	Valine-tRNA ligase	Catalyzes the attachment of valine to tRNA(Val). PMID: 16251277; 16311624
Pum0284s0015.1	VALS.2	VALS1	Valine-tRNA ligase	Catalyzes the attachment of valine to tRNA(Val). PMID: 15875012; 14752164
Pum0258s0021.1	ALA1.1	ALARS, ALATS	Alanine-tRNA ligase	Catalyzes the attachment of alanine to tRNA (Ala).
Pum0666s0009.1	ALA1.2	ALARS, ALATS	Alanine-tRNA ligase	Catalyzes the attachment of alanine to tRNA (Ala).
Pum0191s0009.1	THRS.1		Threonine-tRNA ligase	Catalyzes the attachment of threonine to tRNA (Thr).
Pum0126s0017.1	THRS.2		Threonine-tRNA ligase	Catalyzes the attachment of threonine to tRNA (Thr).
Pum0373s0005.1	SERS.1		Serine-tRNA ligase	Catalyzes the attachment of serine to tRNA (Ser).
Pum0373s0008.1	SERS.2		Serine-tRNA ligase	Catalyzes the attachment of serine to tRNA (Ser).
Pum0246s0005.1	SERS.3		Serine-tRNA ligase	Catalyzes the attachment of serine to tRNA (Ser).

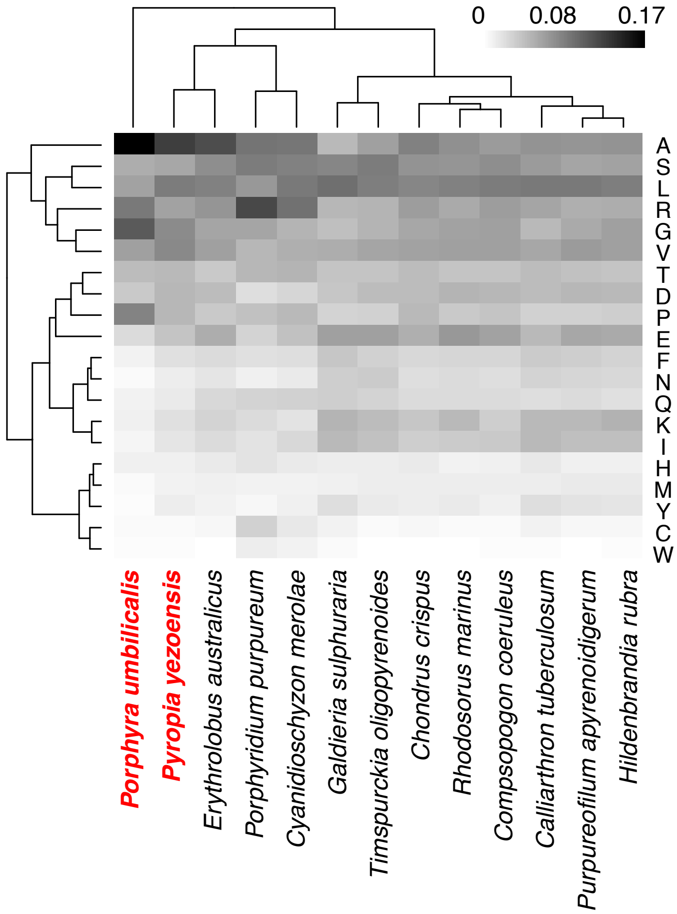


Figure S59. Amino acid composition.

Heat map of amino acid composition of predicted protein coding regions in *P. umbilicalis* and other red algae. *Porphyra* has a distinctive amino acid profile as demonstrated by the cluster analysis. Data from Table S13.

B Vitamins

Algal foods, such as nori, are an excellent dietary source of several vitamins, especially vitamin B_{12} (114). A survey of *Porphyra* alongside four other red algae with sequenced genomes (*Chondrus crispus*, *Galdieria sulphuraria*, *Cyanidioschyzon merolae,* and *Pyropia yezoensis*) confirms none encodes the genes necessary to synthesize vitamin B_{12} *de* novo, as expected since no eukaryote is known to make vitamin B_{12}. Homologs that are found are either likely to be for siroheme biosynthesis (*CobA, CobM*), or chlorophyll synthesis (Aerobic *CobN, CobS*). *Porphyra* does encode two B_{12}-dependent enzymes: B_{12}-dependent methionine synthase (METH) and methylmalonyl-CoA mutase (MCM) (Table S60). However, the presence of B_{12}-independent methionine synthase (*METE*) in *P. umbilicalis* suggests this alga may not have an absolute requirement for vitamin B_{12}, since this isoform of methionine synthase can function in the absence of a B_{12} cofactor, and MCM has been reported in algae that do not strictly need B_{12} under standard growth conditions (126). Interestingly, the thermophilic microalga *C. merolae* that encodes *METE*, lacks both *METH* and *MCM* implying that this alga may have dispensed with using vitamin B_{12} entirely, perhaps because the vitamin is not readily available in the hot, acidic environment it inhabits (19). In any case, the source of vitamin B_{12} in *Porphyra* is likely from associated B_{12}-synthesizing bacteria. *Porphyra* encodes *CobS* and *CobT*, which means it can remodel the pseudocobalamin that some bacteria produce into bioavailable vitamin B_{12} (127). In contrast none of the four other sequenced red algae analyzed here have this ability (Table S60).

Table S60. Analysis of 6 sequenced red algal genomes for genes encoding vitamin B_{12}-related proteins

N.F. not found

GENE	*P. umbilicalis*	*C. crispus*	*C. merolae*	*G. sulphuraria*	*P. purpureum*	*P. yezoensis*
METE	Pum0414s0003.1	Phy0034O3I	CMJ234C	XP_005705621.1	Porphyridium.evm.model.contig_3639.2	PyezoenV1.contig_11982_g2851
METH	Pum0927s0005.1	Phy0034QGV	N.F.	N.F.[a]	Porphyridium.evm.model.contig_3545.3	PyezoenV1.contig_12131_g2904
COBT	Pum0208s0033.1	N.F.	N.F.	N.F.	N.F.	N.F.
COBS	Pum0208s0032.1	N.F.	N.F.	N.F.	N.F.	PyezoenV1.contig_42218_g9563
CBA1	N.F.	N.F.	N.F.	N.F.	N.F.	N.F.
RNRII	N.F.	N.F.	N.F.	N.F.	N.F.	N.F.
MCM	Pum0277s0010.1	N.F.	N.F.	XP_005703451.1	Porphyridium.evm.model.contig_4516.5	PyezoenV1.contig_39938_g9208

[a]a hit to XP_005706691.1 was obtained but since it had no B_{12} binding domain, it is more likely betaine homocysteine S-methyl transferase

We also searched for biosynthesis genes of other B-group vitamins. Thiamine (vitamin B_1) is synthesized through the condensation of a thiazole and pyrimidine moiety to form thiamine monophosphate (TMP) (Figure S61). In eukaryotes, TMP is dephosphorylated and then pyrophosphorylated to make TPP, the biologically active form. However, there are various routes for the synthesis of both the thiazole and pyrimidine moieties. *P. umbilicalis* encodes *ThiC* (Table S62)*,* which is necessary for synthesis of the pyrimidine precursor HMP-P. This is also the route found in bacteria and higher plants, as well as in several other algae, but fungi use an alternative enzyme. In contrast, *Porphyra* does not possess a *THI4* gene, encoding the enzyme responsible for synthesis of the thiazole moiety that is found in green algae and plant*s.* Instead, it encodes several of the genes identified in bacteria (*thiF*, *iscS*, *thiO* and *DXS*, but not *thiG* and *thiS*) suggesting it may be able to synthesize thiazole by this route. However, *P. umbiicalis* appears to lack the gene *thiD,* necessary for the phosphorylation of HMP-P to HMP-PP, suggesting it may do this via an alternative route or scavenge a phosphorylated pyrimidine moiety from the environment. A homolog for TPK was also identified. Incidentally, genes of the bacterial-like thiazole biosynthesis pathway have also been identified in the available stramenopile genomes (including *Phaeodactylum tricornutum*, *Thalassiosira pseudonana, Aureococcus anophagefferens*, *Fragilariopsis cylindrus* and *Pseudo-nitzchia multiseries*), alongside the haptophyte *Emiliania huxleyi* (128, 129), implying they may have originated from a red algal endosymbiont in these organisms.

The biosynthetic pathway for vitamin B_7 (biotin) was also identified in *Porphyra* (Tables S63, S64). Four biotin biosynthesis genes are known from bacteria: *bioF, bioA, bioD* and *bioB.* Three orthologous genes are present in *Arabidopsis thaliana* (*bioF*, *bioA* (*BIO1* in plants)) and *bioB* (*BIO2* in plants), which can synthesize biotin. Homologs for all three of these were present in *P. umbilicalis*, *C. crispus*, *G. sulphuraria*, and *C. merolae* (Table S63). Although several frame-shifts were found within the gene sequence for *bioA,* indicating this is likely a pseudogene. Without a canonical *BioA* gene, it is possible that *Porphyra* is auxotrophic for biotin.

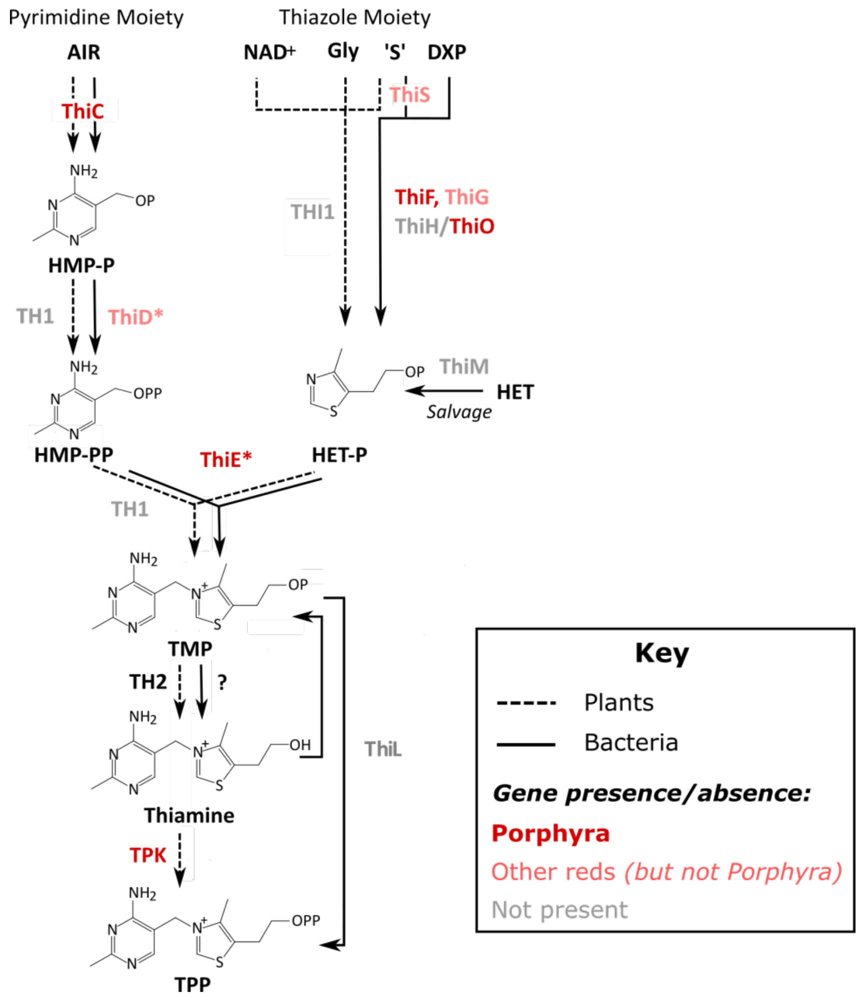


Figure S61. Presence of thiamine pyrophosphate (TPP) genes in red algae.
The presence of TPP genes is indicated for genes in *P. umbilicalis* and other red algae (*C. crispus*, *G. sulphuraria*, *C. merolae, P. purpureum* and *P. yezoensis).* The TPP biosynthetic pathway as described in plants (*A. thaliana*) and bacteria (*E. coli*) is shown, and the presence of genes in red algae indicated. In *A. thaliana* the thiazole moiety (HET-P) is thought to be synthesised from NAD^+, glycine and a sulphur donor ('S') like in fungi. *E. coli* synthesises HET-P via the deoxy-D-xylulose 5-phosphate (DXP) pathway. The pyrimidine biosynthesis pathway is similar in both plants and bacteria, except in *A. thaliana* the enzymatic activities of ThiD and ThiE are combined into a single bifunctional enzyme (encoded by the gene *TH1*) (abbreviations: AIR: 5-aminoimidazole ribonucleotide, HMP-P: 4-amino-2-methyl-5-hydroxymethylpyrimidine

monophosphate, HMP-PP: 4-amino-2-methyl-5-hydroxymethylpyrimidine diphosphate, TMP: thiamine monophosphate).

Table S62. Analysis of six sequenced algal genomes for genes encoding Vitamin B_1 (thiamine) biosynthesis proteins

The table lists gene model identifiers in diverse red algae if present in that sequenced genome. N.F. indicates not found.

	Gene symbol	***P. umbilicalis***	***C. crispus***	***G. sulphuraria***	***C. merolae***	***P. purpureum***	***P. yezoensis***
Pyrimidine branch (*C. reinhardtii*)	**THIC**	Pum0148s0011.1	Phy0034SBL	XP_005707568.1	CMG171C	contig_474.6	contig_5158_g1150
	TH1	N.F.	N.F.	N.F.	N.F.	N.F.	N.F.
	THID	N.F.	Phy0034R0Z	XP_005703802.1	YCMO125C	contig_4432.5	contig_7296
	THIE	Pum0617s0016.1	Phy0034R9P	XP_005705662.1 XP_005706408.1	CMP214C	contig_447.2	N.F.
Pyrimidine branch (fungi)	**NMT1**	N.F.	N.F.	N.F.	N.F.	N.F.	N.F.
Thiazole branch (*C. reinhardtii*)	**THIM**	N.F.	N.F.	N.F.	N.F.	N.F.	N.F.
	THI4	N.F.	N.F.	N.F.	N.F.	N.F.	N.F.
	TPK	Pum0572s0005.1	Y	XP_005708580.1	CMH016C	contig_2016.13	N.F.
Thiazole branch (bacteria)	**THIF**	Pum0144s0011.1	Phy0034O37	XP_005704091.1	CMM289C	contig_3558.5	contig_26628_g6547
	ISCS	Pum2204s0001.1	Phy0034QN0	XP_005703716.1	CMT234C	contig_3527.5	contig_38937
	THIS	N.F.	N.F.	N.F.	CMV092C	N.F.	N.F.
	THIH	N.F.	N.F.	N.F.	N.F.	contig_2123.4	N.F.
	ThiO	Pum0161s0058.1	Phy0034MTJ	XP_005706568.1	CMM298C	contig_2104.14	N.F.
	DXS	Pum0237s0005.1 Pum0237s0013.1	Phy0034PZ7	XP_005704891.1	CMF089C	contig_3433.1	contig_25494_g6284 contig_18029_g4414
	THIG	N.F.	Phy0034MC5	XP_005704987.1	CMV077C	contig_556.3	N.F.
	ThiL	N.F.	N.F.	N.F. N.F.	N.F.	N.F.	N.F.

Table S63. Analysis of 6 sequenced algal genomes for genes encoding Vitamin B_7 (biotin) biosynthesis proteins

The table lists gene model identifiers in diverse red algae if present in that sequenced genome. N.F. indicates not found.

Gene symbol	***P. umbilicalis***	***C. crispus***	***G. sulphuraria***	***C. merolae***	***P. purpureum***	***P. yezoensis***
BIOF	Pum0161s0033.1	Phy0034M27	XP_005709068.1	CMO139C	Porphyridium.evm.model.contig_3624.1	PyezoenV1.contig_23737_g5851
BIOB	Pum0184s0007.1 [a]	Phy0034NVW	XP_005708404.1	CML210C	Porphyridium.evm.model.contig_3587.4	PyezoenV1.contig_32784_g7945
BIOD	N.F.	N.F.	N.F.	N.F.	N.F.	N.F.
BIOA	Pum0159s0004.1 [b]	Phy0034NNJ	XP_005704758.1	CMG023C	Porphyridium.evm.model.contig_4512.5	N.F.

[a] Protein lacks radical SAM domain (PF04055), [b] gene model lacks AAA domain (PF13500). The model can be extended, but frame-shifts within the sequence suggests possible pseudogene

Table S64. B Vitamin synthesis genes

See table S25 for genes involved in synthesis of vitamins C (ascorbic acid) and E (tocopherol)

Porphyra gene identifier	Gene Symbol	Aliases or Synonyms	Defline	Notes
Pum0414s0003.1	METE		Cobalamin-independent methionine synthase; 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	Croft et al. (2005) Nature. 438: 90-93
Pum0927s0005.1	METH		Cobalamin-dependent methionine synthase; 5-methyltetrahydrofolate--homocysteine methyltransferase	
Pum0208s0033.1	COBT		Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	Cobalamin biosynthesis/remodelling
Pum0208s0032.1	COBS		Adenosylcobinamide-GDP ribazoletransferase	Cobalamin biosynthesis/remodelling
Pum0277s0010.1	MCM	MUT	Methylmalonyl-CoA mutase	Cobalamin dependent enzyme
Pum0229s0024.1	PANB	KPHMT1; PANB1	Ketopantoate hydroxymethyltransferase 1	
Pum0053s0022.1	PANC	PS; PTS	Pantoate-beta-alanine ligase; pantothenate synthetase	
Pum0148s0011.1	THIC		Phosphomethylpyrimidine synthase	Thiamine (vitamin B1) biosynthesis protein
Pum0617s0016.1	THIE		Thiamine phosphate synthase (thiamine phosphate pyrophosphorylase)	Thiamine (vitamin B1) biosynthesis protein
Pum0572s0005.1	TPK		Thiamine pyrophosphokinase	
Pum0144s0011.1	THIF		Thiazole biosynthesis adenylyltransferase	Thiamine (vitamin B1) biosynthesis protein
Pum0161s0058.1	ThiO		Thiamine biosynthesis protein	Thiamine (vitamin B1) biosynthesis protein
Pum0237s0013.1	DXS1	DXPS;CLA1;DEF	1-deoxyxylulose-5-phosphate synthase	
Pum0237s0005.1	DXS2	DXPS;CLA1;DEF	1-deoxyxylulose-5-phosphate synthase	
Pum2204s0001.1	ISCS1		Cysteine desulfurase	
Pum0052s0070.1	ISCS2		Cysteine desulfurase	
Pum0161s0033.1	BIOF		7-keto-8aminopelargonic acid synthase	Biotin biosynthesis protein BIOF, 7-keto-8-aminopelargonic acid synthase
Pum0159s0003.1	BIOA		Diaminopelargonic acid synthase/ adenosylmethionine-8-amino-7-oxononanoate aminotransferase	Probable BIOA protein, although gene model lacks AAA domain (PF13500). The model can be extended, but frame-shifts within the sequence suggests possible pseudogene
Pum0184s0007.1	BIOB		Biotin synthase	

References

1. Chan CX *et al.* (2012) *Porphyra* (Bangiophyceae) transcriptomes provide insights into red algal development and metabolism. *J Phycol* 48:1328-1342.
2. Miranda LN *et al.* (2013) Diversity and abundance of the bacterial community of the red macroalga *Porphyra umbilicalis*: Did bacterial farmers produce macroalgae? *PLoS One* 8:e58269.
3. Kim JW *et al.* (2016) Genome analysis of Planctomycetes inhabiting blades of the red alga *Porphyra umbilicalis*. *PLoS One* 11:e0151883.
4. Butler J *et al.* (2008) ALLPATHS: de novo assembly of whole-genome shotgun microreads. *Genome Res* 18:810-820.
5. Gordon D *et al.* (2016) Long-read sequence assembly of the gorilla genome. *Science* 352:aae0344.
6. Chin CS *et al.* (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563-569.
7. Li H (2014) Burrows-Wheeler Aligner (BWA) (Broad Institute), 0.7.6a.
8. McKenna A *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297-1303.
9. Altschul SF *et al.* (1990) Basic local alignment search tool. *J Mol Biol* 215:403-410.
10. Suzek BE *et al.* (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* 23:1282-1288.
11. Team RC (2016) R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria).
12. Smit AF & Hubley R (2008-2015) RepeatModeler Open-1.0.
13. Finn RD *et al.* (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279-285.
14. Thomas PD *et al.* (2003) PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* 13:2129-2141.
15. Smit A, Hubely R & Green P (2013-2015) RepeatMasker Open-4.0.
16. Abrusán G *et al.* (2009) TEclass—a tool for automated classification of unknown eukaryotic transposable elements. *Bioinformatics* 25:1329-1330.
17. Stiller JW *et al.* (2012) Major developmental regulators and their expression in two closely related species of *Porphyra* (Rhodophyta). *J Phycol* 48:883-896.
18. Haas BJ *et al.* (2003) Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res* 31:5654-5666.
19. Matsuzaki M *et al.* (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428:653-657.
20. Collén J *et al.* (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proc Natl Acad Sci USA* 110:5247-5252.
21. Price DC *et al.* (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science* 335:843-847.
22. Bhattacharya D *et al.* (2013) Genome of the red alga *Porphyridium purpureum*. *Nat Commun* 4:1941.

23. Schonknecht G *et al.* (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* 339:1207-1210.
24. Nakamura Y *et al.* (2013) The first symbiont-free genome sequence of marine red alga, Susabi-nori (*Pyropia yezoensis*). *PLoS One* 8:e57122.
25. Anonymous (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Research* 45:D158-D169.
26. Slater GS & Birney E (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6:31.
27. Salamov AA & Solovyev VV (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res* 10:516-522.
28. Yeh RF, Lim LP & Burge CB (2001) Computational inference of homologous gene structures in the human genome. *Genome Res* 11:803-816.
29. Qiu H, Yoon HS & Bhattacharya D (2016) Red algal phylogenomics provides a robust framework for inferring evolution of key metabolic pathways. *PLoS Currents* 8:e4567.
30. Keeling PJ *et al.* (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol* 12:e1001889.
31. Qiu H *et al.* (2015) Evidence of ancient genome reduction in red algae (Rhodophyta). *J Phycol* 51:624–636.
32. Chan CX *et al.* (2011) Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. *Curr Biol* 21:328-333.
33. Parra G, Bradnam K & Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* 23:1061-1067.
34. Smith DR *et al.* (2012) Relative rates of evolution among the three genetic compartments of the red alga Porphyra differ from those of green plants and do not correlate with genome architecture. *Mol Phylogenet Evol* 65:339-344.
35. Hazkani-Covo E, Zeller RM & Martin W (2010) Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet* 6:e1000834.
36. Smith DR, Crosby K & Lee RW (2011) Correlation between nuclear plastid DNA abundance and plastid number supports the limited transfer window hypothesis. *Genome Biol Evol* 3:365-371.
37. Yoon HS *et al.* (2016) Rhodophyta. *Handbook of the Protists*, eds Archibald JM, Simpson AG, & Slamovits CH (Springer International Publishing, Switzerland), pp 1-43.
38. Li L, Stoeckert CJ, Jr. & Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13:2178-2189.
39. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.
40. Capella-Gutierrez S, Silla-Martinez JM & Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972-1973.
41. Notredame C, Higgins DG & Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302:205-217.
42. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312-1313.
43. Merchant SS *et al.* (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318:245-250.

44. Prochnik SE *et al.* (2010) Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science* 329:223-226.
45. Blanc G *et al.* (2010) The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* 22:2943-2955.
46. Blanc G *et al.* (2012) The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biol* 13:R39.
47. Worden AZ *et al.* (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324:268-272.
48. Palenik B *et al.* (2007) The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci USA* 104:7705-7710.
49. Hori K *et al.* (2014) *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat Commun* 5:3978.
50. Lander ES *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921.
51. Mikkelsen TS *et al.* (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* 447:167-177.
52. Simakov O *et al.* (2013) Insights into bilaterian evolution from three spiralian genomes. *Nature* 493:526-531.
53. Srivastava M *et al.* (2008) The *Trichoplax* genome and the nature of placozoans. *Nature* 454:955-960.
54. Srivastava M *et al.* (2010) The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature* 466:720-726.
55. Ryan JF *et al.* (2013) The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science* 342:1242592.
56. King N *et al.* (2008) The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* 451:783-788.
57. Malik S-B *et al.* (2008) An expanded inventory of conserved meiotic genes provides evidence for sex in *Trichomonas vaginalis*. *PLoS One* 3:e2879.
58. Kong Z *et al.* (2015) Kinesin-4 functions in vesicular transport on cortical microtubules and regulates cell wall mechanics during cell elongation in plants. *Mol Plant* 8:1011-1023.
59. Heaslip AT, Nelson SR & Warshaw DM (2016) Dense granule trafficking in *Toxoplasma gondii* requires a unique class 27 myosin and actin filaments. *Molecular Biology of the Cell* 27:2080-2089.
60. Odronitz F & Kollmar M (2007) Drawing the tree of eukaryotic life based on the analysis of 2,269 manually annotated myosins from 328 species. *Genome Biol* 8:R196.
61. Mirarab S *et al.* (2015) PASTA: ultra-large multiple sequence alignment for nucleotide and amino acid sequences. *J Comput Biol* 22:377-386.
62. Han MV & Zmasek CM (2009) phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* 10:356.
63. Marchler-Bauer A *et al.* (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222-226.
64. Wickstead B, Gull K & Richards TA (2010) Patterns of kinesin evolution reveal a complex ancestral eukaryote with a multifunctional cytoskeleton. *BMC Evol Biol* 10:110.
65. Price MN, Dehal PS & Arkin AP (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490.

66. Lawrence CJ *et al.* (2004) A standardized kinesin nomenclature. *J Cell Biol* 167:19-22.
67. DellaPenna D (2005) A decade of progress in understanding vitamin E synthesis in plants. *J Plant Physiol* 162:729-737.
68. Fryer MJ (1992) The antioxidant effects of thylakoid Vitamin E (alpha-tocopherol). *Plant, Cell Environ* 15:381–392.
69. Havaux M *et al.* (2005) Vitamin E protects against photoinhibition and photooxidative stress in *Arabidopsis thaliana*. *Plant Cell* 17:3451-3469.
70. Collakova E & DellaPenna D (2003) The role of homogentisate phytyltransferase and other tocopherol pathway enzymes in the regulation of tocopherol synthesis during abiotic stress. *Plant Physiol* 133:930-940.
71. Kamal-Eldin A & Appelqvist LA (1996) The chemistry and antioxidant properties of tocopherols and tocotrienols. *Lipids* 31:671-701.
72. Ferraces-Casais P *et al.* (2011) Evaluation of bioactive compounds in fresh edible seaweeds. *Food Anal Methods* 5:828–834.
73. Balskus EP & Walsh CT (2010) The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. *Science* 329:1653-1656.
74. Waller RF, Slamovits CH & Keeling PJ (2006) Lateral gene transfer of a multigene region from cyanobacteria to dinoflagellates resulting in a novel plastid-targeted fusion protein. *Mol Biol Evol* 23:1437-1443.
75. Starcevic A *et al.* (2008) Enzymes of the shikimic acid pathway encoded in the genome of a basal metazoan, *Nematostella vectensis*, have microbial origins. *Proc Natl Acad Sci U S A* 105:2533-2537.
76. Miyamoto KT, Komatsu M & Ikeda H (2014) Discovery of gene cluster for mycosporine-like amino acid biosynthesis from Actinomycetales microorganisms and production of a novel mycosporine-like amino acid by heterologous expression. *Appl Environ Microbiol* 80:5028-5036.
77. Yang L-E *et al.* (2014) The P450-type carotene hydroxylase PuCHY1 from Porphyra suggests the evolution of carotenoid metabolism in red algae. *J Integrative Plant Biol* 56:902-915.
78. Lohr M, Schwender J & Polle JE (2012) Isoprenoid biosynthesis in eukaryotic phototrophs: a spotlight on algae. *Plant Sci* 185-186:9-22.
79. Tamura K *et al.* (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725-2729.
80. Mukherjee B & Moroney JV (2001) *Algal Carbon Dioxide Concentrating Mechanisms* (John Wiley & Sons, Ltd).
81. Nakajima K, Tanaka A & Matsuda Y (2013) SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc Natl Acad Sci U S A* 110:1767-1772.
82. Reinfelder JR, Kraepiel AM & Morel FM (2000) Unicellular C_4 photosynthesis in a marine diatom. *Nature* 407:996-999.
83. De Veylder L, Beeckman T & Inze D (2007) The ins and outs of the plant cell cycle. *Nat Rev Mol Cell Biol* 8:655-665.
84. Bisova K, Krylov DM & Umen JG (2005) Genome-wide annotation and expression profiling of cell cycle regulatory genes in *Chlamydomonas reinhardtii*. *Plant Physiol* 137:475-491.

85. Li Y *et al.* (2016) A new class of cyclin dependent kinase in *Chlamydomonas* is required for coupling cell size to cell division. *Elife* 5:e10767.
86. Ma Z *et al.* (2013) Phylogenetic analysis reveals the evolution and diversification of cyclins in eukaryotes. *Mol Phylogenet Evol* 66:1002-1010.
87. Barbier G *et al.* (2005) Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae. *Plant Physiol* 137:460-474.
88. Umeda M, Shimotohno A & Yamaguchi M (2005) Control of cell division and transcription by cyclin-dependent kinase-activating kinases in plants. *Plant Cell Physiol* 46:1437-1442.
89. Cross FR & Umen JG (2015) The Chlamydomonas cell cycle. *Plant J* 82:370-392.
90. Hanschen ER *et al.* (2016) The *Gonium pectorale* genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. *Nat Commun* 7:11370.
91. Fujiwara T *et al.* (2009) Periodic gene expression patterns during the highly synchronized cell nucleus and organelle division cycles in the unicellular red alga *Cyanidioschyzon merolae*. *DNA Res* 16:59-72.
92. Moriyama T *et al.* (2010) Characterization of cell-cycle-driven and light-driven gene expression in a synchronous culture system in the unicellular rhodophyte *Cyanidioschyzon merolae*. *Microbiology* 156:1730-1737.
93. Bouarab K *et al.* (1999) Sulfated oligosaccharides mediate the interaction between a marine red alga and Its green algal pathogenic endophyte. *Plant Cell* 11:1635-1650.
94. Weinberger F & Friedlander M (2000) Response of *Gracilaria conferta* (Rhodophyta) to oligoagars results in defense against agar-degrading epiphytes. *J Phycol* 36:1079–1086.
95. Indrasumunar A *et al.* (2015) Functional analysis of duplicated Symbiosis Receptor Kinase (SymRK) genes during nodulation and mycorrhizal infection in soybean (*Glycine max*). *J Plant Physiol* 176:157-168.
96. Dangl JL & Jones JD (2001) Plant pathogens and integrated defence responses to infection. *Nature* 411:826-833.
97. Conesa A *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.
98. Petersen TN *et al.* (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785-786.
99. Krogh A *et al.* (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567-580.
100. Chevanne D *et al.* (2010) WD-repeat instability and diversification of the *Podospora anserina* HNWD non-self recognition gene family. *BMC Evol Biol* 10:134.
101. Zambounis A *et al.* (2012) Highly dynamic exon shuffling in candidate pathogen receptors ... what if brown algae were capable of adaptive immunity? *Mol Biol Evol* 29:1263-1276.
102. Weis WI, Taylor ME & Drickamer K (1998) The C-type lectin superfamily in the immune system. *Immunol Rev* 163:19-34.
103. Klochkova TA *et al.* (2012) Host–parasite interactions and host species susceptibility of the marine oomycete parasite, *Olpidiopsis* sp., from Korea that infects red algae. *J App Phyc* 24:135-144.

104. O'Rourke D *et al.* (2006) Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. *Genome Res* 16:1005-1016.
105. Boisson-Dernier A, Kessler SA & Grossniklaus U (2011) The walls have ears: the role of plant CrRLK1Ls in sensing and transducing extracellular signals. *J Exp Bot* 62:1581-1591.
106. Hok S *et al.* (2011) An Arabidopsis (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease. *Plant Cell Environ* 34:1944-1957.
107. Basilico C *et al.* (2008) A high affinity hepatocyte growth factor-binding site in the immunoglobulin-like region of Met. *J Biol Chem* 283:21267-21277.
108. Graham LE *et al.* (2016) *Algae* (LJLM Press, Madison, WI) 3rd ed.
109. Viola R, Nyvall P & Pedersen M (2001) The unique features of starch metabolism in red algae. *Proc Biol Sci* 268:1417-1422.
110. Rhodes D, Sims AP & Folkes BF (1980) Pathway of ammonia assimilation in illuminated *Lemna minor*. *Phytochemistry* 19:357-365.
111. Wu H *et al.* (2015) Bioremediation efficiency of the largest scale artificial *Porphyra yezoensis* cultivation in the open sea in China. *Mar Pollut Bull* 95:289-296.
112. Levy I & Gantt E (1990) Development of photosynthetic activity in *Porphyridium purpureum* (Rhodophyta) following nitrogen starvation. *Journal of Phycology* 26:62-68.
113. Karl DM (2014) Microbially mediated transformations of phosphorus in the sea: new views of an old cycle. *Ann Rev Mar Sci* 6:279-337.
114. Wells ML *et al.* (2017) Algae as nutritional and functional food sources: revisiting our understanding. *J Appl Phycol* 29:949-982.
115. Wada H, Gombos Z & Murata N (1994) Contribution of membrane lipids to the ability of the photosynthetic machinery to tolerate temperature stress. *Proc Natl Acad Sci U S A* 91:4273-4277.
116. Blouin N *et al.* (2006) Sensory and fatty acid analyses of two Atlantic species of *Porphyra* (Rhodophyta). *J Appl Phycol* 18:79.
117. Gong Y *et al.* (2011) Characterization of a novel thioesterase (PtTE) from *Phaeodactylum tricornutum*. *J Basic Microbiol* 51:666-672.
118. Alberts B *et al.* (2014) *Molecular Biology of the Cell* (Garland Science, New York) 6th ed.
119. Wu S-Z & Bezanilla M (2014) Myosin VIII associates with microtubule ends and together with actin plays a role in guiding plant cell division. *elife* 3:e03498.
120. Bezanilla M *et al.* (2015) Cytoskeletal dynamics: A view from the membrane. *J Cell Biol* 209:329-337.
121. Gao J *et al.* (2009) FATTY ACID DESATURASE4 of *Arabidopsis* encodes a protein distinct from characterized fatty acid desaturases. *Plant J* 60:832-839.
122. Araki S *et al.* (1987) Positional distribution of fatty acids in glycerolipids of the marine red alga, *Porphyra yezoensis*. *Plant Cell Physiol* 28:761-766.
123. Khozin I *et al.* (1997) Elucidation of the biosynthesis of eicosapentaenoic acid in the microalga *Porphyridium cruentum* (II. Studies with radiolabeled precursors). *Plant Physiol* 114:223-230.
124. De Angelis L *et al.* (2005) Marine macroalgae analyzed by mass spectrometry are rich sources of polyunsaturated fatty acids. *J Mass Spectrom* 40:1605-1608.
125. Hurlock AK *et al.* (2014) Lipid trafficking in plant cells. *Traffic* 15:915-932.

126. Helliwell KE *et al.* (2011) Insights into the evolution of vitamin B_{12} auxotrophy from sequenced algal genomes. *Mol Biol Evol* 28:2921-2933.
127. Helliwell KE *et al.* (2016) Cyanobacteria and eukaryotic algae use different chemical variants of vitamin B_{12}. *Curr Biol* 26:999-1008.
128. McRose D *et al.* (2014) Alternatives to vitamin B_1 uptake revealed with discovery of riboswitches in multiple marine eukaryotic lineages. *ISME J* 8:2517-2529.
129. Bertrand EM & Allen AE (2012) Influence of vitamin B auxotrophy on nitrogen metabolism in eukaryotic phytoplankton. *Front Microbiol* 3:375.