The Plastid Genome of the Cryptomonad *Teleaulax amphioxeia*

Jong Im Kim[1], Hwan Su Yoon[2], Gangman Yi[3], Hyung Seop Kim[4], Wonho Yih[4]*, Woongghi Shin[1]*

1 Department of Biology, Chungnam National University, Daejeon, Korea, 2 Department of Biological Sciences, Sungkyunkwan University, Suwon, Korea, 3 Department of Computer Science, Gangneung-Wonju National University, Wonju, Korea, 4 Department of Marine Biotechnology, Kunsan National University, Kunsan, Korea

* ywonho@kunsan.ac.kr (WY); shinw@cnu.ac.kr (WS)

Abstract

Teleaulax amphioxeia is a photosynthetic unicellular cryptophyte alga that is distributed throughout marine habitats worldwide. This alga is an important plastid donor to the dinoflagellate *Dinophysis caudata* through the ciliate *Mesodinium rubrum* in the marine food web. To better understand the genomic characteristics of *T. amphioxeia*, we have sequenced and analyzed its plastid genome. The plastid genome sequence of *T. amphioxeia* is similar to that of *Rhodomonas salina*, and they share significant synteny. This sequence exhibits less similarity to that of *Guillardia theta*, the representative plastid genome of photosynthetic cryptophytes. The gene content and order of the three photosynthetic cryptomonad plastid genomes studied is highly conserved. The plastid genome of *T. amphioxeia* is composed of 129,772 bp and includes 143 protein-coding genes, 2 rRNA operons and 30 tRNA sequences. The DNA polymerase III gene (*dna*X) was most likely acquired via lateral gene transfer (LGT) from a firmicute bacterium, identical to what occurred in *R. salina*. On the other hand, the *psb*N gene was independently encoded by the plastid genome without a reverse transcriptase gene as an intron. To clarify the phylogenetic relationships of the algae with red-algal derived plastids, phylogenetic analyses of 32 taxa were performed, including three previously sequenced cryptophyte plastid genomes containing 93 protein-coding genes. The stramenopiles were found to have branched out from the Chromista taxa (cryptophytes, haptophytes, and stramenopiles), while the cryptophytes and haptophytes were consistently grouped into sister relationships with high resolution.

Introduction

The genus *Teleaulax* belongs to the class Cryptophyceae and is an ecologically important component of marine phytoplankton communities [1–7]. In addition to its importance in the trophic web, *Teleaulax* is significant in the context of kleptoplasty due to its complex relationship with the grazer ciliate *Mesodinium rubrum* and its indirect relationships with dinoflagellates, such as *Dinophysis* and *Amylax* [8–12].

Since Charles Darwin described an extensive reddish discoloration in the sea south of Valparaiso near the southern boundary of the Peru coastal current [13], the marine ciliate *Mesodinium rubrum* has been identified worldwide as a bloom-forming ciliate [12, 14–21]. *Mesodinium rubrum* contains functional chloroplasts derived from cryptophycean species of the genera *Teleaulax* and *Geminigera* [22–26]. In addition, phototrophic dinoflagellate species of the genera *Dinophysis* and *Amylax* also contain plastids of cryptophyte origin, particularly of *Teleaulax amphioxeia* and *Geminigera cryophila* origin [9, 27–34]. Cryptophyte plastids were acquired by the marine ciliate *M. rubrum* and then transferred to dinoflagellates via feeding. Recently, Kim et al. [9] have revealed that *Dinophysis caudata* feeds on the mixotrophic ciliate *M. rubrum* and secondarily retains the plastids of *T. amphioxeia*, transforming them into stellate compound chloroplasts. Kleptoplasty is a well-known phenomenon occurring in dinoflagellates [27, 29, 30], ciliates [24], and sacoglossan sea slugs [35, 36], which sequester chloroplasts from algal food resources. However, the longevity of a photosynthetically functional plastid in predator cytoplasm depends on the kleptoplast genome itself [37] and not on the lateral transfer of algal nuclear genes to the predator [38, 39]. Therefore, plastid genome analysis of *T. amphioxeia* may be important to better understand the kleptoplastic relationship between *T. amphioxeia* and *M. rubrum*.

Cryptophytes have unique secondary plastids and possess four genomes (host nuclear, mitochondrial, plastid, and nucleomorph genomes). Cryptophyte plastids display four envelope membranes, with a eukaryotic compartment between the outer and inner membrane pairs. Plastid-containing cryptophytes are important models for plastid evolution through secondary endosymbiosis between phagotrophic and photoautotrophic eukaryotes [40, 41], a process that has presumably also occurred in several other protist lineages [42, 43]. Nucleomorphs derived from red algal nuclei persist in the remnant cytosol of engulfed algal cells, between the inner and outer pairs of plastid membranes [44–46]. To date, nine cryptomonad organelle genomes, including 3 plastid, 2 mitochondrial, 3 nucleomorph and 1 nuclear genome, have been sequenced. For example, the plastid [47] and mitochondrial genomes [48] of *Rhodomonas salina* have been completely sequenced, and the nucleomorph and mitochondrial genomes of *Hemiselmis andersenii* have also been published [49, 50]. In addition, the nuclear [46], nucleomorph [51] and plastid genomes [52] of the model cryptomonad species *Guillardia theta* have been sequenced, as well as the nucleomorph [53] and plastid genomes [54] of the nonphotosynthetic cryptomonad *Cryptomonas paramecium*. More recently, the *Chroomonas mesostigmatica* nucleomorph genome was also sequenced [55].

Here, we present the complete plastid genome of *Teleaulax amphioxeia* together with analyses of its genome structure and gene content. This plastid genome sequence is the first to be reported with the full characterization of the plastid genes in the genus *Teleaulax*. Comparative analysis was conducted using the genome of *T. amphioxeia* and three published plastid genomes of the cryptophytes *Cryptomonas paramecium*, *Guillardia theta*, and *Rhodomonas salina*. To identify the taxonomic relationships and evolutionary history of algae with red-algal derived plastid, we reconstructed plastid phylogenies based on 93 protein-coding genes from the currently available genomic data, using 28 plastid genomes, including the genomes of 4 cryptophytes, 4 haptophytes, 12 stramenopiles, and 9 red algal species. Additionally, we demonstrate herein the conserved properties and variability of the plastid genomes among the cryptophyte lineages. Our genetic information and plastid genome comparisons among the cryptophytes provide important insights into both the evolution of organelle genomes and the harmful algae-associated trophic web in marine ecosystems.

Materials and Methods

DNA isolation and sequencing

A culture derived from a single-cell isolate of *Teleaulax amphioxeia* collected from Gomso Bay, Korea (35° 40' N, 126° 40' E), which was established in a previous study [56], was selected for genome sequencing. DNA was extracted from the cultivated sample using a QIAGEN DNEasy Blood Mini Kit (QIAGEN, Valencia, CA, USA), following the manufacturer's instructions. A sequencing library was prepared using an Ion Xpress Plus gDNA Fragment Library Preparation Kit and an Ion OneTouch 200 Template Kit v2 DL (Life Technologies, San Francisco, CA, USA) according to the manufacturer's protocol and sequenced with an Ion Torrent Personal Genome Machine (PGM) at the Yoon laboratory at Sungkyunkwan University (Suwon, Korea) using an Ion PGM Sequencing 200 Kit v2 (Life Technologies, San Francisco, CA, USA).

Genome assembly and plastid contig selection

The data were trimmed (i.e., base = 80 bp, error threshold = 0.05, n ambiguities = 2) using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark) prior to producing a *de novo* assembly with the default options (automatic bubble size, minimum contig length = 1,000 bp). The raw reads were then mapped to the assembled contigs (similarity = 95%, length fraction = 75%), and regions with no evidence of short-read data were removed (up to 1,000 bp). The resulting assembly had an average coverage of ~15x and included one large contig of 191,270 bp that was determined to be the plastid genome based on the following criteria: (1) BLAST searches of commonly known plastid genes against the entire assembly produced hits on this contig with significant e-values (e#10220); and (2) the genome size of 129,772 bp is consistent with the sizes of other photosynthetic cryptophyte plastid genomes, which range from 121 (*Guillardia theta* NC000926) to 136 kbp (*Rhodomonas salina* NC009573). Each contig that contained a plastid sequence was then manually aligned with Genetic Data Environment (MacGDE2.5) program [57], and a consensus sequence was produced. The assemblies were further verified by examining the paired-end distance and depth after re-mapping the reads to the assembled sequence.

Genome annotation and sequence analysis

Databases of protein-coding genes and rRNA and tRNA genes were compiled from all previously sequenced cryptophyte plastid genomes. Preliminary annotations for the protein-coding genes were performed using GeneMarkS (http://opal.biology.gatech.edu/genemarks.cgi) to identify coding sequences (CDSs) and generate a basic option. The final annotation file was evaluated with Geneious Pro 5.1.7 (http://www.geneious.com/) using ORF Finder with the standard genetic code. After the alignments for each gene were completed, they were checked manually, and the corresponding open reading frames (ORFs) in the genome sequences were annotated. Annotations of ORFs with putative functional domains were included in the genome.

To identify the tRNA sequences, the plastid genome was submitted to tRNAscan-SE version 1.2.1 server (http://lowelab.ucsc.edu/tRNAscan-SE/). The genome was searched with the default settings, using the "Mito/Chloroplast" model. To determine the rRNA sequences, a set of known plastid rRNA sequences was extracted from the plastid genome of *Rhodomonas salina* and used as query sequences to search the *R. salina* genome using BLASTn. The annotated sequence was deposited into NCBI GenBank database as KP899713. The graphical gene map was designed with OrganellarGenomeDRAW program (http://ogdraw.mpimp-golm.mpg.de/).

Gene arrangement comparisons

The three published cryptophyte plastid genomes with annotations were downloaded from GenBank [48, 52, 54]. For the structure and arrangement comparisons, the genomes were aligned using Mauve Genome Alignment version 2.2.0 [58] with the default settings. To aid in visualization, we designated the beginning of the *trn*Y and *rpl*19 markers as position 1 in each genome.

Phylogenetic analysis

Phylogenetic analysis was conducted to determine the taxonomic relationships and evolution of the red algal-derived plastid. A dataset was created, combining 93 homologous protein-coding gene sequences in 29 plastid genomes from red algal plastid-bearing organisms, including 4 cryptophytes, 4 haptophytes, 12 stramenopiles and 9 red algae (S1 Table). The sequences of 2 chlorophytes and 1 glaucophyte were used as outgroup taxa to root the tree. The dataset was concatenated into 18,181 amino acids and a single continuous sequence of 56,565 nucleotides to initiate alignment by eye using MacGDE2.5 program (S2 Table).

Maximum likelihood (ML) phylogenetic analyses were performed using RAxML version 8.0.0 [59] with the Le and Gascuel with gamma (LG+GAMMA) model [60] for the amino acid data selected by ProtTest 3 [61] and with the general time-reversible plus gamma (GTR+-GAMMA) model for the nucleotide data [62]. We used 1,000 independent tree inferences and identified the best tree with the-# option within the program. The gamma correction values and the proportion of invariable sites in the combined dataset were obtained automatically by the program (S3 Table). Bootstrap values (MLBS) were calculated using 1,000 replicates with the same substitution model.

Maximum parsimony (MP) and distance (neighbor-joining; NJ) trees were constructed from a combined dataset with PAUP* using a heuristic search algorithm with the following settings: 100 random sequence-addition replicates, tree bisection and reconnection (TBR) branch swapping, MulTrees, all characters unordered and unweighted, and branches with a maximum length of zero collapsed to yield polytomies. The MLBS for the resulting nodes were assessed using bootstrapping with 1,000 iterations on each tree. For NJ analyses, we analyzed the dataset using the Modeltest parameters (S3 Table).

The *dna*X protein encoded by the *T. amphioxeia* plastid was used as a query to identify and retrieve a diverse set of *dna*X, polymerase III gamma/tau and replication factor C proteins from public protein databases. The BLAST search was resulted the conserved protein domain models Cog2812 and TIGR02397. We selected 220 unambiguously aligned amino acid sequences from 90 homologous taxa. Sequences were aligned using MacGDE 2.5 and analyzed using RAxML as described above.

Results and Discussion

The plastid genome of *Teleaulax amphioxeia*

The plastid genome of *Teleaulax amphioxeia* was found to be 129,772 bp in size and is illustrated in Fig 1. The *T. amphioxeia* genome size is similar to those of *Guillardia theta* and *Rhodomonas salina*. Eighty percent of the *T. amphioxeia* plastid genome was predicted to consist of coding regions (Table 1), including structural RNA genes, similar to the percentages of coding regions in *G. theta* (87.7%), *Cryptomonas paramecium* (87.0%) and *R. salina* (80.8%). The proportion of intergenic space in *T. amphioxeia* was 15.5%, which is comparable to those of algae with red-algal derived plastid and other red algal plastid descendants (i.e., haptophytes and stramenopiles). The G+C content was 34.21% for *T. amphioxeia*, which is similar to those of *C.*
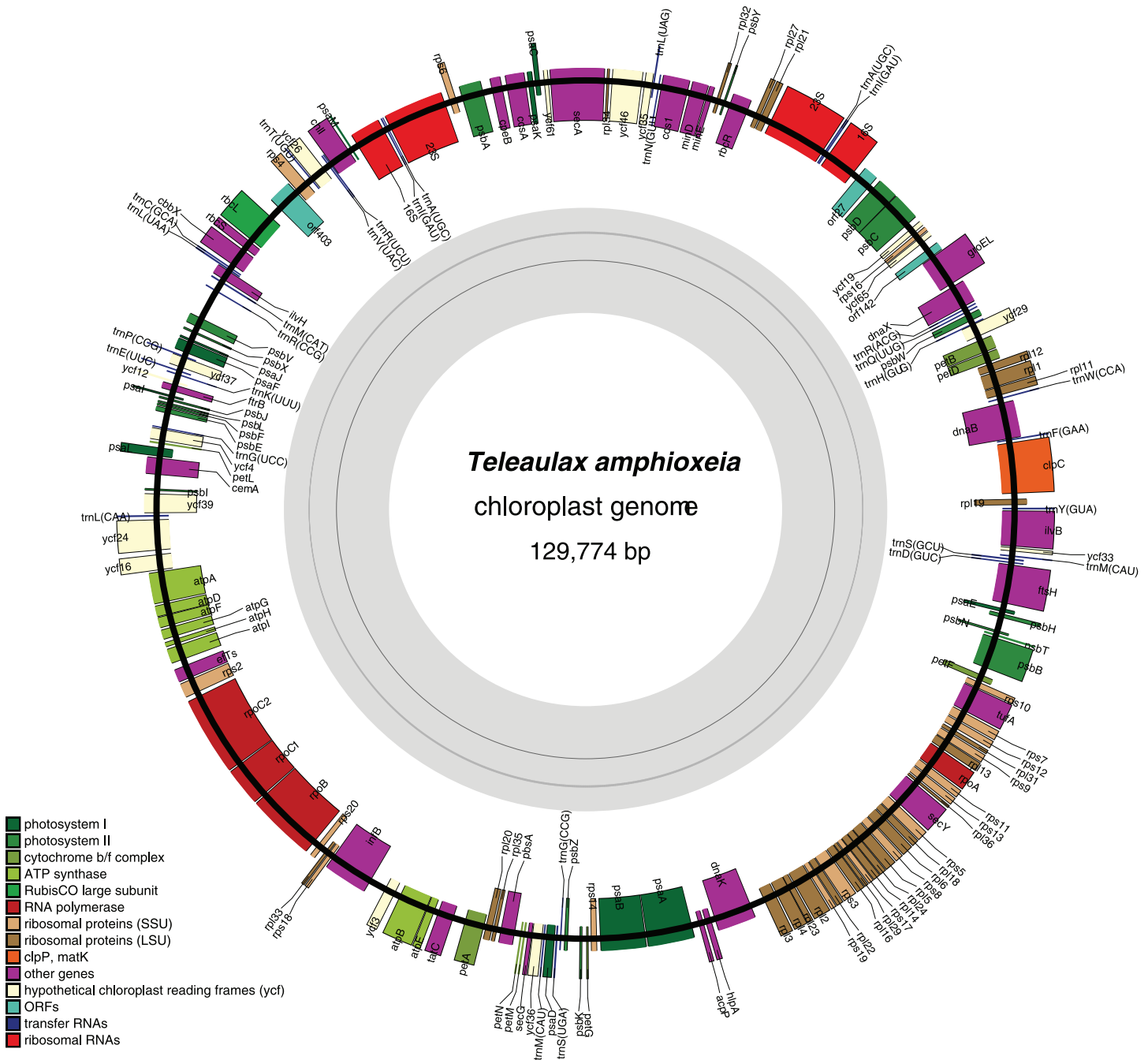
Fig 1. Circular map of the plastid genome of the cryptophyte *Teleaulax amphioxeia*. All of the genes are transcribed in a clockwise direction. Note the dense gene arrangement and the single large intergenic region. The protein-coding genes and ribosomal RNA and transfer RNA genes are labeled inside or outside of the circle. The genes are color coded according to the functional categories listed in the index below the map.

paramecium (38%), *R. salina* (34%), and *G. theta* (32%). The overall G+C content was highly similar to those of other chromists and red algae [52, 63–65].

The *T. amphioxeia* plastid genome was predicted to encode 179 genes, including 2 rRNA operons and 30 tRNA genes (Fig 1 and S4 Table). The ochre termination codon TAA was determined to be used in *T. amphioxeia* 86.7% of the time, and the amber (TAG) and opal (TGA)

Table 1. Characteristics of the cryptophyte plastid genome analyzed in this study.

	Cryptomonas paramecium CCAP977/2a	*Rhodomonas salina* CCMP1319	*Guillardia theta*	*Teleaulax amphioxeia* HACCP CR01
General characteristics				
Size (bp)	77,717	135,854	121,524	129,772
G+C(%)	38%	34%	32%	34.21%
Number of intergenic spaces (%)	9,917 (12.8%)	23,767 (17.5%)	14,792 (12.2%)	20,108 (15.5%)
Number of total genes (including RNAs)	114	183	183	179
Protein-coding genes				
Number of protein-coding genes	82 (92.6%)	146 (80.8%)	147 (87.7%)	143 (80.0%)
Number of ribosomal proteins	42	44	44	44
Start codon usage (no./%)				
ATG	80 (97.6%)	136 (93.2%)	138 (93.9%)	135 (94.4%)
GTG	2 (2.4%)	10 (6.8%)	7 (4.7%)	7 (4.9%)
TTG	-	-	2 (1.4%)	1 (0.7%)
Stop codon usage (no./%)				
TAA	58 (70.7%)	116 (79%)	115 (78.2%)	124 (86.7%)
TAG	12 (14.65%)	26 (18%)	19 (12.9%)	15 (10.5%)
TGA	12 (14.65)	4 (3%)	13 (8.8%)	4 (2.8%)
RNAs				
Number of tRNAs	29	31	30	30
Number of rRNA operons	1	2	2	2
GenBank accession	NC013703	NC009573	NC000926	KP899713
Pseudogene	*atp*F			
Missing	*pet*, *psa*, *psb*			

doi:10.1371/journal.pone.0129284.t001

codons were found to be used 10.5% and 2.8% of the time, respectively. Seven genes contained a valine (GTG) rather than a methionine start codon (*chl*I, *hlp*A, *rpl*23, *rps*3, *rps*8, *rps*13, and *psb*C), and a TTG start codon was present in one gene (*ycf*65). A set of 136 protein-coding genes was shared by all of the plastid genomes evaluated in this study, while 170 genes were unique amongst the three photosynthetic cryptophyte species, except for *C. paramecium*, which is an osmotrophic, colorless species (S1 Table). An additional 107 genes, including 76 protein-coding genes were shared between *T. amphioxeia* and *C. paramecium*.

The four cryptomonads shared a similar tRNA gene set, with 30 tRNAs in *T. amphioxeia*, 29 tRNAs in *C. paramecium*, 30 tRNAs in *G. theta*, and 31 tRNAs in *R. salina* (Table 2), and their tRNAs included redundant isotypes for the amino acids 2 glycine, 2–3 serine, 3 arginine, and 3 leucine and three distinct methionine tRNAs.

Similar to the two photosynthetic cryptophytes, *T. amphioxeia* contained two small (approximately 4.8 kb) and almost identical repeats of rRNA operons encoding 16S, 23S and 5S rRNAs and the two tRNA genes *trn*A (GAT) and *trn*I (TGC). Inverted repeats (IRs) consisting of rRNA operons (and in some cases, a few additional genes) were generally found in the plastid genomes, which represents an ancestral feature [66–68]. These repeats are present in the *G. theta* and *R. salina* genomes [47, 52], as well as in the haptophyte *Emiliania huxleyi* [65] and several diatoms [69]. *C. paramecium* lacks this IR arrangement, and it contains only one rRNA operon in a 16S-*trn*I-*trn*A-23S-5S configuration [54].

Four instances of overlapping genes were found in *T. amphioxeia*, many of which were also identified in the other chromist plastid genomes. The *psb*D-*psb*C overlap found in *T. amphioxeia*

Table 2. tRNA sequences present in the cryptophyte plastid genome.

	trnA (GGC)	trnA (TGC)	trnC (GCA)	trnD (GTC)	trnE (TTC)	trnF (GAA)	trnG (GCC)	trcG (TCC)	trnH (GTG)	trnI (GAT)	trnK (TTT)	trnL (CAA)	trnL (GAG)	trnL (TAA)	trnL (TAG)	trnfM (CAT)	trnM (CAT)	trnN (GTT)	tmP (TGG)	trnQ (TTG)	tmR (ACG)	trnR (CCG)	trnR (CCT)	trnR (TCT)	trnS (CGA)	tmS (GCT)	tmS (GGA)	trnS (TGA)	trnT (GGT)	tmT (TGT)	trnV (GAC)	trnV (TAC)	trnW (CCA)	trnY (GTA)	Total
Cryptomonas paramecium	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		3	1	1	1	1	1	1	1		1	1	1		1		1	1	1	29
Rhodomonas salina	2		1	1	1	1	1	1	1	2	1	1		1	1		3	1	1	1	1	1		2		1	1	1			1	1	1	1	31
Guillardia theta	2		1	1	1	1	1	1	1	3	1	1		1		1	2	1	1	1	1	1		1		1	1	1	1		1	1	1	1	30
Teleaulax amphioxeia	2		1	1	1	1	1	1	1	2	1	1		1			3	1	1	1	1	1		1		1	1	1	1		1	1	1	1	30

existed in all of the sequenced chromist genomes, although the amount of overlap varied. Overlaps involving *atp*D-*atp*F and *rpl*4-*rpl*23 were common in the stramenopiles and cryptophytes, but not in the haptophytes. Single-nucleotide overlaps between *rpl*16-*rpl*29 and *orf*142-*orf*146 were present in *T. amphioxeia*, similar to what was previously found in *R. salina* and *G. theta*.

Gene content and synteny

The *T. amphioxeia* plastid genome was found to contain 143 predicted protein-coding genes (Table 3 and Fig 1). Overall, this genome shows a high degree of syntenic conservation with that

Table 3. List of genes in the *Teleaulax amphioxeia* plastid genome (143 total).

Classification	Genes							
Genetic systems								
Maintenance	*dna*B	*rne*	**dnaX***	*min*D	*min*E	*hlp*A		
RNA polymerase	*rpo*A	*rpo*B	*rpo*C1	*rpo*C2				
Transcription factors	*rbc*R	*ycf*29						
Translation	*inf*B	*tsf*	*tuf*A					
Protein quality control	*clp*C	*dna*K	*fts*H	*gro*EL				
Transport								
Transport	*cem*A	*sec*A	*sec*G	*sec*Y	*suf*B (*ycf*16)	*suf*C (*ycf*24)	**tatC**	
ATP synthesis								
ATP synthesis	*atp*A	*atp*B	*atp*D	*atp*E	*atp*F	*atp*G	*atp*H	*atp*I
Ribosomal proteins								
Large subunit	*rpl*1	*rpl*2	*rpl*3	*rpl*4	*rpl*5	*rpl*6	*rpl*11	*rpl*12
	*rpl*13	*rpl*14	*rpl*16	*rpl*18	*rpl*19	*rpl*20	*rpl*21	*rpl*22
	*rpl*23	*rpl*24	*rpl*27	*rpl*29	*rpl*31	*rpl*32	*rpl*33	*rpl*34
	*rpl*35	*rpl*36						
Small subunit	*rps*2	*rps*3	*rps*4	*rps*5	*rps*6	*rps*7	*rps*8	*rps*9
	*rps*10	*rps*11	*rps*12	*rps*13	*rps*14	*rps*16	*rps*17	*rps*18
	*rps*19	*rps*20						
Metabolism								
Carbohydrates	*rbc*L	*rbc*S						
Lipids	*acp*P							
Nucleotides	-							
Amino acids	*ilv*B	*ilv*H						
Cofactors	*chl*I							
Photosystems								
Phycobilisomes	*cpe*B							
Photosystem I	*psa*A	*psa*B	*psa*C	*psa*D	*psa*E	*psa*F	*psa*I	*psa*J
	*psa*K	*psa*L	*psa*M	*ycf*3	*ycf*4			
Photosystem II	*psb*A	*psb*B	*psb*C	*psb*D	*psb*E	*psb*F	*psb*H	*psb*I
	*psb*J	*psb*K	*psb*L	**psbN***	*psb*T	*psb*V	*psb*W	*psb*X
	*psb*Y	*psb*Z	*ycf*12					
Cytochrome complex	*ccs*1	*ccs*A	*pet*A	*pet*B	*pet*D	*pet*F	*pet*G	*pet*L
	*pet*M	*pet*N						
Redox System	*ftr*B	*pbs*A						
Unknown								
Conserved ORFs	*orf*27	*orf*142	*orf*146	*orf*403	*ycf*19	*ycf*26	*ycf*33	*ycf*35
	*ycf*36	*ycf*37	*ycf*39	*ycf*46	*ycf*65			

doi:10.1371/journal.pone.0129284.t003

of *R. salina* [47]. The gene order was generally well conserved among the four cryptomonad plastid genomes. Large tracts of complete gene order conservation were observed, such as the highly conserved and co-expressed ribosomal protein genes and the *atp* gene cluster (Fig 1).

Many proteins associated with cell and organelle division were found to be encoded within the *T. amphioxeia* plastid genome (Table 3) by genes including *hlp*A (encoding a chromatin-associated architectural protein), *dna*B (encoding a DNA helicase), and *min*D and *min*E (encoding proteins that prevent the creation of DNA-less "minicells" during division [70]). Specifically, *hlp*A, which encodes a histone-like protein, has been previously identified in the two photosynthetic cryptomonads. This gene is also present in the genomes of the red algae *Cyanidioschyzon merolae* [64] and *Galdieria sulphuraria* (NC024665), but it is absent from the haptophyte and stramenopile plastid genomes. The Apicomplexan *hlp*A gene was present in the nuclear genome [71–73]. Additional chaperone protein-encoding genes, such as *gro*EL and *dna*K (a member of the *hsp*70 family), were present in the *T. amphioxeia* plastid genome, and their products presumably participate in protein import and folding [74]. The *fts*H gene, which encodes a protease responsible for the removal of damaged D1 protein from the photosystem II (PSII) complex [37], was identified in the photosynthetic species *T. amphioxeia*, *R. salina* and *G. theta* but not in *C. paramecium*.

The components of a protein translocation system, the *sec* transport system, were maintained (*sec*A, *sec*G, and *sec*Y), and the *suf*B and *suf*C genes were also present in the *T. amphioxeia* plastid genome, the products of which play roles in iron-sulfur cluster assembly [75]. The identification of the *chl*I gene in the *T. amphioxeia* plastid genome may provide additional insights into the role of this magnesium chelatase component in plastid-to-nucleus signaling [76]. The gene encoding the *sec*-independent transport protein *tat*C was also present, as was the proteolytic degradation pathway gene *clp*C. The plastid of *T. amphioxeia* thus appears to have retained the ability to import necessary proteins from the cytoplasm (e.g., proteins linked to cell division) and can mediate their degradation.

The *T. amphioxeia* plastid genome possesses a nearly full complement of the 8 *atp* synthase subunit genes found in the other cryptomonads. These genes showed varying degrees of sequence conservation in the four cryptomonads. A total of 24 *rpl* genes encoding the 50S ribosomal subunit protein and 18 *rps* genes encoding the 30S ribosomal subunit protein were also found in *T. amphioxeia*. Additional genes (including *cpe*B, *ilv*B, *ilv*H, and *inf*B) were present in the plastid genome of the cryptophyte but were absent from those of the stramenopiles and/or haptophytes. The three pseudogenes (*chl*B, *chl*N, and *chl*L) encoding light-independent protochlorophyllide reductase, which is involved in the light-independent synthesis of chlorophyll [77], were identified in the *R. salina* plastid genome but not in the *T. amphioxeia* or *G. theta* genome (Table 3). The reverse transcriptase gene, which is present as an intron within the photosystem gene *psb*N, was identified in *R. salina* [47]; however, the *psb*N gene in the *T. amphioxeia* plastid genome lacked the reverse transcriptase gene.

Photosynthetic genes

The gene encoding the β subunit of phycoerythrin (*cpe*B), which is part of the phycobiliprotein complex in cryptomonads, was present in the plastid genomes of *T. amphioxeia*, *R. salina* and *G. theta* [47, 52] but was missing from that of *C. paramecium* [54]. The photosynthetic regulator and electron transfer gene *ftr*B was present in all of the photosynthetic cryptophyte plastids. The *rbc*L and *rbc*S genes encoding the large and small subunits of ribulose 1, 5-bisphosphate carboxylase/oxygenase were present in all of the cryptophyte plastids, including *C. paramecium*.

The *psa* and *psb* gene families encode the protein subunits of photosystem I (PSI) and PSII, respectively. A total of 11 *psa* and 18 *psb* genes were present in the plastid genomes of the three

photosynthetic cryptomonads [47, 52]. The loss of the *psa* and *psb* genes from the plastid genome of *C. paramecium* accounts for approximately 7.5 kbp of missing plastid DNA [54].

The products of the photosynthetic *pet* gene family form a complex required for oxygenic photosynthesis, particularly for noncyclic electron flow mediated by the cytochrome b6f complex [78]. The 8 *pet* genes were present in the plastid genomes of *T. amphioxeia*, *G. theta*, and *R. salina*; however, they were missing from that of in *C. paramecium*, with the curious exception of *pet*F. In other organisms that have secondarily lost their photosynthetic abilities (e.g., *Euglena longa* and *Aneura mirabilis*), the *pet* genes are either missing or have become pseudogenes [79, 80].

Lateral gene transfer

The most unexpected finding in the *T. amphioxeia* and *R. salina* plastid genomes was the presence of a gene with strong similarity to *dna*X, which encodes the tau/gamma components of bacterial DNA polymerase [47, 81, 82]. Phylogenetic analysis revealed that the *R. salina* and *T. amphioxeia dna*X genes were acquired by lateral gene transfer (LGT), with the direct transfer of *T. amphioxeia dna*X from a cryptophyte alga. The *dna*X proteins of *R. salina* and *T. amphioxeia* were found to be derived from firmicutes, i.e., parasitic mycoplasmas and related organisms [47] (S1 Fig).

Plastid genome rearrangements

Mauve pairwise alignments of the *T. amphioxeia* genome with each of the other four plastid genomes used in this study are shown in Fig 2. All three of the photosynthetic cryptophytes were found to have highly conserved gene arrangements and contents. All of the cryptophyte plastid genes were located in gene clusters that could be readily reconstructed from the *C. paramecium* genome via a small number of inversion events (Fig 2). The three photosynthetic cryptophyte plastid genomes were co-linear. *C. paramecium* had the smallest rearrangement distance, and almost all photosynthetic genes were found to be lost compared with the other plastid genomes; furthermore, it differed from the photosynthetic cryptophyte by only three inversions, suggesting that most of its photosynthetic genes were lost after it acquired phototrophy.

Phylogeny

Phylogenomic analysis was completed using a concatenate of 93 proteins encoded by 29 complete plastid genomes from algae with red-algal derived plastid (including 4 cryptophyte, 4 haptophyte, 12 stramenopile, and 9 rhodophyte genomes) using 3 outgroups (2 chlorophyte and 1 glaucophyte genome). The sequences of chlorophyll c-containing dinoflagellates were not included due to the limited sizes of their plastid genomes. RAxML trees based on 18,180 amino acids (Fig 3A) and 56,569 nucleotides (Fig 3C) differed at the rhodophyte lineages. The cryptophyte clade was strongly supported as a monophyletic clade, which is congruent with gene synteny (Fig 3B). The resulting phylogeny suggested that the cryptophytes had sister relationships with the haptophytes in both trees. In the amino acid-based tree (Fig 3A), the Cyanidiophyceae clade was located at the base of the red-algal derived lineage, but all red algal species in the Florideophyceae and Bangiophyceae clades branched out as sisters of the Cryptophyta/Haptophyta lineage, with the exclusion of the Cyanidiophyceae clade [47, 54, 83]. In contrast with the protein ML tree, a tree based on a combined dataset of 93 protein-coding gene sequences (Fig 3C) showed that the Rhodophyta was a monophyletic clade sistered with taxa of hacrobian lineages.

Previous phylogenetic studies have demonstrated that cryptophyte, stramenopile, and haptophyte plastids are derived from a red alga [66, 84, 85]. Among the three groups
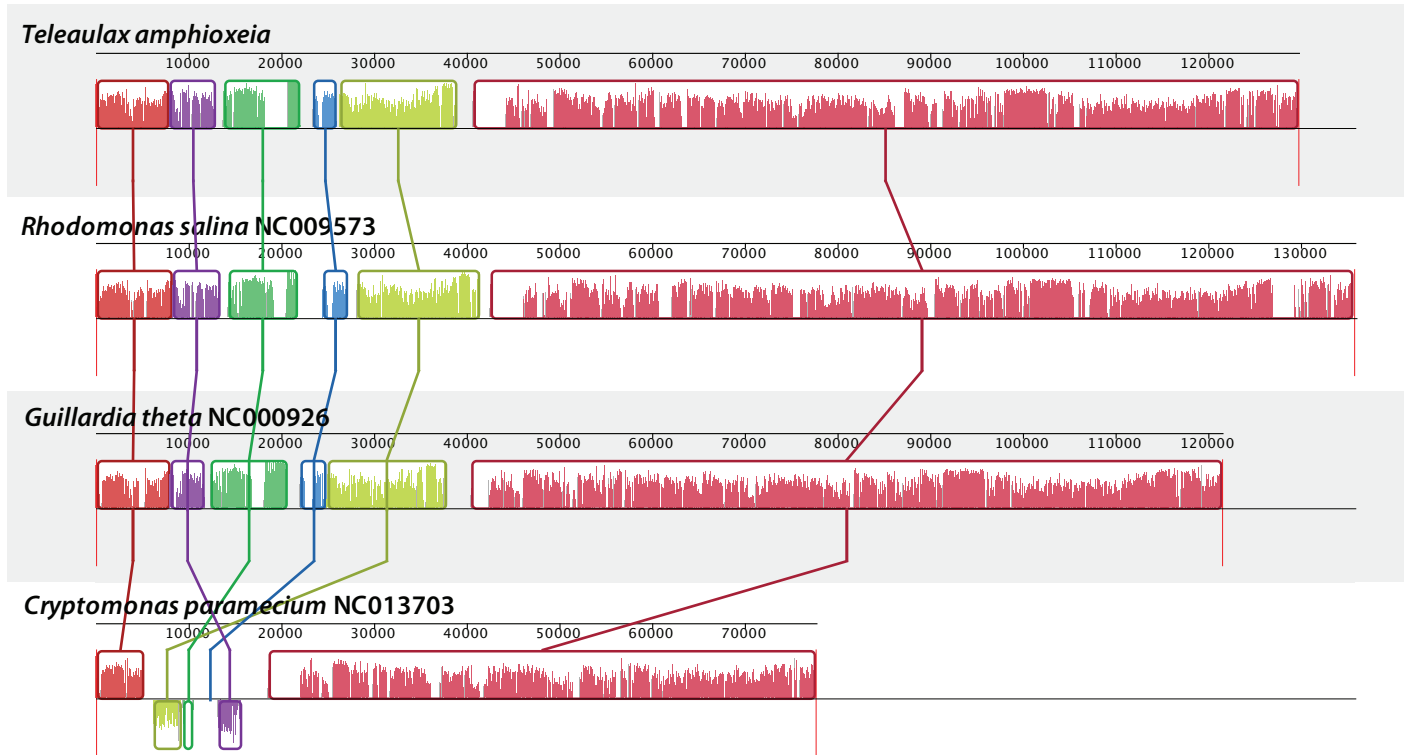
Fig 2. Overview of the red algal plastid genomes. Linearized maps of the complete *T. amphioxeia* plastid genome is compared with those of other cryptophytes. Color-coded syntenic blocks are shown above each genome, and gene maps are shown below each genome. The syntenic blocks above the horizontal line are on the same strand, and those below the line are on the opposite strand. The horizontal bars inside of the syntenic blocks indicate sequence conservation. The block boundaries correspond to sites at which inversion events occurred. On the gene maps, the genes above the horizontal line are transcribed from left to right, and those below the horizontal line are transcribed from right to left. The rRNA operons are shown in red.

(cryptophytes, haptophytes, and stramenopiles), the stramenopiles were branched outside of a cluster of chromist taxa, while the cryptophytes and haptophytes were consistently branched together as the closest relatives. These results differ from those of other analyses of five plastid genes (16S rRNA, *psa*A, *psb*A, *rbc*L, and *tuf*A), which have indicated that the stramenopiles and haptophytes are grouped together [86, 87]. The common ancestry of hacrobian (cryptophyte and haptophyte) plastids is also strongly supported ([47, 44, 83] in this study), consistent with the LGT of *rpl*36 in the cryptophyte and haptophyte plastid genomes as evidence of the sisterhood of these 2 groups and the exclusion of stramenopiles [88]. The findings of nuclear gene analyses also support this interpretation [89, 90]. However, phylogenomic data reported by other studies strongly suggest that cryptophytes and haptophytes have separate origins [91, 92]. These analyses have indicated that the haptophytes are sisters of the SAR (Stramenopile, Alveolate, and Rhizaria) group and that the cryptophytes are grouped together with the katablepharids as a broken "hacrobiana" taxa [92]. According to recent model of serial plastid endosymbioses [93], the cryptophyte plastid is more closely related to the stramenopile plastid than the haptophyte plastid. However, our phylogenies suggest the grouping together of the cryptophytes and haptophytes with moderate to high bootstrap support.

Conclusions

We have determined the plastid genome sequence of the cryptophyte *T. amphioxeia*, which is the first plastid genome reported for the genus *Teleaulax*. As increasing numbers of genomes
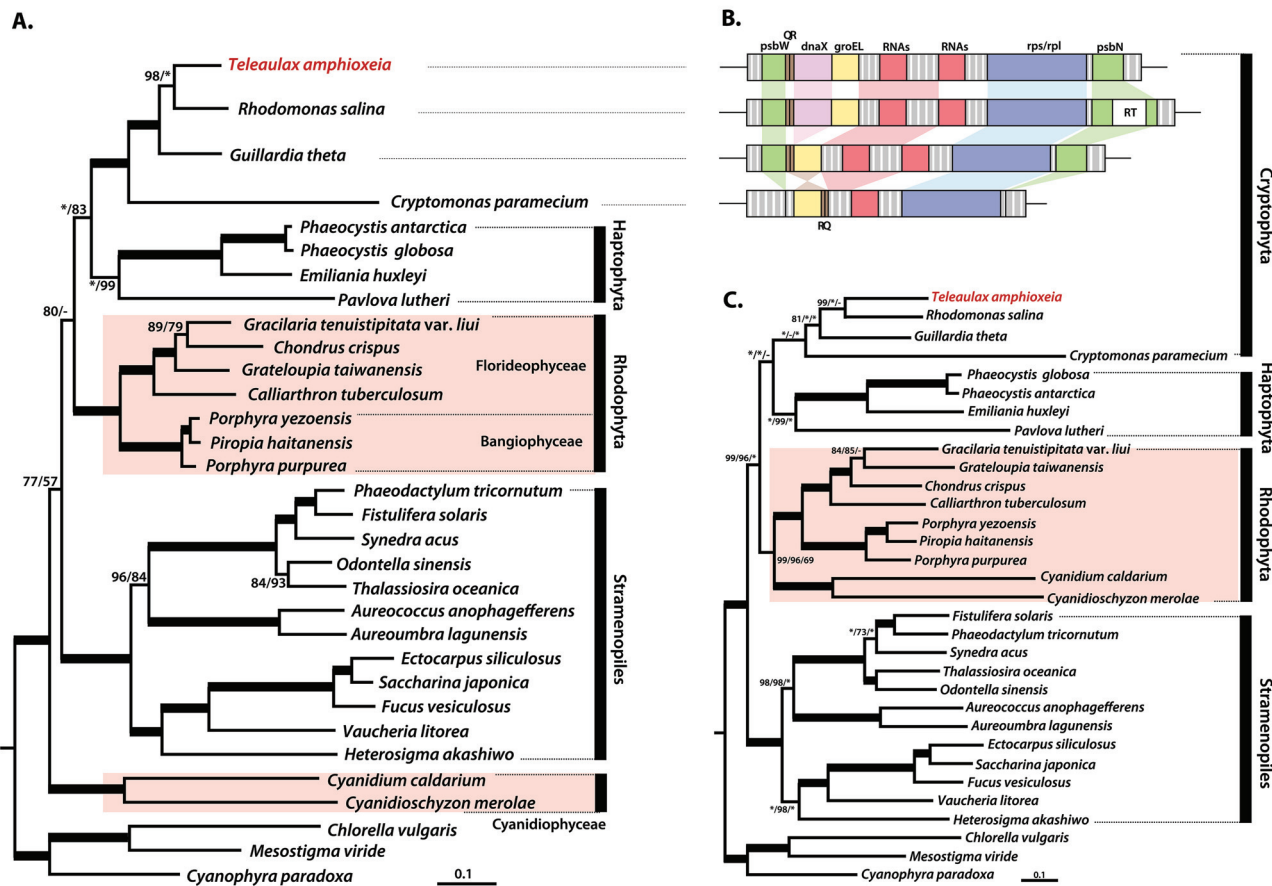
Fig 3. Phylogenetic tree of the cryptophyte plastids. (A) RAxML phylogeny constructed from a concatenate of 93 proteins (18,180 amino acids). The ML and MP bootstrap values are shown above or below the branches. (B) A simple overview of cryptophyte plastid gene synteny. (C) RAxML phylogeny constructed based on 93 protein-coding genes (56,569 nucleotides). The ML, MP and NJ bootstrap values are shown above or below the branches. The bold branches indicate strongly supported values (ML, MP, and NJ = 100%). The asterisks indicate supported values = 100%, and the dashes indicate values with < 50% support. The scale bar depicts the number of substitutions/site.

doi:10.1371/journal.pone.0129284.g003

are annotated and published, comparative genomic analyses of secondary plastids will provide new insights into the patterns and processes of endosymbiosis, particularly in lineages with red-algal derived plastids. The genes that are common to all cryptophyte plastids are likely essential for plastid function and represent a useful starting point for the future annotation of plastid genomes. Several previous studies focusing on cryptophyte plastids have shown the potential of plastid genome research for answering unresolved questions about the history of these lineages, increasing our understanding of the evolution of cryptophyte plastids. The addition of the *T. amphioxeia* plastid genome to the suite of complete plastid genome sequences increases the breadth of plastid genomes that have been sampled to date and will help to identify common trends in organellar genomes. Many studies have shown that the *Teleaulax* species donated its plastid to the ciliate *Mesodinium rubrum* and then to the dinoflagellates *Dinophysis caudata* and *Amylax triacantha* through the trophic web and that these species have retained the acquired plastid and produce water blooms in marine ecosystems. Our *T. amphioxeia*

plastid genome data will provide clues about the complicated plastid relationships between the donor cryptophyte *Teleaulax* and retainers, such as the ciliate *Mesodinium* and the dinoflagellates *Dinophysis* and *Amylax*.

Supporting Information

S1 Fig. The phylogenetic tree constructed from 90 *dna*X homologs. RAxML bootstrap values are shown above the branches with > 50% support. The scale bars indicate the number of substitutions/amino acid site.
(EPS)

S1 Table. Cryptophyte plastid genome alignments.
(XLSX)

S2 Table. Concatenated data set of protein-coding genes for constructing phylogenetic tree.
(XLSX)

S3 Table. Evolutionary models, log likelihood values (-lnL), and model parameters proposed by ProtTest3 for amino acids [61] and Modeltest 3.7 for nucleotides [93].
(DOCX)

S4 Table. The gene contents and arrangement of the *Teleaulax amphioxeia* plastid genome.
(DOCX)

Author Contributions

Conceived and designed the experiments: JIK WS. Performed the experiments: JIK. Analyzed the data: JIK GY. Contributed reagents/materials/analysis tools: HSY HSK WY. Wrote the paper: JIK WS.

References

1. Booth BC, Lewin J, Norris RE. Nanoplankton species predominant in the subarctic Pacific in May and June 1978. Deep Sea Res. 1982; 29: 185–200.

2. Hill DRA. A revised circumscription of *Cryptomonas* (Cryptophyceae) based on examination of Australian strains. Phycologia. 1991; 30: 170–188.

3. Hill D. *Teleaulax acuta* (butcher) Hill (Cryptophyceae). Baltic Sea phytoplankton identification sheet no. 12. Ann Bot Fenn. 1992; 29: 173–174.

4. Novarino G. Observations on some new and interesting Cryptophyceae. Nord J Bot. 1991; 11: 599–611.

5. Novarino G, Mills DK, Hannah F. Pelagic flagellate populations in the southern North Sea, 1988–89. I. Qualitative observations. J Plankton Res. 1997; 19: 1081–1109.

6. Barlow SB, Kugrens P. Cryptomonads from the Salton sea. Hydrobiologia. 2002; 473: 129–137.

7. Seoane S, Laza A, Urrutxurtu I, Orive E. Phytoplankton assemblages and their dominant pigments in the Nervion River estuary. Hydrobiologia. 2005; 549: 1–13.

8. Park M, Kim S, Kim H, Myung G, Kang Y, Yih W. First successful culture of the marine dinoflagellate *Dinophysis acuminata*. Aquat Microb Ecol. 2006; 45: 101–106.

9. Kim M, Kim S, Yih W, Park MG. The marine dinoflagellate genus *Dinophysis* can retain plastids of multiple algal origins at the same time. Harmful Algae. 2012; 13: 105–111.

10. Kim M, Kim KY, Nam SW, Shin W, Yih W, Park MG. The effect of starvation on plastid number and photosynthetic performance in the kleptoplastidic dinoflagellate *Amylax triacantha*. J Eukaryot Microbiol. 2014; 61: 354–363. doi: 10.1111/jeu.12115 PMID: 24734883

11. Nishitani G, Nagai S, Hayakawa S, Kosaka Y, Sakurada K, Kamiyama T, et al. Multiple plastids collected by the dinoflagellate *Dinophysis mitra* through kleptoplastidy. Appl Environ Microbiol. 2012; 78: 813–821. doi: 10.1128/AEM.06544-11 PMID: 22101051

12. Hansen PJ, Nielsen LT, Johnson M, Berge T, Flynn KJ. Acquired phototrophy in *Mesodinium* and *Dinophysis*—a review of cellular organization, prey selectivity, nutrient uptake and bioenergetics. Harmful Algae. 2013; 28: 126–139.

13. Hart TJ. Darwin and 'water-bloom'. Nature. 1943; 152: 661–662.

14. Hart TJ. Red water-bloom in South African seas. Nature. 1934; 134: 459–460.

15. Clemens WA. Red water-bloom in British Columbia waters. Nature. 1935; 135: 473.

16. Bary BM, Stuckey RG. An occurrence in Wellington Harbour of *Cyclotrichium meunieri* Powers, a ciliate causing red water, with some additions to its morphology. Trans Roy Soc N Z. 1950; 78: 86–92.

17. Ryther JH. Occurrence of red water off Peru. Nature. 1967; 214: 1318–1319.

18. Taylor FJR, Blackbourn DJ, Blackbourn J. Ultrastructure of the chloroplasts and associated structures within the marine ciliate *Mesodinium rubrum* (Lohmann). Nature. 1969; 223: 819–821.

19. Taylor FJR, Blackbourn DJ, Blackbourn J. Red-water ciliate *Mesodinium rubrum* and its incomplete symbionts—review including new ultrastructural observations. J Fish Res Board Can. 1971; 28: 391–407.

20. White AW, Sheath RG, Hellebust JA. A red tide caused by the marine ciliate *Mesodinium rubrum* in Passamaquoddy Bay, including pigment and ultrastructure studies of the endosymbiont. J Fish Res Board Can. 1977; 34: 413–416.

21. Lindholm T. *Mesodinium rubrum*—a unique photosynthetic ciliate. In: Jannasch HW, Williams PJ, editors. Advances in aquatic microbiology. New York: Academic Press; 1985. pp. 1–48.

22. Gustafson DE, Stoecker DK, Johnson MD, Van Heukelem WF, Sneider K. Cryptophyte algae are robbed of their organelles by the marine ciliate *Mesodinium rubrum*. Nature. 2000; 405: 1049–1052. PMID: 10890444

23. Johnson MD, Oldach D, Delwiche CF, Stoecker DK. Retention of transcriptionally active cryptophyte nuclei by the ciliate *Myrionecta rubra*. Nature. 2007; 445: 426–428. PMID: 17251979

24. Johnson MD. Acquired phototrophy in ciliates: a review of cellular interactions and structural adaptations. J Eukaryot Microbiol. 2011; 58: 185–195. doi: 10.1111/j.1550-7408.2011.00545.x PMID: 21518077

25. Moeller HV, Johnson MD, Falkowski PG. Photoacclimation in the phototrophic marine ciliate *Mesodinium rubrum* (Ciliophora). J Phycol. 2011; 47: 324–332.

26. Myung G, Kim HS, Park JS, Park MG, Yih W. Population growth and plastid type of *Myrionecta rubra* depend on the kinds of available cryptomonad prey. Harmful Algae. 2011; 10: 536–541.

27. Schnepf E, Elbrächter M. Cryptophycean-like double membrane-bound chloroplast in the dinoflagellate, *Dinophysis* Ehrenb.: evolutionary, phylogenetic and toxicological implications. Bot Acta. 1988; 101: 196–203.

28. Lucas IAN, Vesk M. The fine structure of two photosynthetic species of *Dinophysis* (Dinophysiales, Dinophyceae). J Phycol. 1990; 26: 345–357.

29. Hewes CD, Mitchell BG, Moisan TA, Vernet M, Reid FMH. The phycobilin signatures of chloroplasts from three dinoflagellate species: a microanalytical study of *Dinophysis caudate*, *D. Fortii*, and *D. Acuminata* (Dinophysiales, Dinophyceae). J Phycol. 1998; 34: 945–951.

30. Takishita K, Koike K, Maruyama T, Ogata T. Molecular evidence for plastid robbery (kleptoplastidy) in *Dinophysis*, a dinoflagellate causing diarrhetic shellfish poisoning. Protistologica. 2002; 153: 293–302.

31. Hackett JD, Maranda L, Yoon HS, Bhattacharya D. Phylogenetic evidence for the cryptophyte origin of the plastid of *Dinophysis* (Dinophysiales, Dinophyceae). J Phycol. 2003; 39: 440–448. PMID: 15117231

32. Janson S, Granéli E. Genetic analysis of the *psb*A gene from single cells indicates a cryptomonad origin of the plastid in *Dinophysis* (Dinophyceae). Phycologia. 2003; 42: 473–477.

33. Janson S. Molecular evidence that plastids in the toxin-producing dinoflagellate genus *Dinophysis* originate from the free-living cryptophyte *Teleaulax amphioxeia*. Environ Microbiol. 2004; 6: 1102–1106. PMID: 15344936

34. Park MG, Park JS, Kim M, Yih W. Plastid dynamics during survival of *Dinophysis caudata* without its ciliate prey. J Phycol. 2008; 44: 1154–1168.

35. Cruz S, Calado R, Serôdio J, Cartaxana P. Crawling leaves: photosynthesis in sacoglossan sea slugs. J Exp Bot. 2013; 64: 3999–4009. doi: 10.1093/jxb/ert197 PMID: 23846876

36. de Vries J, Woehle C, Christa G, Wägele H, Tielens AG, Jahns P, et al. Comparison of sister species identifies factors underpinning plastid compatibility in green sea slugs. Proc Biol Sci. 2015; 282: 20142519. doi: 10.1098/rspb.2014.2519 PMID: 25652835

37. de Vries J, Habicht J, Woehle C, Huang C, Christa G, Wägele H, et al. Is *fts*H the key to plastid longevity in sacoglossan slugs? Genome Biol Evol. 2013; 5: 2540–2548. doi: 10.1093/gbe/evt205 PMID: 24336424

38. Wägele H, Deusch O, Händeler K, Martin R, Schmitt V, Christa G, et al. Transcriptomic evidence that longevity of acquired plastids in the photosynthetic slugs *Elysia timida* and *Plakobranchus ocellatus* does not entail lateral transfer of algal nuclear genes. Mol Biol Evol. 2011; 28: 699–706. doi: 10.1093/molbev/msq239 PMID: 20829345

39. Bhattacharya D, Pelletreau KN, Price DC, Sarver KE, Rumpho ME. Genome analysis of *Elysia chlorotica* egg DNA provides no evidence for horizontal gene transfer into the germ line of this kleptoplastic mollusc. Mol Biol Evol. 2013; 30: 1843–1852. doi: 10.1093/molbev/mst084 PMID: 23645554

40. Douglas SE, Murphy CA, Spencer DF, Gray MW. Cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes. Nature. 1991; 350: 148–151. PMID: 2005963

41. McFadden GI. Second-hand chloroplasts: evolution of cryptomonad algae. In: Callow JA, editor. Advances in botanical research. London: Academic Press Limited; 1993. pp. 189–230.

42. Bhattacharya D, Medlin L. The phylogeny of plastids: a review based on comparisons of small-subunit ribosomal RNA coding regions. J Phycol. 1995; 31: 489–498.

43. Delwiche CF, Palmer JD. Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. Mol Biol Evol. 1996; 13: 873–882. PMID: 8754222

44. Gilson PR, Maier UG, McFadden GI. Size isn't everything: lessons in genetic miniaturisation from nucleomorphs. Curr Opin Genet Dev. 1997; 7: 800–806. PMID: 9468790

45. Archibald JM. Nucleomorph genomes: structure, function, origin and evolution. BioEssays. 2007; 29: 392–402. PMID: 17373660

46. Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, et al. Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature. 2012; 492: 59–65. doi: 10.1038/nature11681 PMID: 23201678

47. Khan H, Parks N, Kozera C, Curtis BA, Parsons BJ, Bowman S, et al. Plastid genome sequence of the cryptophyte alga *Rhodomonas salina* CCMP1319: lateral transfer of putative DNA replication machinery and a test of chromist plastid phylogeny. Mol Biol Evol. 2007; 24: 1832–1842. PMID: 17522086

48. Hauth AM, Maier UG, Lang BF, Burger G. The *Rhodomonas salina* mitochondrial genome: bacteria-like operons, compact gene arrangement and complex repeat region. Nucleic Acids Res. 2005; 33: 4433–4442. PMID: 16085754

49. Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, Bowman S, et al. Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. Proc Natl Acad Sci USA. 2007; 104: 19908–19913. PMID: 18077423

50. Kim E, Lane CE, Curtis BA, Kozera C, Bowman S, Archibald JM. Complete sequence and analysis of the mitochondrial genome of *Hemiselmis andersenii* CCMP644 (Cryptophyceae). BMC Genomics. 2008; 9: 215. doi: 10.1186/1471-2164-9-215 PMID: 18474103

51. Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, et al. The highly reduced genome of an enslaved algal nucleus. Nature. 2001; 410: 1091–1096. PMID: 11323671

52. Douglas SE, Penny SL. The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol. 1999; 48: 236–244. PMID: 9929392

53. Tanifuji G, Onodera NT, Wheeler TJ, Dlutek M, Donaher N, Archibald JM. Complete nucleomorph genome sequence of the nonphotosynthetic alga *Cryptomonas paramecium* reveals a core nucleomorph gene set. Genome Biol Evol. 2011; 3: 44–54. doi: 10.1093/gbe/evq082 PMID: 21147880

54. Donaher N, Tanifuji G, Onodera NT, Malfatti SA, Chain PS, Hara Y, et al. The complete plastid genome sequence of the secondarily nonphotosynthetic alga *Cryptomonas paramecium*: reduction, compaction, and accelerated evolutionary rate. Genome Biol Evol. 2009; 1: 439–448. doi: 10.1093/gbe/evp047 PMID: 20333213

55. Moore CE, Curtis B, Mills T, Tanifuji G, Archibald JM. Nucleomorph genome sequence of the Cryptophyte Alga *Chroomonas mesostigmatica* CCMP1168 reveals lineage-specific gene loss and genome complexity. Genome Biol Evol. 2012; 4: 1162–1175. doi: 10.1093/gbe/evs090 PMID: 23042551

56. Yih W, Kim H, Jeong H, Myung G, Kim Y. Ingestion of cryptophyte cells by the marine photosynthetic ciliate *Mesodinium rubrum*. Aquat Microb Ecol. 2004; 36: 165–170.

57. Smith SW, Overbeek R, Woese CR, Gilbert W, Gillevet PM. The genetic data environment: an expandable GUI for multiple sequence analysis. Comput Appl Biosci. 1994; 10: 671–675. PMID: 7704666

58. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004; 14: 1394–1403. PMID: 15231754

59. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. BioInformatics. 2014; 30: 1312–1313. doi: 10.1093/bioinformatics/btu033 PMID: 24451623

60. Le SQ, Gascuel O. An improved general amino acid replacement matrix. Mol Biol Evol. 2008; 25: 1307–1320. doi: 10.1093/molbev/msn067 PMID: 18367465

61. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. BioInformatics. 2011; 27: 1164–1165. doi: 10.1093/bioinformatics/btr088 PMID: 21335321

62. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics 1998; 14:817–8. PMID: 9918953

63. Kowallik KV, Stoebe B, Schaffran I, Kroth-Pancic P, Freier U. The chloroplast genome of chlorophyll a +c-containing alga, *Odontella sinensis*. Plant Mol Biol Rep. 1995; 13: 336–342.

64. Ohta N, Matsuzaki M, Misumi O, Miyagishima SY, Nozaki H, Tanaka K, et al. Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. DNA Res. 2003; 10: 67–77. PMID: 12755171

65. Sánchez Puerta MV, Bachvaroff TR, Delwiche CF. The complete plastid genome sequence of the haptophyte *Emiliania huxleyi*: a comparison to other plastid genomes. DNA Res. 2005; 12: 151–156. PMID: 16303746

66. Stoebe B, Kowallik KV. Gene-cluster analysis in chloroplast genomics. Trends Genet. 1999; 15: 344–347. PMID: 10461201

67. Palmer JD. The symbiotic birth and spread of plastids: how many times and whodunnit? J Phycol. 2003; 39: 4–11.

68. Kim E, Archibald JM. Diversity and evolution of plastids and their genomes. In: Aronsson H, Sandelius AS, editors. The chloroplast-interactions with the environment. Berlin: Springer; 2009. pp. 1–39.

69. Oudot-Le Secq M-P, Grimwood J, Shapiro H, Armbrust EV, Bowler C, Green BR. Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: comparison with other plastid genomes of the red lineage. Mol Genet Genomics. 2007; 277: 427–439. PMID: 17252281

70. Simpson CL, Stern DB. The treasure trove of algal chloroplast genomes. Surprises in architecture and gene content, and their functional implications. Plant Physiol. 2002; 129: 957–966. PMID: 12114552

71. Hall N, Pain A, Berriman M, Churcher C, Harris B, Harris D, et al. Sequence of *Plasmodium falciparum* chromosomes 1, 3–9 and 13. Nature. 2002; 419: 527–531. PMID: 12368867

72. Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, et al. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature. 2005; 438: 1151–1156. PMID: 16372009

73. Pain A, Renauld H, Berriman M, Murphy L, Yeats CA, Weir W, et al. Genome of the host-cell transforming parasite *Theileria annulata* compared with *T. parva*. Science. 2005; 309: 131–133. PMID: 15994557

74. Wang SL, Liu X-Q. The plastid genome of *Cryptomonas phi* encodes an hsp70-like protein, a histone-like protein, and an acyl carrier protein. Proc Natl Acad Sci USA. 1991; 88: 10783–10787. PMID: 1961745

75. Borza T, Popescu CE, Lee RW. Multiple metabolic roles for the nonphotosynthetic plastid of the green alga *Prototheca wickerhamii*. Eukaryot Cell. 2005; 4: 253–261. PMID: 15701787

76. Nott A, Jung H-S, Koussevitzky S, Chory J. Plastid-to-nucleus retrograde signaling. Annu Rev Plant Biol. 2006; 57: 739–759. PMID: 16669780

77. Shi C, Shi X. Characterization of three genes encoding the subunits of light-independent protochlorophyllide reductase in *Chlorella protothecoides* CS-41. Biotechnol Prog. 2006; 22: 1050–1055. PMID: 16889380

78. Blankenship R. Molecular mechanisms of photosynthesis. Oxford: Blackwell Publishing House; 2002.

79. Hallick RB, Hong L, Drager RG, Favreau MR, Monfort A, Orsat B, et al. Complete sequence of *Euglena gracilis* chloroplast DNA. Nucleic Acids Res. 1993; 21: 3537–3544. PMID: 8346031

80. Wickett NJ, Zhang Y, Hansen SK, Roper JM, Kuehl JV, Plock SA, et al. Functional gene losses occur with minimal size reduction in the plastid genome of the parasitic liverwort *Aneura mirabilis*. Mol Biol Evol. 2008; 25: 393–401. PMID: 18056074

81. Blinkova A, Hervas C, Stukenberg PT, Onrust R, O'Donnell ME, Walker JR. The *Escherichia coli* DNA polymerase III holoenzyme contains both products of the *dnaX* gene, tau and gamma, but only tau is essential. J Bacteriol. 1993; 175: 6018–6027. PMID: 8376347

82. Dallmann HG, McHenry CS. DnaX complex of *Escherichia coli* DNA polymerase III holoenzyme. Physical characterization of the DnaX subunits and complexes. J Biol Chem. 1995; 270: 29563–29569. PMID: 7493999

83. Janouškovec J, Horák A, Oborník M, Lukeš J, Keeling PJ. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci USA. 2010; 107: 10949–10954. doi: 10.1073/pnas.1003335107 PMID: 20534454

84. Keeling PJ. Diversity and evolutionary history of plastids and their hosts. Am J Bot. 2004; 91: 1481–1493. doi: 10.3732/ajb.91.10.1481 PMID: 21652304

85. Archibald JM, Keeling PJ. On the origin and evolution of plastids. In: Saap J., editor. Microbial phylogeny and evolution. New York: Oxford University Press; 2005. pp. 238–260.

86. Yoon HS, Hackett JD, Pinto G, Bhattacharya D. The single, ancient origin of chromist plastids. Proc Natl Acad Sci USA. 2002; 99: 15507–15512. PMID: 12438651

87. Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ. The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. Mol Biol Evol. 2007; 24: 54–62. PMID: 16990439

88. Rice DW, Palmer JD. An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters. BMC Biol. 2006; 4: 31. PMID: 16956407

89. Patron NJ, Inagaki Y, Keeling PJ. Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. Curr Biol. 2007; 17: 887–891. PMID: 17462896

90. Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rümmele SE, Bhattacharya D. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of Rhizaria with Chromalveolates. Mol Biol Evol. 2007; 24: 1702–1713. PMID: 17488740

91. Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, et al. Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. Mol Biol Evol. 2010; 27: 1698–1709. doi: 10.1093/molbev/msq059 PMID: 20194427

92. Burki F, Okamoto N, Pombert J.-F, Keeling PJ. The evolutionary history of haptophytes and cryptophytes: phylogenomic evidence for separate origins. Proc Biol Sci. 2012; 279: 2246–2254. doi: 10.1098/rspb.2011.2301 PMID: 22298847

93. Stiller JW, Schreiber J, Yue J, Guo H, Ding Q, Huang J. The evolution of photosynthesis in chromist algae through serial endosymbioses. Nat Commun. 2014; 5: 5764. doi: 10.1038/ncomms6764 PMID: 25493338