

SCIENTIFIC REPORTS



OPEN

Extensive gain and loss of photosystem I subunits in chromerid algae, photosynthetic relatives of apicomplexans

Roman Sobotka^{1,3}, Heather J. Esson², Peter Koník^{1,3}, Eliška Trsková^{1,3}, Lenka Moravcová^{1,3}, Aleš Horák^{2,3}, Petra Dufková² & Miroslav Oborník^{1,2,3}

In oxygenic photosynthesis the initial photochemical processes are carried out by photosystem I (PSI) and II (PSII). Although subunit composition varies between cyanobacterial and plastid photosystems, the core structures of PSI and PSII are conserved throughout photosynthetic eukaryotes. So far, the photosynthetic complexes have been characterised in only a small number of organisms. We performed *in silico* and biochemical studies to explore the organization and evolution of the photosynthetic apparatus in the chromerids *Chromera velia* and *Vitrella brassicaformis*, autotrophic relatives of apicomplexans. We catalogued the presence and location of genes coding for conserved subunits of the photosystems as well as cytochrome *b₆f* and ATP synthase in chromerids and other phototrophs and performed a phylogenetic analysis. We then characterised the photosynthetic complexes of *Chromera* and *Vitrella* using 2D gels combined with mass-spectrometry and further analysed the purified *Chromera* PSI. Our data suggest that the photosynthetic apparatus of chromerids underwent unique structural changes. Both photosystems (as well as cytochrome *b₆f* and ATP synthase) lost several canonical subunits, while PSI gained one superoxide dismutase (*Vitrella*) or two superoxide dismutases and several unknown proteins (*Chromera*) as new regular subunits. We discuss these results in light of the extraordinarily efficient photosynthetic processes described in *Chromera*.

Photosystem I (PSI) and photosystem II (PSII) are large pigment-protein complexes, assembled from tens of subunits and harboring various cofactors, embedded in the thylakoid membranes of oxygenic phototrophs (cyanobacteria, algae and plants). Using energy from photons, PSII generates redox potential that is high enough to oxidise water, a virtually unlimited source of electrons in nature. Electrons withdrawn from water are passed via plastoquinone molecules to cytochrome *b₆f*, where their energy is used to pump protons into the thylakoid lumen. The low-energy electrons are then transferred to the luminal protein plastocyanin. At this point PSI must provide very strong redox potential in order to transfer electrons from plastocyanin to stromal ferredoxin. The redox potential of this electron carrier is sufficient to reduce NADP to NADPH in the stroma and thus provide reducing power for the Calvin-Benson cycle. The proton gradient generated by the photosynthetic electron transport chain powers ATP synthase, an enzyme that converts ADP to ATP.

Over 3.5 billion years photosystems probably evolved from a simple homodimeric structure into complicated and very efficient nano-machines¹. All modern oxygenic photosynthetic organisms, however, share the common core structures of PSI and PSII, whose overall configurations are ancient and apparently of cyanobacterial origin. The PSI core is assembled from two large and closely related proteins, PsaA and PsaB, which form a heterodimer. Together PsaA and PsaB bind about 100 chlorophyll molecules and other cofactors (e.g. phylloquinone and a Fe-S cluster) required for electron transfer. The small stromal subunits PsaC, PsaD and PsaE form the binding site for ferredoxin and PsaC binds another Fe-S cluster serving as a terminal electron acceptor. About 15 small membrane subunits are attached to PsaA/B, most of which are highly conserved and present in the PSI complexes

¹Centre Algatech, Institute of Microbiology, Czech Academy of Sciences, Třeboň, Czech Republic. ²Institute of Parasitology, Biological Centre, Czech Academy of Sciences, České Budějovice, Czech Republic. ³Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic. Roman Sobotka and Heather J. Esson contributed equally to this work. Correspondence and requests for materials should be addressed to R.S. (email: sobotka@alga.cz) or M.O. (email: obornik@paru.cas.cz)

of all lineages of oxygenic phototrophs (PsaF, PsaH, PsaK, PsaI and PsbL). The core of the PSII complex consists of four large proteins - D1, D2, CP47 and CP43 - that bind most of the chlorophyll molecules (~30) and most of the other cofactors. These large subunits are joined by a number of small, single helix proteins (for example, PsbI, PsbH, PsbT, and PsbZ); two small proteins (PsbE and PsbF) together bind a heme molecule. At the luminal site of PSII several hydrophilic proteins are assembled around a manganese cluster, where water molecules are split into oxygen and protons.

PSI composition in different taxa varies mostly due to the chlorophyll-binding, light-harvesting complexes associated with PSI that are present in eukaryotes but absent in cyanobacteria, where PSI is associated with phycobilisomes. The number and arrangement of antenna proteins and their pigment compositions vary between plastid lineages (e.g. refs^{2,3}). There are also differences in numbers of small subunits; red algal-derived plastids, for example, lack PsaG and PsaH subunits⁴. The evolution of eukaryotic (plastid) PSI also eliminated the oligomeric (trimeric) form of PSI common in cyanobacteria; plastid PSIs are typically monomeric⁵. Similar to PSI, the eukaryotic PSII became surrounded by light-harvesting complexes and the number and composition of extrinsic luminal proteins differs in cyanobacteria, algae and plants. However, the typical dimeric structure of plastid PSII remains highly conserved and closely resembles cyanobacterial PSII^{6–8}.

The number of taxa where the differences in photosystem protein composition have been examined and compared is relatively small – studies have been confined mostly to higher plants, *Chlamydomonas*, and cyanobacteria, although research has recently extended to red algae and stramenopiles^{4,9–13}. Considering the tremendous diversity of photosynthetic organisms and the complexity of plastid evolution in eukaryotes, our current understanding of photosystem composition is extremely limited.

Extant plastid diversity is the product of a complicated history involving multiple endosymbiotic events. A single-celled eukaryote engulfed and retained a cyanobacterial prey cell giving rise to glaucophytes, red algae, green algae and plants (primary endosymbiosis). In two distinct secondary endosymbiotic events, green algal prey cells were engulfed by the ancestor of the chlorarachniophytes and by a phago-heterotrophic euglenid, respectively. Secondary endosymbiosis involving a red algal prey cell has occurred at least once in evolutionary history, giving rise to the plastids in stramenopiles, dinoflagellates possessing peridinin, haptophytes, cryptophytes, and apicomplexans^{14–18}; however, hypotheses involving independent secondary¹⁹ or tertiary endosymbioses are becoming more plausible (e.g. ref.²⁰). Endosymbiosis invariably led to the reduction of endosymbiont (plastid) genomes as genes were gradually incorporated into the host nucleus^{21–23}. In non-photosynthetic plastid-containing lineages the loss of photosynthesis has been accompanied by physical reduction of the plastid. An important example is that of the apicomplexans, unicellular parasites that are the causative agents of malaria, toxoplasmosis, and other human and livestock diseases. Most apicomplexans, most notably *Plasmodium* and *Toxoplasma*, host a relict plastid (the “apicoplast”) in their cells^{24–26}. Their closest, extant, photosynthetic relatives are *Chromera velia* and *Vitrella brassicaformis* (hereafter *Chromera* and *Vitrella*), unicellular algae that are associated with stony corals^{27–30}. The fully sequenced plastid genome of *Chromera* revealed some evolutionary oddities, such as two essential thylakoid proteins (PsaA and AtpB) each split into two smaller proteins, and an apparently linear genome^{28,31}. The split PsaA subunit does not impede the function of PSI, which exhibits extremely high photochemical trapping efficiency³².

Transmission electron micrographs of *Chromera* PSI reveal a typical structure which is bound to light-harvesting complexes closely related to the PSI-antennae supercomplexes of red algae^{33,34}. *Chromera* also possesses a clade of unique antennae proteins, termed “*Chromera* Light Harvesting” complexes that exhibit similarities to harvesting complexes from xanthophytes^{33,34}. This is consistent with previous research indicating an ochrophyte origin of chromerid plastids³⁵. The photosynthetic pigment complement of *Vitrella* - chlorophyll-*a* (but not chlorophyll-*c*), violaxanthin, vaucheriaxanthin, β -carotene, and an isofucoaxanthin-like pigment – closely resembles that of the eustigmatophyte *Nannochloropsis limetica*^{27,29}. The pigment complement in *Chromera* is similar to that in *Vitrella*, but *Chromera* lacks vaucheriaxanthin and possesses isofucoaxanthin instead²⁷.

Photosynthesis studies have revealed *Chromera*'s high capacity for adaptation to different light conditions, allowing it to maintain high rates of photosynthesis and growth without succumbing to photoinhibition^{36–38}. In order to expand our understanding of the relationship between photosystem structure and function in chromerid algae, we used biochemical methods to identify the components of PSI. To put our results in an evolutionary context, we analysed the sequences of 55 proteins integral to the structure and function of the four thylakoid membrane complexes from *Chromera*, *Vitrella*, stramenopiles, red algae, green algae, plants, glaucophytes and cyanobacteria.

Results and Discussion

Loss of plastid proteins associated with photosynthetic complexes. To obtain a broader view on the evolution of photosynthetic complexes we retrieved the amino acid sequences of 55 proteins associated with PSI, PSII, cytochrome *b₆f* and ATP synthase from sequence databases (genes and accession numbers are given in Table S1, Supplementary Material). We focused specifically on subunits documented both in cyanobacteria and in eukaryotes. Sequences were obtained from the cyanobacteria *Prochlorococcus marinus*, *Synechococcus elongatus*, *Thermosynechococcus elongatus*, *Anabaena* spp., and *Microcystis* spp.; the chromerids *Chromera* and *Vitrella*; the rhizarians *Paulinella chromatophora* and *Bigelowiella natans*; the glaucophytes *Cyanophora paradoxa* and *Glaucocystis nostochinearum*; the red algae *Gracilaria* spp., *Chondrus crispus* and *Cyanidioschyzon merolae*; the green algae *Chlamydomonas reinhardtii*, *Volvox carteri*, *Micromonas* spp. and *Ostreococcus* spp.; the plants *Arabidopsis thaliana* and *Physcomitrella patens*; the stramenopiles *Aureococcus anophagefferens*, *Ectocarpus siliculosus*, *Nannochloropsis gaditana*, *Phaeodactylum tricorutum*, *Thalassiosira pseudonana* and *Fragilariopsis cylindrica*; the haptophyte *Emiliania huxleyi*; the cryptophyte *Guillardia theta*; and the euglenophyte *Euglena gracilis*. Proteins were scored as either nuclear-encoded, plastid-encoded, or absent (Fig. 1). There were several observable phylogenetic patterns of transfer and loss in certain genes and taxa, particularly when *Paulinella* is excluded from

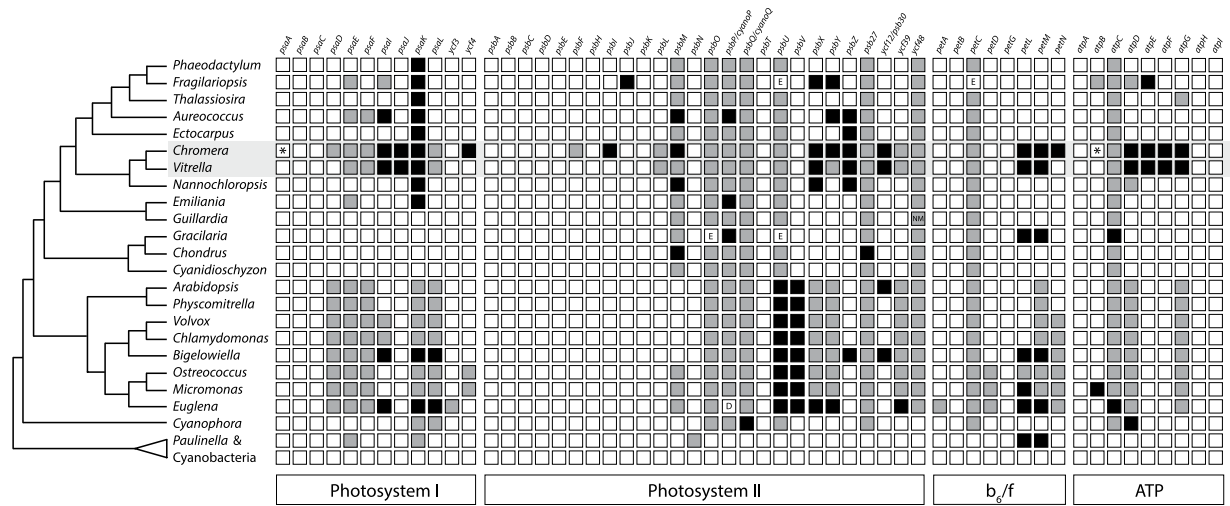


Figure 1. Nuclear transfer and loss of genes coding for thylakoid membrane complex proteins in cyanobacteria and eukaryotes. *Chromera* has experienced a greater degree of gene loss than other eukaryotic taxa. White squares denote plastid-encoded genes; gray squares, nuclear-encoded genes; black squares, absent genes. In *Chromera* both *psaA* and chloroplast *atpB* are split into two segments (indicated with an asterisk). E: protein sequence is derived from EST data. D: protein was directly sequenced. NM: protein is encoded in the nucleomorph.

consideration. *Chromera*, however, appears to have lost an unusually high number of genes: from PSI, *psaI*, *psaJ*, *psaK*, and *ycf4*; from PSII, *psbI*, *psbM*, *psbX*, *psbY*, *psbZ*, and *psb30*; from cytochrome b_6/f , *petL*, *petM*, and *petN*; and from ATP synthase, *atpD*, *atpE*, *atpF*, and *atpG*.

The PsaI protein stabilises PsaL in the cyanobacterium *Synechocystis*³⁹; PsaL and PsaH form a docking site for light-harvesting complex II (LHCII) during state transitions in tobacco⁴⁰ and the stability of this site relies on PsaI⁴¹. PsaI is present in all taxa considered here except for *Chromera*, *Vitrella*, *Aureococcus*, and the secondary green algae *Bigelowiella natans* and *Euglena gracilis* (Fig. 1). While *psaH* is absent from the red lineage (and from cyanobacteria, and thus not considered here⁴), *psaL* is present (although it has undergone transfer to the nucleus in *Chromera* and *Vitrella*; Fig. 1). PsaK, responsible for binding LHCI⁴², is absent in all secondary red algae (Fig. 1). These absences could be explained as a result of the different light harvesting requirements and associated pigment profiles of different plastid lineages^{37,43}; they could, however, be an artefact resulting from our failure to detect divergent protein sequences. PsaJ appears to stabilise PsaF, which in turn facilitates electron transfer from plastocyanin to the PSI reaction center in *Chlamydomonas* and *Arabidopsis*^{44,45}. Interestingly, although *psaJ* is absent in *Chromera* and *Vitrella*, *psaF* is present in the nucleus of both (Fig. 1). PsaF, however, is predicted to influence the binding capabilities of PsaN⁴⁶, a protein responsible for plastocyanin binding in plants⁴⁵ but absent in the red lineage. It seems likely, therefore, that the interactions between PSI subunits differ in red plastids and the more widely studied green lineage. Ycf4 is involved in the biogenesis of PSI in green algae and plants^{47,48} and is absent only in *Chromera*. While we cannot rule out the possibility that *Chromera* Ycf4 is too divergent to detect using BLAST or PSI-BLAST, the protein is not essential in tobacco⁴⁹ and it is therefore conceivable that Ycf4 is not required for PSI biosynthesis in *Chromera*.

PsbI and PsaM are small subunits of PSII thought to be associated with stability^{50,51} and dimerization of the photosystem⁵². PsaX, PsaY, PsaZ and Psa30 are peripheral PSII subunits that were also linked to photosystem stability and photoprotection^{53–56}. The loss of *psbM* in *Chromera* is interesting in light of its role in dimerization (see below); this loss is shared with *Aureococcus*, *Nannochloropsis* and *Chondrus* (Fig. 1). While the nature of PSII oligomerization in *Aureococcus* and *Chondrus* is unknown, monomeric PSII has been observed in *Nannochloropsis oceanica*⁵⁷. Since the absence of the peripheral proteins in *Chromera* does not prevent efficient and adaptable photosynthesis^{36–38}, *Chromera* must possess novel mechanisms, and possibly novel PSII subunits, to enable this efficiency. While no other taxon considered here lacks all of the four peripheral proteins missing in *Chromera*, eight other taxa – *Fragilariopsis*, *Aureococcus*, *Vitrella*, *Nannochloropsis*, *Arabidopsis*, *Bigelowiella*, and *Euglena* – lack one or more of them, suggesting that the protein infrastructure required for maintaining photosystem stability and light tolerance is evolutionarily flexible. Nevertheless, *Chromera* exhibits the highest degree of PSII gene loss: six gene losses in total.

PetL, PetM and PetN are small subunits of the cytochrome b_6/f complex. PetL appears to be involved in maintaining stability of the complex and functional conformation of the Rieske protein (PetC). While it has been shown to be non-essential for survival or photosynthesis in tobacco, the green alga *Chlamydomonas*, and the cyanobacterium *Synechocystis*, its loss can reduce b_6/f stability and the efficiency of electron transport^{58–60}. PetM is a peripheral protein⁶¹ whose suppression prevents the accumulation of the b_6/f complex in young tobacco leaves⁶²; in *Synechocystis*, on the other hand, interruption of the *petM* gene has no observable effect on b_6/f but does decrease PSI and phycobilisome content⁶³. PetN is required for the stability of the b_6/f complex in tobacco and knockout mutants are incapable of carrying out photosynthesis^{60,64}. The absence of *petL* and *petM* in *Chromera*,

Vitrella, *Bigelowiella*, *Euglena* and *Paulinella* (Fig. 1), combined with the relatively minor effects of their suppression as described above, could indicate the presence of redundant, as yet undescribed stabilizing proteins in these taxa. *Chromera*, however, is the only taxon lacking *petN*, implying either a highly divergent sequence (entirely possible given *PetN*'s small size⁶⁴), or a novel mechanism for maintaining *b₆f* stability unique to *Chromera*.

The ATP synthase complex also exhibits extensive loss in chromerids (Fig. 1). Of particular interest are four genes encoding stabilizing and structural proteins in ATP synthase, absent in both *Chromera* and *Vitrella*: *atpD*, (also absent from the glaucophyte *Cyanophora paradoxa*), *atpE*, *atpF* (absent in *Nannochloropsis*), and *atpG*. The absence of these genes, present in almost all other eukaryotes surveyed here, should have serious and extensive consequences for the structure and function of ATP synthase – for example, in *Arabidopsis*, knockout of *atpD* resulted in seedling lethality⁶⁵. The electron transport chain still functions in *Chromera* and *Vitrella*, however, and, in the case of *Chromera*, does so efficiently^{36–38}. It seems extremely likely, therefore, that *Chromera* and *Vitrella* merely possess ATP synthase subunits with highly divergent amino acid sequences that have eluded previous similarity-based searches and annotation efforts⁶⁶.

Phylogenetic analysis of photosynthetic complex proteins. In order to provide a phylogenetic framework for interpreting our results, we constructed individual alignments for 55 proteins and concatenated these into a single alignment for phylogenomic analysis. Highly variable sites were excluded to form two datasets from this primary alignment: one site-rich (10 465 sites) and one site-poor (4872 sites). The alignments were analysed using both maximum likelihood (RAxML⁶⁷) and Bayesian (PhyloBayes 3.3⁶⁸) methods. Although their resulting topologies differed, both analyses recovered the red lineage as a clade to the exclusion of cyanobacteria, glaucophytes, and the green lineage (Fig. S1, Supplementary Material). *Chromera* and *Vitrella* formed a well-supported clade within the stramenopiles in both analyses and were recovered together as a sister group to the eustigmatophyte *Nannochloropsis gaditana* in the site-rich analysis with moderate/high support (Fig. S1A). In the site-poor analysis *Nannochloropsis* was the sister taxon to the remaining stramenopiles and the chromerids, but this relationship was unsupported (Fig. S1B).

The branching of chromerids within the stramenopiles – and specifically with *Nannochloropsis* – is consistent with pigment composition²⁹ and phylogenomic analyses of plastid genomes performed previously³⁵. It also supports suggestions that a single endosymbiotic event involving a red algal prey cell is not the best explanation for extant red plastid diversity^{20,69–72}. On the contrary, the distribution of secondary red plastids among eukaryotes is better explained as a result of multiple secondary endosymbiotic events involving a red algal endosymbiont^{19,70} or the lateral inheritance of a secondary red plastid via tertiary and/or quaternary endosymbiotic events in the relevant lineages²⁰. If an ancestor of “chromodellids” (i.e. the clade comprised of colpodellids and chromerids – neither group is itself monophyletic^{23,73}) and apicomplexans obtained its plastid via higher order endosymbiosis involving a recent ancestor of *Nannochloropsis* (after its divergence from other stramenopiles), the topology of the true plastid tree would resemble our topology (which admittedly excludes the dinoflagellates) (Fig. S1). A common origin of the eustigmatophyte and chromerid/apicomplexan plastids is further supported by the similar pigment compositions of *Nannochloropsis limnetica* and *Vitrella*; namely, both possess chlorophyll *a* (and lack chlorophyll *c*, as does *Chromera*), violaxanthin, vaucherixanthin, and β -carotene²⁹. It is important to note, however, that both *Chromera* and *Vitrella* are prone to long-branch attraction in phylogenetic analyses – this is so severe in plastid genomes that *Chromera* sequences have occasionally been omitted from analyses in order to minimise the risk of recovering inaccurate topologies due to long-branch attraction³⁵. Moreover, the sequence of hypothesised evolutionary events required to produce the conserved SELMA translocation machinery, active in the outer two membranes of secondary red plastids, is arguably unlikely to have evolved independently in multiple eukaryotic lineages⁷⁴. Nevertheless, ML phylogeny presented in this work is consistent with the late acquisition of a eustigmatophyte endosymbiont in the common ancestor of chromerids and apicomplexans.

Phylogenetic analysis inferred from proteins of photosystems suggests that the genes coding for subunits of photosynthetic machinery have been lost multiple times over evolutionary history, and are therefore not strictly necessary for a functioning photosynthetic electron transport chain. *Chromera* and *Vitrella*, in particular, have sustained several losses that have not occurred in the majority of other red plastids; these genes – namely *psaI*, *psaJ*, *ycf4*, *psbI*, *psbY*, *ycf12/psb30*, *petL*, *petM*, *petN*, *atpD*, *atpE*, *atpF* and *atpG* – are therefore likely to have been lost in a recent common ancestor of the chromerid plastids. This would constitute a relatively rapid loss of photosystem proteins when compared to gene loss patterns in other plastid lineages. It is intuitively tempting to explain the extensive gene losses in chromerids in terms of the evolution of parasitism and predation in their apicomplexan and colpodellid relatives, respectively. Indeed, the acquisition of a parasitic life cycle in the apicomplexan lineage is associated with gene loss rather than the development of evolutionary novelties^{23,75}, so the reduction of the photosynthetic apparatus could have been the precursor to the multiple losses of photosynthesis that have occurred in the chromodellid and apicomplexan lineage.

Protein mass spectrometry of *Chromera* and *Vitrella* thylakoid membranes. Given the extensive loss of thylakoid complex proteins in both chromerids, we questioned whether the photosynthetic complexes in these algae are merely reduced or whether they contain novel, non-canonical subunits. Using 2D Clear-Native/SDS electrophoresis (CN/SDS-PAGE) combined with mass-spectrometry (MS) we demonstrated previously that the PSI subunit PsaA and ATP synthase subunit AtpB in *Chromera* are split into two smaller proteins³¹. Nevertheless, the identification of most of the protein spots was not possible because of the lack of genomic data. Annotations of the fully sequenced nuclear and chloroplast genomes, however, are now available^{28,66}.

In order to characterise the organization of *Chromera* photosystems we used three different detergent systems to solubilise membrane complexes: a) n-dodecyl- α -maltoside (α -DM) that is recognised as a very mild detergent⁷⁶, b) a conventional β -DM detergent that is generally less mild than α -DM⁷⁶, and c) a relatively harsh mixture of β -DM and SDS to induce the partial disruption of protein complexes. After the separation of

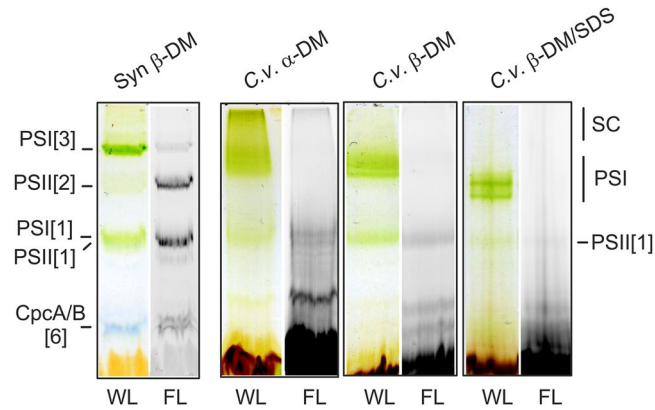


Figure 2. CN-PAGE of *Chromera* (*C.v.*) membrane complexes solubilised by α -DM, β -DM or β -DM with 0.4% SDS. For each line 15 μ g of chlorophyll was loaded and the gel scanned under white light illumination (WL). Chlorophyll fluorescence emitted by PSII and by ‘free’ antenna proteins (FL) was detected by Fuji LAS 4000 after excitation by blue light. SC marks a PSI-PSII supercomplex observed only in the α -DM solubilised membranes (see Fig. 3); PSI marks a monomeric PSI supercomplex with (fucoxanthin-chlorophyll) light-harvesting antennae. Membranes (3 μ g of chlorophyll) isolated from the cyanobacterium *Synechocystis* 6803 and solubilised with β -DM (*Syn* mem) were used to demonstrate the mobility of photosynthetic complexes: PSI[1] (~350 kDa) and PSI[3] (~1 MDa), monomer and trimer of PSI, respectively; PSII[2] (~700 kDa), dimer of PSII; CpcA/B[6], ~100-kDa heterohexamers of CpcA and CpcB phycobiliproteins.

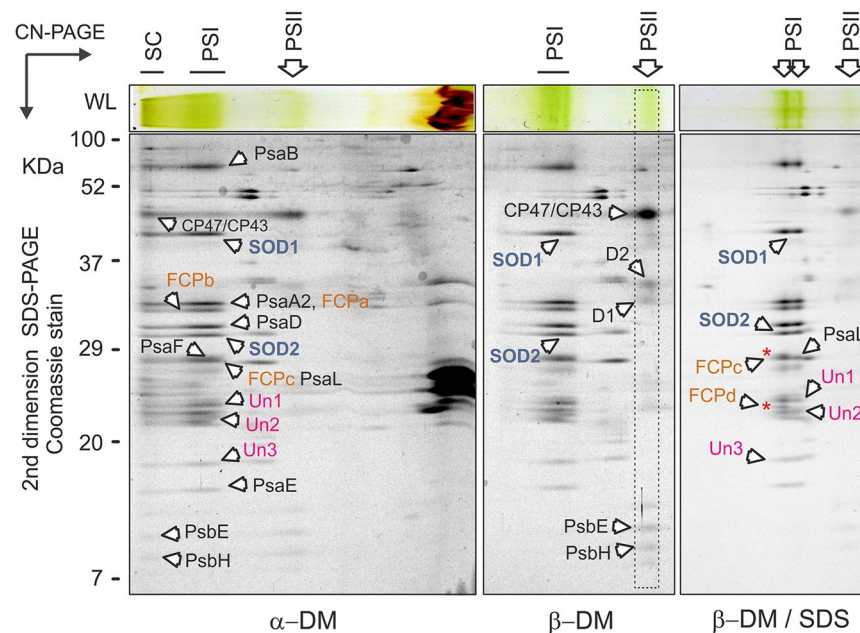


Figure 3. 2D-PAGE of membrane complexes from *Chromera* solubilised by different detergents. Gel strips from CN-PAGE (see Fig. 2) were further separated in a second dimension by SDS-PAGE and stained with Coomassie blue; only the upper part of the gel is shown for the β -DM and β -DM/SDS samples. Individual spots were identified by MS. The PSI-PSII supercomplex (SC) observed after solubilisation with mild α -DM detergent was almost completely disrupted by β -DM solubilisation. Further addition of SDS resulted in the release of two antenna proteins from PSI (marked by red asterisks), however two superoxide dismutases (SOD1 and SOD2) and three other putative PSI subunits (Un1-3) remained bound to PSI core. All identified protein spots are listed in Table 1.

solubilised membrane complexes via CN-PAGE we obtained a different pattern for each detergent system (Fig. 2). Interestingly, a large assembly containing both PSI and PSII (see also Fig. 3) can be observed at the top of the CN gel after separation of the α -DM-solubilised complexes. It is notable that the chlorophyll fluorescence of PSII present in this putative supercomplex is strongly quenched, indicating that PSI acts as an energy trap for PSII fluorescence (Figs 2 and 3). The observed assembly is not preserved in β -DM and the individual PSI and PSII

Protein	Annotation ^a	Predicted mass (Da)	Identified tryptic peptides/theoretical	
			Membranes	PSI
PsaB	YP_003795270.2 (P)	102938	13/49	13/49
AtpB2^b	YP_003795258.2 (P)	61968		2/40
CP47	YP_003795264.2 (P)	61162	7/23	—
CP43	YP_003795274.1 (P)	53151	6/18	—
PsaA2	YP_003795268.2 (P)	49083	3/20	4/20
SOD1	Cvel_7136 (N)	41820	9/33	6/33
D2	YP_003795275.1 (P)	38569	6/14	—
D1	YP_003795256.1 (P)	38569	5/11	—
FCPa	Cvel_9928 (N)	31699	7/88	8/22
FCPb	Cvel_13255 (N)	32398	7/18	7/18
PsaD	Cvel_25206 (N)	30567	7/22	10/22
SOD2	Cvel_3019 (N)	29877	11/20	11/20
PsaF	Cvel_23764 (N)	35801	6/26	—
PsaL	Cvel_23765 (N)	29657	2/15	4/15
FCPc	Cvel_3431 (N)	26189	3/16	4/16
PsaA1	YP_003795268.2 (P)	32332	—	1/11
Un1	Cvel_30781 (N)	45879	3/37	5/37
FCPd	Cvel_1169 (N)	24034	6/18	—
Un2	Cvel_3217 (N)	22554	8/18	8/18
Un3	Cvel_6643 (N)	22517	2/18	5/18
PsaE	Cvel_5788 (N)	14402	4/11	4/11
PsbE	YP_003795269.1 (P)	9657	2/6	—
PsbH	YP_003795330.1 (P)	9348	3/3	—

Table 1. List of *Chromera* proteins as identified by MS analysis of individual spots on 2D-PAGE of membrane complexes (Fig. 3) and the purified PSI (bold; Fig. 5). *a* – CryptoDB database (<http://cryptodb.org/cryptodb/>) or NCBI database (<https://www.ncbi.nlm.nih.gov/>); N = nuclear encoded, P = plastid encoded. *b* – probably a contaminating protein.

complexes are therefore better resolved. Adding SDS led to faster mobility of PSI due to the dissociation of some weakly bound subunits and the partial loss of PSII (Fig. 2).

Protein complexes were further separated in a second dimension by SDS-PAGE and individual putative subunits (co-migrating with PSI/PSII bands) were analysed by mass spectrometry (MS). As noted above, both PSI and PSII subunits can be found in a large structure at the top of the native gel with α -DM (Fig. 3), which suggests that in *Chromera* photosystems are tightly associated. Surprisingly, two different proteins showing high homology to Fe⁺-superoxide dismutase enzymes (marked SOD1 and SOD2 in Fig. 3) were apparently co-migrating with PSI. Four SOD enzymes are annotated in the *Chromera* genome - two are MnSODs and two are FeSODs (SOD1/2). In order to confirm their putative identity, we queried the sequences of all four using SignalP⁷⁷ and TargetP^{78,79} (see Supplementary Methods). The two FeSODs possess bipartite sequences consistent with plastid targeting (Supplementary Table S2). Sequences were also queried using HECTAR and HECTAR[^]SEC⁸⁰. While only SOD2 (Cvel_3019) was identified as plastid-targeted by full HECTAR analysis, SOD1 (Cvel_7136) was predicted to possess both a signal peptide and a relatively high score for plastid targeting (Supplementary Table S3). Furthermore, HECTAR[^]SEC identified signal peptides in both sequences (Supplementary Table S4). A phylogenetic analysis of Fe- and MnSODs indicated that the *Chromera* FeSODs are not monophyletic; unfortunately, support values were too low to determine their evolutionary origin with certainty (Supplementary Fig. S3). To clarify whether SOD1/2 are true SOD enzymes, we measured SOD activity in a native gel after separation of *Chromera* membrane complexes. Indeed, we detected clear SOD activity co-localised with PSI. This activity was specifically inhibited by peroxide but not by KCN (Fig. 4). This implies that at least one of the two SOD proteins is an active FeSOD but, most likely, SOD1 and SOD2 work together as a heterodimer (see discussion).

In addition to SOD1 and SOD2, we identified three other putative PSI subunits, but none of these proteins (Un1-3, Fig. 3; Table 1) had any homologs revealed by BLASTing against the NCBI database. Their protein profiles on CryptoDB showed only signal peptides, transmembrane domains, and low complexity regions; Un1 also possessed prosite domain PS51257, a prokaryotic membrane lipoprotein lipid attachment site profile. Finally, four different fucoxanthin-chlorophyll antenna homologs (FCPa-d) appeared to be associated with PSI. After treatment with SDS two antenna proteins (FCPc, FCPd) were released from the PSI complex but otherwise this PSI appeared very stable and SOD1/2 proteins as well as Un1-3 proteins co-migrated with PSI even after solubilisation with SDS. In contrast to PSI we did not find any unexpected proteins associated consistently with PSII with comparable masses of those in *Chromera* or cyanobacterial monomeric PSIIs using CN-PAGE (~350 KDa; Fig. 2).

Thylakoid complexes (including antenna proteins) isolated from *Vitrella* tended to smear on CN-gel (Fig. 5) and thus we failed to obtain CN-gel of α -DM solubilised membranes. However, MS analysis of β -DM and β -DM/

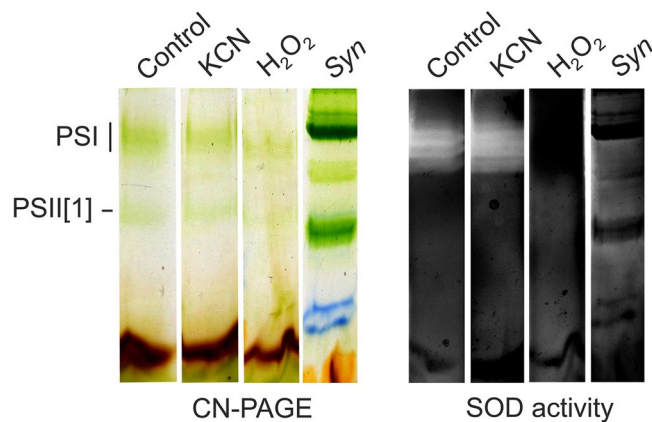


Figure 4. Detection of FeSOD activity in *Chromera* membrane complexes separated by CN-PAGE. Membrane were solubilised by β -DM and for each line 15 μ g of chlorophyll was loaded. Membranes (3 μ g of chlorophyll) isolated from the cyanobacterium *Synechocystis* 6803 were used as a negative control. After separation, the gel was stained for SOD activity essentially as described in ref.⁹⁶. FeSOD enzymes are inhibited by peroxide but not by KCN, which however inhibits MnSOD enzymes. ZnSODs are inhibited neither by peroxide or KCN⁹⁶.

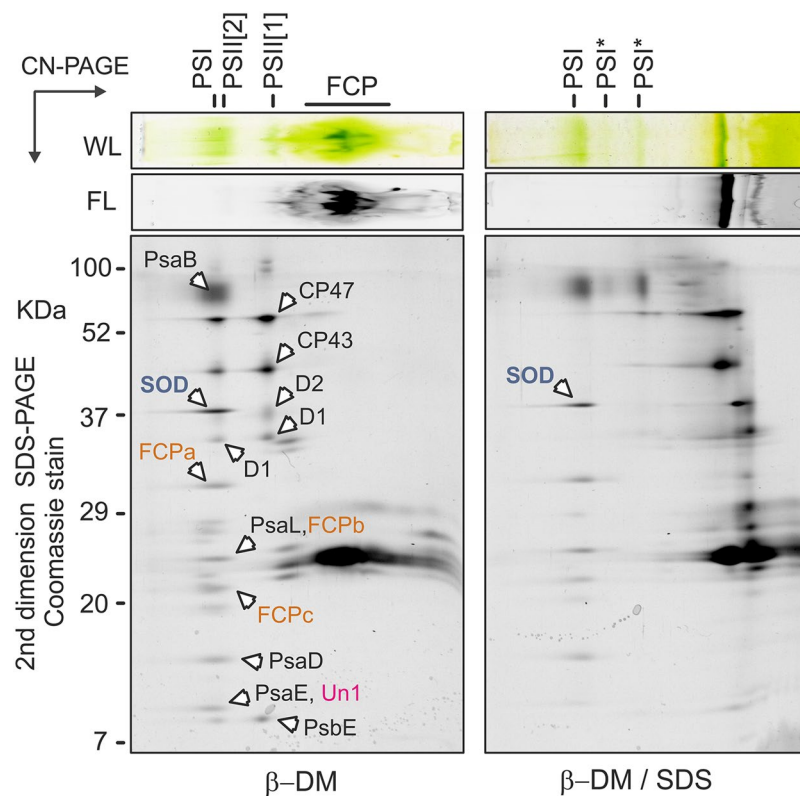


Figure 5. 2D-PAGE of membrane complexes from *Vitrella* solubilised by β -DM and β -DM with 0.4% SDS. The gel was stained by Coomassie blue and individual spots were identified by MS (listed in Table 2). PSII[1] and PSII[2]; monomeric and dimeric PSII, respectively. PSI* marks partially disassembled PSI complexes.

SDS gels identified most spots belonging to PSI and PSII complexes unambiguously (Fig. 5; Table 2). In contrast to *Chromera*, the *Vitrella* PSII forms a dimer with a mass very similar to monomeric PSI – a pattern typical for plant photosystems⁷⁶. The absence of a PSII dimer in *Chromera* is probably related to the loss of PsbM and PsbI subunits⁵²; these proteins are coded in the *Vitrella* genome (Fig. 1). Similarly with *Chromera*, a FeSOD enzyme co-migrated with *Vitrella* PSI complex (Table 2) and this putative PSI-SOD interaction was preserved even after adding of SDS (Fig. 5). In addition, we detected an unknown protein co-migrating with *Vitrella* PSI (Table 2).

Subunit composition of the purified *Chromera* PSI. To confirm tight binding of FeSODs and other novel subunits to chromerid PSI we purified this complex from *Chromera* by a combination of sucrose gradient

Protein	Annotation ^a	Predicted mass (Da)	Identified tryptic peptides/theoretical
PsaB	ADJ66611 (P)	85581	1/39
SOD	Vbra_10461 (N)	37353	5/28
FCPa	Vbra_17392 (N)	31914	3/16
PsaL	Vbra_18375 (N)	25913	1/11
FCPb	Vbra_9733 (N)	23552	2/16
FCPc	Vbra_1573 (N)	21039	6/11
PsaD	ADJ66635 (P)	13697	5/13
PsaE	Vbra_21597 (N)	13990	3/9
Un1	Vbra_13060 (N)	12117	2/4
CP47	ADJ66608 (P)	58872	12/30
CP43	ADJ66597 (P)	52667	4/31
D2	ADJ66652 (P)	41915	3/17
D1	ADJ66640 (P)	39086	4/15
PsbE	YP_003795269.1 (P)	12117	2/4

Table 2. List of proteins of *Vitrella* identified by MS on 2D CN/SDS PAGE of membrane complexes (see Fig. 4). *a* - CryptoDB database (<http://cryptodb.org/cryptodb/>) or NCBI database (<https://www.ncbi.nlm.nih.gov/>).

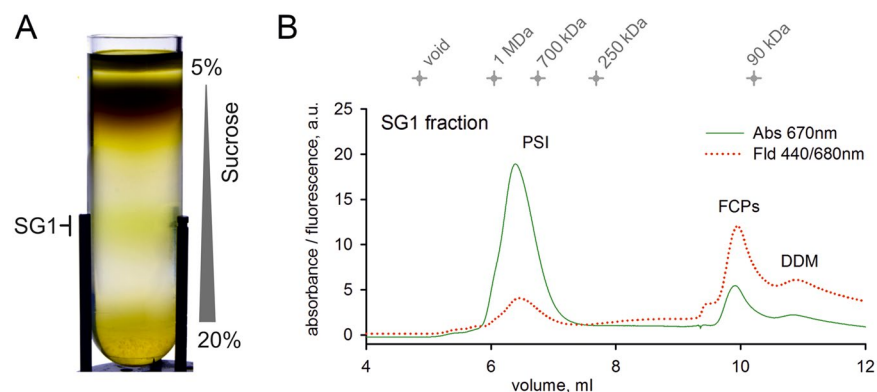


Figure 6. Isolation of the PSI complex from *Chromera*. (A) Solubilised membranes were loaded on a sucrose gradient and centrifuged at 75 000 x g for 18 hours; the SG1 (PSI) fraction was then collected and separated on a size-exclusion column (B). Eluted proteins and complexes were detected by a diode-array detector and a fluorescence detector set to 440/680 nm (excitation/emission wavelengths). Chlorophyll absorbance (670 nm) is shown by a green line and chlorophyll fluorescence is shown by a red dotted line; FCP, free antenna proteins. PSI complex, represented by peak 6-7.5 mL, was collected; see also Supplementary Fig. S3.

and size-exclusion chromatography (Fig. 6A,B). The isolated PSI had very low chlorophyll fluorescence (Fig. 6B) and exhibited an absorbance spectrum typical for algal (*Nannochloropsis*) PSI (Supplementary Fig. S3)¹². The mass of *Chromera* PSI, as estimated from size-exclusion chromatography and CN-PAGE, was about 750 KDa (Fig. 6B, Supplementary Fig. S3). The migration of isolated PSI on CN-PAGE was similar to the PSI band in the β -DM-solubilised membrane complexes (compare Fig. 2 and Supplementary Fig. S3) implying that the purified complex is fairly intact.

The isolated PSI was separated on a 2D gel and all protein spots were cut and analysed by MS. Table 1 summarises identified PSI subunits. Although we did not detect PsaF and the FCPc antenna found previously in the separated membrane complexes (Fig. 3; Table 1) we confirmed that the SOD1/2 proteins and the three uncharacterised proteins Un1-3 are firmly attached to the PSI core. In addition, there was at least one more subunit with a mass of about 20 KDa which we failed to identify (Fig. 7). In the total membrane fraction and in the isolated PSI, all identified PSI subunits appear as abundant as the typical PSI subunits. We can therefore conclude that the *Chromera* PSI complex lost PsaI and PsaK, as indicated by our *in silico* analysis, but contains five new genuine subunits.

Although a new, small PSI protein (10 kDa) has been found in diatoms¹⁰, the addition of five new and relatively large subunits to PSI is unprecedented, particularly if none of these extra subunits seem to be related to antenna proteins⁸¹. In addition, after mild membrane solubilisation, the *Chromera* photosystems appear to be organised as a large assembly (Fig. 3). The existence of large molecular machinery in *Chromera* containing PSI and PSII together with b_6f and ATP synthase would provide an attractive explanation for the observed structural changes in PSI including the PsaA split, the PSII monomerisation (lack of PsbM and PsbI), the loss of peripheral PSII subunits (PsbX/Y/Z/30) as well as for the loss of b_6f subunits and the modification of ATP synthase (AtpB

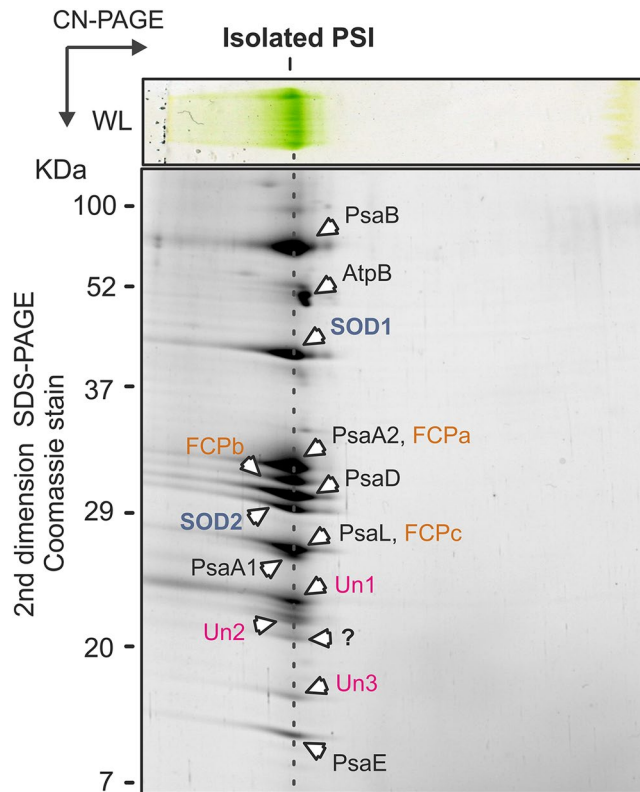


Figure 7. 2D-PAGE of *Chromera* PSI complex purified by a combination of sucrose gradient and size-exclusion chromatography as described in Fig. 6. The gel was stained with Coomassie blue and individual spots were identified by MS (listed in Table 1).

split, loss of AtpD/E/F/G). Indeed, it is a question how such a respirasome-like organisation would complicate the battery of regulatory mechanisms employed by oxygenic phototrophs to balance the excitation of photosystems and electron/proton fluxes⁸². It is known, however, that under high light conditions the majority of PSI complexes in *Arabidopsis thaliana* binds physically to PSII⁸³ and supercomplexes containing PSI, cytochrome *b₆f* and respiratory complex I are also well characterised in plants and green algae⁸⁴. The formation of megadalton protein assemblies between photosystems and other thylakoid protein complexes therefore appears to be common in eukaryotic phototrophs. The packing of membrane complexes in *Chromera* might be much tighter due to structural changes in photosynthetic machinery and these lower dynamics are compensated for by the re-wiring of regulatory mechanisms.

SODs are crucial protective enzymes in oxidative phototrophs^{85,86}. Together with catalases or peroxidases they detoxify superoxide anion radicals generated on the acceptor side of PSI. Although this reaction can be deleterious, it is generally utilised by autotrophs for the so-called water-water cycle as a sink for the excess electrons and protons produced by PSII. During this cycle O_2^- generated by PSI is converted by SODs and peroxidases back to water while oxidising NADPH⁸². Under stress conditions in particular this cycle can account for a large electron flux⁸⁷. Plant SODs are found in the vicinity of PSI⁸⁸ and it is worth noting that two different plastid FeSOD enzymes form a heterodimer⁸⁹. It is possible that such a heterodimer is a component of PSI in *Chromera*, whereas a FeSOD homodimer, as observed in the closely related *Plasmodium falciparum* cytosolic FeSODs⁹⁰ (see Supplementary Fig. S2), is probably attached to *Vitrella* PSI (Fig. 5). Monomeric FeSODs are not found in nature⁹¹. The fact that *Chromera* FeSODs are transformed into regular PSI subunits indicates that the water-water cycle plays a prominent role in the regulation of photosynthesis. An interesting possibility is that chromerids can consume a significant amount of oxygen using the water-water cycle. Quigg and colleagues suggest that the very high carbon assimilation rates observed in *Chromera* are facilitated by the activation of the oxygen consumption process, which maintains high RuBISCO efficiency³⁶. Organisation of the photosynthetic apparatus into a large supercomplex with extremely rapid electron transfer combined with a robust water-water cycle could provide a framework for the efficient and adaptable photosynthesis observed in *Chromera*³⁷. Although more biochemical and physiological work is needed to elaborate this hypothesis, these results, together with previous reports^{31,92} contribute to our growing understanding of the metabolic changes undergone in the ancestors of apicomplexans as they made the transition from alga to parasite.

Materials and Methods

Phylogenetic Analysis. NCBI and other databases (see Supplementary Table S1) were queried using blastp and tblastn from the ncbi+ package to retrieve genes coding for thylakoid membrane complex subunits (Supplementary Table S1). Blast results were parsed and respective homologues were extracted from our custom

databases using perl scripts (available upon request). Datasets for each protein were aligned using MAFFT⁹³. For each dataset, ambiguously aligned regions were manually deleted and the maximum likelihood (ML) phylogeny was inferred using RAxML⁶⁷ with the gamma-corrected LG substitution matrix. The topology with the highest likelihood score was selected using the rapid bootstrapping algorithm (500 replicates). The branching support was assessed using thorough (“slow”) non-parametric bootstrap from 500 replicates using the model described above. Alternatively, Bayesian inference (BI) of single-gene phylogenies was performed using Phylobayes 3.3f⁸⁸ using a combination of the empirical profile mixture model (C40) and the LG model. For each dataset, two chains were run until they converged (i.e. maximum observed discrepancy was 0.3 or lower) and the effective sample size of model parameters reached 100. Posterior probabilities computed after discarding the first 1/5th of generations represent statistical support for branches.

A total of 55 protein datasets were then concatenated into a single dataset and analysed as described above with GTR for ML analysis and CAT + GTR for BI.

Gene transfer and loss. Information for the location of each gene was gathered from its source database (see Supplementary Table S1). In the case of ambiguities, sequences were searched against an in-home BLAST database of relevant plastid genomes compiled from sequences publicly available on NCBI to ensure accurate annotation. Due to the divergent nature of chromerid plastids^{28,35} and extensive endosymbiotic gene transfer to the nucleus (Fig. 1), we have performed additional Position-Specific Iterated (PSI) BLAST screen against the complete set of predicted proteins available for *C. velia* and *V. brassicaformis* (<http://cryptodb.org/>) for all the chromerid genes missing from the standard blastp/tblastn searches (Fig. 1). Multiple sequence alignments of these genes (with the taxon sampling as defined by Fig. 1) were used as a ‘query’ in standalone PSI-BLAST (-in_msa option). We allowed the program to automatically estimate the number of iterations necessary for run convergence and set the significance threshold (e-value) to 1e-10.

Strain cultivation. *Chromera* cells were grown in glass tubes in artificial seawater medium with supplementation of *f/2* nutrients under continuous illumination with 100 μmol of photons $\text{s}^{-1} \text{m}^{-2}$ (white light bulb) at 28 °C and bubbled with air. *Vitrella* cells were grown at 28 °C in a rotating 250-mL Erlenmeyer flask under irradiance of 20 μmol of photons $\text{s}^{-1} \text{m}^{-2}$.

Two dimensional electrophoresis and protein mass spectrometry. Cells were harvested and cellular membranes isolated as previously described⁹⁴. Obtained membranes were solubilised with α -DM or β -DM for 1 h on ice (α/β -DM/chlorophyll = 20 w/w). Analysis of native membrane complexes was performed using CN electrophoresis as described by Wittig and colleagues⁹⁵. Individual proteins in membrane complexes were resolved in the second dimension by SDS-PAGE in a 12–20% linear gradient polyacrylamide gel containing 7 M urea and protein identification by LC-MS/MS were carried out essentially as described previously³¹.

Isolation of *Chromera* PSI. *Chromera* membranes prepared as described above were solubilised with β -DM for 1 h on ice (β -DM/chlorophyll = 20 w/w). After removing insoluble parts by centrifugation (65,000 \times g, 20 min) solubilised protein complexes were separated by ultracentrifugation (141 000 \times g for 18 h) on a 5–20% gradient of sucrose in thylakoid buffer containing 0.04% β -DM. A prominent green band with a migration expected for PSI was taken and concentrated 10-fold on a 100-kD cutoff microconcentrator (Millipore). The concentrated sample was immediately injected onto an Agilent-1200 HPLC machine and separated on a Yara 3000 column (Phenomenex) using 25 mM HEPES buffer, pH 7.5, containing 0.1% β -DM at a flow rate of 0.2 mL min^{-1} at 10 °C. Fractions were collected using an Agilent-1200 fraction collector. The column was calibrated using photosynthetic complexes of known size isolated from the cyanobacterium *Synechocystis* 6803: trimeric PSI (~1 MDa), dimeric PSII (~700 kDa), monomeric PSII lacking oxygen-evolving complex (~250 kDa) and the purified His-tagged CP43 assembly module (~90 kDa).

Measurement of SOD activity. The SOD activity staining was performed as described in ref.⁹⁶. Briefly, the strips from CN-gel were incubated with 0.1% nitroblue tetrazolium solution in dark for 20 min. After washing with water the gel was further incubate with a solution of 28 μM riboflavin and 28 mM TEMED in 0.1 M potassium phosphate buffer (pH 7.0) in dark for 20 min. The washed gel was finally illuminated with a white-light box for 20 min to develop SOD activity bands. In order to identify different SOD species a gel-strip was incubated in 8 mM hydrogen peroxide prior to incubation with nitroblue tetrazolium to inhibit both ZnSOD and FeSOD activities. The ZnSOD activity only was inhibited by adding of 8 mM KCN into riboflavin solution.

References

- Rutherford, A. W. & Faller, P. Photosystem II: evolutionary perspectives. *Philos. Trans. R. Soc. London B.* **358**, 245–253 (2003).
- Green, B. R. After the primary endosymbiosis: An update on the chromalveolate hypothesis and the origins of algae with Chl c. *Photosynth. Res.* **107**, 103–115 (2011).
- Nielson, J. A. D. & Durnford, D. G. Structural and functional diversification of the light-harvesting complexes in photosynthetic eukaryotes. *Photosynth. Res.* **106**, 57–71 (2010).
- Alboresi, A. *et al.* Conservation of core complex subunits shaped the structure and function of photosystem I in the secondary endosymbiont alga *Nannochloropsis gaditana*. *New Phytol.* **213**, 714–726 (2017).
- Dudkina, N. V., Folea, I. M. & Boekema, E. J. Towards structural and functional characterization of photosynthetic and mitochondrial supercomplexes. *Micron* **72**, 39–51 (2015).
- Nagao, R. *et al.* Purification and characterization of a stable oxygen-evolving Photosystem II complex from a marine centric diatom. *Chaetoceros gracilis*. *Biochim. Biophys. Acta* **1979**, 160–166 (2010).
- Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**, 55–60 (2011).

8. Wei, X. *et al.* Structure of spinach photosystem II-LCHII supercomplex at 3.2 Å resolution. *Nature* **534**, 69–74 (2016).
9. Amunts, A. & Nelson, N. Plant photosystem I design in the light of evolution. *Structure* **17**, 637–650 (2009).
10. Grouneva, I., Rokka, A. & Aro, E.-M. The thylakoid membrane proteome of two marine diatoms outlines both diatom-specific and species-specific features of the photosynthetic machinery. *J. Proteome Res.* **10**, 5338–5353 (2011).
11. Grouneva, I. *et al.* Phylogenetic viewpoints on regulation of light harvesting and electron transport in eukaryotic photosynthetic organisms. *Planta* **237**, 399–412 (2013).
12. Basso, S. *et al.* Characterization of the photosynthetic apparatus of the eustigmatophycean *Nannochloropsis gaditana*: Evidence of convergent evolution in the supramolecular organization of photosystem I. *Biochim. Biophys. Acta* **1837**, 306–314 (2014).
13. Yang, H., Liu, J., Wen, X. & Lu, C. Molecular mechanism of photosystem I assembly in oxygenic organisms. *Biochim. Biophys. Acta* **1847**, 838–848 (2015).
14. Cavalier-Smith, T. Principles of protein and lipid targeting in secondary symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. *J. Euk. Microbiol.* **46**, 347–366 (1999).
15. Archibald, J. M. The puzzle of plastid evolution. *Current Biology* **19**, R81–R88 (2009).
16. Keeling, P. J. The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu. Rev. Plant Biol.* **64**, 583–607 (2013).
17. Stiller, J. W. Toward an empirical framework for interpreting plastid evolution. *J. Phycol.* **50**, 462–471 (2014).
18. Archibald, J. M. Genomic perspectives on the birth and spread of plastids. *Proc. Natl. Acad. Sci. USA* **112**, 10147–10153 (2015).
19. Falkowski, P. G. *et al.* The evolution of modern eukaryotic phytoplankton. *Science* **305**, 354–360 (2004).
20. Petersen, J. *et al.* *Chromera velia*, endosymbiosis and the Rhodoplex Hypothesis – plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). *Genome Biol. Evol.* **6**, 666–684 (2014).
21. Martin, W. & Herrmann, R. G. Gene transfer from organelles to the nucleus: How much, what happens, and why? *Plant Physiol.* **118**, 9–17 (1998).
22. Rogers, M. B., Gilson, P. R., Su, V., McFadden, G. I. & Keeling, P. J. The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol. Biol. Evol.* **24**, 54–62 (2007).
23. Janouškovec, J. *et al.* Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proc. Natl. Acad. Sci. USA* **112**, 10200–10207 (2015).
24. McFadden, G. I., Reith, M. E., Munholland, J. & Lang-Unnasch, N. Plastid in human parasites. *Nature* **381**, 482 (1996).
25. Köhler, S. *et al.* A plastid of probable green algal origin in apicomplexan parasites. *Science* **275**, 1485–1489 (1997).
26. Oborník, M., Janouškovec, J., Chrudimský, T. & Lukeš, J. Evolution of the apicoplast and its hosts: From heterotrophy to autotrophy and back again. *Int. J. Parasitol.* **39**, 1–12 (2009).
27. Moore, R. B. *et al.* A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* **451**, 959–963 (2008).
28. Janouškovec, J., Horák, A., Oborník, M., Lukeš, J. & Keeling, P. J. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proc. Natl. Acad. Sci. USA* **107**, 10949–10954 (2010).
29. Oborník, M. *et al.* Morphology, ultrastructure and life cycle of *Vitrella brassicaformis* n. sp., n. gen., a novel chromerid from the Great Barrier Reef. *Protist* **163**, 306–323 (2012).
30. Cumbo, V. R. *et al.* *Chromera velia* is endosymbiotic in larvae of the reef corals *Acropora digitifera* and *A. tenuis*. *Protist* **164**, 237–244 (2013).
31. Janouškovec, J. *et al.* Split photosystem protein, linear-mapping topology and growth of structural complexity in the plastid genome of *Chromera velia*. *Mol. Biol. Evol.* **30**, 2447–2462 (2013).
32. Belgio, E. *et al.* High photochemical trapping efficiency in Photosystem I from the red clade algae *Chromera velia* and *Phaeodactylum tricorutum*. *Biochim. Biophys. Acta* **1858**, 56–63 (2017).
33. Tichý, J. *et al.* Light harvesting complexes of *Chromera velia*, photosynthetic relative of apicomplexan parasites. *Biochim. Biophys. Acta* **1827**, 723–729 (2013).
34. Pan, H., Šlapeta, J., Carter, D. & Chen, M. Phylogenetic analysis of the light-harvesting system in *Chromera velia*. *Photosynth. Res.* **111**, 19–28 (2012).
35. Ševčíková, T. *et al.* Updating algal evolutionary relationships through plastid genome sequencing: did alveolate plastids emerge through endosymbiosis of an ochrophyte? *Sci. Rep.* **5**, 10134, <https://doi.org/10.1038/srep10134> (2015).
36. Quigg, A. *et al.* Photosynthesis in *Chromera velia* represents a simple system with high efficiency. *PLOS One* **7**, e47036, <https://doi.org/10.1371/journal.pone.0047036> (2012).
37. Kotabová, E. *et al.* Novel type of red-shifted chlorophyll a antenna complex from *Chromera velia*. I. Physiological relevance and functional connection to photosystems. *Biochim. Biophys. Acta* **1837**, 734–743 (2014).
38. Mann, M. *et al.* Unusual features of the high light acclimation of *Chromera velia*. *Photosynth. Res.* **122**, 159–169 (2014).
39. Xu, Q. *et al.* Mutational analysis of photosystem I polypeptides in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* **270**, 16243–16250 (1995).
40. Lunde, C., Jensen, P. E., Haldrup, A., Knoetzel, J. & Scheller, H. V. The PSI-H subunit of photosystem I is essential for state transitions in plant photosynthesis. *Nature* **408**, 613–615 (2000).
41. Plöschinger, M. *et al.* The low molecular weight protein Psal stabilizes the light-harvesting complex II docking site of photosystem I. *Plant. Physiol.* **172**, 450–463 (2016).
42. Jensen, P. E., Gilpin, M., Knoetzel, J. & Scheller, H. V. The PSI-K subunit of photosystem I is involved in the interaction between light-harvesting complex I and the photosystem I reaction center core. *J. Biol. Chem.* **275**, 24701–24708 (2000).
43. Durnford, D. G. *et al.* A phylogenetic assessment of the eukaryotic light-harvesting antenna proteins, with implications for plastid evolution. *J. Mol. Evol.* **48**, 59–68 (1999).
44. Hippler, M., Drepper, F., Farah, J. & Rochaix, J.-D. Fast electron transfer from cytochrome *c*₆ and plastocyanin to photosystem I of *Chlamydomonas reinhardtii* requires Psal. *Biochem* **36**, 6343–6349 (1997).
45. Haldrup, A., Simpson, D. J. & Scheller, H. V. 2000. Down-regulation of the PSI-F subunit of photosystem I (PSI) in *Arabidopsis thaliana*. *J. Biol. Chem.* **275**, 31211–31218 (2000).
46. Hansson, A. *et al.* Knock-out of the chloroplast-encoded PSI-J subunit of photosystem I in *Nicotiana tabacum*. *FEBS J.* **274**, 1734–1746 (2007).
47. Boudreau, E., Takahashi, Y., Lemieux, C., Turmel, M. & Rochaix, J.-D. The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. *EMBO J.* **16**, 6095–6104 (1997).
48. Ozawa, S. *et al.* Biochemical and structural studies of the large Ycf4-Photosystem I assembly complex of the green alga *Chlamydomonas reinhardtii*. *Plant Cell* **21**, 2424–2442 (2009).
49. Krech, K. *et al.* The plastid genome-encoded Ycf4 protein functions as a nonessential assembly factor for photosystem I in higher plants. *Plant Physiol.* **159**, 579–591 (2012).
50. Künstner, P., Guardiola, A., Takahashi, Y. & Rochaix, J.-D. A mutant strain of *Chlamydomonas reinhardtii* lacking the chloroplast photosystem II *psbI* gene grows photoautotrophically. *J. Biol. Chem.* **270**, 9651–9654 (1995).
51. Schwenkert, S. *et al.* PsbI affects the stability, function, and phosphorylation patterns of Photosystem II assemblies in tobacco. *J. Biol. Chem.* **281**, 34227–34238 (2006).
52. Kawakami, K. *et al.* Roles of PsbI and PsbM in photosystem II dimer formation and stability studied by deletion mutagenesis and X-ray crystallography. *Biochim. Biophys. Acta* **1807**, 319–325 (2011).

53. Funk, C. Functional analysis of the PsbX protein by deletion of the corresponding gene in *Synechocystis* sp. PCC 6803. *Plant Mol. Biol.* **44**, 815–827 (2000).
54. Sugiura, M. *et al.* Psb30 contributes to structurally stabilise the Photosystem II complex in the thermophilic cyanobacterium *Thermosynechococcus elongatus*. *Biochim. Biophys. Acta.* **1797**, 1546–1554 (2010).
55. von Sydow, L. *et al.* The PsbY protein of *Arabidopsis* Photosystem II is important for the redox control of cytochrome b_{559} . *Biochim. Biophys. Acta* **1857**, 1524–1533 (2016).
56. Inoue-Kashino, N., Kashino, Y. & Takahashi, Y. Psb30 is a photosystem II reaction center subunit and is required for optimal growth in high light in *Clamydomonas reinhardtii*. *J. Photochem. Photobiol. B.* **104**, 220–338 (2011).
57. Bina, D., Gardian, Z., Herbštová, M. & Litvin, R. Modular antenna of photosystem I in secondary plastids of red algal origin: a *Nannochloropsis oceanica* case study. *Photosynth. Res.* **131**, 255–266 (2017).
58. Takahashi, Y. *et al.* The chloroplastycf7 (*petL*) open reading frame of *Chlamydomonas reinhardtii* encodes a small functionally important subunit of the cytochrome b_6f complex. *EMBO J.* **15**, 3498–3506 (1996).
59. Schneider, D., Vokmer, T. & Rögner, M. PetG and PetN, but not PetL, are essential subunits of the cytochrome b_6f complex from *Synechocystis* PCC 6803. *Res. Microbiol.* **158**, 45–50 (2007).
60. Schwenkert, S. *et al.* Role of the low-molecular-weight subunits PetL, PetG, and PetN in assembly, stability, and dimerization of the cytochrome b_6f complex in tobacco. *Plant Physiol.* **144**, 1924–1935 (2007).
61. Stroebel, D., Choquet, Y., Popot, J. -L. & Picot, D. An atypical haem in the cytochrome b_6f complex. *Nature* **426**, 413–418 (2003).
62. Hojka, M. *et al.* Inducible repression of nuclear-encoded subunits of the cytochrome b_6f complex in tobacco reveals an extraordinarily long lifetime of the complex. *Plant Physiol.* **165**, 1632–1646 (2014).
63. Schneider, D., Berry, S., Rich, P., Seidler, A. & Rögner, M. A regulatory role of the PetM subunit in a cyanobacterial cytochrome b_6f complex. *J. Biol. Chem.* **276**, 16780–16785 (2001).
64. Hager, M., Biehler, K., Illerhaus, J., Ruf, S. & Bock, R. Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b_6f complex. *EMBO J.* **18**, 5634–5842 (1999).
65. Maiwald, D. *et al.* Knock-out of the genes coding for the rieske protein and the ATP-synthase δ -subunit of *Arabidopsis*. Effects on photosynthesis, thylakoid protein composition, and nuclear chloroplast gene expression. *Plant Physiol.* **133**, 191–202 (2013).
66. Woo, Y. H. *et al.* Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *eLife* **4**, e06974, <https://doi.org/10.1371/journal.pone.0047036> (2015).
67. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–90 (2006).
68. Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286–8 (2009).
69. Bodyl, A. Do plastid-related characters support the Chromalveolate Hypothesis? *J. Phycol.* **41**, 712–719 (2005).
70. Bodyl, A., Stiller, J. W. & Mackiewicz, P. Chromalveolate plastids: direct descent or multiple endosymbioses? *Trends Ecol. Evol.* **24**, 119–121 (2009).
71. Baurain, D. *et al.* Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. *Mol. Biol. Evol.* **27**, 1698–1709 (2010).
72. Zimorski, V., Ku, C., Martin, W. F. & Gould, S. B. Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* **22**, 38–48 (2014).
73. Gile, G. H. & Slamovits, C. H. Transcriptomic analysis reveals evidence for a cryptic plastid in the colpodellid *Voromonas pontica*, a close relative of chromerids and apicomplexan parasites. *PLOS ONE* **9**, e96258, <https://doi.org/10.1371/journal.pone.0096258> (2014).
74. Gould, S. V., Maier, U. -G. & Martin, W. F. Protein import and the origin of red complex plastids. *Curr. Biol.* **25**, R515–R521 (2015).
75. Danne, J. C., Gornik, S. G., MacRae, J. I., McConville, M. J. & Waller, R. F. Alveolate mitochondrial metabolic evolution: Dinoflagellates force reassessment of the role of parasitism as a driver of change in apicomplexans. *Mol. Biol. Evol.* **30**, 123–139 (2012).
76. Pagliano, C., Barera, S., Chimirri, F., Saracco, G. & Barber, J. Comparison of the α and β isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim. Biophys. Acta* **1817**, 1506–1515 (2012).
77. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
78. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1–6 (1997).
79. Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequences. *J. Mol. Biol.* **300**, 1005–1016 (2000).
80. Gschloessl, B., Guermeur, Y. & Cock, J. M. HECTAR: A method to predict subcellular targeting in heterokonts. *BMC Bioinformatics* **9**, 393 (2008).
81. Busch, A. & Hippler, M. The structure and function of eukaryotic photosystem I. *Biochim. Biophys. Acta* **1807**, 864–877 (2011).
82. Miyake, C. Alternative electron flows (water-water cycle and cyclic electron flow around PSI) in photosynthesis: molecular mechanisms and physiological functions. *Plant Cell Physiol.* **51**, 1951–1963 (2010).
83. Yokono, M., Takabayashi, A., Akimoto, S. & Tanaka, A. A megacomplex composed of both photosystem reaction centres in higher plants. *Nat. Comm.* **6**, 6675 (2015).
84. Yadav, K. N. S. *et al.* Supercomplexes of plant photosystem I with cytochrome b_6f , light-harvesting complex II and NDH. *Biochim. Biophys. Acta* **1858**, 12–20 (2017).
85. Asada, K. The water-water cycle in chloroplasts: Scavenging of active oxygens and dissipation of excess photons. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 601–639 (1999).
86. Sae-Tang, P. *et al.* Overexpressed superoxide dismutase and catalase act synergistically to protect the repair of PSII during photoinhibition in *Synechococcus elongatus* PCC 7942. *Plant Cell Physiol.* **57**, 1899–1907 (2016).
87. Hirotsu, N., Makino, A., Ushio, A. & Mae, T. Changes in thermal dissipation and the electron flow in the water-water cycle in rice grown under conditions of physiologically low temperature. *Plant Cell Physiol.* **45**, 635–644 (2004).
88. Ogawa, K., Kanematsu, S., Takabe, K. & Asada, K. Attachment of CuZn-superoxide dismutase to thylakoid membranes at the site of superoxide generation (PSI) in spinach chloroplasts: detection of immune-gold labelling after rapid freezing and substitution method. *Plant Cell Physiol.* **36**, 565–573 (1995).
89. Myoung, F. *et al.* A heterocomplex of iron superoxide dismutases defends chloroplast nucleoids against oxidative stress and is essential for chloroplast development in *Arabidopsis*. *Plant Cell* **20**, 3148–3162 (2008).
90. Boucher, I. W. *et al.* The crystal structure of superoxide dismutase from *Plasmodium falciparum*. *BMC Struct. Biol.* **6**, 20, <https://doi.org/10.1186/1472-6807-6-20> (2006).
91. Muñoz, I. G., Moran, J. F., Becana, M. & Montoya, G. The crystal structure of an eukaryotic iron superoxide dismutase suggests intersubunit cooperation during catalysis. *Prot. Sci.* **14**, 387–394 (2005).
92. Kořeny, L., Sobotka, R., Janoušek, J., Keeling, P. J. & Oborník, M. Tetrapyrrole synthesis of photosynthetic chromerids is likely homologous to the unusual pathway of apicomplexan parasites. *Plant Cell* **23**, 3454–3462 (2011).

93. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–66 (2002).
94. Kaňa, R. *et al.* Violaxanthin inhibits nonphotochemical quenching in light-harvesting antenna of *Chromera velia*. *FEBS Lett.* **590**, 1076–1085 (2016).
95. Wittig, I. & Schägger, H. Features and applications of blue-native and clear-native electrophoresis. *Proteomics* **8**, 3974–3990 (2008).
96. Kuo, W., Huang, C., Shih, C. & Jinn, T. Cellular extract preparation for superoxide dismutase (SOD) activity assay. *Bio Protoc.* **3**, e811 (2013).

Acknowledgements

This work was supported by grants from the Grant Agency of the Czech Republic (P501-12-G055) to M.O. and the Czech Ministry of Education (projects LO1416 and CZ 1.05/2.1.00/19.0392). We would like to thank Martin Lukeš for his help with *Vitrella* cells.

Author Contributions

R.S. and M.O. designed the experiments and led the study. H.J.E. and P.D. performed the comparison of proteins of photosystems. R.S., E.T. and L.M. performed the biochemical analyses. P.K. performed the mass spectrometry analyses. A.H. and H.J.E. performed the phylogenetic analyses. H.J.E., R.S. and M.O. drafted the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-017-13575-x>.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017