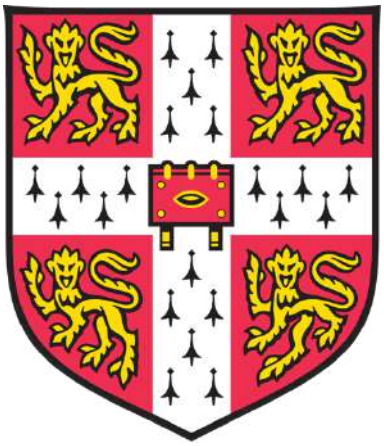
Role of the small subunit of Rubisco in the green algal phylogeny and Carbon Concentrating Mechanism expression



Myriam Madeleine Marthe Goudet

Lucy Cavendish College

University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

April 2020

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the Biology Degree Committee.

Myriam Goudet

April 2020

Myriam Madeleine Marthe GOUDET

Abstract: Role of the small subunit of Rubisco in the green algal phylogeny and Carbon Concentrating Mechanism expression

Photoautotrophic organisms globally fix 111-117x10^{15} grams of carbon per year and around half of this global net primary production is aquatic (Behrenfeld *et al*., 2001; Field *et al*., 1998), with green algae a major contributor to this global carbon fixation. However, aquatic environments have some limitations The concentration of CO_2 is often 2,200 times lower in water than in air, and diffusion is also 8,000 times slower. In addition, Rubisco, which catalyses the first major step of carbon fixation, converting atmospheric CO_2 into precursors of energy-rich molecules, exhibits slow catalytic rates, low affinity for CO_2 and competition from O_2 for the active sites. Therefore, most green algae have developed a Carbon Concentrating Mechanism (CCM). In eukaryotic micro-algae, the Rubisco micro-compartment is called the pyrenoid and together with active inorganic carbon transporters and strategically located carbonic anhydrases, elevated CO_2 within the pyrenoid improves photosynthetic efficiency. Most photosynthetic organisms have an hexadecameric Rubisco holoenzyme (L_8S_8), composed of eight ~55-kDa large subunit (LSU), encoded by a chloroplast gene (rbcL) and eight ~15-kDa small subunit (SSU), encoded by a nuclear gene family (*RbcS*) in Form I Rubisco. The CCM has been particularly well-defined in the model unicellular chlorophyte *Chlamydomonas reinhardtii* and recent studies showed that for full CCM induction, a key protein linker EPYC1 and its interaction with Rubisco SSU were necessary.

The overall goal of this study was to use a phylogenetic approach, firstly to investigate SSU structure across the green algal phylogeny, and also to explore CCM diversity in two specific groups of species. This study used a variety of methodologies combining physiological experiments, biochemistry, imaging and bioinformatic analyses. The results firstly showed the presence of two different Rubisco SSU structures within the green algae. Secondly, the Rubisco catalytic properties found in streptophyte algae closely related to land plants (streptophytes) reflect the strength of any CCM and pyrenoid leakiness, whereas Rubisco in extant land plants reflects more recent selective pressures associated with the terrestrial atmospheric environment. This research also provides evidence for diversity of CCM expression in two closely related genera (*Chlamydomonas* and *Chloromonas*), ranging from species expressing a CCM and pyrenoid, or a CCM without a pyrenoid, to neither pyrenoid or CCM. This study provides the first preliminary analyses of five different genomes confirming multiple independent origins of the pyrenoid in green algae but has also allowed an initial comparison of the molecular components essential for pyrenoid formation across these species.

Acknowledgements

Firstly, I would like to thank my supervisor Prof. Howard Griffiths, not only for being a wonderful mentor but for his patience, his incredible knowledge and for always pushing me to see the bigger picture. I truly believe that to make a change in someone else life you only need one person. Thank you for trusting me when no one else wanted to give me a chance. However, my PhD would not have been the same journey without Dr. Moritz Meyer. Thanks to him for his support, our discussions, his guidance and way more. In a team, there are leaders, defenders and coaches but the most important players are those in the background. Moritz was definitively one of them.

Secondly, I would like to thank Dr. Douglas Orr, Prof. Elizabete Carmo-Silva and the rest of the photosynthesis group in Lancaster University for their tremendous help. They provided me their high level of expertise, a stimulating environment to work and a lot of patience, making Lancaster my second home.

We never achieve a goal without help and support. Therefore, I would like to thank Dr. Gita Yadav and her student Citu for their help. They have been my second brain when nothing was working and strong shoulders to rely on. I would also like to thank Sebastian Eves-van den Akker for his expertise in genome sequencing. Thank you to Yi, Stéphanie, Sophie, Jessica and all the new PhD students in 108 for their friendships, making the physiological ecology lab (and Wanne's lab), an incredible place to work.

Some of the results in this thesis are also the results of a great collaboration with the Cambridge Advanced Imaging Centre (CAIC) and the Goodwin laboratory. Thank you to Karin Muller and Lyn Carter for their technical support for the imaging and to James Rolfe for the isotope data.

Acknowledgements need to go to the Cambridge Trust, Lucy Cavendish College and the National Environment Research Council (NERC) for the 3.5 years funding to complete my PhD. I always felt very privileged to receive these different awards.

Finally, this PhD taught me once again that life is like a bicycle, to stay balanced you need to keep going. For keeping me balanced, I would like to thank Cambridge University Women's Boat Club including my three Blue Boats, Rob Baker, Sally, Lucy, Pat Marsh, and all the others. We survived waves and losses, overcame deceptions to turn them into victories and joy. I am feeling so grateful to be surrounded by such amazing people. I have also been very lucky to have an incredible cheerleading team including Ellie, Jo Ryan, Poly,

April and the rest of Lucy Cavendish College Boat Club, without forgetting my French cheerleaders…. They know who they are. Most importantly, a plant never grows without strong roots, so thank you from the bottom of my heart to my family for being my water to grow and to my husband, Hadi, for being the sunshine in my life.

Content

List of Figures

List of Tables

List of Appendices

List of Abbreviations and Acronyms

1KP	The1000 plant project
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CA	Carbonic Anhydrase
CAM	Crassulacean Acid Metabolism
CBB	Calvin-Benson-Bassham cycle
CCM	Carbon Concentrating Mechanism
CO_2	Carbon dioxide
C_3	C_3 metabolism
C_4	C_4 metabolism
DIW	Deionized Water
EPYC1	Essential Pyrenoid Component 1
GYA	Giga Years Ago
LCI	Low CO_2-inducible
LSU	Large Subunit of Rubisco
MCC	Maximum Clade Credibility
MCMC	Markov Chain Monte-Carlo
Mya	Million years ago
NADP(H)	Nicotinamide Adenine Dinucleotide Phosphate
O_2	Oxygen
LRT	Likelihood Ratio Test
PSI	Photosystem I
PSII	Photosystem II
rbcL	Rubisco Large Subunit gene
RCA	Rubisco activase
RbcS	Rubisco Small Subunit gene
Rubisco	Ribulose-1,5-biphosphate carboxylase/oxygenase
SSU	Small Subunit of Rubisco

Chapter 1: General Introduction

1.1 Green algae: A very diverse group of organisms

1.1.1 Origins

Green algae are one of the major groups of oxygenic photosynthetic eukaryotes. Including between 6,000 and 8,000 species (Chapman, 2013), they are now widespread and abundant, and have colonized all the Earth's aquatic environments and many terrestrial habitats and soils. Crucial for our modern ecosystems, algae have shaped the Earth for hundreds of millions of years (Falkowski *et al*., 2004; O'Kelly, 2007; Leliaert *et al*., 2011). Green algae conquered the land between 500 and 450 million years ago (Mya) (Gensel *et al*., 2008; Kenrick *et al*., 2012; Lenton *et al*., 2016); a key event in the history of life since it led to a dramatic drawdown of CO_2 and an increase in the oxygen concentration of the atmosphere (Berner, 2003). The ancestry of land plants is well established and supported by molecular data (Manhart, 1994; Bhattacharya & Ehlting, 1995; Kranz & Huss, 1996; Friedl, 1997). Green algae arose from an initial primary endosymbiosis event, which occurred between 1 and 1.5 billion years ago (Gya) (Hedges *et al.*, 2004; Yoon *et al.,* 2004). During this endosymbiotic event, a heterotrophic host cell captured a cyanobacterium (prokaryote) that became stably integrated and ultimately became incorporated as a plastid (Archibald, 2009; Keeling, 2010) (Figure 1.1). The subsequent diversification of this new photosynthetic eukaryote gave rise to the green lineage (green algae) but also to the two other groups of algae: the red algae and the glaucophytes. This first diversification of these algae with a primary plastid in aquatic environments occurred during the "Boring billion" (1,800-800 Mya) (Becker, 2013) when the climate on Earth was stable. Other endosymbiosis events occurred later, diversifying widely into multiple lineages.

The phylogeny of green algae is now well resolved (Leliaert *et al*., 2012; Del Cortona *et al*., 2020, Leebens-Mack *et al.*, 2019). Following the primary endosymbiosis, the hypothetical ancestral flagellate diversified into two main lineages (Figure 1.2). On one side the chlorophytes diversified early as prasinophytes in the oceans and then as core chlorophytes

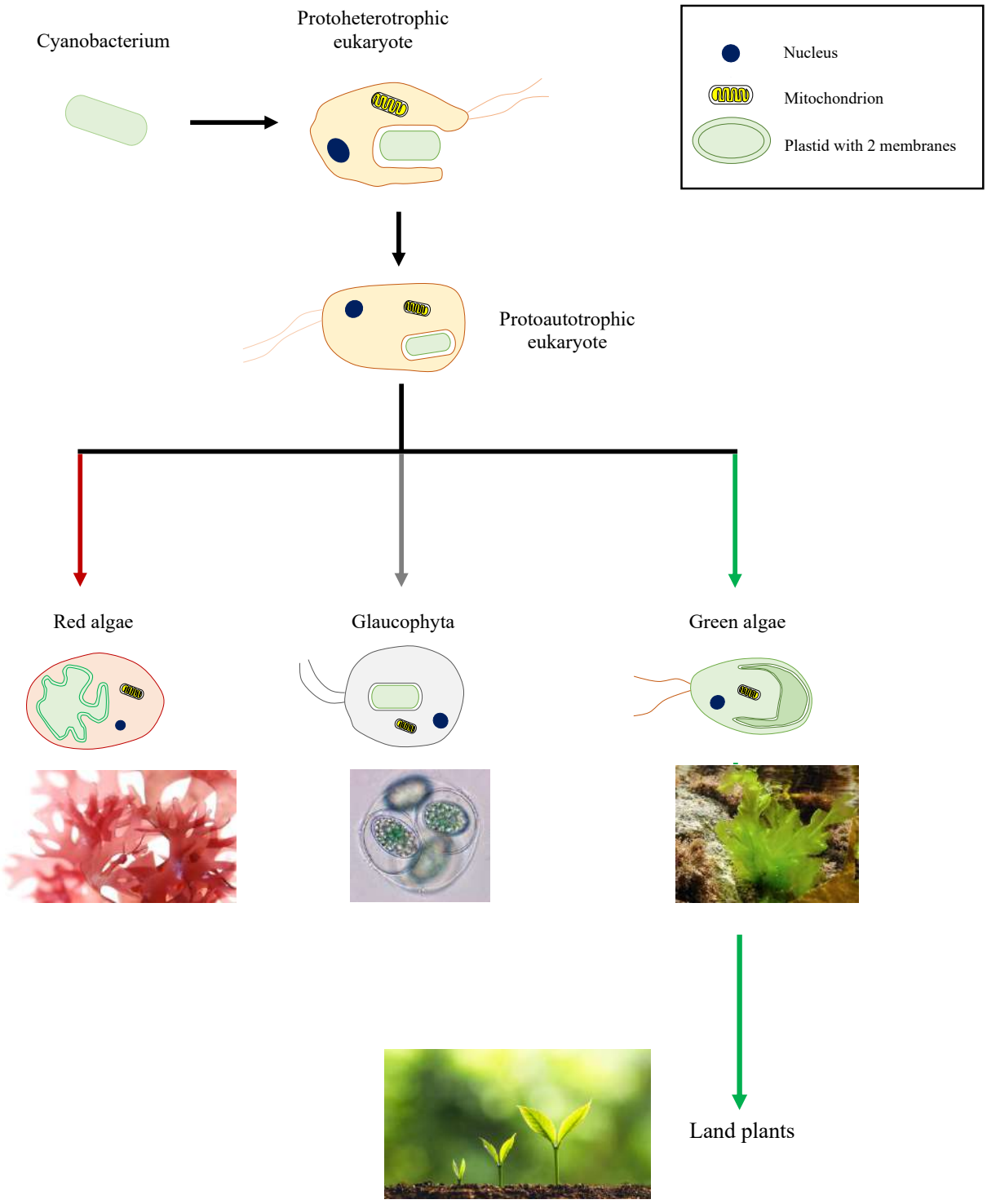


Figure 1.1 The primary endosymbiosis event (adapted from Keeling, 2004) which gave rise to the three different lineages of algae with a primary chloroplast. A heterotrophic host cell (in yellow at the top with flagella) captured a cyanobacterium (prokaryote, in green at the top left) that became stably integrated and turned into a plastid (top middle). The first endosymbiosis gave rise to three different lineages: the red algae, the glaucophytes and the green algae. Later on, other endosymbiotic events occurred giving rise to other lineages not described in this figure, such as the diatoms.

in both fresh and marine environments. It now includes most extant green algae with hundreds of species that play important roles as primary producers in marine and fresh water ecosystems (McCourt *et al*., 2004). On the other side, the streptophytes (which include both land plants and charophytes) diversified in fresh waters and then colonized land between 500-450 Mya. Such adaptation to fresh waters is thought to be a key element for this later land colonization (Becker, 2013). The split between chlorophytes and streptophytes probably occurred 936 Mya (Becker, 2013) but more recent analyses showed that this split could be older with an estimate dated around 1,000 Mya (early Neoproterozoic; Del Cortona *et al*., 2020). Interestingly, the diversification of the streptophytes coincided with the Sturtian glaciation (Gradstein *et al*., 2004) and a fall in atmospheric carbon dioxide concentration (Pierrehumbert *et al*., 2011). Although streptophytes are an ancient lineage, they are not as diversified as the chlorophytes.

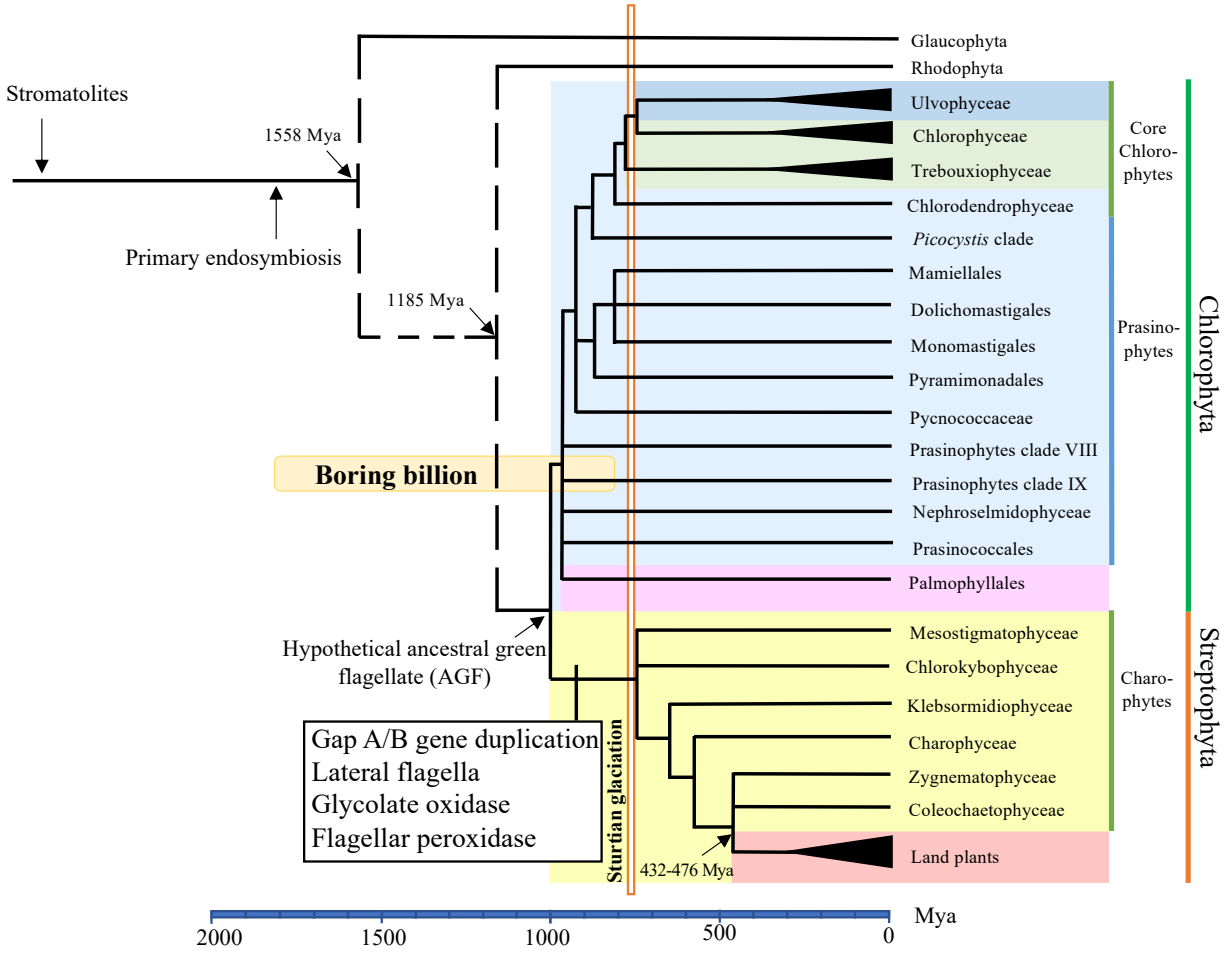


Figure 1.2 Evolutionary relationships of algae arising from the primary endosymbiosis and major glaciation events which occurred during the diversification of the green algae lineages modified from Leliaert *et al*. (2012) and Becker (2013). Evolutionary hypotheses for the streptophyte algae (morphological and molecular characters) are indicated.

1.1.2 Common features

Green algae are unified within the same group because they share common features. The typical green colour of this group is due to the presence of chlorophyll *a* and *b,* whilst the pigments such as carotenes and xanthophylls are also present. Following the primary endosymbiosis, chloroplasts are enclosed by a double membrane, (Figure 1.1) and within the chloroplasts, thylakoids are grouped in lamellae that contain the different chlorophylls. Across green algae, most of them have stiff cell walls with a fibrillar matrix composed of cellulose (Leliaert *et al.,* 2012) and some cells possess structurally similar flagella, though they may be of different lengths. To go further, Melkonian (1984) showed that the flagellar transition zone is characterized by a stellate structure, which is a nine-pointed star linking nine pairs of microtubules.

1.1.3 Diversity

Despite these common features, as a result of a long diversification, green algae show a great diversity of morphology, size and ecology (Figure 1.3). They are present in all aquatic environments, from lakes and rivers to extreme environments. *Dunaliella salina* has been found in hyper-saline conditions where water bodies can contain more than 10% salt (Phadwahl & Singh, 2003) whereas *Klebsormidium flaccidum* has been discovered in acidic waters with extreme concentrations of heavy metals (Zettler *et al*., 2002). The presence of algae in such habitats is a sign of highly developed adaptations which allow these species to overcome the different environmental pressures. However, most green algae are particularly abundant in freshwaters, although prasinophytes are almost exclusively found in marine water. Seaweeds, such as *Ulva*, *Caulerpa* and *Codium,* are present in coastal habitats (Leliaert *et al*., 2012), some of which are responsible for the proliferation of algae along coasts (Dion *et al.,* 1998). One of the most striking examples of this green tide is due to *Ulva armoricana* in Brittany, France (Coat *et al*., 1998). Surprisingly, some green algae can be exclusively terrestrial (Trentepohliales; López-Bautista *et al*., 2006) or develop symbiosis with fungi to form lichen. The genus *Trebouxia* (Chlorophytes) is probably the most common photobiont in lichens (Honegger, 2018).

Green algae can exhibit a great diversity of sizes, morphology and a range of motile and non-motile forms. Most of the green algae are microscopic (*Ostreococcus, Chlamydomonas*), but others can grow thalli up to one meter in length (*Chara*) and look very similar to plants. In addition, species can be branched (*Draparnaldia*) or unbranched

(*Oedogonium*). Finally, they can live in colonies such as *Volvox*, which can include up to 50,000 cells.

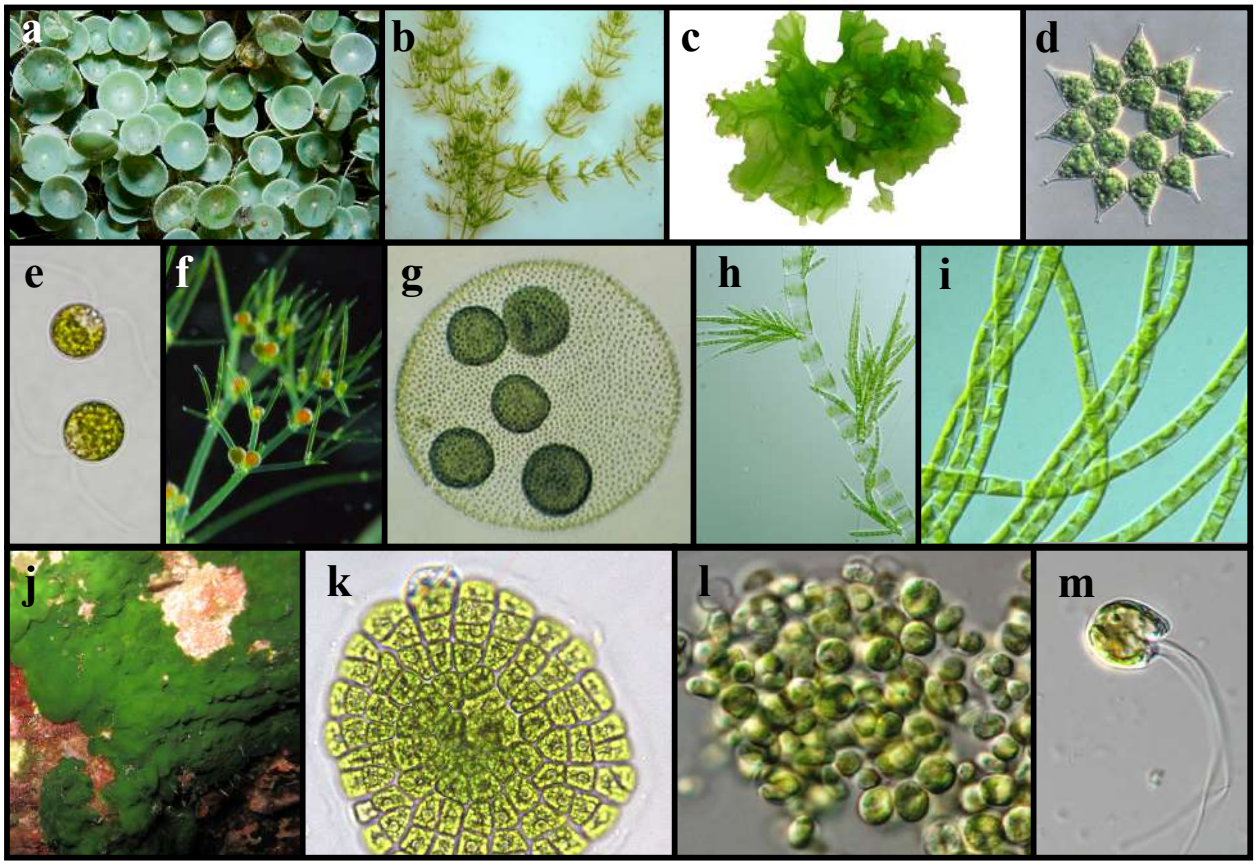


Figure 1.3 Morphological diversity among green algae. **a**: *Acetabularia acetabulum* (Polyphysacea) Mediterranean marine species; **b**: *Chara braunii* (Characeae) European freshwater species; **c**: *Ulva lactuca* (Ulvaceae) marine worldwide species; **d**: *Pediastrum* (Hydrodictyaceae) photoautotrophic nonmotile algae in freshwaters; **e**: *Dunaliella salina* (Dunaliellaceae) halophile micro-algae; **f**: *Nitella gracilis* (Characeae); **g**: *Volvox globator* (Volvoceae) form spherical colonies in freshwater habitats; **h**: *Draparnaldia plumosa* (Chaetophoraceae) composed of a chain of cells arranged in one row; **i**: *Klebsormidium flaccidum* (Klebsormidiaceae) filamentous charophyte algae; **j**: *Palmophyllum crissum* (Palmophyllaceae) marine species; **k**: *Coleochaete* (Coleochaetaceae) form flat, sprawling discs on solid surfaces in freshwater streams; **l**: *Coccomyxa* (Trebouxiophyceae) small elliptical shape; **m**: *Cymbomonas* (Pyramimonadales) marine prasinophycean green algae.

1.2 Aquatic photosynthesis

Stromatolites are thought to be the first evidence of life on Earth (Schopf, 1992). Around 3.5 Gya old, they were most likely photosynthetic organisms (Schopf, 1993; Figure 1.2). Therefore, photosynthesis is not only one of the oldest biological process (Falkowski & Raven, 2007), it allowed the development of aerobic life on Earth (Blankenship & Hartmann, 1998) by generating breathable oxygen in the atmosphere. Studies suggested the rise of oxygen (O_2) was directly caused by the evolution of oxygenic photosynthesis (Blankenship & Hartmann, 1998; Soo *et al*., 2017). Photosynthesis is a biological process which converts light energy (via the Photosystem I and II coupled with the Calvin Benson-Bassham cycle) into chemical bond energy that is stored in the form of organic compounds (Falkowski & Raven, 2007). This process would not be possible without an enzyme called Ribulose-1,5-biphosphate carboxylase/oxygenase (Rubisco), which is involved in the first major step of carbon fixation (see paragraph on Rubisco, below). The origin and nature of the earliest forms of photosynthesis remain unclear. Granick (1965) suggested that the first photosynthetic organisms had a primitive reaction centre based on the transport of electron/proton (Fe^{2+}/$FeOH^{+}$) through membranes. Other studies hypothesized an origin of photosynthesis derived from phototactic systems (when whole organisms move towards or away from a stimulus of light; Nisbet *et al*., 1995) or ultraviolet protection systems (Mulkidjanian & Junge, 1997), but none of these hypotheses have been properly tested. However, Rubisco is known to be older than the Great Oxidation Event which occurred ~2.3 Gya (Tabita *et al.,* 2007; 2008, 2008b), and the increase of oxygen concentration in the atmosphere deeply affected the activity of Rubisco as it conducts both carboxylation and oxygenation reactions (the latter leading to photorespiration). The higher concentration of O_2 started to compete with CO_2 for the active sites, dramatically reducing photosynthetic efficiency.

At current atmospheric O_2 concentrations and 20ºC, the oxygenase activity of Rubisco accounts for up to 25% of catalytic activity. In warm and dry conditions C_3 plants, which include most temperate crops (wheat, rice, pulses etc..), can lose up to 50% of their potential carbohydrate yield due to photorespiration % (Andersson, 2008; Bauwe *et al*., 2010; Zhu *et al*., 2010). Early algae perhaps excreted the photorespiratory 'waste product' as glycolate, and then developed the photorespiratory cycle to recapture some lost C and N (Colman *et al*., 1974). Despite photorespiration, more than 90% of inorganic carbon converted into biomass is fixed by Rubisco. Globally, photoautotroph organisms fix 111-117x10^{15} grams

of carbon per year, and around half of this global net primary production is aquatic (Behrenfeld *et al*., 2001; Field *et al*., 1998). Cyanobacteria and unicellular eukaryotic algae are the main actors in this process across all of the aquatic environments (Falkowski & Raven, 2007). Photosynthesis in aquatic and terrestrial environments follows the same basic catalytical processes. The development of aquatic photosynthesis probably coincided with a fall of atmospheric CO_2 from a concentration ~100 fold higher than in the present day atmosphere to approximatively half of the present level (Berner, 2001). Additional physical constraints on aquatic photosynthesis have resulted in some differences from terrestrial photosynthesis. First of all, the diffusion of CO_2 is around 8,000 times slower in water than in the air with additional limitations imposed by surface boundary layers (Raven *et al*., 1985; Raven & Richardson, 1986; Borges & Frankignoulle, 2002; Yamano *et al*., 2015). In addition, the presence of carbon (C_i) is more often under other forms such as bicarbonate (HCO_3^-) or carbonate (CO_3^{2-}), depending on pH or alkalinity of the medium (Figure 1.4), which reduces significantly the concentration of CO_2 compared to the air (400 µmol mol^{-1} of air *vs* 9-10 µmol mol^{-1} of water). In acid environments, the vast majority of C_i is in the CO_2 form, whereas in more alkaline water, C_i is mostly in the form of HCO_3^- (Moroney & Ynalvez, 2007). Finally, at lower ambient temperatures, O_2 is relatively less soluble than CO_2.

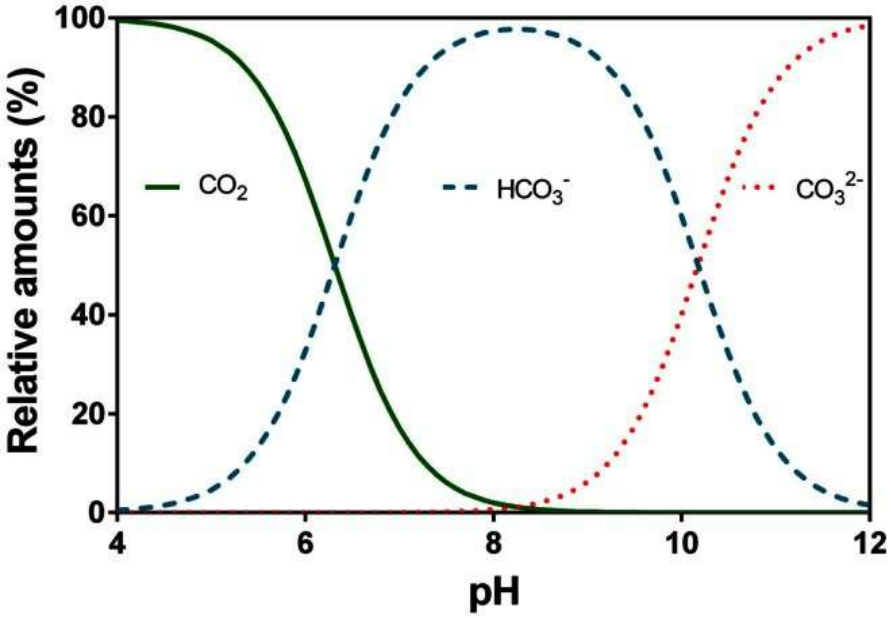


Figure 1.4 The aquatic bicarbonate buffer system, showing relative amounts (%) of Carbon dioxide (CO_2), bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}) in water in function of pH (Pedersen *et al*., 2013).

1.3 Carbon Concentrating Mechanism

Most aquatic photosynthetic organisms developed a biophysical Carbon-Concentration Mechanism (CCM) in order to overcome the aquatic limitations on carbon availability. These CCMs are primarily based around the active accumulation of inorganic carbon around Rubisco in order to saturate active sites with CO_2, thereby suppressing oxygenase activity and minimising photorespiration. This process concentrates C_i against a free energy gradient, thereby increasing the CO_2/O_2 ratio around Rubisco active sites and providing high capacity for net organic carbon production at low external Ci level (Renberg *et al*., 2010). The high affinity for Ci of unicellular algae (Berry *et al*., 1976) along with a lack of measurable photorespiration (Lloyd *et al*., 1977) led to the hypothesis that algae do not rely on the simple diffusion of CO_2 for photosynthesis.

Three major ecological roles have been identified for CCMs (Beardall & Giordano, 2002): *i)* improvement of CO_2 supply and competitive advantages when inorganic carbon and/or CO_2 in the environment are present in low concentration, *ii)* improving resource use efficiency when nutrients (N, P, Fe or S) are in short supply and *iii)* energy dissipation.

Decreasing CO_2 concentration in the atmosphere due to the Great Oxygenation Event appeared to be the main reason why most of the aquatic photosynthetic organisms developed a CCM when the ratio between CO_2/O_2 shifted in favour of O_2. Interestingly, CCM probably appeared multiple times and in different lineages independently (Rasmussen *et al*., 2008). However, precisely dating when CCM appeared is difficult, mainly due to the absence of fossils. Times with low CO_2 concentrations in the atmosphere have been clearly identified and five episodes (Figure 1.5) provided conditions where the development of CCM could have occurred: 2.4-2.1 Gya, 750, 650, 320-270 Mya and finally during the Pleistocene (last 2.4Mya) (Raven *et al*., 2011, 2012). However, Griffiths *et al.* (2017) went one step further and based on the higher solubility of CO_2 in seawater compared to O_2, they took into account times when aqueous O_2 concentration overtook CO_2 concentration in seawater. Therefore, they estimated a potential origin of the CCM around the late Silurian/early Devonian (~420 Mya). However, a more conservative approach (Badger & Price, 2003) estimated an origin of CCM (cyanobacterial and eukaryotic CCM) following the "Devonian Drop" (~400Mya) and during the Carboniferous (~360 Mya). The lack of direct evidences (Raven *et al*., 2012) and the absence of more powerful tools than the molecular clock make the appearance of CCM difficult to estimate. Works on diatoms and haptophytes showed (Young *et al*., 2012) episodes of positive selection on Rubisco (rbcL) during low CO_2 episodes (between 1 and 0.5 Gya) and could be correlated with the origin of CCM.

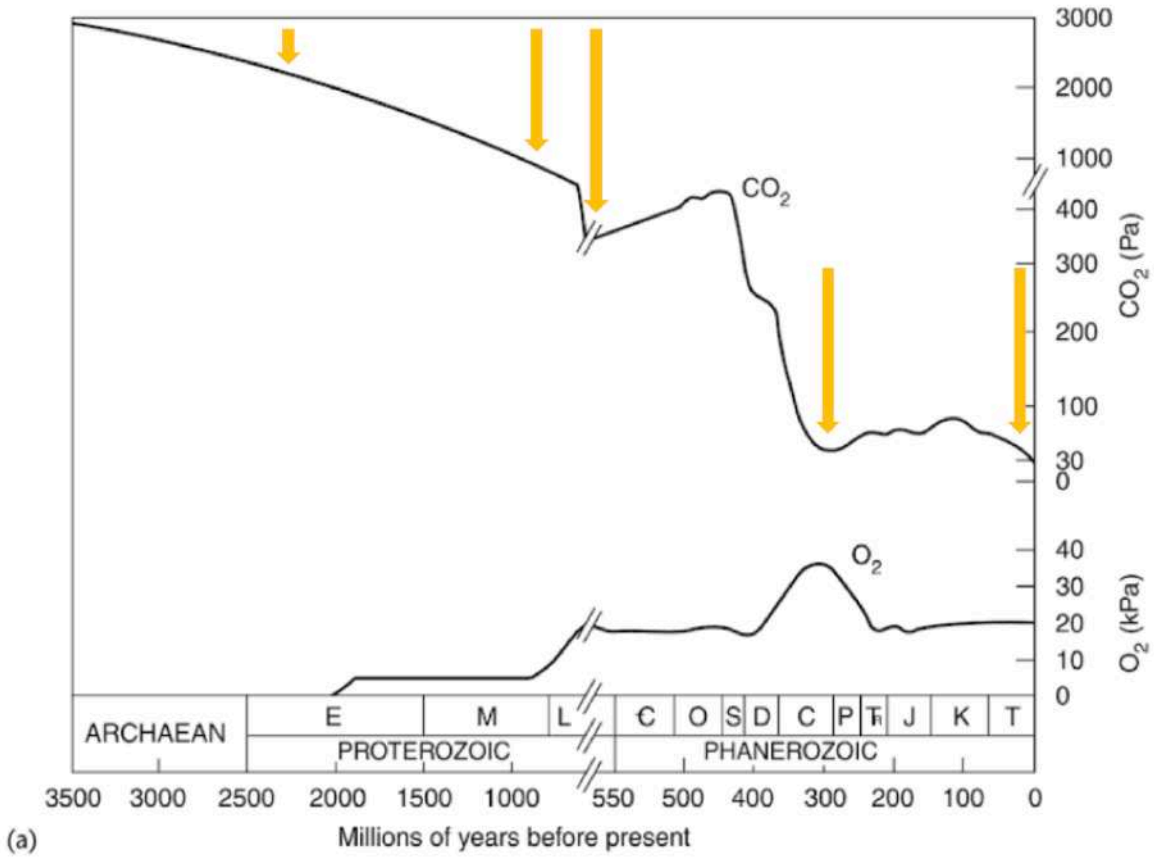


Figure 1.5 Reconstruction of variations in the partial pressures of CO_2 and O_2 in the atmosphere through geological time from Falkowski & Raven (2013). Episodes of low CO_2 conditions providing conditions for the development of CCM are indicated with a yellow arrow.

In unicellular organisms, the induction of CCMs is dependent on numerous ecological imperatives (Maberly & Gontero, 2017). Among them, the four main factors are: availability of inorganic carbon and light, the temperature and any nutrient limitations (Beardall & Giordano, 2002). Physiological evidence for Ci uptake and CCM expression have been shown multiple times with different methods in the green algae and the model organism *Chlamydomonas reinhardtii*. First of all by measuring the whole-cell photosynthesis rates (oxygen evolution, Badger *et al*., 1980; Meyer *et al*., 2012; Mitchell *et al*., 2014) but also by direct measurement of Ci uptake and Ci concentration inside the cell (Moroney & Tolbert, 1985; Sültemeyer *et al*., 1989, 1991; Asamiziu *et al*., 2000). Further studies, including mutants grown in low (0.04% CO_2; air level) and high CO_2 (5% CO_2) conditions gave us more insight on the underlying mechanisms allowing Ci uptake (Spalding *et al*., 1983; Vance & Spalding, 2005; Wang & Spalding, 2006). Such observations have also been

described in other eukaryotes such as *Chlorella vulgaris* (Shiraiwa & Miyachi, 1985), *Chlorella emersonii* (Beardall *et al*., 1982) or in diatoms (Burkhart *et al*., 2001) but also in cyanobacteria (Bacteria; Badger & Gallagher, 1987). Light availability also appeared to be a key component of CCM induction. Recent work (Mitchell *et al*., 2014) showed that algal CCM are induced during the dark to light transition in synchronised cells in *Chlamydomonas reinhardtii*. In addition, works on *Chlorella vulgaris*, *Anabaena variabilis* or *Dunaliella tertiolecta* showed that species when grown under light-limited conditions showed reduce affinity for Ci (Shiraiwa & Miyachi, 1983; Beardall, 1991; Young & Beardall, 2005) without complete repression of CCM capacity.

Interestingly, improving photosynthesis by locally concentrating CO_2 is an adaptation which occurred multiple times under different forms in the history of life (Figure 1.6). There are two forms of biophysical CCMs, which are characterised as a simple compartment within the chloroplast or in cyanobacterial cell. Carboxysomes are found in cyanobacteria, whilst pyrenoids are at the heart of most algal CCMs. The presence of a compartment has the advantage of maintaining a high CO_2: bicarbonate concentration in a very localized part of the chloroplast, saturating Rubisco and protecting CO_2 from leakage. Later on, land plants developed: Crassulacean Acid Metabolism (CAM) and C_4 metabolism, biochemical CCMs analogous to the biophysical CCM. These two metabolisms use two different strategies to improve photosynthesis efficiency. In C_4 metabolism Rubisco is spatially partitioned from the initial PEP carboxylation, whilst in CAM Rubisco is temporally separated (Figure 1.6). Most of the aquatic photosynthetic organisms developed a biophysical CCM, whilst hornworts are the only land plants with a pyrenoid-based CCM (Vaughn *et al*., 1990; Villareal & Renner, 2012). Cyanobacteria have always carboxysomes associated with their CCM, but not all algae have a pyrenoid. The mechanism and the molecular physiology of the cyanobacterial CCM and associated carboxysome structures are better understood than the algal CCM components mainly due the ease of working with a prokaryotic system (Espie & Kimber, 2011; Price, 2011) compared to the algal CCM, despite a greater interest in the model green alga *Chlamydomonas reinhardtii*.

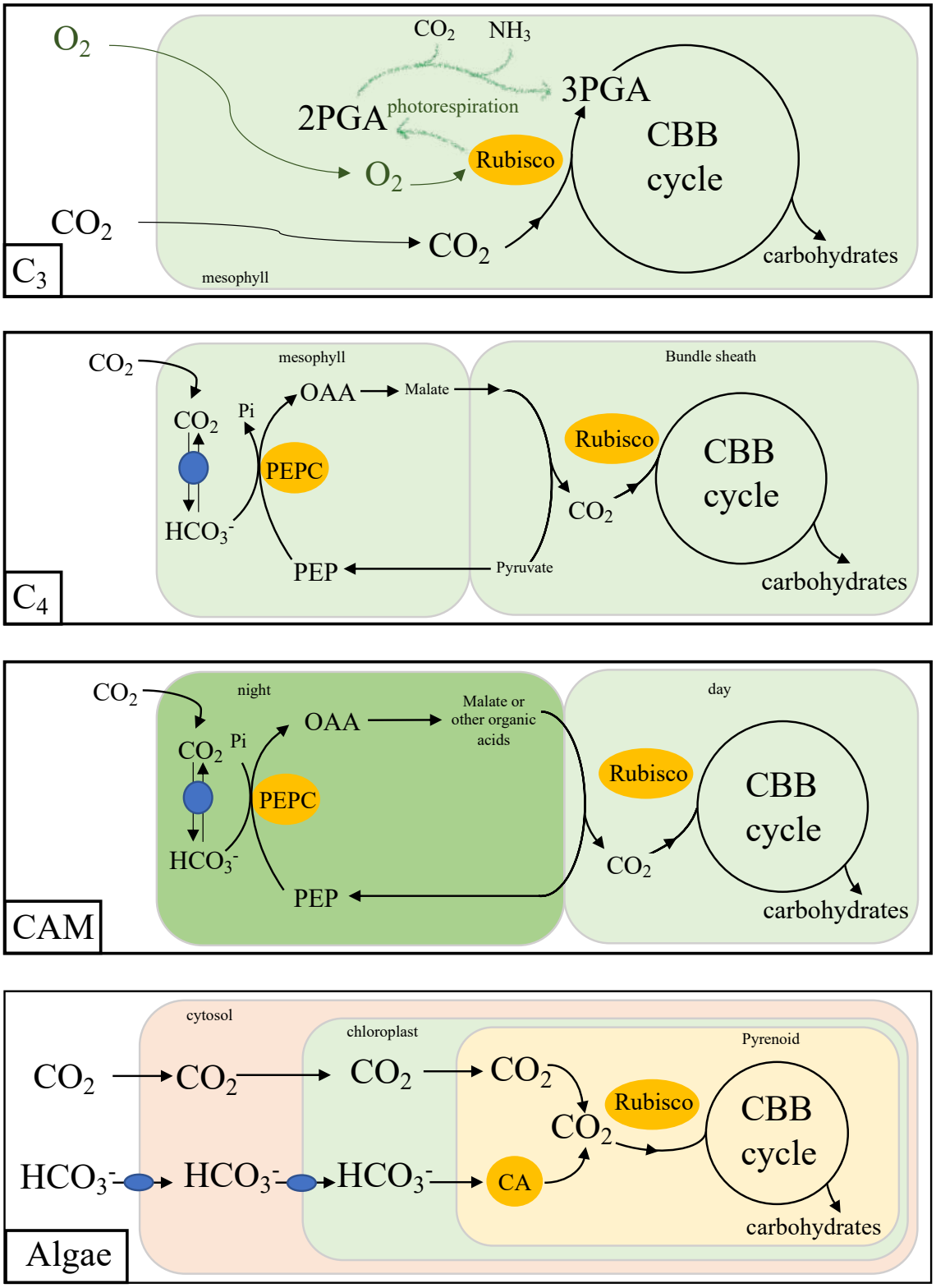


Figure 1.6 The four different photosynthesis pathways found in Viridiplantae. The biophysical CCM found in algae is characterised in the bottom square and aims to maintain a high bicarbonate concentration in a very localized part of the chloroplast saturated in Rubisco and where bicarbonate is protected from any leakage. The carboxysome found in cyanobacteria follows the same principle except that cyanobacteria have no chloroplasts. The two biochemical CCMs found in land plants (C_4 and CAM) are described in the two middle squares.

1.4 The pyrenoid, the algal biophysical CCM

The first observation of a pyrenoid dates back to the middle of the 18th century (drawing of *Conferva jugalis*, now Spirogyra, in Flora Danica, Müller) but the word «pyrenoid» started to be used by Schmitz in 1882. A pyrenoid can generally be visualised as a microcompartment with a dense aggregation of Rubisco, and usually bisected by thylakoid membranes (Figures 1.7-8). Pyrenoids are not delineated by a specific membrane, but some form of starch sheath may provide additional demarcation. Although the presence of a pyrenoid is the marker of the presence of CCM, not all the eukaryotic algae with a CCM have a pyrenoid. The most striking example is the comparison between the *Chlamydomonas* and *Chloromonas* genera, that will be detailed further in this study. Phylogenetically close to each other, the genus *Chloromonas* include species with neither a pyrenoid nor a CCM (*Chloromonas clathrata*), but also species without pyrenoid but with a CCM (*Chloromonas serbinowii*, Morita *et al*., 1998, 1999). However, the algal pyrenoid is a widespread trait found in different type of algae from unicellular, colonial to filamentous species but coastal algae are generally pyrenoid less (at the exception of *Ulva*). Precisely estimating the contribution of the pyrenoid to the global net primary production is difficult because it requires a good estimation of species abundance.

Pyrenoids can be defined as a Rubisco containing micro-compartment within the algal chloroplast that is an integral component of the CCM. When fully induced, the algal CCM involves 3 main components: *i)* a pyrenoid in the chloroplast, which is the site of Ci accumulation and Rubisco aggregation, characterised by a Rubisco matrix traversed by a network of thylakoid tubules, which in most cases is surrounded by a starch sheath (Engel *et al*., 2015). *ii)* inorganic transporters, which actively transport Ci into the cell and *iii)* conversion of accumulated Ci to CO_2 by carbonic anhydrase activity (CA), (Moroney & Ynalvez, 2007; Wang *et al*., 2015; Mackinder *et al.*, 2017).

Pyrenoids and other biochemical CCMs are likely to be the result of convergent evolution meaning that they independently appeared multiple times in different lineages, similarly to the 62 independent origin of C_4 metabolism (Sage *et al*., 2011). In addition, hornwort pyrenoids have been lost at least five times during the last 100 million years (Villareal & Renner, 2012) supporting the theory that CCMs are not inherited from a single common ancestor. It is difficult to estimate when the pyrenoid appeared for the first time (See paragraph CCM) but is probably linked to the limited CO_2 concentrations and diffusive supply in aquatic environments. Badger & Price (2003) suggested that the development of

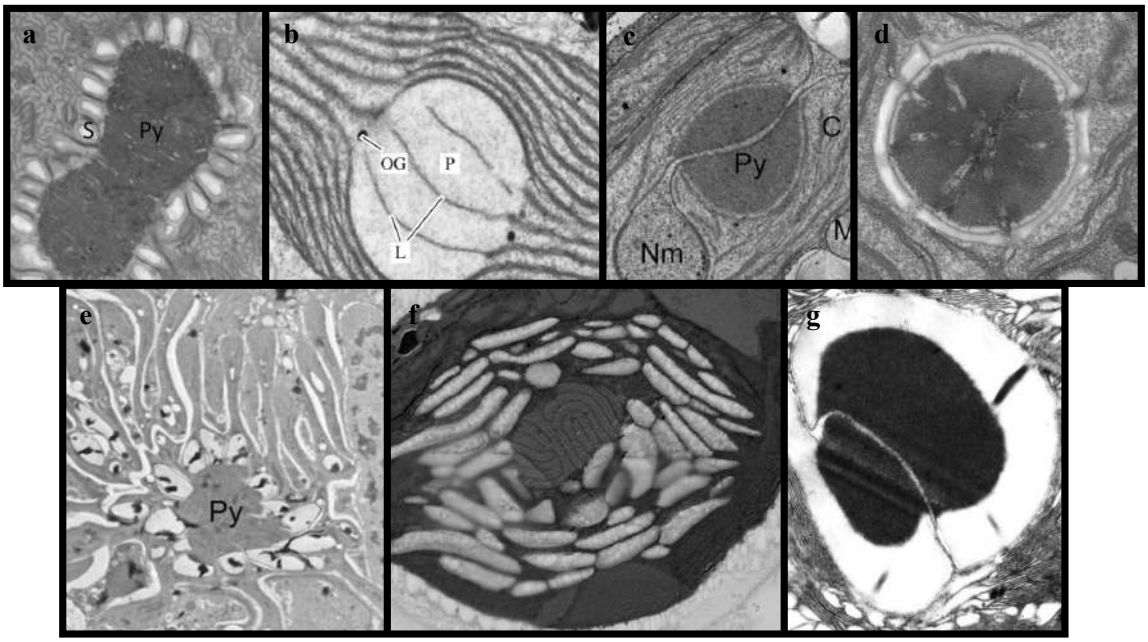


Figure 1.7 Morphological diversity of microalgal pyrenoid matrix and associated network of thylakoid membranes. **a.** Pyrenoid of *Zygnema sp.* (Zygnemataceae; Holzinger *et al.*, 2018). **b.** Pyrenoid of *Aulacoseira baicalensis* (Diatom, Bedoshvili *et al*., 2009). **c.** Pyrenoid of *Viridiuvalis adhaerens* (Chlorarachniophyte, Shiratori *et al.,* 2017). **d.** Pyrenoid of *Chlamydomonas reinhardtii* (Chlamydomonaceae, courtesy Dr. Moritz Meyer). **e.** Pyrenoid of *Penium margaritaceum* (Desmidiales, Raimundo *et al*., 2018). **f.** Pyrenoid of *Chlorokybus atmophyticus* (Chlorokybaceae). **g.** Pyrenoid of *Rhizoclonium ramosum* (Cladophorales, Zhao *et al*., 2016). Py/P= pyrenoid, S= Starch sheath, OG= osmiophilic globules, L= lamella, Nm= nucleomorph, C= chloroplast, M= mitochondrial profile.

pyrenoids in algal lineages with a primary chloroplast occurred before the Carboniferous (359.2 to 299 Mya). Such estimation is also supported by Griffiths *et al.* (2017; see previous paragraph) and by Raven *et al.* (2017). Using a different approach, Raven *et al.* (2017) determined the CO_2 concentration needed to saturate photosynthesis using diffusive CO_2 supply alone and looked for similar conditions in the history of Earth. This estimated period also matches the results found in Griffiths *et al.* (2017).The multiple evolution of the pyrenoid probably explains why pyrenoids are so diverse, with different shapes and morphologies. Single or multiple pyrenoids can be found per cell (Figure 1.7). They are highly variable both across species and have plastic responses to a change in growth conditions. Pyrenoids can be embedded within the chloroplast (Figure 1.7) or projected from a bulge at the inner face of the chloroplast. The simplest pyrenoids are electron-dense proteinaceous aggregates of Rubisco without any external structures («naked pyrenoid», e.g.

Rhodella), but these are relatively rare. More common are pyrenoid matrices traversed by one or multiple membranes derived from thylakoids, typically found in *Chlamydomonas reinhardtii*, the green alga adopted as a study model, where the trans-pyrenoidal membranes form a stellar knot at the heart of the pyrenoid (Figure 1.8A). In recent years *Chlamydomonas* has been extensively used to develop new tools to understand pyrenoid structure and function.

Successive discoveries have allowed the structure of the pyrenoid to be defined (Figure 1.8A): the localisation of the pyrenoid at the base of the chloroplast, the dense matrix of Rubisco and the starch plates were revealed by electron microscopy (Ohad *et al*., 1967; Goodenough & Levine, 1970; Goodenough, 1970). Inorganic pumps such as LCI1 (Low CO_2 inducible membrane protein 1; Yamano *et al*., 2015), HLA3 (plasma membrane localized ABC-type bicarbonate transporter; Ohnishi *et al*., 2010) and LCIA (chloroplast envelope anion channel; Duanmu *et al*., 2009) have been identified which facilitate the transport of bicarbonate (HCO_3^-) across the plasma membrane and the chloroplast envelope and CA convert bicarbonate to CO_2. The network of thylakoids anchors Rubisco to the centre of the pyrenoid and minitubules allow diffusive exchange of carboxylation substrates. Immunolocalization with anti-Rubisco antibody determined that 99% of Rubisco is aggregated into the pyrenoid when CCM is fully induced (low CO_2 conditions) (Lacoste-Royal & Gibbs, 1987; Süss *et al.*, 1995) but also showed the presence of Rubisco activase (RCA1), a Rubisco chaperone which releases tightly bound inhibitors from the catalytic sites of Rubisco (Mckay & Gibbs, 1991; Parry *et al*., 2008). Mackinder *et al*. (2016), using mass spectrometry of isolated pyrenoids induced in low CO_2 conditions, revealed the presence of another protein in the pyrenoid: a putative linker EPYC1, which allows a normal pyrenoid size and morphology and binds to Rubisco. More recently, two studies (Rosenzweig *et al*., 2017; Mackinder *et al*., 2017) showed that the Rubisco-EPYC1 complex acts as a fluid matrix and also that during cell division the pyrenoid divides equally and that in some cases pyrenoid are formed *de novo*. Finally, up to 89 different proteins are implicated in the *Chlamydomonas* pyrenoid (Mackinder *et al*., 2016; Zhan *et al*., 2018) .

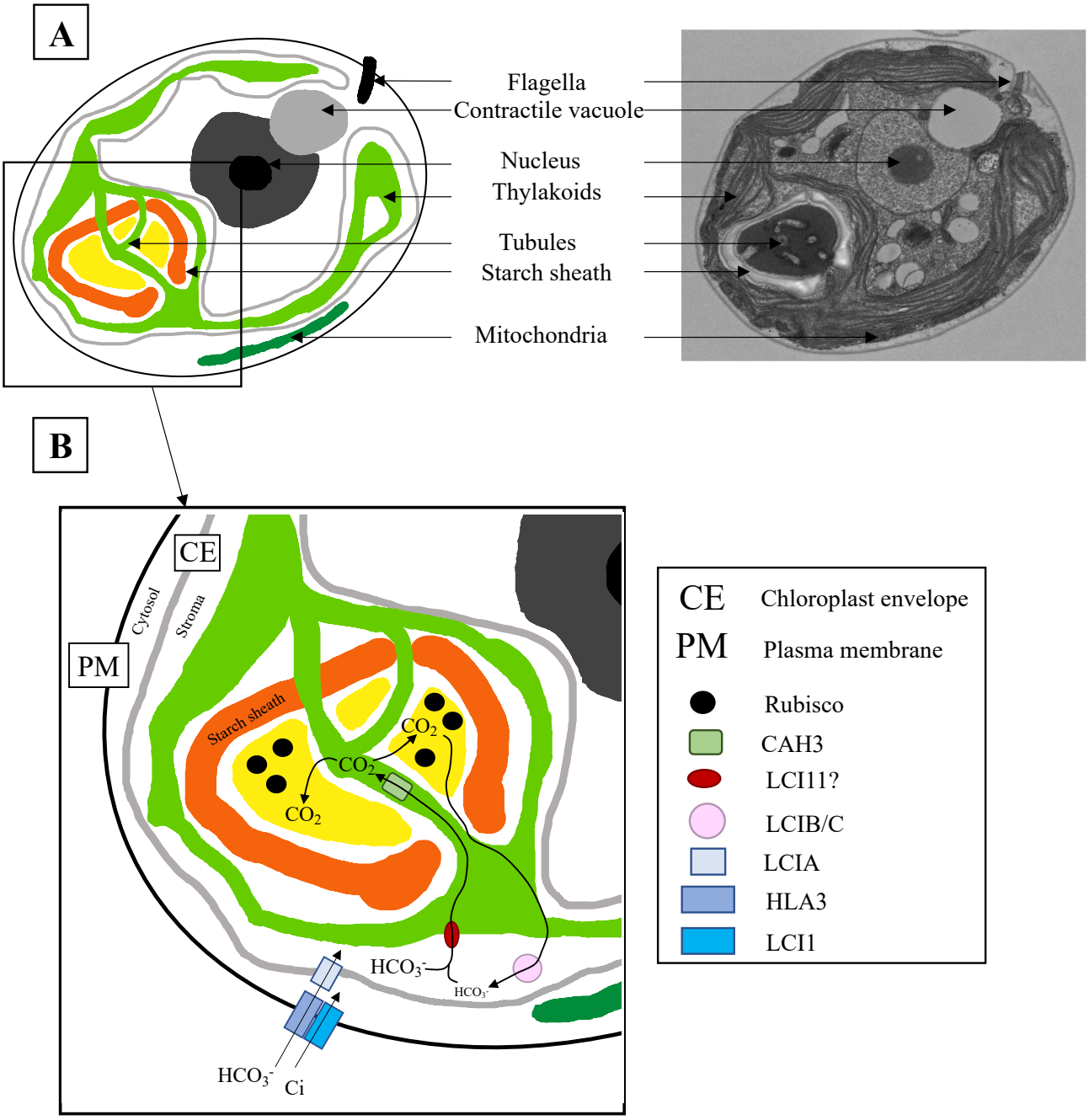


Figure 1.8 A. Schematic and transmission electron micrograph of wild type *Chlamydomonas reinhardtii* grown under low CO_2 conditions (0.04% CO_2). Image courtesy of Dr. Moritz Meyer. **B.** Model of the algal CCM and its mechanism (adapted from Meyer & Griffiths, 2015). CO_2 crosses both the plasma membrane and the chloroplast envelope by simple diffusion. Two carbon pumps allow the transport of bicarbonate (HCO_3^-) through the two different membranes via HLA3 and LCI1 and then via LCIA. Bicarbonate is converted to CO_2 via the carbonic anhydrases CAH3 located inside the lumen of the thylakoids tubules (green). Finally, CO_2 is packed in the centre of the pyrenoid and prevented from leaking with the LCIB/C protein complexes.

1.5 Rubisco: the most abundant enzyme on Earth

Ribulose-1,5-biphosphate carboxylase/oxygenase (Rubisco) is the enzyme of photosynthesis catalysing the first major step of carbon fixation. Rubisco is the most abundant enzyme on Earth (Ellis, 1979; Bar-On & Milo, 2019), indeed it has been estimated that there is 5 kg of Rubisco for every person on Earth (Philips & Milo, 2009). The abundance of photosynthetic organisms on Earth, but most importantly Rubisco inefficiency can explain such values. Despite perhaps 3.6 billion years of evolution, Rubisco still exhibits a slow catalytic rate and a low affinity for atmospheric CO_2. In addition, Rubisco has this dual role of carboxylase and oxygenase, which slows down the photosynthetic process and leads to a competition for the active sites. Consequently, land plants allocate as much as 50% of their leaf nitrogen to Rubisco. However, such statement has been recently challenged with Rubisco's catalytic performance comparison with other chemically related enzymes (Bathellier *et al.,* 2018). This study, supported by other observations made by Cummins *et al.* (2018), not only shows that Rubisco is not especially slow, but also pushes us to reassess the assumptions and knowledges we have on Rubisco.

Table 1.1 Summary of the different Rubisco forms and structures described in the literature.

Rubisco forms	**Rubisco structure**	**Organisms**
I	L_8S_8	Cyanobacteria, most of eukaryotic algae and Embryophytes
II	$(L_2)_{1-8}$	Prokaryotes, dinoflagellates
III	$(L_2)_5$	Archaea
IV (Rubisco-like Protein)	L_2	Proteobacteria, archaea (*e.g. Chlorobium tepidum; Bacillus subtilis,* Tabita *et al*., 2007).

Classically, Rubisco is composed of both large (~55-kDa; LSU) and small (15-kDa; SSU) subunits to form a hexadecameric protein structure (Figure 1.9B). LSUs form dimers and include the active sites. However, four different Rubisco structural forms (I, II, III and IV) have evolved (Table 1.1; Tabita, 1999; Tabita *et al*., 2007). The number of LSU and SSU is the main, but not only, difference between the different forms. Form IV or Rubisco-like Protein, initially thought to be another version of Form III, is mainly found in proteobacteria, archaea and in some algae such as *Chlorobium tepidum* (phototrophic bacterium) or *Bacillus*

subtilis (Tabita *et al*., 2007). It only consists of a dimer of LSUs (L_2 structure) but is incapable of catalysing RuBP CO_2 fixation because of the absence of many essential active sites (Hanson & Tabita, 2001; Tabita *et al*., 2008a). Form III is only found in archaea and is made of five LSU dimers [$(L_2)_5$ structure] (Maeda *et al*., 1999) and probably represents the most ancient form of the enzyme. Interestingly, despite the absence of small subunits, the crystal structure of the form III Rubisco in the methanogenic archaeon *Methanococcoides burtonii* $(L_2)_5$ revealed a 29 amino acids insertion near the C-terminus, which folds as a separate domain in the structure. Located in a similar position to the SSUs in L_8S_8 Rubisco, such insertions would play a role in the assembly process and therefore would be an inbuilt SSU mimic that concentrate L_2 dimers (Gunn *et al.,* 2017). Form II is mainly found in some prokaryotes and dinoflagellates (Morse *et al*., 1995; Tabita, 1999; Whitney & Andrews, 1998) and is made of multiple dimers (from 2 to 8). Finally, Form I of Rubisco is the most widespread and the most abundant form. Present in all cyanobacteria, in most of the eukaryotic algae and all the Embryophytes. It is characterized by a typical L_8S_8 structure: 8 LSUs forming 4 dimers capped by 8 (2 sets of 4) SSUs (Figure 1.9B). In plants, Rubisco biogenesis requires a number of assembly chaperones. Folding of LSUs are mediated by cylindrical a chloroplast chaperonin Cpn60 and a cofactor Cpn20, whilst the final assembly of the L_8S_8 complex, with the addition of the SSUs, includes chaperones such RbcX, Rubisco accumulation factors 1 (Raf11) and 2 (Raf2) and a Bundle sheath defective 2 (BSD2) (Spreitzer & Salvucci, 2002; Wilson & Hayer-Hartl, 2018). In addition, a fully functioning Rubisco molecule is highly dependent on the presence of Rubisco activase. As mentioned in the previous paragraph, Rubisco activase is one the four proteins found at the heart of the pyrenoid. It has two main roles: *i)* activates Rubisco by facilitating carbamylation of Rubisco in the presence of RuBP but also *ii)* relieving inhibition by tight binding inhibitors (Salvucci & Ogren, 1996).

However, four subclasses of Form I Rubisco have been described (Tabita *et al*., 2008). Form IA and IB are found in the green lineage (Cyanobacteria and Viridiplantae) whereas Forms IC and ID are found in the red lineage which includes red algae, Stramenopiles, Haptophyta and Cryptophyta. Form IA and IB have SSUs nuclear encoded whereas both LSUs and SSUs are located in a single operon on the chloroplast genome in the forms IC and ID. Interestingly, pyrenoids are exclusively associated to Form I of Rubisco and Form II in dinoflagellates. In addition, Form I and II large subunits only share 30 % sequence identity (Tabita, 1999). Nowadays, the typical L_8S_8 structure has been crystallised multiple times (Taylor *et al*., 2001; Andersson & Backlund, 2008).

The origin of Rubisco is still unclear. It appears clear that the history of Form I Rubisco is linked to the evolution of oxygenic photosynthesis (Andersson & Backlund, 2008). Studies suggested that Rubisco evolved before the Great Oxygenation Event without having any constraints by O_2 as explained above. The increase of O_2 concentration in the atmosphere following the Oxygenation Event forced Rubisco to learn how to discriminate CO_2 from O_2. The new competition for the active sites appears to be the origin of Rubisco inefficiency. Tawfik *et al*. (2014) showed that the cost of such competition turned Rubisco to reduce its catalytic rate but also to become more specialist and therefore having a slower catalysis.

However from what did Rubisco arise? Three main hypotheses have been suggested: *i)* Rubisco evolved from a non CO_2 fixing ancestor (Rubisco-like Protein) *ii)* Rubisco evolved in a non-autotrophic context *iii)* Rubisco evolved from a simple enzyme complex (Erb & Zarzycki, 2018).

Despite an almost complete similarity between the different Form I Rubiscos, natural variations in kinetic properties have been highlighted between photosynthetic organisms (Jordan & Ogren, 1981). Kinetic properties of Rubisco can be defined by different parameters. The ability of Rubisco to discriminate O_2 from CO_2 is called specificity factor (Ω) and gives the relative rate of carboxylation to oxygenation. Ω is defined by the equation:

Equation 1 $$\Omega = \frac{V_c.K_o}{V_o.K_c}$$

Where V_c and V_o are the maximum velocities for carboxylation and oxygenation and K_c and K_o are the relative Michaelis-Menten constants for CO_2 and O_2 respectively. Rubisco activity and specificity are tightly linked to each other (Tcherkez *et al.,* 2006; Savir *et al*., 2010; Studer *et al*., 2014; Shih *et al*., 2016;). Specificity factor measurements revealed a general increase of the specificity from cyanobacteria to C_3 plants (Cyanobacteria<green algae<C_4<C_3). Cyanobacteria have a low specificity (Ω), with low values (around 40; Jordan & Ogren, 1981) and green algae between 59 and 64 whereas C_3 and C_4 metabolisms exhibit higher Ω with higher values (around 78 for C_4 plants and mid-80 for C_3 plants). The absence of a biophysical CCM in C_3 plants is correlated with high Ω values, whereas the presence of CCMs in C_4 and CAM plants shifts Ω to lower values. Furthermore, Meyer & Griffiths (2013) but also Tortell (2000) and Young *et al*. (2012) suggested that the selection on V_c in RuBisCO has been relaxed due to the saturating CO_2 environment that the CCM provided for such a long period of time (C4<green algae<cyanobacteria). This hypothesis could

explain why there are differences in terms of affinity and specificity for CO_2 between cyanobacteria and land plants (the more recent organisms).

1.6 The small subunit of Rubisco, a neglected subunit.

Although Rubisco is probably one of the most studied enzymes on Earth, the small subunit of Rubisco has often been neglected. Despite being among the first plant nuclear genes cloned and sequenced for expression (Dean *et al.,* 1989), the small subunit remains poorly characterised compared to rbcL. With thousands of sequences on GenBank (Benson *et al*., 2012), rbcL is often used, among other markers, for phylogeny reconstructions. Chloroplast encoded, rbcL is indeed more conserved than the shorter *RbcS*. Two main reasons can explain the lack of interest for the small subunit. First of all, studies showed that Rubisco was perfectly able to fulfil its role of enzyme without small subunit (Andrew, 1988; Lee & Tabita, 1990). Secondly, the absence of a small subunit in Form III of Rubisco confirms that SSUs are not essential in photosynthesis.

If not essential in photosynthesis, where do SSUs come from and what do they look like? With 4 SSUs located on each side of the Rubisco molecule, SSUs are made of four β-strands (A-D) and two α-helices (Figure 1.9A). The β-strands A and B delimit the βA-βB loop, which form the «central solvent» channel with the three others β-loops of the three other small subunits on each side of Rubisco. The diameter of the central solvent channel is strongly linked to the loop length (Karkehabadi *et al.,* 2005; Andersson & Backlund, 2008; van Lun *et al*., 2011). The βA-βB loop length is one of the most variable features between Rubisco enzymes (Esquivel *et al*., 2013). Different loop lengths can be observed across the different photosynthetic organisms. Prokaryotes and non-green algae have a loop length 10 residues long whereas higher plants have a loop length 22 residues long (Figure 1.9B,C). Green algae have the longest known loop with 28 residues in *Chlamydomonas reinhardtii* (Spreitzer, 2003). Comparison of the two Rubisco crystallography structures between *Chlamydomonas reinhardtii* (Taylor *et al.,* 2001; PDB number: 1GK8) and *Spinacia oleracea* (Spinach, Taylor & Andersson, 1997; PDB number: 1RCX) gives the most striking example of this loop length variation (Figure 1.9B,C). With loop length 5 amino acids longer, *Chlamydomonas reinhardtii* shows a very narrow channel whereas the channel of *Spinacia oleracea* is significantly wider (Figure 1.9B). There are several hypotheses regarding the origin of the SSUs. Comparisons of cyanobacteria with

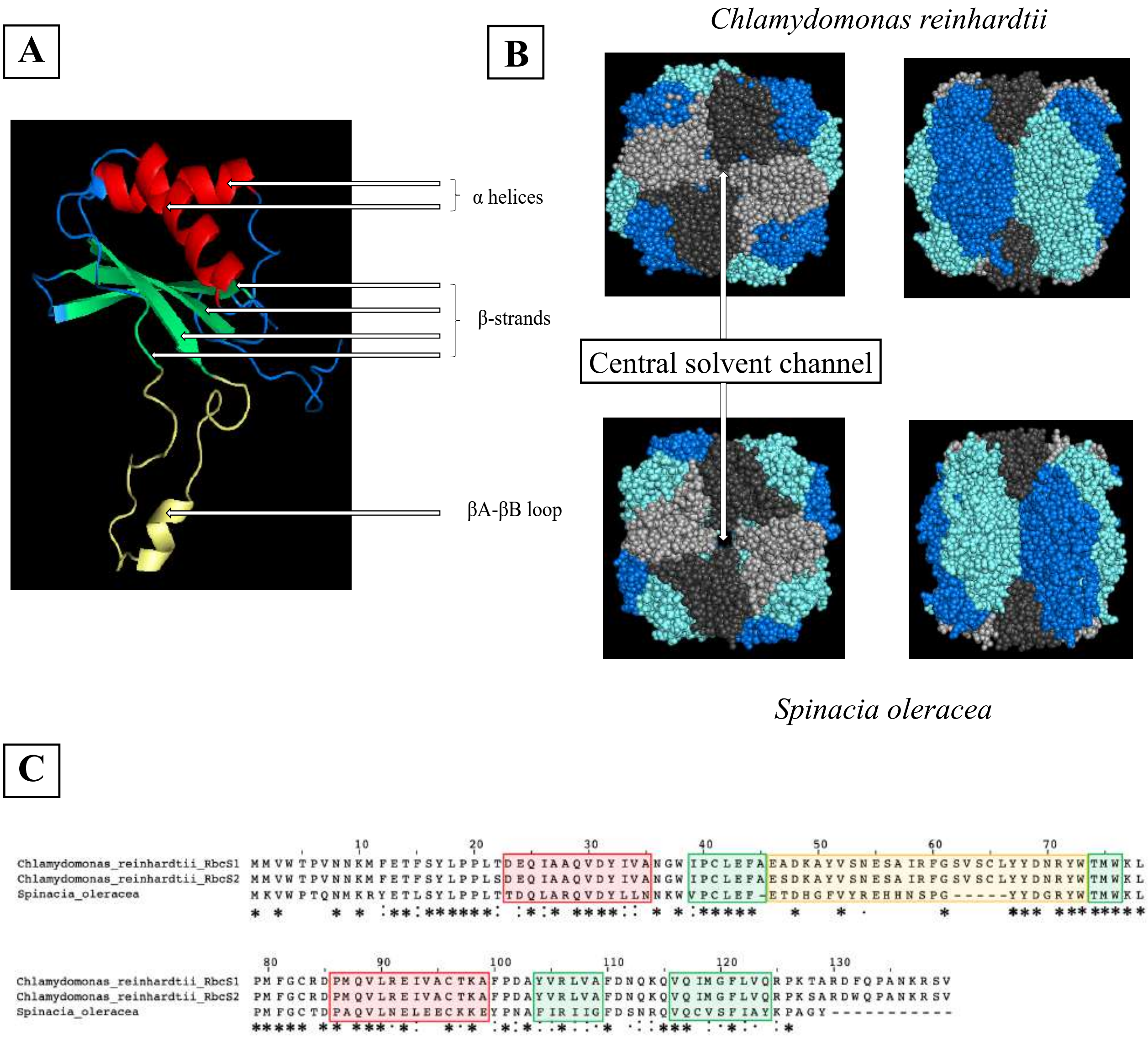


Figure 1.9 The secondary structure of the small subunit of Rubisco and its βA-βB loop length variation between *Chlamydomonas reinhardtii* and *Spinacia oleracea.* **A**. 3D structure of the small subunit of Rubisco in *Chlamydomonas reinhardtii* obtained with PyMOL (DeLano, 2002). The two surface-exposed α-helices A and B coloured in red. The four β-strands are represented by the green arrows. The βA-βB loop is coloured in yellow and includes a short α-helix and forms the central solvent channel with the three other loops of the three other small subunits. **B.** Crystallography structures of *Chlamydomonas reinhardtii* (top) and *Spinacia oleracea* (bottom). LSUs are coloured in dark and light blue and SSUs in dark and light grey. The change of diameter of the central solvent channel is easily observable. Spinacia oleracea exhibits a large channel (bottom left figure) whereas *Chlamydomonas reinhardtii* shows a very narrow channel (top left figure). **C**. Alignments of the two copies of *RbcS* in *Chlamydomonas reinhardtii* and *Spinacia oleracea*. The alignment shows that the diameter of the solvent channel varies due to the absence of 5 amino acids in *Spinacia oleracea* (from 62th to 66th sites: Serine-Valine-Serine-Cysteine-Leucine in *Chlamydomonas reinhardtii).* * indicates positions which have a single fully conserved residues; «:» indicates a site belonging to group exhibiting strong similarity (strong score >0.5); «.» indicates sites belonging to a group from weak similarity (weak score =<0.5).

green algae highlighted sequence similarities between SSU and a possible carboxysomal protein (ccmM gene product) in *Synechococcus* (Price *et al*., 1993). This observation led to thehypothesis that SSUs may have evolved from a protein involved in assembling carboxysomes (Kaplan & Reinhold, 1999). In addition, *RbcS* undertook a transfer from the chloroplastic to the nuclear genome. The reason why such transfers occur are probably multifactorial (Martin & Herrmann, 1998). Retaining genes in organelles permits the regulation of its expression by the reduction-oxidation (chemical reaction in which the oxidation states of atoms are changed) state of its gene products (CoRR hypothesis; Allen & Raven, 1996; Allen, 2017). Another hypothesis suggested that all the genes have the potential to be expressed in the nucleus but that the resulting proteins would be too difficult to be imported across the membranes surrounding plastids (Palmer, 1997; Race *et al.,* 2000) whereas Doolittle (1998) hypothesised that the use of base or codon usage might prevent nuclear expression of some organellar genes, locking them in chloroplasts (Doolittle, 1998; Race *et al.,* 1999). However, relocating genes to the nucleus offers multiple advantages. First of all, it offers economy of resources because a cell with only one genetic system needs fewer genes and resources devoted to protein synthesis. Secondly, gene regulation on the nuclear genome increases the rate of recombination and thus improves the capacity of the organisms to delete all the deleterious mutations. However, the higher rate of recombination of the nuclear genome means that genes are more likely to undergo duplication events, explaining why *RbcS* in now part of a gene family. *RbcS* exists in multiple copies in several organisms. It is encoded 22 times in wheat (Sasanuma, 2001) and *Chlamydomonas reinhardtii* has 2 copies (*RbcS1* and *RbcS2*; Goldschmidt-Clermont & Rahire, 1986) which differ by only 4 amino acids. The presence of multiple copies suggests that selection did not favour any sequences in particular and raise the possibility that all the sequences have their own functions (Spreitzer, 2003).

It is now clear that SSUs have several roles. First of all, many studies on hybrids (Read *et al*., 1992; Kanevski *et al*., 1999), chimeric (Spreitzer *et al.,* 2005; Karkehabadi *et al*., 2005) and mutant SSU enzymes (Kostov *et al*., 1997; Du *et al*., 2000; Spreitzer *et al.*, 2001; Genkov *et al*., 2006;) showed that SSUs could strongly influence both carboxylation catalytic efficiency and Ω. However, how SSUs influence these parameters is still not fully understood as SSU residues are not in direct contact with the active sites. Genkov and Spreitzer (2009) with a mutagenesis approach showed that none of the conserved SSU residues were essential for Rubisco assembly nor function but that many of these residues

affected LSUs catalytic efficiency (identified via so-called suppressor deletions). However, another feature within SSUs appeared to play not only a role in the Rubisco kinetic variation among organisms with a L_8S_8 Rubisco, but also in stabilizing the Rubisco holoenzyme. Understanding the role of the central solvent channel in regulating Rubisco activity has been a recurrent question. The βA-βB loop is known to be important as an assembly domain (Du *et al*., 2000; Flachmann & Bohnert, 1992). In swapping the very conserved Arg53 with a glutamic acid, Flachmann and Bohnert blocked the assembly of pea SSU with LSU in the chloroplast proving that the βA-βB loop was essential for the holoenzyme assembly. The influence of the loop on Rubisco activity and Ω was demonstrated with direct mutagenesis experiments in *Chlamydomonas reinhardtii* (Spreitzer *et al*., 2001) but also in *Synechococcus* (Wasmann *et al*., 1989). The conclusions made by these two studies highlighted the importance of the interface between SSUs and LSUs at the entrance of the central solvent channel, which would contribute significantly to the differences in catalytic properties between algal and land plant enzymes. Despite such studies, the exact role of the channel is still not understood. More recently, the central solvent channel has been thought to partition CO_2 and O_2 to the Rubisco active sites (Esquivel *et al*., 2013).

More than just influencing the Rubisco kinetic properties, SSUs have been shown to play a role in regulating Rubisco aggregations and formation of the algal pyrenoid (Genkov *et al*., 2010; Meyer *et al*., 2012). With a mutagenesis approach, Meyer *et al*., (2012) showed that *Chlamydomonas reinhardtii* mutant with a native LSU and higher plant SSU (known not to have CCM) lost its pyrenoid. Even more than proving that SSUs were important in the pyrenoid occurrence, this showed that the two SSU α-helices were sufficient and necessary for pyrenoid formation. However, the specific residues causing this recruitment and the potential interactions with other partners have not been identified.

1.7 Aims and hypotheses of this study

The work done in Griffiths laboratory in the last 20 years has mainly focused on trying to identify the components that explain pyrenoid occurrence. Understanding the pyrenoid occurrence has only been possible by extensive research on the green algal model, *Chlamydomonas reinhardtii.* Generally, aquatic photosynthesis is the result of pyrenoid occurrence, CCM activity, Rubisco and Rubisco kinetics, which are therefore intrinsically

linked to each other. However, no study to date have tried to combine these different components in a broader context such as the green algae lineage.
In addition, some specific questions remain unanswered. Firstly, following Meyer's work (2012) the residues of the two SSU α-helices involved in pyrenoid occurrence have still not been identified. Secondly, the interactions between Rubisco structure, occurrence of the pyrenoid and Rubisco kinetics have never been fully characterised across chlorophytes and streptophyte algae. In particular, the way Rubisco kinetic properties vary across the ancestral green algae to land plants has not been investigated, and selective pressures on *RbcS* have never been tested. The green algae lineage provides an incredible opportunity to understand CCM and pyrenoid expression in a phylogenetic context since there are natural differences in CCMs among green alga. As previously explained (see paragraph "Pyrenoid" above), among green algae, most of organisms have been shown to express a CCM associated with a pyrenoid, but some have only a CCM and others neither a CCM nor pyrenoid. Such observations raised questions such as: What are the genetic differences between these three states? Does CCM expression link to an interactions between LSUs and SSUs?

Following these different observations, the work completed for this study aimed to give us more insight in our understanding of the evolution of green algae but also on the photosynthetic processes during the transition to terrestrial plants.

Specifically, following this **General Introduction**, a method chapter and four data chapters will be developed as follows:
Chapter 2 describes all the technical and experimental methods used in this dissertation.

Chapter 3 addresses the possible interactions between Rubisco SSU structure and phylogeny, and the occurrence of any reported CCM or pyrenoid across the green algae lineage but also evaluates the selective pressure on *RbcS*.

Chapter 4 defines the co-evolution between Rubisco kinetics and Carbon Concentrating Mechanisms (CCM) across chlorophytes and streptophyte algae through the measurement of key Rubisco kinetic properties and characterization of the CCM for selected members of the streptophyte algae.

Chapter 5 explores the physiological diversity of CCMs in more detail through the two genera *Chlamydomonas* and *Chloromonas* (5 species in total) and sets out how these two genera can be good model organisms to further our understanding of the pyrenoid occurrence.

Chapter 6 develops the findings of Chapter 5 through a comparative analysis of whole genome sequences of the 5 species in order to characterise the genetic components of CCM expression.

Finally, the **General Discussion** combines and draws together the different findings of these chapters in order to summarise the new insights on the evolution of the small subunit of Rubisco interactions with CCM expression, Rubisco kinetics and Rubisco structure.

Chapter 2: Materials & Methods

2.1 Physiological analyses

2.1.1 Growth of algae strains

Six streptophyte algae representing the main streptophytes lineages (Appendix 1) were ordered from the Culture Collection of Algae at Göttingen (Table 2.1,2) (*Chlorokybus atmophyticus*: SAG 34.98 (Chlorokybophyceae), *Klebsormidium subtile*: SAG 384-1 (Klebsormidiophyceae), *Cosmarium subtumidum*: SAG 612-12 and *Onychonema laeve*: SAG 1.93 (Desmidiales), *Spirogyra sp.*: SAG 170.8 and *Coleochaete scutata*: SAG 110.8). The wild-type *Chlamydomonas reinhardtii* strain was kindly provided by Dr. Cindy Chan (Chan, 2018; strain CC-4533, Li *et al.*, 2016). The *Chlamydomonas* strains: *Chlamydomonas augustae* and *Chlamydomonas mutabilis* were ordered from the Culture Collection of algae at The University of Texas at Austin (UTEX; Table 2.1) (UTEX LB 1969 and UTEX 578 respectively). All the *Chloromonas* strains were also ordered from UTEX (Table 2.1): *Chloromonas serbinowii* (UTEX LB 0492), *Chloromonas clathrata* (UTEX LB 1970) and *Chloromonas rosae* (UTEX B 1337).

All the strains were maintained both on solid and in liquid media. All the media were prepared following the algae collection recommendations (Table 2.1, Appendices 2 to 4) except for *Chl. reinhardtii* which was maintained in TAP medium (pH 7.4). For the experiments, which required CCM induction and CCM measurements, strains were directly transferred from liquid media to Tris-phosphate (TP; pH 7.4) medium and bubbled with air from a compressed gas cylinder under two conditions. Induction of CCM was performed by switching gas supply of 5% (v/v) CO_2 in air to ambient air supply (0.04% v/v CO_2).

One litre of TAP medium was prepared with 1 ml of Hutner's trace elements (Hutner *et al*., 1950), 10 ml 2M Tris stock (121.1g Tris Base, 75 ml concentrated HCl made up to 500ml), 10ml phosphate stock (7.17g K_2HPO_4, 3.63g KH_2PO_4 made up to 500 ml), 10 mL of 1 M acetate stock (27.2g $NaAc.3H_2O$ made up to 1L), 40 ml of Beijerinck's Solution (8g NH_4Cl, 1g $CaCl_2.2H_2O$, 2g $MgSO_4.7H_2O$ made up to 1L) and made up to 1L with DIW. TP medium consisted of TAP medium without acetate.

Table 2.1 List of the species names, recommended growth medium and library collection with their associated accession numbers (Ag: on agar) used in this study.

Species name	growth medium	library collection	accession number
Chlorokybus atmophyticus	ES Ag	SAG	34.98
Klebsormidium subtile	ESF1 Ag	SAG	384-1
Cosmarium subtumidum	ESP Ag	SAG	612.12
Onychonema laeve	MiEB12 Ag	SAG	1.93
Spirogyra sp.	MiEB12	SAG	170.9
Coleochaete scutata	ES	SAG	110.8
Chlamydomonas augustae	MB3N	UTEX	LB 1969
Chlamydomonas mutabilis	MB3N	UTEX	578
Chloromonas rosae	MB3N	UTEX	B 1337
Chloromonas clathrata	MB3N	UTEX	LB 1970
Chloromonas serbinowii	MB3N	UTEX	LB 492

Table 2.2 Classification and habitat description of the six streptophyte algae.

Species name	Systematic classification	Habitats	References
Chlorokybus atmophyticus	Chlorokybophyceae, Charophytes	Present in soils, almost exclusively in subaerial habitats	Algaebase
Klebsormidium subtile	Klebsormidiophyceae, Charophytes	Occurs in freshwaters, but can be found in soil and moist substrate	Mikhailyuk *et al*., 2015
Cosmarium subtumidum	Desmidiales, Zygnematophyceae, Charophytes	Periphytic, metaphytic, common in freshwaters, occasionally subaerial or in basic eutrophic freshwater	Taniguchi *et al*., 2003
Onychonema laeve	Desmidiales, Zygnematophyceae, Charophytes	Freshwater	Algaebase
Spirogyra sp.	Zygnematales, Zygnematophyceae, Charophytes	Widespread in freshwater, reported from all continents, occurs frequently in stagnant but aerobic habitats in floating or submerged mats	Algaebase
Coleochate scutata	Coleochaetophyceae, Charophytes	Common in freshwater periphyton	Graham *et al*., 2012

2.1.2 Oxygen evolution for photosynthetic affinity for inorganic carbon

Apparent affinity for inorganic carbon (Ci) was determined by oxygen evolution (Badger *et al*., 1980) and as described in Mitchell *et al*. (2014). All 12 algal strains were grown in Tris-phosphate medium (TP) and cells in log phase were harvested by centrifugation at 3,234 *g* for 5 minutes at 20 °C and resuspended in 1 mL of 25mM HEPES-KOH (pH 7.5). Cells were added to a sealed Clark oxygen electrode chamber (Rank Brothers, Cambridge, UK) attached to a circulating water bath at 25°C, stirred and exposed to white light (200-300 µmol photons $m^{-2}s^{-1}$) to allow consumption of internal Ci. When net oxygen evolution ceased, aliquots of HCO_3^- were added to the cells every 30 seconds. The rate of oxygen evolution was recorded every second using a PicoLog 1216 data logger (Pico Technologies, St Neots, UK). Cumulative concentrations of HCO_3^- after each addition were as follow: 2.5, 5, 25, 50, 100, 250, 500, 1000, and 2000 µM for the cultures grown in low CO_2 condition. Five extra concentrations were added in cultures grown under high CO_2 condition in order to reach the maximum rate of oxygen evolution (2500, 3000, 4000, 4500 and 5000 µM). The Excel spreadsheet created for the PhD work of Dr. Madeline Mitchell (Mitchell, 2014) and SigmaPlot (Appendix 5, Systat Software, San Jose, CA) were used to fit the Michaelis-Menten kinetics equation to the curves of external inorganic carbon versus photosynthetic rate. In order to support the presence of a CCM in low CO_2 conditions, Student's t-tests (Student, 1908; Appendix 6) were performed to statistically test for the differences between low and high CO_2 results but also the different affinities between the different species grown in the same condition. Finally, to allow comparison between the different strains, O_2 evolutions (y axis) were transformed in maximum photosynthetic rate (%) where the maximum velocity rate was considered as 100% of the photosynthetic capacity.

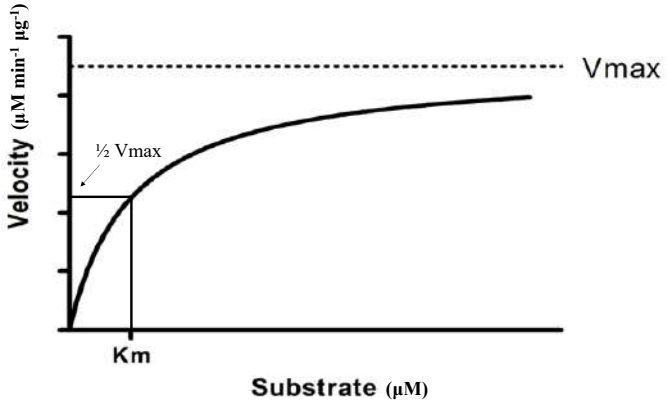


Figure 2.1 Michalis-Menten saturation curve for an enzyme reaction showing the relation between substrate concentration and reaction rate.

2.1.3 Chlorophyll extraction and measurements

Chlorophyll concentrations were measured for the normalization of oxygen evolution measurements. Following measurements of oxygen evolution, cells were harvested and resuspended in 1 mL of 100% methanol. After centrifugation (1 min, max speed), the concentrations of chlorophylls *a* and *b* were determined by measuring absorbance at 470, 653, 665, 666 and 750 nm with a spectrometer (UV 300 UV-visible spectrometer, Spectronic Unicam) and obtained using the equation of Porra *et al*. (1989).

2.1.4 Determination of isotopic ($\delta^{13}C$) for composition of organic matter

All the analyses were performed at the Godwin Laboratory for Paleoclimate Research at the University of Cambridge. Algae cultures were harvested by centrifugation at 3,234 *g* for 5 minutes at 20°C (Eppendorf, Centrifuge 5804 R), resuspended in 0.1M HCl to remove inorganic carbon and washed several times with deionized water (DIW). Samples were stored overnight at -80°C and then dried in a freeze drier for 24 hours. Samples where then weighed (0.5 mg) in triplicate into 3mm x 5mm tin capsules (Experimental Microanalysis Ltd., Okehampton, UK). The results were reported with reference to the international standard VPDB with a precision better than +/- 0.08 per mil for $^{12}C/^{13}C$.

2.2 Microscopy methods

2.2.1 Fixing and embedding for pyrenoid morphologies

Pyrenoid morphologies were examined using blockface imaging by Scanning Electronic Microscopy (SEM). All 12 algal strains were grown in Tris-phosphate medium and bubbled with ambient air. Cells in log phase were harvested by centrifugation at 3,234 *g* for 5 minutes at 20°C. Sample preparation and imaging were undertaken at the Cambridge Advanced Imaging Centre (CAIC) with Karin Müller. Algal cells were collected by centrifugation (1,000 ×*g*, 5 min with slow deceleration, at room temperature), resuspended and fixed in 1.5 mL of 2% glutaraldehyde, 2% formaldehyde in 0.05 M sodium cacodylate buffer (pH 7.4, containing 2 mM $CaCl_2$) at 4°C overnight. Samples were then washed five times with 0.05 M sodium cacodylate buffer (pH 7.4) and osmicated for 3 days at 4°C in 1% osmium tetroxide (OsO_4), 1.5% potassium ferricyanide ($K_3Fe(CN)_6$) and 0.05 M sodium cacodylate buffer (pH 7.4). Cells were treated with 0.1% thiocarbohydrazide for 20 min at room temperature in the dark, then osmicated a second time for 1 h at room temperature in 2% of

$OsSO_4$ (in DIW) and treated with bulk stain (2% uranyl acetate in 0.05 M maleate buffer pH 5.5) for 3 days at 4°C. Samples were washed 5 times in DIW between each of the last three steps. Two dehydration steps were then undertaken. Firstly, the samples went through a series of ethanol solutions of 50%, 70%, 95% and 100% (3x for 5 min in each). Secondly, samples were dehydrated twice in 100% dry ethanol, twice in 100% dry acetone and 3 times in 100% dry acetonitrile for at least 5 min for each step. Samples were mixed in an equal volume of Quetol resin mix [12 g Quetol 651, 15.7g Nonenyl Succinic Anydride (NSA), 5.7 g methyl-5-Norbornene-2,3-Dicarboxylic Anhydride (MNA)] and 100% dry acetonitrile overnight. Samples were incubated in pure Quetol resin mix for 3 days, then samples were exchanged into fresh Quetol resin mix [containing 0.5 g Benzyldimethylamine (BDMA)] every day. Finally, cells were spun for 10 min at 13,000g in the resin and were put into a curing oven at 60°C for 48 hours. Samples were then removed from the Eppendorf tubes using a hacksaw and mounted on small aluminium SEM stubs using conductive epoxy resin. Resins were hardened at 60°C for 30 min. Stubs were sputtered with 35 nm gold using an EmiTech-Quorum sputter coater (EMITECH-Quorum Technologies, Kent, UK). A Leica Ultracut ultramicrotome (Leica, Austria) was then used to smooth the blockfaces. Stubs were sputtered again with 30 nm carbon using a Quorum 150T carbon coater. Finally, blockfaces were imaged using a FEI Verios 460 scanning electron microscope run at 4 keV accelerating voltage/0.2 nA probe current using the concentric backscatter detector (10 □s dwell time, 2 line integrations) and a working distance of about 3.5 mm.
Maps were imaged overnight using FEI MAPS software using a pixel resolution of 1536 x 1024, a horizontal field width of 15.9 μm/tile, an x-y tile overlap of 15 %/20 % and the MAPS default stitching profile.

2.2.2. Image analyses

Cell morphologies were measured with ImageJ (Schrödinger, 2010) and measurements were treated with Microsoft Excel.

2.3 Molecular biology

2.3.1 Genomic DNA extraction

All the strains were grown in their recommended medium (Table 2.1) and bubbled with ambient air. Cells in log phase were harvested by centrifugation at 3,234 *g* for 5 min at 20°C

and resuspended in 500 µL of CTAB Extraction Buffer (2% w/v CTAB, 100 mM Tris-HCl pH=8, 1.4M NaCl, 20mM EDTA-Na_2 pH=8 and 2% v/v β-mercaptoethanol). 500 µL of phenol:chloroform:isoamylalcohol (25:24:1) (Sigma-Aldrich) was added to the different samples after incubation (1 hour at 65°C). New mixes were then vortexed and spun down at 12,000 rpm for 20 min. Supernatants were transferred to new 2 mL Eppendorf tubes and DNA was precipitated by adding 0.7x the volume of supernatant with isopropanol to each sample. DNA precipitation was collected by centrifugation at maximum speed for 15 min at 4°C. Supernatants were then discarded, pellets washed with 500 µL of 80% ethanol and samples spun down at 8,000 *g* for 5 min at 20°C. After drying, pellets were resuspended in 30 µL of di-ionized water (DIW).

2.3.2 Genomic DNA quantification

Quantification of DNA was made with Qbit® Fluorometer (Life Technologies, Carlsbarg,CA). Two ratios of absorbance (260/280nm and 260/230nm) were measured with Nanodrop (Thermo Scientific NanoDrop Products) to assess the purity of DNA.

2.4 Biochemistry

2.4.1 Rubisco purification

Strains were cultured in an incubator shaker (Incubator shaker, Innova 42, New Brunswick Scientific) in their recommended medium within large, 2 L flasks, under constant light at room temperature and bubbled with ambient air. Due to the low concentration of Rubisco in algae (Losh *et al*., 2013; Valegård *et al*., 2018) a minimum of 30 g wet paste per sample was cultured in order to have enough material for the following experiments.

Algal cells were broken using an Emulsiflex-C5 high pressure homogenizer (Avestin Inc., Ottawa, Canada) kindly loaned by Biopharma Group (Winchester, UK). Cell pastes were suspended in *ca.* 200 mL of extraction buffer containing 10 mM $MgCl_2$, 50 mM Bicine, 10 mM $NaHCO_3$, 1 mM DTT, 1 mM ε-aminocaproic acid, 1 mM benzamidine, 0.1 M phenylmethylsulfonyl fluoride, and 200 µL of protease inhibitor cocktail (Sigma, UK). Total soluble proteins were extracted via centrifugation at 22,000 ×*g* for 12 minutes (min) at 4°C (with an Avanti centrifuge, Beckman-Coulter). After this initial centrifugation step, PEG_{4000} (60% w/v) and 1 M $MgCl_2$ were added to the supernatant and the rest of the purification carried out as described previously (Orr & Carmo-Silva 2018). Peak fractions containing

Rubisco (based on CABP binding [Sharwood *et al*., 2016]) were concentrated using Amicon Ultracel-15 concentrators (100 kDa MWCO, Merck-Millipore, UK). Aliquots were snap-frozen in liquid nitrogen and stored at -80°C.

2.4.2 Rubisco catalytic properties

Rubisco activity for the six streptophyte algae was determined by incorporation of $H^{14}CO_3$ into acid-stable products at 25°C and pH=7 as described in Prins *et al*. (2016) with some modifications. Purified Rubisco was diluted using desalting buffer (Orr & Carmo-Silva, 2018) and then desalted using a G-25 MidiTrap column (GE Healthcare, UK). Samples were allowed to activate on ice for 45 mins prior to assaying. All the purified Rubisco were assumed to be functionally active. Carboxylation activity was measured at nine different concentrations of CO_2 (8, 16, 24, 36, 68, 100, 180, 280 and 400 μM) and with O_2 concentrations of 0 and 21%. In order to ensure that the activity measured was entirely due to Rubisco, three controls were performed: CO_2 fixation (acid-stable ^{14}C) was measured in reaction solutions lacking RuBP or $NaHCO_3$, and following total inhibition of Rubisco by prior treatment with an excess of the tight-binding inhibitor 2-carboxyarabinitol-1,5-bisphosphate (CABP). Radioactive content of ^{14}C-labelled compounds was measured in 0.4 ml aqueous solutions to which were added 3.6 ml Gold Star Quanta Scintillation cocktail (Meridian Biotechnologies, UK), in a Tri-Carb 2250 CA Liquid Scintillation Analyser (Perkin-Elmer, USA). Turnover number (k_{cat}: mol product mol active site^{-1} s^{-1}) was calculated from the corresponding V_{max} value (V_c : μmol acid-stable ^{14}C mg Rubisco^{-1} min^{-1}).

2.4.3 Rubisco quantification

Rubisco quantification was via [^{14}C]CABP binding assay as described Sharwood *et al*. (2016). Rubisco was incubated for 25 min after adding [^{14}C]CABP. Each quantification was performed in duplicate. Radioactive content of ^{14}C-labelled compounds was measured using scintillation counting as described above.

2.5 Bioinformatic analyses

2.5.1 *RbcS* analyses

2.5.1.1 Data collection

All the *RbcS* sequences used in this research were kindly been provided by "The 1000 plants project" (1KP; Leebens-Mack *et al*., 2019; Carpenter *et al*., 2019). Data consisted of two different data sets (DNA and protein sequences) and were treated in two phases. The main phylogeny of *RbcS* was firstly based on the protein sequences, whilst the DNA sequences were used for the analyses of selection. The protein data set consisted of 2,674 protein sequences, including 239 species of algae (171 green algae, 28 red algae, 35 Chromista and 5 Glaucophyta) collected in the field, provided and extracted by Michael Melkonian (University of Cologne).

2.5.1.2 Multiple alignment of *RbcS* protein sequences

Only the sequences belonging to the green algae were analysed. The two copies of *RbcS* of *Chl. reinhardtii* and one sequence of *Coccomyxa subellipsoidea* available on GenBank were also added to the dataset. All the protein sequences were manually and individually screened. Sequences showing cross-contamination (Carpenter *et al*., 2019), or which were too short or incomplete, were removed. The dataset did not allow to unambiguously identify *Rbc*S isoforms. Although it is generally taken that all photosynthetic members of the Viridiplantae have multiple copies of the *RbcS* gene, conservatively only one sequence was used in the analysis for each species, except when the data was sourced from independently sequenced genomes (e.g. for *Asteromonas*). A total of 187 protein sequences belonging to 113 species (31 streptophyte algae, 10 prasinophytes, 72 chlorophytes) were then aligned with Clustal Omega (Sievers *et al*., 2011; Appendix 7).

2.5.1.3 Selection of the best-fit models of evolution

To obtain a robust gene phylogeny, the most appropriate model of evolution (which best fits the data) needs to be identified. Models of evolution can also be called models of substitutions and indicate the probability of change from a given amino acid/nucleotide to another. The selection of the best model is made based on the estimation of the quality of each model relative to each of the other models (Bozdogan, 1987). Three different ways are used to estimate the quality of the models: The Akaike Information Criterion (AIC), the

Akaike Information Criterion C (AICc; which is a simple correction of the previous model but for finite sample sizes) and finally the Bayesian Information Criterion (BIC). ProTest v2.4 (Abascal *et al*., 2005) was used to select the best model of protein evolution.

2.5.1.4 Protein phylogeny reconstruction

BEAST 2 provides a platform for Bayesian phylogenetic analysis of molecular sequences (Bouckaert *et al*., 2014). It does not require an outgroup but needs strong prior parameters for the evolutionary model, the molecular clock, the model of speciation and the rate of evolution. The evolutionary model describes the different probabilities of change from one nucleotide to another, the molecular clock uses the mutation rate of nucleotides/amino acids to estimate when species diverged and thus influences branch lengths in the trees. The models of speciation make predictions about the shape of the phylogenetic tree connecting extant species (Steel & McKenzie, 2001). The rate of evolution is also an important parameter because it varies due to different selective constraints that are acting on the different sites (Stamatakis, 2014), consequently the usual phylogenetic approach is to consider a rate of evolution heterogeneous among sites.

The rate of evolution for each site is modelled as a random variable drawn from a specified prior distribution. The most common model used to describe these variations is the gamma distribution (where 4 rate categories provide accurate approximations for dataset of medium size). Finally, BEAST 2 uses the Markov Chain Monte Carlo (MCMC) method to reconstruct phylogenies. MCMC is a simulation algorithm which explores all the tree possibilities given the prior parameters (Yang & Rannala, 2012). Consequently, the chain parameters also need to be set.

The original input file was created with Beauti v2.3.1 (Drummond & Rambaut, 2007) using a LG model of evolution (Le & Gascuel, 2008) with frequencies of transversion and transition all equal, a gamma distribution model with 4 categories, a relaxed molecular clock with rates for each branch drawn from a log normal distribution and a Yule model of speciation with a random starting tree. The Yule model is the simplest stochastic model of speciation and assumes that at any time, each of the extant species are equally likely to give rise to one species, or, in contrast, are equally likely to go extinct. Three independent chains were run, each of length $8x10^7$ with parameter values sampled every $10x10^2$ steps. Chain convergences were checked using Tracer v1. (Drummond & Rambaut, 2007). Posterior parameters were summarised with Tree Annotator v1.8.2 (Drummond & Rambaut, 2007)

using a maximum clade credibility (MCC) tree, a posterior probability limit of 0.5 and a mean node height. Finally FigTree v1.4.2 (Rambaut, 2007) was used for tree visualizations.

2.5.1.5 Phylogenetic analyses

2.5.1.5.1 Systematic analyses

Terminal branches were coloured as a function of their divisions (chlorophyta, prasinophytes and streptophyte algae) according to the 1KP classification (Mickael Melkonian; http://www.onekp.com/samples/list.php?set=algae) in order to see how the species were distributed across the phylogeny of *RbcS*.

2.5.1.5.2 Scoring for pyrenoid presence/absence

Scoring for pyrenoid/presence of the pyrenoid in the different species of the phylogeny of *RbcS* was based on bibliographic analyses. Presence of a pyrenoid was either confirmed by electron microscope image or simple description in articles (Appendix 8). The absence of pyrenoid was often simply mentioned in the species description. However, for some species the diagnostic remained unknown due to the absence of description in the literature.

2.5.1.5.3 Scoring for βA-βB loop length

Loop length was determined based on the multiple alignment of the protein sequences. The number of residues in the loop was counted according to Spreitzer (2003) and mapped on the phylogeny of *RbcS.*

2.5.1.6 Tests for selective pressure on *RbcS*

2.5.1.6.1 Analyses of positive selection

2.5.1.6.1.1 Theory

Positive selection can be defined by the relative rates of synonymous and non-synonymous substitutions (Miyata *et al*., 1979; Li *et al*., 1985). Synonymous substitutions are substitutions (d_S) of one base by another that does not alter the amino acid sequence after translation, whereas non synonymous substitution (d_N) alters the amino acid sequence. Therefore, substitution rate is a function of selective pressure on the proteins and is defined

by the ratio between d_N and d_S (d_N/d_S) and can also be called omega (ω). When, ω is less than 1 (ω<1), the selection is defined as «purifying selection» and removes the amino acids with either a direct negative impact on the protein. In contrast when ω>1, mutations are advantageous and will be fixed at a higher rate than synonymous mutations. Finally when, ω=1 evolution is considered as neutral.

2.5.1.6.1.2 Identification of residues under positive selection

To test the importance of two SSU α-helices for pyrenoid formation in *Chl. reinhardtii* (Meyer *et al*., 2012), the Codon-based package (codeml) implemented in PAML v4.9 (Yang & Nielsen, 1998; Yang, 1998, 2007) was used to detect residues under positive selection across the green algae lineage. In addition, the presence of a CCM is not universal across the green algae so the branch model also implemented in PAML was used to detect branches under positive selection. The DNA phylogenetic tree was reconstructed using BEAST v2.3.1 with 135 cDNA *Rbc*S sequences of green algae from the 1KP, with a GTR model of protein evolution (Tavaré, 1986) and the same gamma distribution, molecular clock and model of speciation previously used. Three independent chains were run, each of length $5x10^7$ steps, parameters values and trees were sampled every $10x10^2$ steps. Chain convergences, posterior parameters and tree visualization were analysed with the same method explained above. All the analyses were run using "user tree" run mode, meaning that both the multiple alignment and the phylogenetic tree were used to test for positive selection. Several models of codon evolution that allow for variations in ω (d_N/d_S) among codons were tested (Site models) using Likelihood Ratio Tests (LRTs) (Neyman & Pearson, 1928) and as described in Kapralov & Filatov (2007). Positive selection was evaluated by contrasting a null model (H0) that does not allow variations between sites with a more general model (H1) where this condition is allowed (Pie, 2006). The level of significance was assessed using the likelihood ratio statistic [calculated by multiplying twice the difference in likelihood scores (2ΔlnL) of each model]. This ratio was then compared with a $\chi 2$ distribution with the number of freedom (which is the number of values in the final calculation and which can vary during the analysis). The different degrees of freedom (df) were calculated by the difference in the numbers of parameters in the two tested models.

For this research, residues under positive selections were detected using three different LRTs: M0-M3, M7-M8 and M8a-M8 each fitting different number of parameters. The first model comparison aimed to show variations between sites where M0 had 1 ratio fitting a

single ω_0 averaged over all sites (H0) and where M3 fit 3 discrete classes of sites, each with different ω_0 (H1). M7-M8 and M8a-M8 aimed to detect residues under positive selection where M7 (model considering beta-distributed selective pressures on the sites and allows 10 site classes each with $\omega<1$) and M8a (β and $\omega=1$, H0) were the null models. Both models were compared to M8 (β and $\omega>1$, H1), which allows 11 sites classes, one of which allows for $\omega>1$. An equilibrium codon frequency in codon substitution model was assumed from the average nucleotide frequencies at the three codon positions. Kappa (transition/transversion rate ratios) and ω were estimated but α was fixed. Positions with a probability superior to 0.9 were considered to be under positive selection. Bayes empirical Bayes (BEB) (Yang *et al*., 2005) was used to calculate posterior probabilities for site classes and to identify sites under positive selection in the cases where likelihood ratio tests were significant according to PAML manual (Yang, 2007).

2.5.1.6.1.3 Identification of branches under positive selection

Branch models were used to test for positive selection across branches. The null model allowed for variations in ω among branches ($0<d_N/d_S<1$ and $d_N/d_S=1$ for both foreground and background branches) and also included two additional classes of codons with fixed $d_N/d_S=1$ on foreground branches but restricted as $0<d_N/d_S<1$ and $d_N/d_S=1$ for background branches. The alternative model allowed $0<d_N/d_S<1$ and $d_N/d_S=1$ for both foreground and background branches but also included two additional classes of codons under positive selection with $d_N/d_S>1$ on foreground branches with restriction as $0<d_N/d_S<1$ and $d_N/d_S=1$ on background branches. Branches leading to species without pyrenoid were labelled as foreground branches (allows positive selection) and the rest of the branches were considered as background branches (with no positive selection). The level of significance was tested as described above.

2.5.1.6.2 Analyses of relaxed selection

Relaxation of selective strength is characterised by a reduction in the efficiency or intensity of natural selection and can lead to evolutionary innovation but also to lineage extinction or loss of function (Wertheim *et al*., 2014). In order to test for relaxed selection in *RbcS*, the RELAX program (Wertheim *et al*., 2014) implemented in Hyphy (Pond & Muse, 2005) was used. RELAX aims to test whether the strength of natural selection has been relaxed or intensified among specific sets of branches. Similar to PAML, it compares two different

models (H0 and H1) where H0 includes the reference branches and H1 the tested branches. RELAX introduces a new parameter called k (where $k \geq 0$), which describes the relation between the reference and tested branches. H0 constrains $k=1$, whereas H1 set k as a free parameter. Then, LRT is used to compare the two models. If k is greater than one, then selection has been intensified along the tested branches whereas $k<1$ indicates relaxed selection along the tested branches.

Relaxed selection was tested using the same nuclear phylogenetic tree of *RbcS* previously obtained for PAML analyses. Five different tests were performed (Figure 2.2) ten times each to avoid false positive results. In the first test, branches including all the streptophyte algae were labelled as test branches (H1) and the rest of the tree considered as reference branches (H0). The second test tested for the opposite, with all the branches leading to chlorophytes considered as test branches (H1) and the rest of the tree, reference branches (H0). The next three tests were performed on the basal branches of the tree. The third test was performed on the basal branches of the tree leading to the two main clusters (H1) with the rest of the tree was labelled as reference (H0). Finally, the fourth and fifth tests were performed on the same branches, but branches were tested individually.

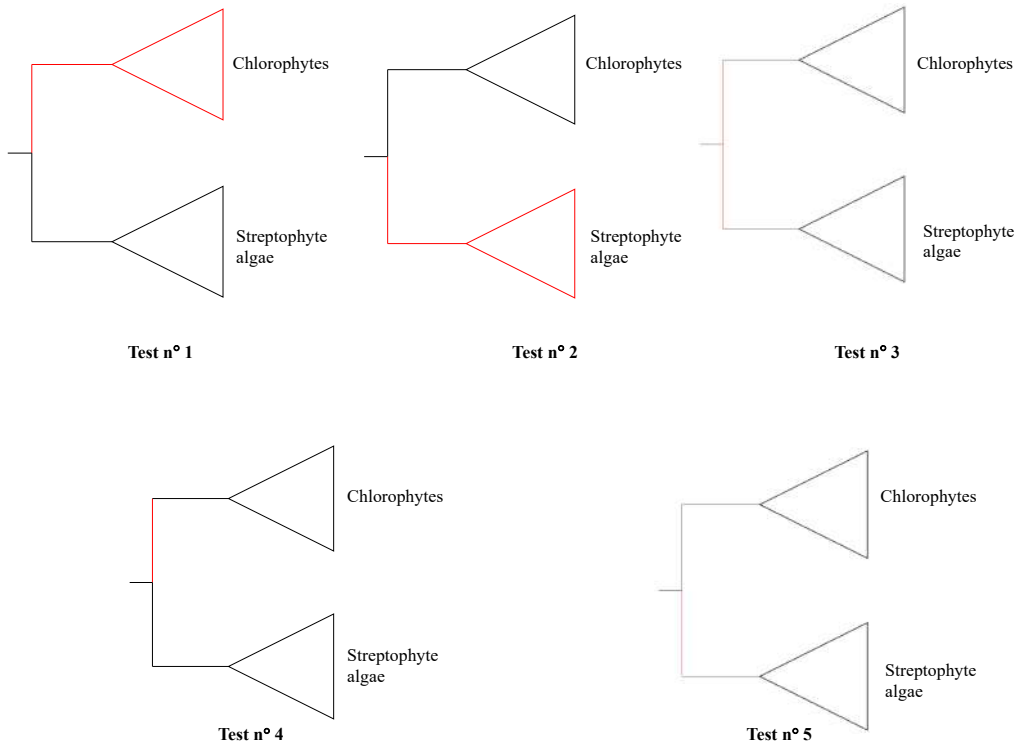


Figure 2.2 Illustration of the five different tests performed with RELAX (Wertheim *et al.*, 2014) on the phylogeny of *RbcS* (simplified for this figure). Red branches represent the tested branches (H1) and the black branches were considered as reference branches (H0).

2.5.2 Whole genome sequencing

2.5.2.1 Re-sequencing

Previously extracted genomic DNA (see paragraph 2.3.1) was re-sequenced by BGI using BGISeq technology. Paired-end libraries were prepared using BGISEQ library construction. Read lengths were 100 bp long. Resulting reads were assembled using SPAdes (Bankevich *et al*., 2012). *K*-mer frequency was set to 55, 77 and 99.

2.5.2.2 *De novo* sequencing

Genomic DNA was extracted as explained above and *de novo* sequenced by the Norwegian Sequencing Center, CEES in Oslo using PacBio SMRT cell. Long reads were assembled using wtdbg2 (Ruan & Li, 2019).

2.5.2.3 Hybrid assembly

Both sequencing technologies were combined in order to obtain new *de novo* assembly including both long and short reads. The combination of both technologies increases the accuracy of the genomes. Hybrid assemblies were obtained using Burrows-Wheeler Aligner (BWA; Li & Durbin, 2009). The quality of the new assemblies was assessed with QUAST (Gurevich *et al*., 2013.

2.5.2.4 Chloroplast genome reconstruction

Annotations of the five genomes is ongoing, therefore, in this study limited analyses were conducted on the new genomes. In addition, all the following analyses were made on the hybrid (combination of the short and the long reads) genomes. Despite being closely related, the chloroplastic genome of *Chl. reinhardtii* could not be used to reconstruct the full chloroplast genome of the five new algal strains and therefore only coding sequences (CDS) were extracted. Nodes containing the different CDS were identified using BLAST (Kent, 2002) based on the CDS genes of *Chl. reinhardtii* (NC005353) and *Chloromonas perforata* (KT625416). CDS genes from the five new strains were then manually extracted using Geneious (Kearse *et al.,* 2012).

Table 2.3 List of the green algae species used to reconstruct the chloroplastic phylogeny of green algae, their systematic classification, their accession numbers on GenBank and their pyrenoid diagnostic. Species without pyrenoid are highlighted in light grey.

Species name	Systematic	Accession number	Pyrenoid presence/ absence	References
Arabidopsis thaliana	Land plant	NC000932	Presence	
Bathyococcus prasinos	Prasinophytes	FO082259	Absence	Eikrem & Throndsen, 1990
Botryococcus braunii	Chlorophytes	NC_025545	Presence	Wolf & Cox, 1981
Bryopsis plumosa	Chlorophytes	NC_026795	Presence	Ogawa, 1988
Chaetosphaeridium globosum	Streptophyte algae	AF494278	Presence	Moestrup, 1974
Chara vulgaris	Streptophyte algae	DQ229107	Presence	
Chlamydomonas augustae	Chlorophytes	WGS of this study	Presence	See Chapter 4
Chlamydomonas mutabilis	Chlorophytes	WGS of this study	Presence	See Chapter 4
Chlamydomonas reinhardtii	Chlorophytes	NC_005353	Presence	Meyer *et al*., 2012; Mackinder *et al*., 2016
Chlorella mirabilis	Chlorophytes	NC_025528	Presence	Dempsey *et al.,* 1980
Chlorella sorokiniana	Chlorophytes	KJ742376	Presence	Dempsey *et al.,* 1980
Chlorokybus atmophyticus	Streptophyte algae	DQ422812	Presence	See Chapter 3
Chloromonas clathrata	Chlorophytes	WGS of this study	Absence	See Chapter 4
Chloromonas perforata	Chlorophytes	KT625416	Absence	Buchheim *et al*., 1997
Chloromonas rosae	Chlorophytes	WGS of this study	Absence	See Chapter 4
Chloromonas serbinowii	Chlorophytes	WGS of this study	Absence	See Chapter 4
Closterium baillyanum	Streptophyte algae	NC_030314	Presence	Leyon, 1954
Coccomyxa sp	Chlorophytes	MF805805	Absence	Crespo *et al.,* 2009
Coleochaete scutata	Streptophyte algae	NC_030358	Presence	See Chapter 3
Cosmarium botrytis	Streptophyte algae	NC_030357	Presence	Gerrath, (1968)
Cylindrocystis brebissonii	Streptophyte algae	NC_030359	Presence	Croasdale & Grönblad, 1964

Cymbomonas tetramitiformis	Prasinophytes	KX013545	Absence	Moestrup *et al.*, 2003
Entransia fimbriata	Streptophyte algae	NC_030313	Presence	Cook, 2004
Floydiella terrestris	Chlorophytes	NC_014346	Presence	
Gloeotilopsis sarcinoidea	Chlorophytes	KX306821	Presence	
Igniatius tetrasporus	Chlorophytes	NC_034712	Presence	Watanabe & Nakayama, 2007
Interfilum terricola	Chlorophytes	NC_025542	Presence	Mikhailyuk *et al.*, 2008
Klebsormidium flaccidum	Streptophyte algae	NC_024167	Presence	Mikhailyuk *et al.*, 2014
Leptosira terrestris	Chlorophytes	EF506945	Presence	Stewart *et al.*, 1973
Lobosphaera incis	Chlorophytes	KM821265	Presence	
Mesostigma viride	Streptophyte algae	AF166114	Presence	Buchmann & Becker, 2009
Mesotaenium endlicherianum	Streptophyte algae	NC_024169	Presence	West, 1904
Microthamnnion kuetzingianum	Chlorophytes	NC_025537	Absence	Watson, 1975
Monomastix sp	Prasinophytes	FJ493497	Presence	Belcher & Swale, 1961
Neocystis brevis	Chlorophytes	NC_025535	Presence	
Nephroselmis olivacea	Prasinophytes	AF137379	Presence	Suda *et al.*, 2004
Nitella hyalina	Streptophyte algae	KX306884	Presence	Osterhout, 1945
Oltmannsiellopsis viridis	Chlorophytes	DQ291132	Presence	Chihara *et al.*, 1986
Oogamochlamys gigantae	Chlorophytes	NC_028580	Presence	Pröschold *et al.*, 2001
Ostreococcus tauri	Prasinophytes	NC_008289	Absence	Meyer & Griffiths, 2013
Palmophyllum crassum	Chlorophytes	NC_033387	Presence	
Parachlorella kessleri	Chlorophytes	FJ968741	Presence	Juárez et al., 2011
Pediastrum angulosum	Chlorophytes	NC_037919	Presence	Wilcox & Floyd, 1988
Pedinomonas minor	Chlorophytes	FJ968740	Presence	Moestrup, 1991
Pedinomonas tuberculata	Chlorophytes	KM462867	Presence	Manton & Parke, 1960
Phacotus lenticularis	Chlorophytes	NC_028587	Presence	Hepperle & Krienitz, 1996
Picocystis salinarum	Prasinophytes	NC_024828	Absence	Lewin et *al.*, 2000

Prasinococcus sp	Prasinophytes	KJ746597	Presence	Guillou *et al*., 2004
Prasinoderma coloniale	Prasinophytes	NC_024817	Presence	Hasegawa *et al*., 1996
Prototheca wickerhamii	Chlorophytes	KJ001761	Absence	Joshi *et al*., 1975
Pycnococcus provasolii	Prasinophytes	FJ493498	Presence	Guillard *et al*., 1991
Pyramimonas parkae	Prasinophytes	KX013546	Presence	Pearson & Norris, 1975
Roya obtusa	Streptophyte algae	NC_030315	Presence	Ljunggren & Oja, 1961
Scenedesmus obliquus	Chlorophytes	DQ396875	Presence	Miyachi *et al*., 1986
Scherffelia dubia	Chlorophytes	NC_029807	Absence	Melkonian & Preisig, 1986
Spermatozopsis similis	Chlorophytes	MG778500	Absence	Preisig & Melkonian, 1984
Spirogyra maxima	Streptophyte algae	NC_030355	Presence	See chapter 3
Stichococcus bacillaris	Chlorophytes	NC_025527	Presence	Massalski *et al*., 2001
Stigeoclonium helveticum	Chlorophytes	DQ630521	Presence	Stewart *et al.*, 1973
Symbiochloris handae	Chlorophytes	KM462860	Presence	
Tetraselmis sp	Chlorophytes	KU167097	Presence	Chengwu & Hongjun, 2018; Hori et al., 1986
Trentepohlia odorata	Chlorophytes	NC_043776	Absence	Algaebase
Ulva lactuca	Chlorophytes	NC_042255	Presence	Stewart *et al.*, 1973
Watanabea reniformis	Chlorophytes	NC_025526	Absence	Hanagata *et al.*, 1998
Zygnema circumcarinatum	Streptophyte algae	AY958086	Presence	Pichrtová *et al.*, 2013

2.5.2.5 Phylogeny of the chloroplastic CDS

Two phylogenies were built based on chloroplastic genes. The first phylogeny was built using the five newly sequenced strains from *Chlamydomonas* and *Chloromonas* and 57 other species found on GenBank (Table 2.3). Species selected for this phylogeny were chosen in order to maximise the number of pyrenoid free species whilst also covering all the main lineages of the green algae phylogeny (Lelieart *et al*., 2012; Leebens-Mack *et al*., 2019). *Arabidopsis thaliana* (land plant; NC_000932) was used as an outgroup to root the second tree.

To test for the monophyly of the genus *Chloromonas* and secondly to infer the multiple and independent origin of the pyrenoid in green algae, another phylogeny was built with a smaller sample size including only *Chlamydomonas* and *Chloromonas* strains. *Cosmarium botrytis* (Streptophyte algae; NC030357) was used as an outgroup to root the tree. For both phylogenies, only the most common 44 chloroplastic genes found across all the 64 species were used (Table 2.4). However, because not all the species included the 44 CDS genes, when absent the gene of interest was replaced by "N" of the length of the gene in the alignment. In the same way as in paragraph 2.5.1.3 jModelTest (Posada, 2008) was performed on the entire dataset (64 species) to select the best model of evolution and sequences were aligned using MAFFT (Katoh *et al*., 2005). After alignment, phylogenies were built using RAxML (Stamatakis, 2014) and 100 bootstraps.

Table 2.4 List of the 44 CDS genes used to build the two chloroplastic phylogenies and their functions

CDS name	Function
atpA	ATP synthase CF1 alpha subunit
atpB	ATP synthase CF1 beta subunit
atpE	ATP synthase CF1 epsilon subunit
atpF	ATP synthase CF0 B subunit
atpH	ATP synthase CF0 C subunit
atpI	ATP synthase CF0 A subunit
ccsa	Cytochrome c biogenesis protein
ChlB	Light-independent protochlorophyllide reductase subunit B
ChlL	Photochlorophyllide reductase subunit L
petA	Cytochrome f
petD	Cytochrome b6/f complex subunit IV
petL	Cytochrome b6/f complex subunit VI
psaA	Photosystem I P700 chlorophyll a apoprotein A1
psaC	Photosystem I subunit VII
psaJ	Photosystem I subunit IX
psbA	Photosystem II protein D1
psbB	Photosystem II 47 kDa protein
psbD	Photosystem II protein D2
psbE	Photosystem II protein V
psbF	Photosystem II protein VI
psbH	Photosystem II protein H
psbI	Photosystem II protein I
psbK	Photosystem II protein K
psbM	Photosystem II protein M
psbN	Photosystem II protein N
psbZ	Photosystem II protein Z
rbcL	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
rpl2	Ribosomal protein L2
rpl5	Ribosomal protein L5
rpl14	Ribosomal protein L14
rpl16	Ribosomal protein L16
rpl20	Ribosomal protein L20
rpl23	Ribosomal protein L23
rpl36	Ribosomal protein L36
rpoA	RNA polymerase alpha subunit
rpoC2	RNA polymerase beta' subunit
rps4	Ribosomal protein S4
rps7	Ribosomal protein S7
rps8	Ribosomal protein S8

rps11	Ribosomal protein S11
rps14	Ribosomal protein S14
rps19	Ribosomal protein S19
tufA	Elongation factor Tu
ycf3	Photosystem I assembly protein Ycf3

2.5.3 Genome comparison

2.5.3.1 Rubisco modelling and interactions between rbcL/*RbcS*

Rubisco homology modelling was performed using Chimera v1.13 (Pettersen *et al*., 2004). The Rubisco structure of *Chl. reinhardtii* was used as a template (1GK8 in Protein Data Bank; Taylor *et al*., 2001). Full rbcL sequences were used for homology modelling whereas only the *RbcS* α-helices showing the most consistency between the different sequencing methods were used for modelling. Homology modelling (Janson *et al*., 2017) was run 10 times for each complex of rbcL/*RbcS* and the best model (the one with the lowest DOPE score) was selected. In order to remove potential mistakes (e.g. Van der Waals clashes) FoldX v4.0 (Schymkowitz *et al.*, 2005) was used to repair the new Rubisco structure. 3D structures were visualized with PyMol v1.3 (Schrödinger, 2010).

Interactions between small and large subunits were identified with the software PyMol v1.3 and the script "Interface Residues" (Schrödinger, 2010; Appendix 9) whereas surface exposed residues were detected using the script "Surface residues" provided by the author (Schrödinger, 2010) was used (Appendix 10).

2.5.3.2 Screening for the 88 essential genes for pyrenoid formation

Following Luke Mackinder's work (Rubisco and CCM protein interactome; Mackinder *et al*., 2017), Jonika's group work as well as an analysis of transcriptome data in synchronised cells (Mitchell *et al.,* 2014; Zones *et al.,* 2015), and Gita Yadav's work (5C genes, unpublished data, personal communication), 88 genes essential for pyrenoid formation in *Chl. reinhardtii* were put together (Table 2.5) and blasted (BLAST; Kent, 2002) against the five new sequenced genomes containing both the short and long reads (hybrid genomes). Only the presence or the absence of these targeted genes were reported.

Table 2.5 List of the 88 genes essential for pyrenoid formation in *Chlamydomonas reinhardtii,* their common names and their functions when possible.

Gene ID	Gene Name	Function
	BST4	
Cre01.g014350.t1.2	PRX5	Peroxiredoxin, type II
Cre01.g027150.t1.1		
Cre01.g030900.t1.1		CoA ligase / OSB-CoA synthetase
Cre01.g045902.t1.1		
Cre01.g051500.t1.2		Uncharacterized thykaloid lumenal polypeptide
Cre01.g054850.t1.2		
Cre02.g073850.t1.2		
Cre02.g078507.t1.2	PF13326	Photosystem II Pbs27 (PSII_Pbs27)
Cre02.g097800.t1.1		
Cre02.g105650.t1.2		
Cre02.g111550.t1.1		Kinase
Cre02.g120100	RbcS1	Rubisco SSU1
Cre02.g120150	RbcS2	Rubisco SSU2
Cre02.g120250.t1.1	STT7	Interact with CAH3
Cre02.g143450.t1.2	PTHR36738:SF1	Expressed protein
Cre03.g146167.t1.1	TEF10a	Predicted protein
Cre03.g151650.t1.1	SMM	
Cre03.g156600.t1.2	GluTRBP	Glutamyl-tRNA reductase binding protein
Cre03.g162800.t1.2		
Cre03.g179800.t1.2	LCI 24	
Cre03.g183850.t1.2	FDX6	Ferredoxin
Cre03.g185550.t1.2		
Cre03.g188700.t1.2		
Cre03.g189800.t1.2	CYN38	Peptidyl-prolyl cis-trans isomerase, cyclophilin-type
Cre03.g191250.t1.2	LCI 34	
Cre04.g223050.t1.2		
Cre04.g223300.t1	CCP1	binds weakly to Rubisco, found in Zhan/Lemaire proteome
Cre04.g229300.t1.1	RCA1	Rubisco activase 1
Cre05.g248450.t1.2		
Cre06.g259100.t1.1	SAGA1 analog	
Cre06.g259900.t1.2	ATPc	ATP synthase gamma chain, chloroplastic
Cre06.g261750		
Cre06.g273700.t1.2		
Cre06.g283750.t1.2		
Cre06.g295450.t1.2		

Cre06.g298300.t1.1		
Cre06.g307500.t1.1	LCI C	
Cre06.g309000.t1.2		
Cre07.g330250.t1.2	PSAH	Subunit H of photosystem I
Cre08.g362900.t1.1	PSBP4	Lumenal PsbP-like protein
Cre08.g372450.t1.2	PSBQ	Oxygen evolving enhancer protein 3
Cre09.g389615.t1.1		
Cre09.g394150.t1.1		
Cre09.g394473.t1.1	LCI 9	Low CO_2 inducible protein
Cre09.g394621.t1.1	SAGA like 1	
Cre09.g396950.t1.1		Candidate Na+/HCO3- transporter from screens
Cre09.g415700.t1.2	CAH3	Carbonic anhydrase
Cre09.g416800.t1.2		
Cre09.g416850.t1.2		Potential kinase, Rubisco physical interactor RBMP2
Cre10.g423500.t1.2		
Cre10.g430150.t1.2		
Cre10.g436550	EPYC1 /LCI 5	
Cre10.g439350.t1.2	PTHR17130:SF24 - GAN	
Cre10.g440000.t1.1		
Cre10.g440050.t1.2	CSP41A	Bind to Rubisco
Cre10.g444700.t1.1	SBE3	Starch branching enzyme
Cre10.g452800.t1.2	LCI B	
Cre11.g467712.t1.1	SAGA 1	
Cre12.g484200.t1.2	GGPS1	
Cre12.g485050.t1.2		
Cre12.g494850.t1.2	ADK3	Adenylate kinase 3
Cre12.g497300.t2.1		
Cre12.g507300.t1.2		
Cre12.g509050.t1.1	PSBP3	
Cre12.g519300.t1.2	TEF9	predicted protein
Cre12.g524300.t1.2		
Cre12.g524500.t1.2		
Cre12.g531050.t1.1		
Cre12.g560950.t1.2	PSAG	
Cre13.g574000.t1.1		Putative voltage-gated bicarbonate transporter from screens
Cre13.g577100.t1.2	ACP2	Acyl carrier protein
Cre13.g578650.t1.1		
Cre13.g581850.t1.2		Kinase
Cre14.g616600.t1.2	FZL	

Cre14.g626700.t1.2	Fd/FDX1	Ferrodoxin
Cre16.g651050.t1.2	CYC6	cytochrome c6
Cre16.g652800.t1.2		
Cre16.g658400.t1.2	FDX2	Ferrodoxine
Cre16.g659050.t1.1		
Cre16.g662150.t1.2	CCB1/CPLD51	cytochrome b6f complex assembly
Cre16.g662600.t1.2	BST1	
Cre16.g663450.t1.2	LCI 11	
Cre17.g721500.t1.2	STA2	Starch synthase, chloroplastic/amyloplastic
Cre17.g724300.t1.2	PsaK	Photosystem I reaction center subunit psaK
Cre17.g725500.t1.2		
Cre17.g740950.t1.2	LHL4	High intensity light-inducible lhc-like gene
	rbcL	

2.5.3.3 Screening for EPYC1

EPYC1 was screened separately using Mackinder's methods (Mackinder *et al*., 2016) since EPYC1 is a strongly disordered protein. The five new complete genomic sequences were translated into protein sequences (6 frames) using EMBOSS-6.6.0 (Rice *et al.,* 2000). Protein sequences were then analysed for tandem repeat using Xstream (Newman & Cooper, 2007) with the same setting used in Mackinder *et al*. (2016): Min Period 40; Max Period 80; Min Copy 3, Min TR Domain 75 and Min Seq Content 0.7.Gene Infinity Protein Isoelectric Point Calculator (http://www.geneinfinity.org/sms/sms_proteiniep.html) was used to determine the pI of the Xstream hits. The disorder profiles of these hits were then calculated using VLXT (Romero *et al.*, 2001) and the presence of transmembrane domains was detected using TMHMM v2.0 (Krogh *et al*., 2001). As mentioned in Mackinder's paper, hits with an oscillating disorder profile with a frequency between 40-80 were classified as potential EPYC1-like Rubisco linker proteins.

Chapter 3: Role of the Small subunit of Rubisco in the green algal phylogeny

Most of this chapter can be found as part of: Rubisco and Carbon Concentration Mechanism (CCM) co-evolution across Chlorophytes and Streptophytes (Goudet MMM., Orr DJ., Melkonian M., Müller KH., Meyer MT., Carmo-Silva E. & Griffiths H. Submitted to New Phytologist: 19 Nov. 2019; Decisioned: 31 Jan. 2020; manuscript under revision for resubmission; accepted: 23 March 2020)

3.1 Introduction

Photoautotrophic organisms globally fix 111-117x10^{15} grams of carbon per year and around half of this global net primary production is aquatic (Behrenfeld *et al*., 2001; Field *et al*., 1998), with green algae a major contributor to this global carbon fixation. Green algae are classified into two major groups: chlorophytes and streptophytes, the latter demonstrating a wide range of ultrastructural and developmental traits closely related to land plants. Despite the existence of terrestrial green algae (Warren *et al.*, 2019), both groups remain subject to key limitations in the aquatic milieu (low CO_2 diffusion and availability, light limitation; Borges & Frankignoulle, 2002; Yamano *et al*., 2015).

Green algal inter-relationships have been resolved through numerous molecular phylogenies, including the chloroplast gene (*rbcL*) encoding the large subunit (LSU) of the primary carboxylase Rubisco (ribulose 1,5-bisphosphate carboxylase/oxygenase). An early split after the primary endosymbiosis saw the diversification of the hypothetical ancestral flagellate into two main lineages (Leliaert *et al*., 2011; 2012). First, the chlorophytes, which diversified early as prasinophytes in marine waters, which then gave rise to the core chlorophytes (chlorophytes without prasinophytes, Fig. S1, Supporting Information) in fresh or marine waters. Second, the streptophyte algae, which diversified in fresh water and some subaerial/terrestrial habitats (Harholt *et al*., 2016). The split between chlorophyte and streptophyte probably occurred during the Neoproterozoic (between 1,000 and 541 million years ago; Becker, 2013; Del Cortona *et al*., 2020). Extant photosynthetic chlorophyte and

streptophyte algae (as well as non-algal streptophytes, i.e. land plants) have a form 1B Rubisco. Selection pressures on the Rubisco catalytic properties are driven by the availability and diffusive supply of inorganic carbon, the CO_2:O_2 ratio and the development of any carbon concentrating mechanism (CCM) which improves the operating efficiency of Rubisco in many aquatic photosynthetic microorganisms (Tortell, 2000; Young *et al*., 2012; Meyer & Griffiths, 2013; Griffiths *et al*., 2017; Rickaby & Hubbard, 2019). The origins of the algal CCM could be related to equimolar CO_2:O_2 concentrations in surface waters around 500 million years ago (Griffiths *et al*., 2017).

The challenge for inorganic carbon delivery within aquatic environments is that bicarbonate (HCO_3^-) or carbonate (CO_3^{2-}) are often much more prevalent, and under current conditions, the concentration of CO_2 is often ~2,000 times lower in water than in air, and diffusion is 8,000 times slower (Raven *et al*., 1985; Falkowski & Raven, 2007; Young *et al*., 2012). A CCM is typically associated with active transport of bicarbonate across membranes, and catalytic conversion to CO_2 within a chloroplast microcompartment, the pyrenoid (Meyer *et al*., 2017). Although the presence of a pyrenoid is a robust marker of the presence of a CCM, not all the eukaryotic algae with a CCM have a pyrenoid (Morita *et al.,* 1999; Raven *et al.*, 2005).

The CCM has been particularly well-defined in the model unicellular chlorophyte *Chlamydomonas reinhardtii*, where the pyrenoid is present with a clearly defined starch sheath, and the associated inner Rubisco matrix transversed by knotted thylakoid tubules, thought to be involved in the delivery of CO_2 within the matrix (Meyer & Griffiths, 2013; Engel *et al*., 2015; Mackinder *et al*., 2017; Meyer *et al*., 2017; Mukherjee *et al*., 2019). The CCM is inducible following transfer from elevated to ambient CO_2, and a key linker protein (EPYC1) has been associated with the recruitment of Rubisco to the pyrenoid (Mackinder *et al*., 2016; Freeman-Rosensweig *et al*., 2017). This recruitment ultimately involves interactions with the Rubisco Small Subunit (SSU) (Wunder *et al*., 2018; Atkinson *et al*., 2019), presumably at the level of surface exposed α-helices (Meyer *et al*., 2012). However, there has been little systematic analysis of the extent to which some form of carbon accumulation mechanism occurs across this chlorophyte clade, or comparative physiological and molecular studies on CCM characteristics or Rubisco kinetic properties, and whether these traits are captured across chlorophyte, prasinophyte and streptophyte algal lineages in *Rbc*S.

Chlamydomonas reinhardtii has also been used as a model organism to explore the interactions between Rubisco LSU, SSU and catalytic properties. The eight identical 55-kDa LSUs assemble as four dimers, while two sets of four 15-kDa SSUs, top and tail the Rubisco holoenzyme. A central 'solvent channel' runs through Rubisco and the width of its aperture is dependent on the length of the βA-βB loop in each set of four SSUs capping the LSU octamer (Spreitzer, 2003) and interacting residues between LSUs and SSUs affect Rubisco operating efficiency and catalytic properties (Spreitzer *et al*., 2005).

The overall aim of this study was to address the possible interactions between Rubisco SSU structure and phylogeny, and occurrence of any reported CCM or pyrenoid across the green algae.

Specifically, this study sought to *(i)* develop a phylogeny for *Rbc*S sequences in green algae as compared to a consensus phylogeny (e.g. Leliart *et al*., 2012; Leebens-Mack *et al.,* 2019), and compare the distribution of pyrenoid and CCM across the algal clades; *(ii)* to identify whether any selection pressure on residues within the SSU were associated with the broader phylogeny or were related to CCM activity. Our results reveal that a change in Rubisco SSU secondary structure (namely the βA-βB loop) is a distinctive trait of the division between core chlorophytes and streptophyte algae. This study also provides additional insights for selection pressures driving the evolution of green algae and photosynthetic processes, particularly during the transition to terrestrial plant life forms.

3.2 Results

3.2.1 The length of the βA-βB loop drives the phylogeny of *Rbc*S

The protein phylogeny of *Rbc*S was originally constructed to identify any residues specific to species with a pyrenoid as a determinant of CCM activity. Despite the low number of variable sites and the low posterior probabilities of some of the most recent nodes attributable to the brevity of the sequence, *RbcS* recapitulated at the phylum level the green lineage phylogeny (e.g. Leliart *et al*., 2012; Leebens-Mack *et al.,* 2019). However, the present study found that species without a pyrenoid were dispersed throughout the whole *RbcS* phylogeny. Therefore, specific residues in the SSU α-helices (Meyer *et al*., 2012) were not sufficient to explain the pyrenoid occurrence across the entire phylum (Figure 3.1). A

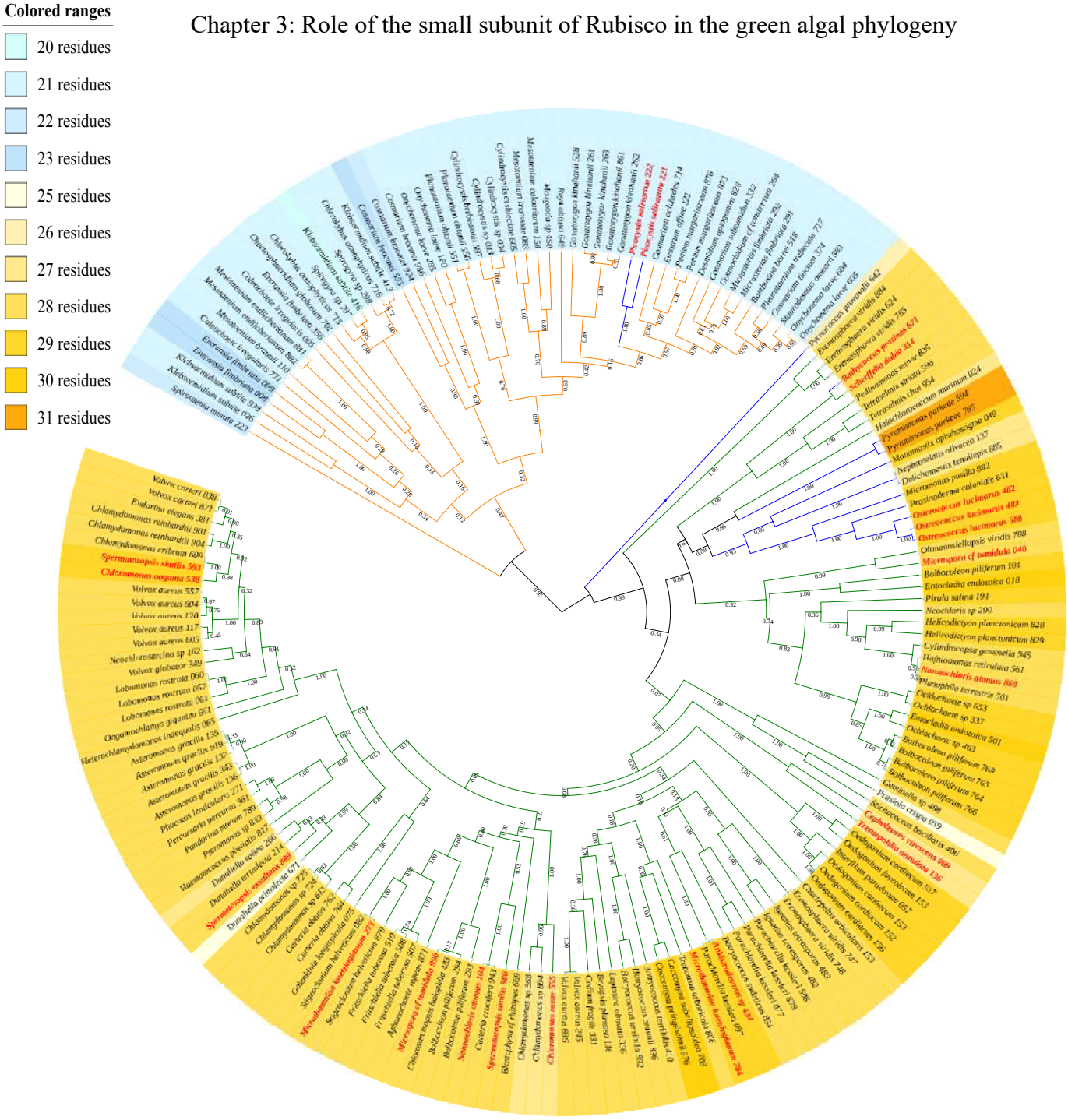


Figure 3.1 Protein phylogeny of the small subunit of Rubisco (*RbcS*) in green algae built with BEAST 2 (Bouckaert *et al*., 2014). Branches were colored according to the different phylum [chlorophytes: green (with prasinophytes in blue); streptophyte algae: orange]. Species lacking pyrenoids are indicated in bold red font. Length of the βA-βB loop was mapped onto each species and highlighted by the colour chart in the top left corner (species with a βA-βB loop length superior or equal to 25 residues are highlighted in the different shade of orange whereas species with a loop length inferior to 25 are highlighted in the different shade of blue). The phylogeny is clustered in two main clades. The first includes all the chlorophytes (green branches) and some prasinophytes (blue branches) and shows a loop length greater than, or equal to 25 residues. The second cluster includes all the streptophyte algae (orange branches) and the remaining prasinophytes (blue branches) with a loop length lower than 25 residues. Species without a pyrenoid (red font) are distributed across the phylogeny and not clustered together. Posterior probabilities are shown along branches.

closer examination of the solvent-exposed residues (available for possible interactions with the Rubisco linker EPYC1) of the amino acids and their electrostatic properties in the two α-helices, hypothesised to be the key elements for the formation of a pyrenoid (Meyer *et al*., 2012; Mackinder *et al*., 2016), varied in their distribution (Figure 3.2). For example, *Spermatozopsis similis* (pyrenoid-less) exhibited α-helices identical to *Chl. reinhardtii* (pyrenoid-positive), and *Chloromonas oogama* (pyrenoid-less) differed by only one residue (Figure 3.2). The absence of any consistent pattern which could differentiate pyrenoid-less from pyrenoid-positive species suggests that neither the specific residues in the two α-helices and their properties nor the solvent-exposed residues, can singlehandedly explain pyrenoid occurrence in green algae, as we had hypothesized.

However, the *Rbc*S phylogeny did systematically differentiate streptophyte algae and core chlorophytes, which were clustered separately into two sister clades (Figure 3.1). Prasinophytes clustered with the core chlorophytes, except *Picocystis salinarum*. The phylogenetic differentiation in *Rbc*S clearly coincided with differences in the βA-βB loop length. Core chlorophytes and prasinophytes consistently showed a βA-βB loop length of 25 or more residues, whereas the vast majority of streptophyte algae exhibited a βA-βB loop length of less than 23 residues with 52 of the 58 sequences having a βA-βB loop 21 residues long. The short loop of *P. salinarum* (21 residues) matches that of *Picocystis* sp. (draft genome; Junkins *et al.* 2019). The nested position within streptophyte algae could be due to this singular property, although the overall short length of *RbcS* and low bootstrap values at internal branches were likely additional factors. Interestingly, the phylogeny of *RbcS* remained very similar without the βA-βB loop with the presence of the same three separate clusters (core chlorophytes, prasinophytes and streptophyte algae; Appendix 11) showing that the βA-βB loop length was not the only driver of the phylogeny of *RbcS*. *Picocystis salinarum* was once again clustered again with the streptophyte algae suggesting that the low support of the node leading to this species in the tree cannot entirely explain its clustering with the streptophyte algae. However, the difference in loop length between core chlorophytes and streptophyte algae clearly revealed different Rubisco structures between these two groups. With a wider central solvent channel due to the shorter βA-βB loop, streptophyte algae have a Rubisco structure more similar to that in land plants as embryophytes (Spreitzer, 2003).

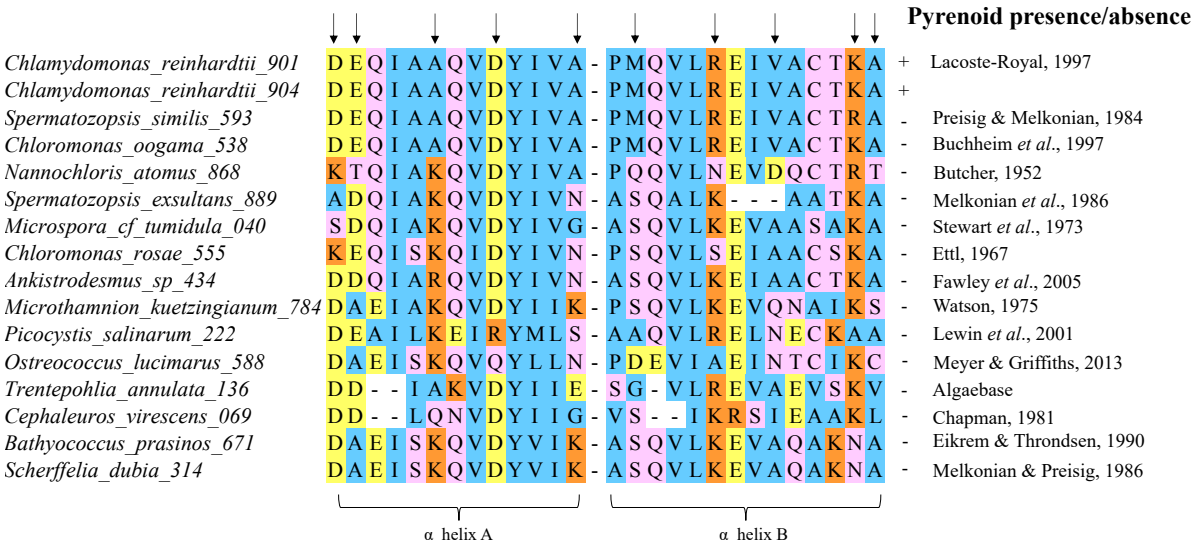


Figure 3.2 Comparison of the amino acids composition of the two Rubisco SSU α-helices for species without pyrenoid and compared to *Chlamydomonas reinhardtii* (pyrenoid positive). Acid and polar residues are in yellow, basic and polar residues are in orange, non-polar neutral residues are in blue and polar neutral residues are in pink. Residues with a solvent-exposed side chain are indicated with a black arrow according to Meyer *et al*., 2012.

3.2.2 *Rbc*S is neither under positive or relaxed selection

As an additional test for residues under positive selection in *Rbc*S, in association with a CCM or at the level of the SSU α-helices, 135 DNA sequences from green algae were used (Appendix 12). One Likelihood Ratio Test (LRT) for d_N/d_S heterogeneity across codons (M0-M3) was successfully performed and was significant, indicating expected heterogeneity in selective pressure across *Rbc*S molecules (2ΔlnL =2312.99, *P*-value<0.0001, df=8) (Table 3.1). Two LRTs were also performed to test for the presence of codons under positive selection (M7-M8 and M8-M8a) and both comparisons rejected models with positive selection (Table 3.1). The model M7 (which allows for 10 site classes, each with a ω >1) was selected in favour of the model M8 (11 sites classes with one of which allows for ω>1) and was consequently not significant (2ΔlnL= -0.00049, *P*-value=0.5, df=2). The more stringent comparison between the model M8a (which is similar to M7 but which allows for an extra class of codons with d_N/d_S=1) and M8 was also not significant (2ΔlnL= -0.07013, *P*-value=0.5, df=1) confirming the absence of codons under positive selection in *Rbc*S. The absence of residues under positive selection suggests that the appearance of new residues would not confer selective advantages in *Rbc*S, and particularly

at the level of the α-helices (consistent with observations arising from Figure 3.1 and 3.2, described above).

Table 3.1 Results of the three Likelihood Ratio Tests (LRTs) for positive selection using the site-models (M0-M8) (codeml) implemented in PAML (Yang, 2007) and their associated parameters.

	Number of classes (ω)	N[a]	Length (bp)[b]	LRT (2ΔlnL)	p-value (P<0.05)	df[c]
M0	1	135	462	2312.99077	<0.0001	8
M3	5	135	462			
M7	10	135	462	0	0.5	2
M8	11	135	462			
M8a	11	135	462	0	0.5	1
M8	11	135	462			

a: Number of sequences analysed
b: length of *RbcS* sequences analysed
c: degrees of freedom

Branches under positive selection were successfully tested with the branch-model implemented in PAML. The LRT for heterogeneity across branches (H0-H1) was significant (2ΔlnL=9.358, *P*-value=0.0011, df=1) (Table 3.2). However, background and foreground omega showed values less than 1, implying positive selection was absent among foreground branches (ωα=0.082; ωβ=0.16 <1). These results suggest that the presence of variation in across branches in *Rbc*S*,* but not significant enough to show positive selection, or any correlation with CCM occurrence.

Finally, tests for Relaxed selection were performed in order to see if the loss of amino acids at the level of the βA−βB loop observed during the transition from the chlorophytes to streptophyte algae could be due to Relaxed selection. All the tests were successfully performed (Appendix 13). All the *K* values were inferior to 1 (Table 3.3), except for the first test with a K=1.512, however, any of the tests showed significant p-values (p>0.005) meaning absence of relaxation in *RbcS*. Therefore, the change of βA−βB loop length observed in streptophyte algae could not be associated to presence of relaxed selection in *RbcS.*

Table 3.2 Results of the three LRTs for positive selection using the branch-models (H0-H1) (codeml) implemented in PAML (Yang, 2007) and their associated parameters.

	dN/dS	LRT (2ΔlnL)	p-value (P<0.05)	df
H0	ω=0.08445			
H1	ω^a=0.08262	9.358	0.0011	1
	ω^b=0.16371			

a: omega for background branches
b: omega for foreground branches

Table 3.3 Summary of the different parameters obtained for test Relax Selection (Wertheim *et al*., 2015) after 10 replicates. *K* represents the selection intensity parameter, *p* is the p-value and *LR* the Likelihood Ratio Test.

	K	p	LR	
All the streptophyte algae labelled (test n°1)	1.512	0.711	-98.38	Relaxation not significant
All the chlorophytes labelled (test n°2)	0.98	0.80	-30.52	Relaxation not significant
Basal branches of the tree labelled (test n°3)	0.85	0.78	-6.31	Relaxation not significant
Basal branch of the chlorophytes labelled (test n°4)	0.76	0.78	-6.91	Relaxation not significant
Basal branch of the streptophyte algae labelled (test n°5)	0.93	0.90	-9.20	Relaxation not significant

3.3 Discussion

3.3.1 Rubisco SSU residues do not systematically equate to a CCM.

There was no immediately apparent correlation between SSU amino-acid sequence and pyrenoid occurrence/inferred CCM activity across the newly-created phylogeny of *Rbc*S for green algae. Our expectation was based on (*i*) the observations that the *Rbc*S α-helices are important for pyrenoid formation in *Chl. reinhardtii* (Meyer *et al.,* 2012), as well as (*ii*) recent *in vitro* and *in vivo* experiments showing that the SSU is needed to interact with the *Chlamydomonas* Rubisco linker EPYC1 (Wunder *et al*., 2018; Atkinson *et al*., 2019).Whether streptophyte pyrenoids assemble with an EPYC1 analogue is currently unknown. Based on the primary sequence alone, there are no EPYC1 homologues outside the Chlamydomonadales, so it would seem that other Rubisco aggregation mechanisms may occur in more distantly related lineages, perhaps through interactions with other elements of the SSU and/or the LSU, which is the *modus operandi* in some cyanobacterial carboxysomes (Long *et al*., 2011; Oltrogge *et al*., 2019; Wang *et al*., 2019). It would be interesting to determine whether the widespread occurrence of some form of pyrenoid across green algae was due to multiple independent origins of the algal CCM (Meyer *et al.,* 2017), as found in C_4 and CAM pathways (Sage *et al*., 2011). However, the absence of a pyrenoid does not always equate to lack of a CCM (Giordano *et al.*, 2005), particularly in *Chloromonas*, which is closely related to *Chlamydomonas* (Morita *et al.,* 1999; Nozaki *et al.,* 2002; Pröschold *et al*., 2001; Meyer *et al.,* 2017), and although the underlying mechanisms of carbon accumulation of such species remain unknown there is also a consistent relationship between carbon isotope composition and CCM activity in those closely related species (and will be investigated in subsequent Chapters in this Thesis).

Overall, detailed alignments of the *Rbc*S α-helix residues did not discriminate between pyrenoid-positive and pyrenoid-negative species (Figure 3.1 and 3.2). The two *Chlamydomonas Rbc*S isoforms (Goldschmidt-Clermont & Rahire, 1986) show inverse patterns of gene expression across the day-night cycle (Zones *et al*., 2015). For the present study, it was not possible to establish the functionality of *Rbc*S paralogues in terms of CCM expression (See Materials & Methods). Therefore, determining the exact number of copies, and their sequence specificity, for each of the pyrenoidless species would provide additional confirmation for the absence of specific residues essential for pyrenoid formation in green

algae. An extensive evaluation of positive selection also showed no significant shifts in *Rbc*S amino acid residues associated with the CCM across the phylogeny (Table 3.1) whereas 13 residues under positive selection have been detected in *Rbc*S in angiosperms (Yamada *et al*., 2019). The absence of positive selection along branches leading to a pyrenoid could be an artefact of the small number of species *lacking* a pyrenoid within the green algae (Figure 3.1), or indeed those possessing some form of a CCM but lacking a pyrenoid structure (see above). A possible alternative explanation is that all green algae retained a pyrenoid-competent Rubisco SSU (as also supported by *in vitro* assays; Wunder *et al.*, 2018; Atkinson *et al*., 2019) but that the absence of a pyrenoid is rather determined by the lack (ancestral or through secondary loss) of a Rubisco linker, of similar or different ancestry as the *Chl. reinhardtii* EPYC1 (Mackinder *et al.*, 2016). Here too, future comparative proteomic studies with pyrenoidless algal CCMs will help resolve this question .

3.3.2 Streptophyte algal Rubisco SSU structure is similar to land plants

The phylogeny of *Rbc*S revealed a Rubisco structure in streptophyte algae similar to that of embryophytes, with SSUs possessing a shorter βA−βB loop and therefore a central solvent channel with a similar open structure as that shown for embryophytes (Spreitzer, 2003). Although the shorter loop in land plants has been well described (Spreitzer, 2003) and was probably thought to be a consequence of the transition from the aquatic environment to land, the presence of a similar structure in the streptophyte algae has not been previously reported. The phylogeny of *Rbc*S showed that this loss of amino acids is more ancient, and probably occurred during the split between chlorophytes and streptophyte algae, which occurred somewhere between 736 Mya (Becker, 2013) and 1,000 Mya (early Neoproterozoic; Del Cortona *et al*., 2020). The Rubisco structural change was not an isolated event at this time. The split between chlorophytes and streptophytes coincides with the appearance of multiple new traits (Hori *et al*., 2014; Nishiyama *et al*., 2018) such as lateral flagella, a flagellar peroxidase and also a Gap A/B gene duplication (McCourt *et al*., 2004; Petersen *et al*., 2006; Finet *et al*., 2010). Interestingly, the photorespiratory pathway has been shown to differ between chlorophytes and streptophyte algae. Chlorophytes use a mitochondrial glycolate dehydrogenase, which produces NADH and H^+ whereas streptophytes use a peroxisomal glycolate oxidase which produces H_2O_2 for the conversion of glycolate to glyoxylate (Stabenau & Winkler, 2005).

In addition, the similar tree topologies of *RbcS* with or without the βA-βB loop gave us more insight on the overall evolutionary history of *RbcS*. Other residues outside the □□□□□ loop within the amino acid sequences could help to differentiate chlorophytes from streptophyte algae. One explanation could come from the tight relationship between rbcL and *RbcS*. RbcL is known to be a good phylogenetical marker and co-evolution between *RbcS* and rbcL have been shown in land plants (Pei *et al*., 2013; Yamada *et al.,* 2019). Therefore, residues in *RbcS* co-evolving with rbcL could also reflect the three divisions.

The role of the SSU and of the βA−βB loop in particular is not entirely understood but the central solvent channel may facilitate channelling of substrates and products to and from the active sites (Esquivel *et al*., 2013). Spreitzer (2001; 2002) demonstrated the importance of the loop for holoenzyme assembly and showed that direct mutagenesis within the □□□□□ loop changed Rubisco catalytic properties. However, these studies did not investigate the relationship to presence of absence of the pyrenoid in green algae. Direct substitution of a non-surface exposed residue, distant from the solvent channel, R71A, decreased Rubisco specificity and increased K_c and K_o values in *Chl. reinhardtii* (Spreitzer *et al*., 2001) whereas suppressor substitutions of two SSU residues nearer the solvent channel, N54V and A57V, increased *Vc*, the specificity and the thermal stability of the large subunit L290F mutant enzyme (Du *et al.,* 2000). In addition, Spreitzer *et al*. (2005) demonstrated that the interface between SSU/LSU, far from the active sites, contributes to different catalytic properties between *Chl. reinhardtii* and *Spinacia oleracea*. Despite the change in Rubisco SSU structure between chlorophytes and streptophytes, and effect on solvent channel width and possible "suppressor" interactions between LSU and SSU (Spreitzer *et al*., 2001, 2005), there was a continued need for CCMs across the entire phylogeny (Figure 3.1) which is reflected in the catalytic properties of the streptophyte algae.

In conclusion, this study has highlighted that Rubisco SSU structure effectively differentiates between streptophytes and core chlorophytes, with a transition occurring in the prasinophyte clade which contains mostly species with a long βA−βB loop. Otherwise, the *Rbc*S phylogeny recaptures the latest consensus green algal phylogenies built from many marker genes, including *rbcL* (Leebens-Mack *et al*., 2019).

Chapter 4: Rubisco and Carbon Concentration Mechanism (CCM) co-evolution across Chlorophytes and Streptophytes

This chapter can be found as part of: Rubisco and Carbon Concentration Mechanism (CCM) co-evolution across Chlorophytes and Streptophytes (Goudet MMM., Orr DJ., Melkonian M., Müller KH., Meyer MT., Carmo-Silva E. & Griffiths H. Submitted: 19 Nov. 2019; Decisioned: 31 Jan. 2020; manuscript under revision for resubmission; accepted: 23 March 2020).

4.1 Introduction

Following Chapter 3, it appears that there has been little systematic analysis of the extent to which some form of carbon accumulation mechanism occurs across this chlorophyte clade, or comparative physiological and molecular studies on CCM characteristics or Rubisco kinetic properties, and whether these traits are retained across chlorophyte, prasinophyte and streptophyte algal lineages in *Rbc*S.

As previously explained in the General Introduction, natural variation in Rubisco catalytic properties exists among photosynthetic organisms (Jordan & Ogren, 1981). A shift in the catalytic parameters towards a higher turnover rate per active site (k_{cat}) and higher affinity for CO_2 (K_c) has been observed from cyanobacteria, chlorophyte to land plants (Badger *et al*., 1998; Meyer & Griffiths, 2013). However, it has also been suggested that selective pressures on V_C and K_c could have been relaxed due to the saturating CO_2 environment provided by a CCM over evolutionary time (Tortell *et al.,* 2000; Young *et al*., 2012; Meyer & Griffiths, 2013).

Surprisingly, no model organisms for physiological studies have been identified in streptophyte algae, despite the previous interest in using species with giant algal cells to characterise carbon uptake mechanisms (Lucas & Berry, 1985) and the recently published genome of *Chara braunii* (Nishiyama *et al.,* 2018). In addition, only few Rubisco catalytic properties are available for a few green alga species including *Euglena gracilis* (Yokota *et*

al., 1989), *Coccomyxa sp.* (Palmqvist *et al*., 1995) or *Scenedesmus obliquus* (Jordan & Ogren, 1981; Badger *et al*., 1998) but none of them are streptophyte alga. Recent measurements have largely focussed on embryophytes (Kapralov *et al*., 2010; Galmes *et al.,* 2014, 2015, 2016; Hermida-Carrera *et al*., 2016; Orr *et al*., 2016; Prins *et al*., 2016) or core chlorophytes (Jordan & Ogren, 1981; Spreitzer, 2003; Spreitzer *et al*., 2005).

Therefore, this chapter set out *i)* to define key Rubisco catalytic properties for selected streptophyte algae, as compared to *Chlamydomonas reinhardtii* and *ii)* to determine whether the catalytic properties of Rubisco across contrasting streptophyte algal groups reflected the overall phylogeny or specific activity of a CCM at the whole organism level.

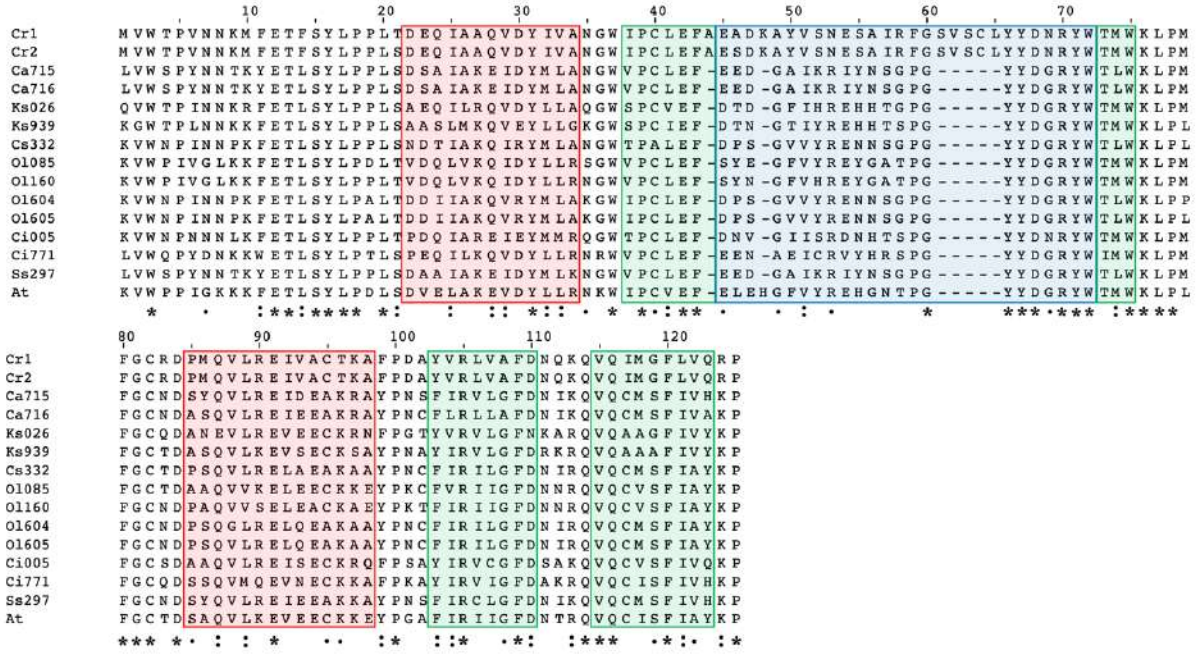


Figure 4.1 Subset alignment of sequences from the 1KP of the representative streptophyte algae Rubisco small subunit (*RbcS*) and their primary structures compared to the two copies of *RbcS* in *Chlamydomonas reinhardtii* (Chlorophytes, *Cr1* and *Cr2*) and *Arabidopsis thaliana* (*At*, land plants). *Ca* (*Chlorokybus atmophyticus*), *Ks* (*Klebsormidium subtile*), *Cs* (*Cosmarium subtumidum*), *Ol* (*Onychonema laeve*), *Ci* (*Coleochaete irregularis*) and *Ss* (*Spirogyra sp.*). Red boxes indicate residues of the two α-helices, green boxes indicate residues of the four β sheets and the blue box includes all the residues of the βA-βB loop. The multiple alignment clearly shows the absence of five amino acids from the sites 61 to 66 compared to the chlorophyte *Chlamydomonas reinhardtii*. * indicates positions which have a single fully conserved residues; «:» indicates a site belonging to group exhibiting strong similarity (strong score >0.5); «.» indicates sites belonging to a group from weak similarity (weak score =<0.5).

4.2 Results

4.2.1 Streptophyte algae share Rubisco catalytic properties with both chlorophytes and embryophytes

A more detailed investigation of Rubisco catalytic properties was undertaken in order to explore whether any evolutionary progression towards land plant characteristics was evident in streptophyte algae. The multiple alignment of *Rbc*S in six representative streptophyte algae selected for this component of the study confirmed the deletion of five amino-acids in this group compared to *Chl. reinhardtii* (Figure 4.1; Spreitzer, 2003). This shortens the loop between the first and the second β-sheets, reducing the constriction at the entry of the holoenzyme's solvent channel. Rubisco catalytic properties at 25°C for the six green algae are shown in Table 4.1, including *Chl. reinhardtii* which was used as a control, and to compare this analytical system with previous measurements for this species, albeit of different genetic parentage and analytical methods (Jordan & Ogren, 1981; Satagopan & Spreitzer, 2008). The absence of measurements for *Chlorokybus atmophyticus* was due to many unsuccessful attempts at Rubisco extraction. In *Chl. reinhardtii*, Rubisco catalytic properties varied slightly from previous measurements (Satagopan & Spreitzer, 2008; Jordan & Ogren, 1981) but remained in the same range. Michaelis-Menten constant for carboxylation (K_c) showed similar values (39.6 and 34 μM) whereas the Rubisco turnover rate (k_{cat}) was somewhat higher in this study compared to the value found in Satagopan & Spreitzer (2008). The streptophyte algae did not show a clear systematic shift from chlorophyte towards land plant catalytic properties despite similar Rubisco SSU structural changes. Of the five streptophyte algae, only *Klebsormidium subtile* and *Onychonema laeve* showed a higher affinity for CO_2 (lower K_c values), closer to land plant values (e.g. *Arabidopsis thaliana*; 10.7 μM) with K_c values of 18.7 and 27.3 μM respectively (Table 4.1). *Cosmarium subtumidum*, *Spirogyra* sp. and *Coleochaete scutata* had a relative low affinity for CO_2 with K_c values in the range of the core chlorophytes or slightly higher (45.3, 49.1 and 43.1 μM respectively).

Table 4.1 Kinetic parameters of Rubisco at 25°C in streptophyte algae in comparison to *Chlamydomonas reinhardtii* (Chlorophytes) and *Arabidopsis thaliana* (land plant) previously measured using the same protocol (Atkinson *et al.*, 2017). Species are ordered from the most distant species (*Chlamydomonas reinhardtii*, Chlorophytes, Chlorophyceae) from land plants to the closest (*Coleochaete scutata*, Coleochaetophyceae, Streptophytes). K_c is the Michaelis-Menten constant for carboxylation under 0% O_2; K_c^{air} is the Michaelis-Menten constant for carboxylation under 21% O_2 and k_{cat} is the Rubisco turnover rate. Values are means ± SEM.

Species name	**n[a]**	**k_{cat} (S^{-1})**	**K_c (µM)**	**K_c^{air} (µM)**	**k_{cat} /K_c**	**k_{cat} /K_c^{air}**
Chlamydomonas reinhardtii	3	3.25±0.18	39.6 ± 5.1	50.9 ± 7.0	0.086±0.015	0.067±0.011
Klebsormidium subtile	6	3.79±0.67	18.7 ± 1.4	28.8 ± 2.1	0.228±0.070	0.144±0.040
Cosmarium subtumidum	4	2.51±0.45	45.3 ± 13.1	55.6 ± 12.7	0.061±0.008	0.040±0.006
Onychonema laeve	4	2.39±0.44	27.3 ± 5.5	53.8 ± 12.9	0.088±0.003	0.052±0.010
Spirogyra sp	5	4.90±0.32	49.1 ± 8.0	56.9 ± 4.3	0.108±0.015	0.086±0.010
Coleochaete scutata	4	1.67±0.29	43.1 ± 9.8	62.6 ± 14.6	0.047±0.013	0.032±0.009
Arabidopsis thaliana (Atkinson *et al*., 2017)		4.1 ± 0.1	10.7 ± 0.7	15.8 ± 1.0	-	0.25 ± 0.01

a: number of replicates

The catalytic turnover rate (k_{cat}) showed a trend towards lower values. *Onychonema laeve* and *Cosmarium subtumidum*, both members of the Zygnematophyceae, had similar k_{cat} values (2.39 and 2.51 s^{-1} respectively). *Spirogyra* sp. appeared to be an exception with a high k_{cat} value compared to the other streptophyte algae (4.90 s^{-1}), similar to *A. thaliana* (4.1 s^{-1}, Atkinson *et al*., 2017). *Coleochaete scutata* showed the lowest k_{cat} of all the streptophyte algae (1.67 s^{-1}). Higher K_c is usually correlated to a higher k_{cat} and a lower specificity factor (Badger, 1987; von Caemmerer & Quick, 2000; Tcherkez *et al*., 2006; Savir *et al*., 2010; Tcherkez, 2013). *Klebsormidium subtile* presented the highest value for carboxylation catalytic efficiency (k_{cat}/K_c^{air}) (0.14 s^{-1} μM^{-1}), and whilst this was the highest streptophyte algae value determined, it remained well below that of land plants like *A. thaliana* (Atkinson *et al.*, 2017). The remaining streptophyte algae displayed lower efficiency, with *Coleochaete scutata* showing the lowest efficiency (0.032 s^{-1} μM^{-1}).

4.2.2 Streptophyte algae have a CCM, albeit leaky in some species

Oxygen evolution measurements, pyrenoid imaging and $\delta^{13}C$ were used to fully characterise CCM activity in the different streptophyte algae and to investigate whether CCM activity was associated with Rubisco catalytic properties. Oxygen evolution was used to determine the whole cell affinity for inorganic carbon and therefore the extent of any inducible carbon concentrating mechanism (Table 4.2,3; Figure 4.2). The photosynthetic $K_{0.5}$ (Ci) value (Table 4.2; Figure 4.2) of the wild-type *Chl. reinhardtii* under low CO_2 showed a strong affinity for Ci (54 μM Ci), similar to previous values in the literature and in the range of photosynthetic responses of cells expressing a CCM of 10-100 μM Ci (Mitchell *et al*., 2014; Wang *et al*., 2014). *Klebsormidium subtile*, *Spirogyra sp.* and *Coleochaete scutata* showed a whole cell affinity for Ci in the range of *Chl. reinhardtii* with $K_{0.5}$ ranging from 45 to 54 μM Ci, consistent with a fully functional CCM. *Chlorokybus atmophyticus*, *Cosmarium subtumidum* and *Onychonema laeve* exhibited a c. 20% lower apparent affinity for CO_2 compared to the other species ($K_{0.5}$ 62, 64 and 62 μM Ci respectively) suggestive of CCM activity. Student's t-tests statistically confirmed the lower affinities of these three species compared to *Spirogyra sp.*, *Coleochaete scutata*, *Cosmarium subtumidum* and *Chlamydomonas reinhardtii* (p-values<0.05; Table 4.4) except when compared to *Klebsormidium subtile* (p-value=0.83; 0.69 and 0.9). Photosynthetic $K_{0.5}$ (Ci) values of all the species grown under high CO_2 confirmed the absence of CCM activity under such

conditions (Table 4.3; Figure 4.2), and thereby the inducible character of the CCM in all species under examination.

Table 4.2 Whole cell affinity for inorganic carbon in the six streptophyte algae representative species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under low CO_2 conditions (0.04% CO_2) and their associated $\delta^{13}C$ for organic matter. Species are ordered from the most distant species from land plants (*Chlamydomonas reinhardtii*, Chlorophytes, Chlorophyceae) to the closest (*Coleochaete scutata*, Coleochaetophyceae, Charophytes). Values are means ± SEM.

Species name	**$K_{0.5}$(Ci) (µM)**	**$\delta^{13}C$ (‰)**
Chlamydomonas reinhardtii (n=3)	54 ± 23	-18.86 ± 0.01
Chlorokybus atmophyticus (n=3)	62 ± 26	-18.36 ± 0.02
Klebsormidium subtile (n=3)	53 ± 2	-21.18 ± 0.02
Cosmarium subtumidum (n=3)	64 ± 32	-15.80 ± 0.03
Onychonema laeve (n=3)	62 ± 40	-21.31 ± 0.03
Spirogyra sp. (n=3)	48 ± 38	-17.85 ± 0.04
Coleochaete scutata (n=3)	45 ± 23	-18.50 ± 0.09

Table 4.3 Whole cell affinity for inorganic carbon in the six streptophyte algae representative species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under high CO_2 conditions (5% CO_2) and their associated $\delta^{13}C$ for organic matter. Species are ordered from the most distant species (*Chlamydomonas reinhardtii*, Chlorophytes, Chlorophyceae) away from land plants to the closest (*Coleochaete scutata*, Coleochaetophyceae, Streptophytes).Vales are means ± SEM.

Species name	**$K_{0.5}$(Ci) (µM)**	**$\delta^{13}C$ (‰)**
Chlamydomonas reinhardtii (n=3)	149 ± 24	-58.78 ± 0.05
Chlorokybus atmophyticus (n=3)	120 ± 50	-45.43 ± 0.03
Klebsormidium subtile (n=3)	356 ± 282	-61.39 ± 0.03
Cosmarium subtumidum (n=3)	261 ± 124	-62.00 ± 0.14
Onychonema laeve (n=3)	264 ± 54	-45.89 ± 0.08
Spirogyra sp. (n=3)	566 ± 153	-58.79 ± 0.03
Coleochaete scutata (n=3)	599 ± 164	-42.90 ± 0.18

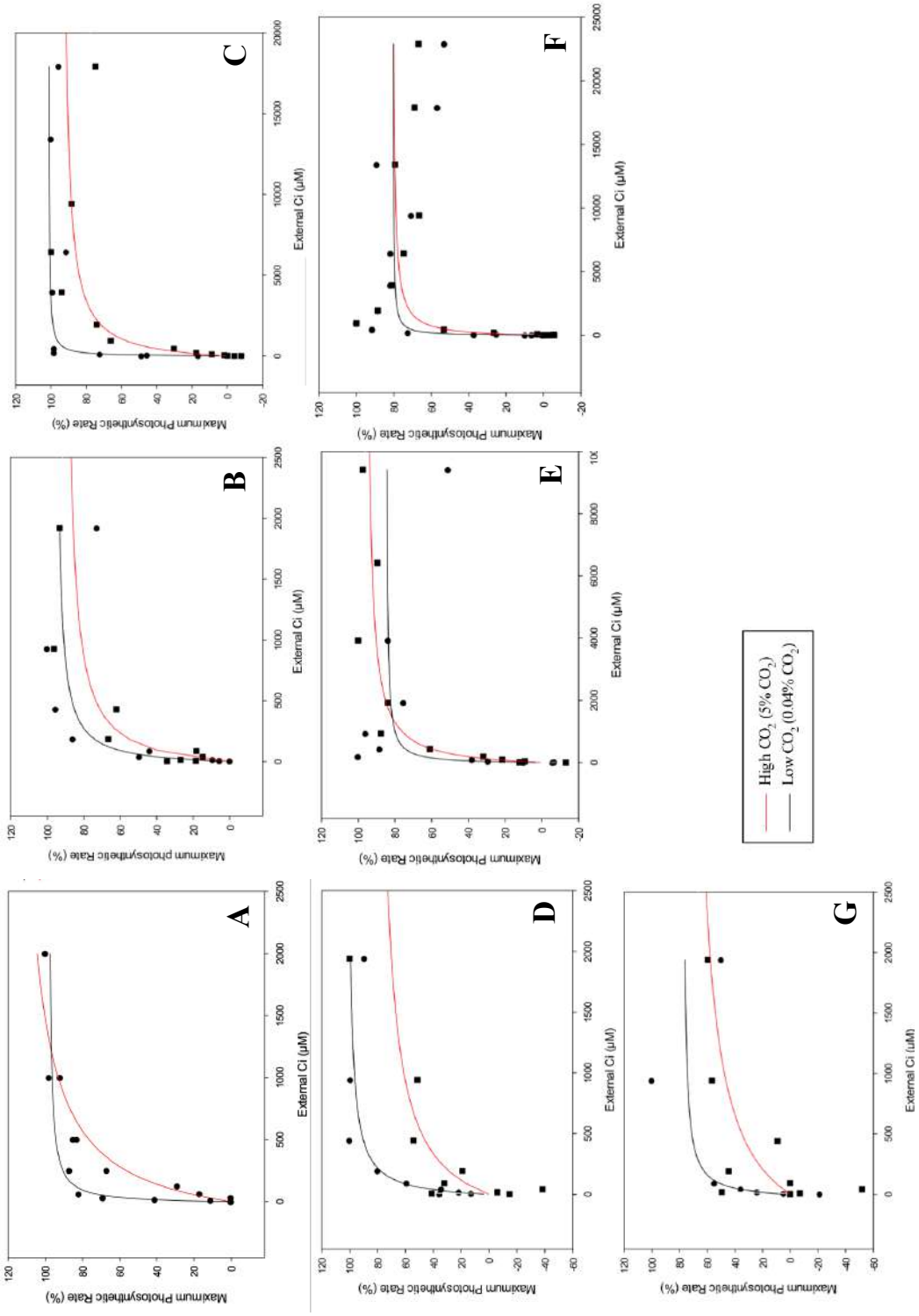


Figure 4.2 Oxygen evolution activity of the six streptophytes algae grown under low (0.04% CO_2; black curves, black dots) and high (5% CO_2; red curves, black squares) CO_2 conditions compared to *Chlamydomonas reinhardtii* (chlorophytes). **A**. *Chlamydomonas reinhardtii*; **B**. *Chlorokybus atmophyticus*; **C**. *Coleochaete scutata*; **D**. *Klebsormidium subtile*; **E**. *Onychonema laeve*; **F**. *Cosmarium subtumidum*; **G**. *Spirogyra sp.*

Table 4.4 Summary of the t-tests performed on the $K_{0.5}$ values to compare species with an apparent whole-cell lower affinity with the other streptophyte algae.

T-tests between	t	df	p-value
***Chlorokybus atmophyticus* VS**			
Klebsormidium subtile	0.2	4	0.8303
Spirogyra sp	27.3	4	0.0000
Coleochaete scutata	18.0	4	0.0001
Chlamydomonas reinhardtii	13.6	4	0.0002
***Cosmarium subtumidum* VS**			
Klebsormidium subtile	0.4	4	0.6936
Spirogyra sp	24.9	4	0.0000
Coleochaete scutata	18.7	4	0.0000
Chlamydomonas reinhardtii	14.3	4	0.0001
***Onychonema laeve* VS**			
Klebsormidium subtile	0.1	4	0.9014
Spirogyra sp	18.5	4	0.0001
Coleochaete scutata	14.9	4	0.0001
Chlamydomonas reinhardtii	8.9	4	0.0009

Stable carbon isotope composition ($\delta^{13}C$) for organic matter was also used as a second proxy for CCM activity in the different species (Meyer *et al*., 2008) (Table 4.2). *Coleochaete scutata, Chlorokybus atmophyticus*, *Spirogyra* sp. and *Cosmarium subtumidum* appeared to be isotopically enriched at -15.8 to -18.8‰ (Table 4.2), values close to *Chl. reinhardtii* (-18.9‰) and close to the upper range typically seen in C_4 terrestrial plants and consistent with a fully-functioning CCM (Raven *et al*., 1982). On the other hand, *Klebsormidium subtile* and *Onychonema laeve* were somewhat isotopically depleted ($\delta^{13}C$ of -21.1 and -21.3‰ respectively) compared to the other species, with values intermediate between typical C_3 and C_4 plants (O'Leary, 1988, $\delta^{13}C$ of C_3 plants range from -25 to -29‰ and $\delta^{13}C$ of C_4 plants range from -12 to -16‰) and consistent with a CCM phenotype prone to leakiness (retro-diffusion of CO_2: Meyer *et al*., 2008) or limited carbon accumulation capacity.

Taken together, these observations reveal that Rubisco catalytic properties correlate to some extend with the strength of CCM activity. Similarly to *Chl. reinhardtii*, the three streptophytes algae *Cosmarium subtumidum*, *Spirogyra* sp. and *Coleochaete scutata*

revealed a fully functioning CCM (low whole-cell affinity, $K_{0.5}$, and low carbon isotope discrimination) but lower Rubisco catalytic affinity for inorganic carbon (high K_c values), whereas *Klebsormidium subtile* and *Onychonema laeve* have a less effective CCM but higher affinity for inorganic carbon in terms of Rubisco catalytic properties (low K_c values). Therefore, in the presence of a less-effective CCM, Rubisco catalytic properties for *Klebsormidium subtile* and *Onychonema laeve* show a systematic shift towards values more typically associated with land plants.

Finally, electronic microscopy was used to diagnose the presence/absence of a pyrenoid in the algal material used in the present study, as an additional diagnostic for an active biophysical CCM. The presence of a pyrenoid was successfully confirmed for all the species except for *Coleochaete scutata* for which tissue embedding was unsuccessful. Presence and morphology of a pyrenoid in that species had been previously published (McKay *et al*., 1991). CCM activities were supported by the presence of a pyrenoid in all the species (Figure 4.3). *Cosmarium subtumidum* (Figure 4.3b), *Onychonema laeve* (Figure 4.3d), *Coleochaete scutata* (Figure 4.3f) and *Spirogyra* sp. (Figure 4.3e) exhibited pyrenoid morphologies similar to *Chl. reinhardtii* with a spheroidal electron dense matrix traversed by multiple tubules, and a single layered peripheral starch sheath. (Figure 4.3; Appendices 14 and 15). There were, however, differences in the fine structure (starch sheath thickness and continuity, density of thylakoid tubules network) that perhaps provide clues to the variability in $K_{0.5}$ and $\delta^{13}C$ measurements. *Klebsormidium subtile* lacked a peripheral starch sheath (Figure 4.3a; Appendix 16) although a starch sheath may occur in *Klebsormidium subtile* dependent on growth stage or light intensity (M. Melkonian, unpublished observations). *Chlorokybus atmophyticus* had multiple layers of short starch plates surrounding the matrix (Figure 4.3c; Appendix 17). The network of cross-pyrenoidal tubules was regular and dense in *Cosmarium* and *Chlorokybus* (Figure 4.3b, c).

Overall, the results show that Rubisco catalytic properties are CCM dependent. However, at this stage, it remains difficult to differentiate limitations in carbon uptake versus leakiness of CO_2 as the selective pressure operating on Rubisco, and more detailed physiological experiments are warranted to fully characterize these contrasting processes.

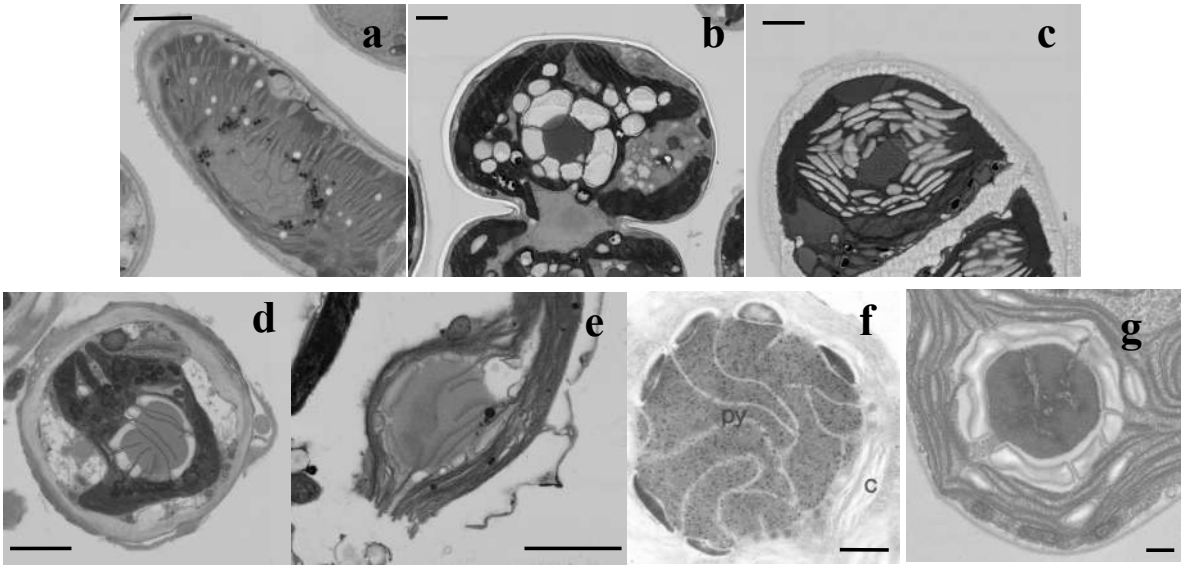


Figure 4.3 Electron micrographs (EM) images of the six representative streptophyte algae and of *Chlamydomonas reinhardtii* (a: *Klebsormidium subtile*, b: *Cosmarium subtumidum*, c: *Chlorokybus atmophyticus*, d: *Onychonema laeve*, e: *Spirogyra sp.*, f: *Coleochaete scutata;* McKay *et al*., 1991 (py=pyrenoid) , g: *Chlamydomonas reinhardtii*). Three distinct pyrenoid morphologies can be observed in the streptophytes: Rubisco matrix enclosed by one layer of starch plates (b, d, e, f and g); pyrenoid enclosed by multiple starch grains (c); and pyrenoid without observable starch sheaths (a). Bars: 2 µm (a to e) and 0.5 µm (f and g).

4.3 Discussion

4.3.1 Rubisco catalytic properties in green algae depend on CCM efficiency

The above observations led to the investigation of Rubisco catalytic properties within the streptophyte algae and their associated physiological CCM activity. Streptophyte algae are difficult to investigate physiologically. Oxygen electrode measurements were also extremely challenging (Table 4.2-3; Figure 4.2).

Despite the clear structural change associated with the βA−βB loop length (see Chapter 3), Rubisco catalytic properties remained generally similar to chlorophytes (Table 4.1) without systematic shift towards values associated with land plants (Satagopan & Spreitzer, 2008; Kapralov *et al*., 2010; Atkinson *et al*., 2017). Over the six streptophyte algae, only *Klebsormidium subtile* and *Onychonema laeve* showed K_c values in this lower range. Direct mutagenesis has shown the importance of the SSU βA−βB loop in Rubisco catalytic properties (See Introduction, paragraph 1.6) but the data in the present study suggested that they were more influenced by the effectiveness of the CCM, consistent with systematic changes in carbon isotope composition ($\delta^{13}C$: Table 4.2). Carbon isotopes have been used to infer leakiness of CCMs found in algae and hornworts (Meyer *et al*., 2008). Although whole cell inorganic carbon (Ci) uptake affinity was similar for all species under ambient growth conditions ($K_{0.5}$, Table 4.2).

The weaker CCM activities (identified through more negative $\delta^{13}C$ values: Table 4.3) in *Klebsormidium subtile* and *Onychonema laeve*, were associated with the highest affinity of Rubisco for CO_2 (K_c, Table 4.1). The importance of the CCM in shaping the adaptation within Rubisco catalytic properties has been a long-standing hypothesis (Tortell, 2000; Young *et al.,* 2012; Meyer *et al*., 2013, Galmes *et al*., 2014, 2016, 2019; Griffiths *et al*., 2017), consistent with the shifts seen in C_4 Rubisco (Jordan & Ogren, 1981; Sage, 2002; Kubien *et al*., 2008). Our results show that Rubisco catalytic properties for this representative range of streptophyte algae are adapted to the presence of the CCM.

A strong CCM (uptake and conversion of inorganic carbon) or reduced retrodiffusion (leakiness) is partly consistent with pyrenoid presence for these two species (with either a naked pyrenoid or simple starch sheath: Figure 4.3 a, d, respectively). In addition, *Klebsormidium subtile* has often been reported to be a cosmopolitan species, colonising a great variety of aquatic and terrestrial habitats (Table 2.2; Materials & Methods; Hoffmann, 1989; Rindi *et al.,* 2011; Mikhailyuk *et al*., 2015). The Rubisco catalytic properties found in

Klebsormidium subtile would place this species as an intermediate between obligate aquatic green algae and land plants. The future study of real subaerial algae such as *Klebsormidium flaccidum* or *Mesotaenium endlicherianum* would allow a more complete understanding of the photosynthetic adaptation to life on land. In the absence of the liquid boundary layer impeding CO_2 diffusion on land which could affect Rubisco catalytic properties (Raven *et al.,* 1985; Sáez *et al*., 2017), the naked pyrenoid in *Klebsormidium subtile* would account for the more land-plant-like Rubisco catalytic properties and a reliance on direct diffusive CO_2 supply.

The co-evolution of Rubisco and CCMs has been demonstrated in multiple non-green photosynthetic organisms (Badger *et al*., 1998).). Diatoms and haptophytes, which possess Form 1D Rubisco, are known to carry most of the oceanic photosynthesis (Delwiche & Palmer, 1997; Yoon *et al*., 2002; Falkowski *et al*., 2004). In these groups, Rubisco affinity for CO_2 (K_c) exhibits larger variations, exceeding those of C_4 plant Rubisco suggesting a large diversity of CCM strengths (Young *et al*., 2016; Heureux *et al*., 2017). In addition, the CO_2:O_2 ratio around the active site led to the suggestion that pyrenoids could have an oxygen exclusion function (McKay & Gibbs, 1991; Griffiths *et al.*, 2017). In land plants, Rubisco catalytic properties have been shown to be linked to changes in the atmospheric CO_2:O_2 ratio over time as well as temperature, in addition to leaf architecture, morphology and conductance (Beerling *et al*., 2001; Franks & Beerling, 2009; Haworth *et al.,* 2011; Galmes *et al*., 2014; 2015; Sharwood *et al.,* 2016; Conesa *et al*., 2019). As the atmospheric CO_2:O_2 ratio decreased over time, Galmes *et al.,* (2014) showed that land plants developed a Rubisco that was more efficient at carboxylation (higher k_{cat}/K_c ratio) with increased affinity for CO_2 (lower K_c) but slower carboxylation rate (k_{cat}). Alongside this changes in catalytic properties, the proportion of soluble protein present as Rubisco increased, counteracting somewhat the effect of decrease in carboxylation rate (Galmes *et al.*, 2014). Furthermore, higher temperatures increase maximum carboxylase turnover rate (k_{cat}) of Rubisco and decrease CO_2 affinity (Bernacchi *et al.,* 2001; *Galmes et al.*, 2015, 2016).

4.4 Conclusion

In conclusion, the Rubisco catalytic properties in streptophyte algae suggests that the activity or strength of any CCM, which may have arisen because of limitations in bulk CO_2 delivery to Rubisco in the aquatic milieu, is associated with the retention of a lower affinity (high K_c)

Rubisco. We showed that the extent of adaptation which occurs should either cause CCM activity to be reduced, or indeed lost during the transition to land, as the reliance on gaseous diffusion to deliver CO_2 to Rubisco began to increase. Overall, this study shows rather than being intransigent and slow, Rubisco catalytic properties adapt to local conditions of CO_2 availability in green algae. This is consistent with the changes seen in Rubisco from C_4 (Jordan & Ogren, 1981; Sage, 2002; Kubien *et al*., 2008) and CAM plants (Griffiths *et al*., 2008), which have been associated with operating within a CCM for the past 5-10 million years. Based on this study, the selective pressures driven by local conditions of photosynthetic CO_2 supply are more likely to explain the shifts in Rubisco catalytic properties during life on land, rather than any long term transition seen through cyanobacteria, algae to land plants, and highlights a dynamic relationship between Rubisco catalytic properties and CCM activity.

Chapter 5: Comparative analysis of the morphology and the physiology of two closely related genera: *Chlamydomonas* and *Chloromonas*.

5.1 Introduction

Chapters 3 and 4 focused on the links between Rubisco SSU, Rubisco kinetics and CCM in a phylogenetic context. The co-evolution between Rubisco kinetics and CCM gave us more insight into photosynthesis in green algae. However, the phylogeny of *RbcS* did not help us to understand either pyrenoid occurrence or to identify specific residues within the 2 α-helices. Newly identified components, such as EPYC1, that interact directly with Rubisco are essential to obtain a normal pyrenoid, (Sections 1.4, 3.1, Mackinder *et al.,* 2016). Analysis of mutants and crystallographic work is underway to establish, how this linker protein interacts with Rubisco SSU in *Chlamydomonas reinhardtii*. As green algae comprise a broad range of species expressing different levels of CCM and/or a pyrenoid they are ideal models in which to understand the natural variation in CCM occurrence. However, the physiological mechanisms behind this range of CCM expression is still unknown.

The overall aim of this chapter was therefore to explore this diversity of CCM by characterising the physiology of organisms exhibiting these natural variations and establish which could be good candidates for future engineering work.

5.1.1 The absence of pyrenoid, a long-standing observation

Used as a taxonomic marker since the middle of the 20th century, it was not until the 21st century that physiologists really appreciated the importance of the pyrenoid in aquatic photosynthesis. The pyrenoid is a widespread trait across the green alga species, however, the lack of pyrenoid was observed in a minority of genera (*Chloromonas*, *Coccomyxa* or *Spermatozopsis*). Multiple independent origins of the pyrenoid have been demonstrated in hornworts (Villarreal &

Table 5.1 Name, algal collection numbers and habitat of the species used in this chapter.

Species name	Habitats
Chlamydomonas reinhardtii	Stagnant water, damp soil, freshwater, seawater, snow algae
Chlamydomonas augustae UTEX LB 1969	Snow, Cascade Mountains, Oregon, USA
Chlamydomonas mutabilis UTEX 578	Soil extract
Chloromonas serbinowii UTEX LB 492	Temperate, Bloomington, Indiana, USA, Yellowwood fish ponds
Chloromonas rosae UTEX B 1337	Soil, High Tatra Mountains, Czech Republic
Chloromonas clathrata UTEX LB 1970	Snow, Todd Lake, Cascade Range, USA

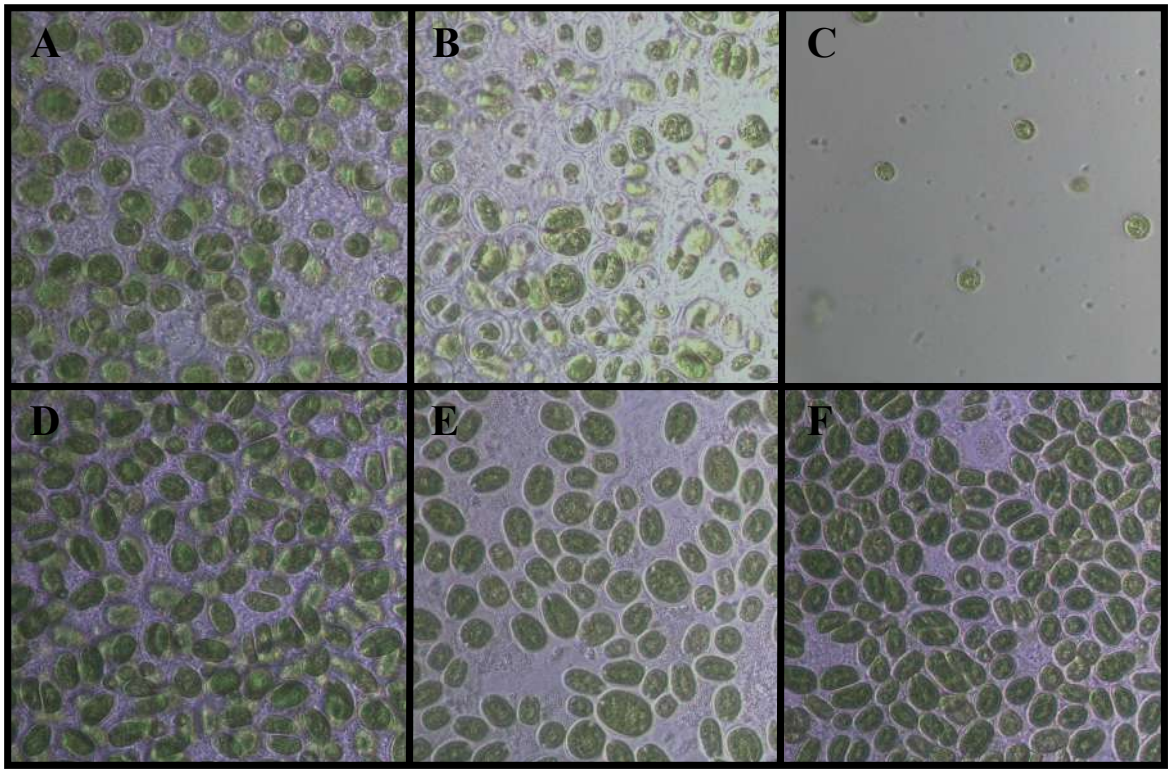


Figure 5.1 Optical microscopy images of the 5 different strains of *Chlamydomonas* and *Chloromonas* (**A**: *Chlamydomonas augustae*; **B**: *Chlamydomonas mutabilis*; **C**: *Chlamydomonas reinhardtii*; **D**: *Chloromonas serbinowii*; **E**: *Chloromonas rosae*; **F**: *Chloromonas clathrata*). Pictures taken with Nikon 40X/0.75 DIC M/N2.

Renner, 2012), the only land plants expressing a pyrenoid, whilst C_4 and CAM photosynthesis probably evolved well over 100 times in land plants (Section 1.4; Sage *et al.,* 2011; Edwards, 2019). Although a few reports (Nozaki *et al.,* 2002; Meyer & Griffiths, 2013) have started to demonstrate the independent origins of the algal pyrenoid, no study to date has reported the origins of the pyrenoid across the full green algae lineages.

The lack of pyrenoid in some species is now widely accepted (Raven *et al.,* 2005) but the underlying mechanisms which could explain how these organisms deal with the absence of CCM (*e.g.* inorganic carbon uptake) and thus rely exclusively on diffusive entry of CO_2, remain unknown.

5.1.2 *Chlamydomonas* and *Chloromonas*: two closely related genera

The Chlamydomonales is one of the larger orders in the green algae with more than 1,700 species (Fritsch & West, 1927). Among all these species, two genera are of particular interest: *Chloromonas and Chlamydomonas.* The genus *Chlamydomonas* is well known because it includes the alga study model *Chlamydomonas reinhardtii* and a great variety of other species such as *Chlamydomonas augustae* or *Chlamydomonas radiata* (Buchheim *et al*., 1997). The unicellular and biflagellate *Chloromonas* are morphologically fairly similar to *Chlamydomonas*, but conspicuously lack a pyrenoid (Ettl, 1967; 1970, 1983; Iyengar & Desikachary, 1981). *Chloromonas* are usually found in cold habitats or as snow algae [Ettl, 1970; Hoham 1975, 1977, 1979 (Table 5.1)] whereas *Chlamydomonas* species are quite ubiquitous, colonising habitats from stagnant water and damp soil, to fresh and seawater (Table 5.1).

The relationship between these two genera has been subjected to multiple studies (Buchheim *et al.,* 1997; Pròschold *et al*., 2001; Nozaki *et al*., 2002). The polyphyly of *Chlamydomonas* and *Chloromonas* is now well established, but all the phylogenetic reconstructions were based on the analyses of single markers (18S and three chloroplastic genes) (Figure 5.2) leading to different tree topologies. However, Nozaki *et al.* (2002) were the first to look at these different phylogenetic trees in taking into account the pyrenoid occurrence in the different strains (Figure 5.2; C-D). They not only showed once again that *Chlamydomonas* and *Chloromonas* were polyphyletic, they highlighted that the pyrenoid could have been lost and regained several times. In addition, Morita *et al.* (1998, 1999) showed that the absence of pyrenoid was not necessarily linked to the absence of CCM. Therefore, the observations made by Morita first followed by Nozaki's work led to a greater interest in *Chlamydomonas* and *Chloromonas* strains to study and understand pyrenoid occurrence in genera

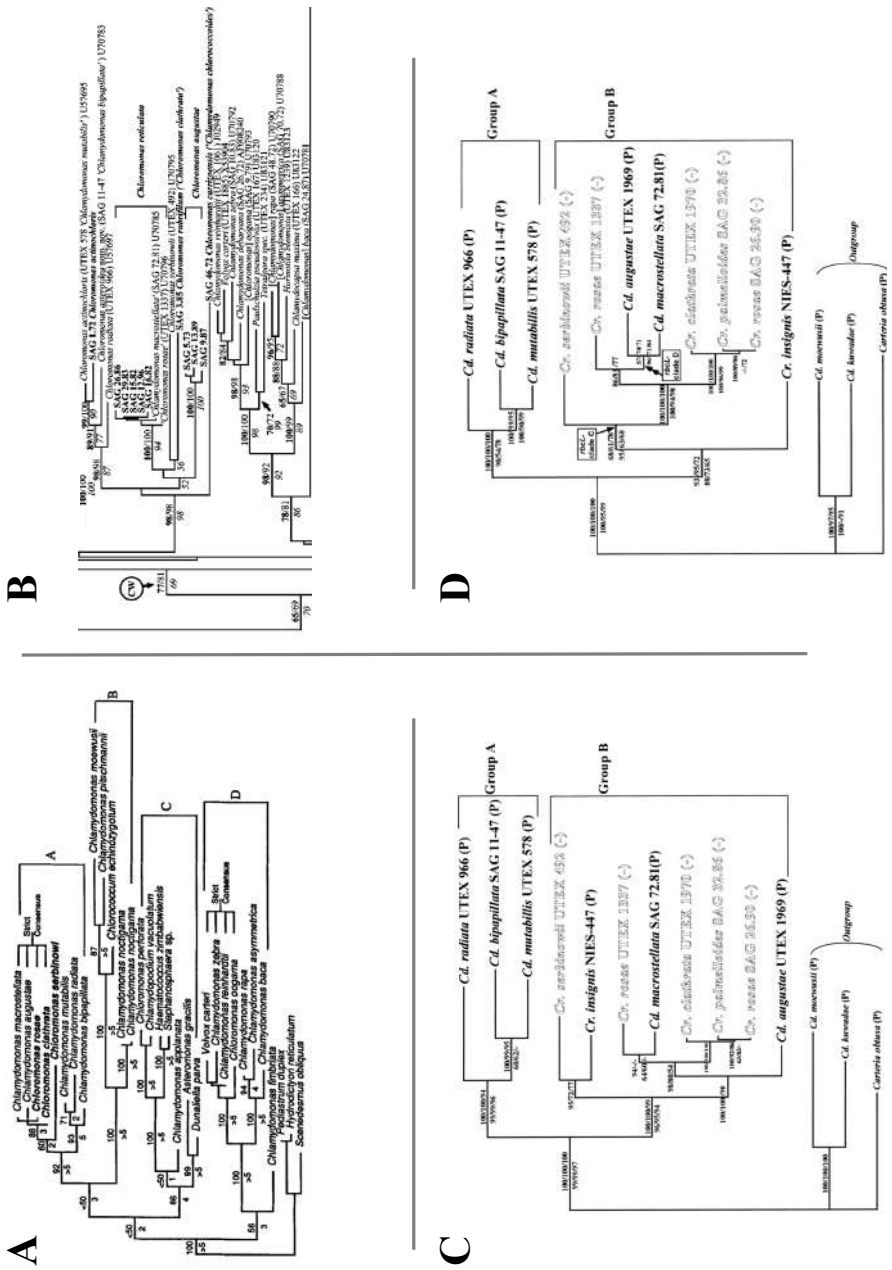


Figure 5.2 Summary of the different phylogenetic reconstructions of *Chlamydomonas* and *Chloromonas* strains in the literature. **A**: Buchheim *et al.*, 1997 based on 18S **B**: Pröschold *et al.*, 2001 **C**: Nozaki *et al.*, 2002 built with *atpB* **D**: Nozaki *et al.*, 2002 built with *rbcL*.

phylogenetically close to each other.

5.1.3 Objectives of this study

Inorganic carbon uptake pathways in green algae have been subjected to a lot of publications over the years (*Ulva lactuca*; Axelsson *et al*., 1995; *Chlorella saccharophila*, Gehl *et al.,* 1990; *Scenedesmus obliquus*; Palmqvist *et al*., 1988). Algae have generally developed different strategies in order to optimize their carbon uptake mechanisms. The availability of the CO_2 in the aquatic environments is deeply related to local ecology, history of the aquatic microorganisms but also how these organisms interact with each other (Raven, 1991). Therefore, parameters such as temperature, salinity, photon flux density but also organism size can deeply affect the ways that these aquatic organisms develop to optimise photosynthesis (Raven, 1991). The classic algal CCM diagram published by Moroney (Moroney & Tolbert*,* 1985) and as shown in the General Introduction (Figure 1.8) describes an inorganic carbon uptake system mainly based on the presence of carbon pumps, carbonic anhydrases and simple diffusion through the membranes. However, organisms can organise their uptake systems around other structures. Periplasmic carbon uptake systems have been found in the genus *Chara* (Lucas *et al.,* 1983; Ray *et al*., 2003), a streptophyte algae. Associated to another structure called *charasomes*, this mechanism allows this genus to grow in environment with high to medium pH. In *Chara corallina* (Mimura *et al*., 1993), periplasmic carbon uptake system have been characterized as H^+ATPase that pumps protons out of the cell and creating a gradient of pH and electrical potential differences across the plasma membrane (Taiz & Zeiger, 1998) whereas in *Chara tomentosa* this periplasmic system is more likely to be H^+ extrusion with the help of proton pumps, combined with carbonic anhydrases and membrane transport of CO_2 (Ray *et al.,* 2003).

Despite many studies on inorganic carbon uptake, until now, very few studies have tried to understand the mechanisms behind the total absence of CCM and/or pyrenoid. With the environmental constraints that aquatic environments provide, it remains difficult to understand how some species survive without expressing an identified CCM. Do they reflect possible other adaptions or different Ci uptake pathways that those found in species expressing a CCM? The most recent and relevant studies are 15 to 20 years old (Morita *et al.,* 1998, 1999; Raven *et al.,* 2005) and thus do not take into account the latest discoveries on the essential components for the expression of the pyrenoid in *Chl. reinhardtii* (Mackinder *et al.,* 2016, 2017; Zhan *et al.,* 2018). In addition, Raven (1991) emphasized the importance of diffusive supply potential based on cell size and assumed boundary layer to

understand the mechanisms of inorganic carbon uptake in algae. Do the species without pyrenoid/CCM have different morphological characteristics which could favour inorganic carbon uptake?

Therefore, to further our knowledge of CCM/pyrenoid expression the aim of this chapter is to fully characterise the CCM activity of 2 wild-type *Chlamydomonas* and 3 wild-type *Chloromonas* strains (Table 5.1), whilst also including measurement of the wild type *Chl. reinhardtii* as a reference. Following the approach used in Chapter 4, apparent affinity for inorganic carbon ($K_{0.5}$) was determined by oxygen evolution (Badger *et al*., 1980) in high (5% CO_2) and low (0.04% CO_2) CO_2 conditions. Carbon isotope ratio analyses were made in order to characterise the CCM effectiveness and finally pyrenoid occurrence/morphologies were examined by scanning electronic microscopy (SEM). Ratio between pyrenoid size and whole cell size was also measured.

5.2 Results

As previously explained in the Material & Methods Chapter, oxygen evolution measurements, pyrenoid imaging and $\delta^{13}C$ were used to fully characterise CCM activity in five different *Chlamydomonas/Chloromonas* species grown in high (5% CO_2) and low (0.04% CO_2) CO_2 concentrations [*Chlamydomonas augustae* (Skuja, 1943), *Chlamydomonas mutabilis* (Gerloff, 1940), *Chloromonas serbinowii* (Wille, 1903), *Chloromonas rosae* (Ettl, 1970) and *Chloromonas clathrata* (Ettl, 1970), Table 5.2 and Figure 5.1].

5.2.1 Differences in photosynthetic affinity for inorganic carbon

The rate of photosynthetic oxygen evolution under high and low CO_2 concentrations was used to determine the whole cell affinity for inorganic carbon and therefore the extent of any inducible carbon concentrating mechanisms.

In all the *Chlamydomonas* strains, photosynthetic affinity ($K_{0.5}$) for Ci in low CO_2 cells was higher than in cells grown under high CO_2 (Table 5.2 and Figure 5.3), statement supported by significant statistical differences between low and high CO_2 values. *Chl. reinhardtii* (results also showed in Chapter 4) and *Chl. augustae* showed strong affinity for Ci under low CO_2 (54 and 57 µM Ci respectively) consistent with a fully functional CCM (supported by statistical differences between the $K_{0.5}$ values in low and high CO_2 conditions; *p-values*= $3.2e^{-8}$ and 0.001 respectively) whereas *Chl. mutabilis* exhibited lower $K_{0.5}$ compared to the

other two *Chlamydomonas* species (84 µM Ci) suggestive of some CCM activity (Table 5.2 and Figure 5.3; *p-value* $K_{0.5}$ low CO_2 versus $K_{0.5}$ high CO_2= 0.034). Photosynthetic affinity for Ci under high CO_2 showed poor affinity for Ci with $K_{0.5}$ values on average 6.9 times higher compared to the low CO_2 values (149, 462 and 835 µM Ci respectively) (Table 5.2 and Figure 5.3) and therefore an absence of CCM activity under such conditions. These results suggested that the three *Chlamydomonas* species possess a low-CO_2 inducible CCM. *Chloromonas* strains showed more diverse results. Overall, all the strains showed a poor photosynthetic affinity for Ci under low CO_2 with values ranging from 214 to 306 µM Ci. However, *Chr. serbinowii* and *Chr. rosae* exhibited even lower affinity for CO_2 under high CO_2 (822 and 871 µM Ci). Significant differences between low and high CO_2 $K_{0.5}$ values (*p-values*= $2.01e^{-7}$ and 0.01) were confirmed with the t-tests and are consistent with the presence of small intracellular Ci pools under low CO_2 conditions for these two strains. In contrast, *Chr. clathrata* did not show clear differences between photosynthetic affinity ($K_{0.5}$) under low and high CO_2 (214 ± 54 vs 537±124 µM Ci; *p-value*= 0.22) with the two oxygen evolution curves superposed on each other (Figure 5.3), supporting the absence of Ci pool in this strain.

Table 5.2 Whole cell affinity for inorganic carbon in the 5 *Chlamydomonas/Chloromonas* representative species and in *Chlamydomonas reinhardtii* (Chlorophytes) grown under low CO_2 (0.04% CO_2) and high CO_2 conditions (5% CO_2). Values are means ± SEM. n=number of replicate. Significance: p<0.0001 ****; 0.0001<p<0.001 ***; 0.001<p<0.01 **; 0.01<p<0.1 *; p>0.1 NS (not significant)

Species name	**$K_{0.5}$(Ci)(µM) Low CO_2 (0.04% CO_2)**	**$K_{0.5}$(Ci)(µM) High CO_2 (5% CO_2)**	**T-test (Low VS High CO_2)**
Chlamydomonas reinhardtii (n=3)	54 ± 23	149 ± 24	*p-value*= $3.2e^{-8}$****
Chlamydomonas augustae UTEX LB 1969 (n=3)	57 ± 20	462 ± 72	*p-value*=0.001**
Chlamydomonas mutabilis UTEX 578 (n=3)	84 ± 15	835 ± 233	*p-value*=0.034*
Chloromonas serbinowii UTEX LB 492 (n=3)	237 ± 70	822 ± 297	*p-value*=$2.01e^{-7}$****
Chloromonas rosae UTEX 1337 (n=3)	306 ± 81	871 ± 213	*p-value*=0.01*
Chloromonas clathrata UTEX LB 1970 (n=3)	214 ± 54	537 ± 127	*p-value*=0.22 NS

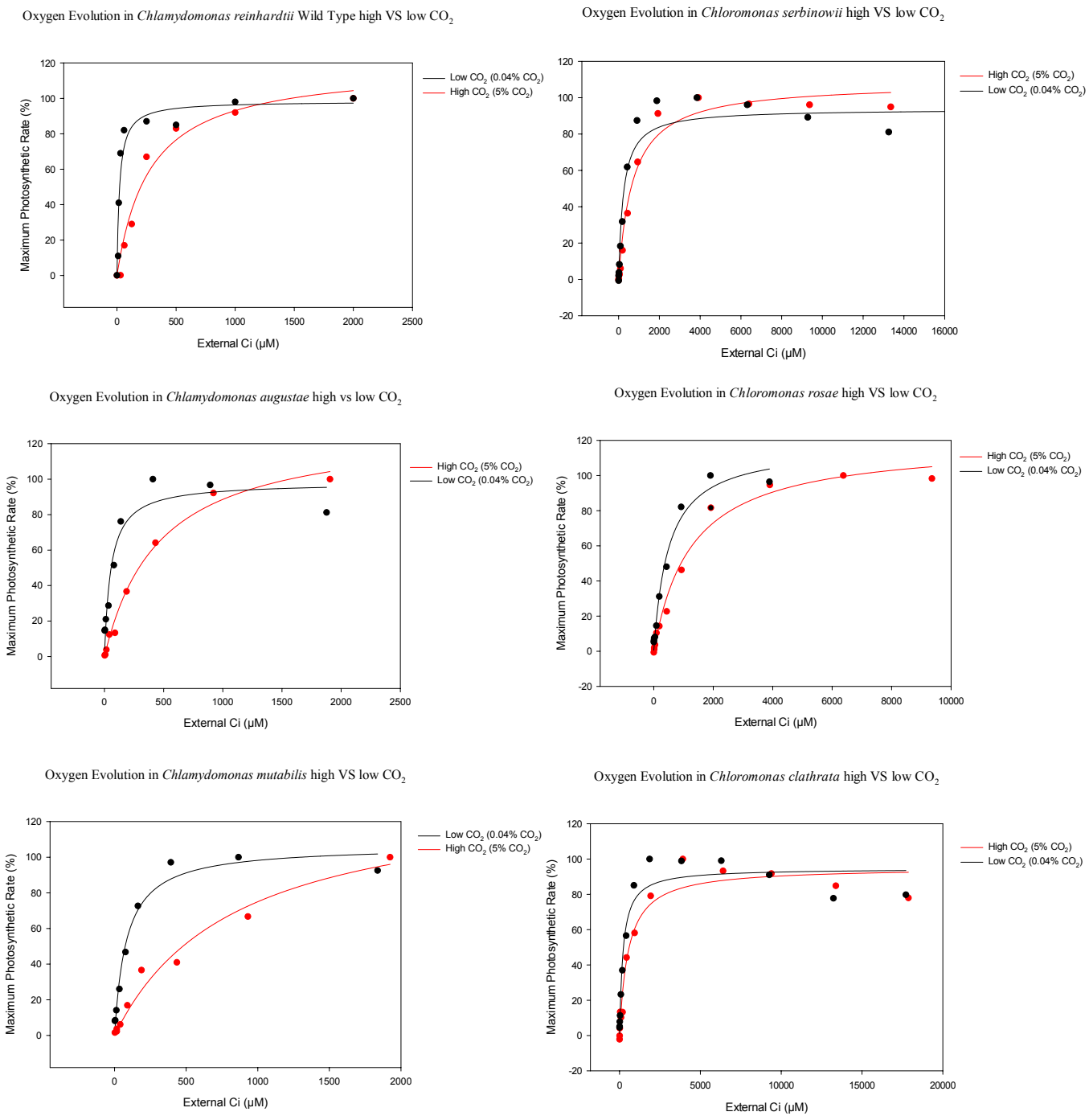


Figure 5.3 Oxygen evolution activity of in the 5 *Chlamydomonas/Chloromonas* representative species and in *Chlamydomonas reinhardtii* (Chlorophytes) grown under low CO_2 (black curves; 0.04% CO_2) and high CO_2 conditions (red curves; 5% CO_2)

5.2.2 When present, pyrenoids exhibit different morphologies

Scanning electronic microscopy (SEM) was used to confirm the presence/absence of a pyrenoid in the different strains as an additional diagnostic for an active biophysical CCM. The presence of a pyrenoid was successfully confirmed for the *Chl. reinhardtii* and *Chl. mutabilis* strains, (Figure 5.4 A-B; Appendix 18) but not for *Chl. augustae* for which tissue embedding was unsuccessful. However, presence and morphology of the pyrenoid were confirmed based on Morita *et al*., 1999 in which the same strain was used. Two different pyrenoid morphologies were observed in the *Chlamydomonas* strains (Figure 5.4A to C). The pyrenoid was enclosed by one layer of starch plate in *Chl. reinhardtii* (Figure 5.4A) whereas pyrenoids were enclosed by multiple starch grains (Figure 5.4B-C) in *Chl. augustae* and *Chl. mutabilis*. Although all the *Chlamydomonas* strains showed a pyrenoid in low CO_2 conditions, none of the three C*hloromonas* strains exhibited a pyrenoid (Figure 5.4D to F; Appendices 19 to 21), consistent with the general observations made on *Chloromonas* (Morita *et al*., 1998).

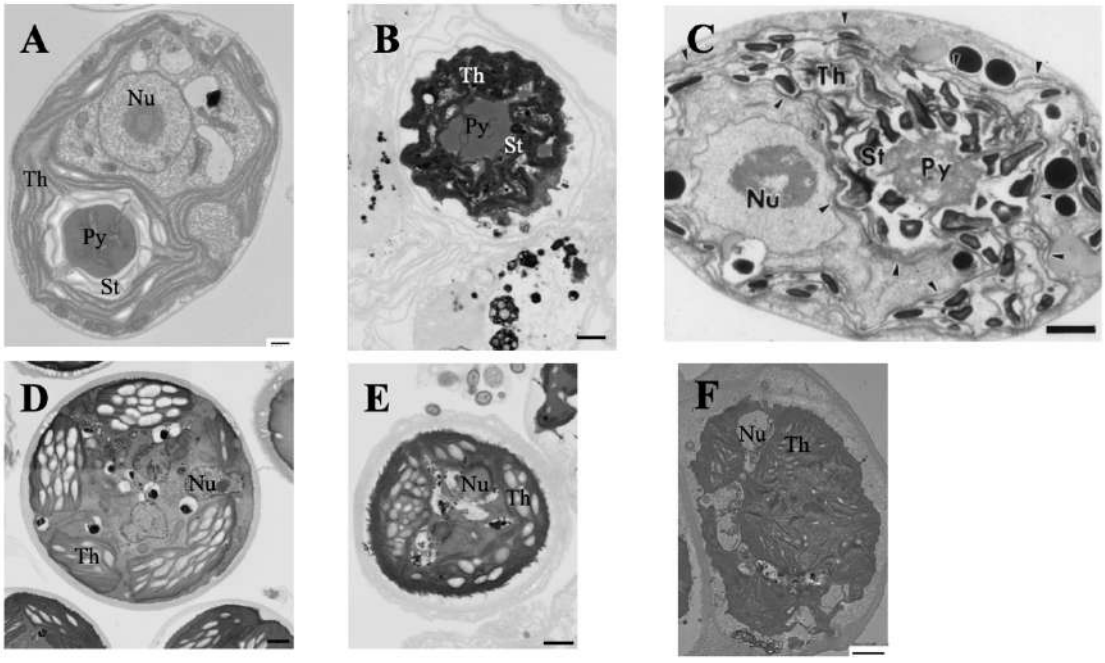


Figure 5.4 Scanning Electron Microscopy (SEM) images of the 5 representative *Chlamydomonas/Chloromonas* and of *Chlamydomonas reinhardtii* (**A**: *Chlamydomonas reinhardtii*, **B**: *Chlamydomonas mutabilis*, **C**: *Chlamydomonas augustae*; Morita *et al*., 1999, **D**: *Chloromonas serbinowii*, **E**: *Chloromonas rosae*, **F**: *Chloromonas clathrata*). Pyrenoids can be observed in the three different *Chlamydomonas* species. Two different pyrenoid morphologies can be observed. Pyrenoid can be enclosed by one layer of starch plate (A) or enclosed by multiple starch grains (B, C). Scale bars: 500 nm (A and F) and 1 µm (B to E). Py: pyrenoid; Th: Thylakoids, St: Starch sheaths, Nu: Nucleus

5.2.3 Stable carbon isotope composition ($\delta^{13}C$) for organic matter reflects CCM strength

Stable carbon isotope composition ($\delta^{13}C$) for organic matter was also used as a second proxy for CCM activity in the six different species (Table 5.3). Interestingly, *Chl. reinhardtii, Chl. mutabilis* and *Chl. augustae* were isotopically enriched in cells grown under low CO_2 (-18.85, -17.2 and -18.8 ‰), with values close to the upper range seen in C_4 terrestrial plants and consistent with a fully-functioning CCM (Raven *et al*., 1982). In contrast, *Chr. clathrata, Chr. serbinowii* and *Chr. rosae* were isotopically depleted (-25.87, -23.64 and -25.33 ‰ respectively) compared to the *Chlamydomonas* strains with values found between C_3 and C_4 plants (O'Leary, 1988) and found in organisms with a CCM phenotype prone to leakiness (see Chapter 4) or limited carbon accumulation capacity. Interestingly, despite being isotopically depleted *Chr. serbinowii,* CCM + pyrenoidless, exhibited a $\delta^{13}C$ value intermediate between (-23.64‰) species with CCM and pyrenoid+ (*Chl. augustae* and *mutabilis*) and species without CCM and without CCM (*Chr. clathrata*). Generally, the $\delta^{13}C$ values lined up with the photosynthetic activity measurement previously found (see Chapter 4).

Table 5.3 $\delta^{13}C$ for organic matter for the 5 representative *Chlamydomonas/Chloromonas* and in *Chlamydomonas reinhardtii*. Values are means ± SEM, n= number of replicate.

	$\delta^{13}C$ (‰) Low CO_2 (0.04% CO_2)	**$\delta^{13}C$ (‰) High CO_2 (5% CO_2)**
Chlamydomonas reinhardtii (n=3)	-18.85 ± 0.01	-58.78 ± 0.05
Chlamydomonas augustae UTEX LB 1969 (n=3)	-18.80 ± 0.03	-60.34 ± 0.02
Chlamydomonas mutabilis UTEX 578 (n=3)	-17.20 ± 0.02	-57.33 ± 0.10
Chloromonas serbinowii UTEX LB 492 (n=3)	-23.64 ± 0.30	-63.04 ± 0.02
Chloromonas rosae UTEX 1337 (n=3)	-25.33 ± 0.01	-54.96 ± 0.03
Chloromonas clathrata UTEX LB 1970 (n=3)	-25.87 ± 0.01	-59.09 ± 002

5.2.4 *Chlamydomonas* and *Chloromonas* exhibit different cell morphologies

Cell morphologies were compared to see if *Chlamydomonas* and *Chloromonas* developed morphological adaptations other than the presence of pyrenoids. Measurements were successfully obtained. All the strains expressing a CCM showed similar cells surface area (between 39.4 and 76.2 μm^2) whereas *Chr. clathrata* (CCM- pyr-) exhibited a significantly smaller cell surface area (12.58 μm^2; Table 5.4). When present, the volume of pyrenoid was between 10 and 20% relative to the total cell size, consistent with the previous findings in the literature (Rawat *et al*., 1996). Total volume of thylakoids relative to the total cell surface were all similar and did not show any differences between the different physiological mechanisms. Finally, cell wall widths showed significant differences between the two *Chlamydomonas* and *Chloromonas* genera. *Chloromonas* had wider cell walls (0.328, 0.704 and 0.544 μm), 3 to 7 fold larger than in *Chlamydomonas* (0.022; 0.068 and 0.130 μm) suggesting that the total absence of CCM led to smaller cell size in *Chr. clathrata,* but also that the absence of pyrenoid in *Chloromonas* strains were associated with thicker cell walls. Such observations could be consistent with the presence of carbonic anhydrases or inorganic pumps to convert directly HCO_3^- from the extracellular bulk into CO_2 but could also reflect a mechanism to avoid leakiness.

Table 5.4 Summary of different cell measurements: total cell surface area, estimation of the volume of the pyrenoid and of thylakoids compared to total cell sizes and cell wall widths for the 5 strains compared to *Chlamydomonas reinhardtii* based on SEM images (Appendices 15-18).

Species names	replicate (n)	Total surface area (µm2)	SD	SE	% area pyrenoid	% area thylakoids	Cell wall width (µm)	SD	SE
Chlamydomonas reinhardtii	**3**	53.49	4.90	2.83	12.33%	45.54%	0.022	0.006	0.003
Chlamydomonas augustae	**1**	54.95	NA	NA	19.22%	38.63%	0.068	NA	NA
Chlamydomonas mutabilis	**3**	39.40	14.57	8.41	11.24%	59.90%	0.130	0.026	0.015
Chloromonas serbinowii	**10**	76.25	21.93	6.94		51.18%	0.328	0.036	0.011
Chloromonas rosae	**10**	46.92	11.81	3.73		48.68%	0.704	0.207	0.069
Chloromonas clathrata	**1**	12.582	NA	NA		62.00%	0.544	NA	NA

5.3 Discussion

Due to their close phylogenetical relatedness (See Introduction and Figure 5.2) and the presence of multiple CCM activities across different strains, *Chlamydomonas* and *Chloromonas* are good study models to understand pyrenoid occurrence in green algae. However, the physiological mechanisms explaining the absence of pyrenoid and/or CCM are not fully understood. Therefore, we explored the diversity of CCM activity for five poorly characterised *Chlamydomonas* and *Chloromonas* species by measuring the photosynthetic affinity for CO_2, pyrenoid imaging and through analysis of organic matter $\delta^{13}C$. These novel measurements were then compared to *Chl. reinhardtii*, in which the CCM activity is already well described (Badger *et al.,* 1980; Sültemeyer *et al.,* 1989; Meyer *et al.,* 2012; Mitchell *et al.* 2014).

5.3.1 Absence of pyrenoid is not necessarily linked to absence of CCM

Although easier to work with compared to streptophyte algae (See Chapter 4), measurements of photosynthetic affinity by O_2 electrode remained challenging but a good approach to test for the presence of CCM. The results clearly highlighted that the absence of pyrenoid (supported by the EM images in the *Chloromonas* strains) was associated with a clear shift towards a lower affinity for CO_2 (higher $K_{0.5}$ values) in low CO_2 conditions. Despite this shift in the $K_{0.5}$ values compared to the *Chlamydomonas* strains, the *Chloromonas* strains showed a diversity in their mechanisms of Ci uptake with mainly 2 states (CCM + pyr- and CCM- pyr-). Only the combination of the O_2 evolution measurements, $d^{13}C$ analyses and Morita's observations allowed us to establish the states of the different *Chloromonas* strains. The presence of intracellular Ci pools were therefore observed in *Chr. rosae* and *Chr. serbinowii* despite the absence of pyrenoid whereas an absence of CCM associated with an absence of pyrenoid was concluded in *Chr. clathrata*. Overall, *Chlamydomonas/Chloromonas* showed a diversity in their mechanisms of Ci uptake with three main states:

- Presence of CCM associated with a pyrenoid (*Chlamydomonas* strains)
- Presence of CCM without a pyrenoid (*Chr. rosae* and *Chr. serbinowii*)
- Absence of CCM, absence of pyrenoid (*Chr. clathrata*)

The diversity of inorganic carbon acquisition in algae has been the subject of intensive studies in the eighties and the nineties, however, the present data did not bring enough

information to make conclusions on the type of inorganic carbon uptake. Species with a CCM rely either on diffusive CO_2 uptake, active transport of HCO_3^- and CO_2 or on an external carbonic anhydrase to facilitate $HCO3^-$ uptake (Colman *et al*., 2002). *Chl. reinhardtii* and *Dunaliella tertiolecta* have been shown to rely on active Ci transport at the chloroplast envelope (Amoroso *et al.,* 1998; Moroney & Chen, 1998) whereas no active transport of Ci was detected in *Chlorella ellipsoide* (Rotatore &Colman, 1991a).

The absence of CCM/pyrenoid is quite a rare trait (See Chapter 3) and potentially explains why there is a lack of studies on these organisms (Raven, 2010): " *the reason for emphasizing algae which lack CCMs is that present evidence suggests that they are in a small minority, so that each discovery of new species has rarity value*". Based on the observations made in this chapter, it appears that the absence of pyrenoid is not necessarily linked to the absence of some sort of CCM which emphasizes even more the need to understand the physiological mechanism used in these species. Absence of CCM has been detected in *Coccomyxa* (Palmqvist *et al*., 1994a, 1994b, 1995, 1997) but also in some red algae (Raven & Beardall, 1981; Raven *et al*., 1982; MacFarlane & Raven, 1985, 1989, 1990; Maberly, 1990; Johnston *et al*., 1992; Mercado & Niell, 1999; Murru & Sandgren, 2004). Interestingly, the lack of CCM is not necessarily linked to the genus, as *Euglena mutabilis* never express a CCM and therefore always depend on diffusive exchange of CO_2 (Colman & Balkos, 2005), whereas the photosynthetic protozoan *Euglena gracilis* can express a CCM when grown at high pH and has therefore a CCM based on the active influx of HCO_3^-. There are also entire clades of algae that cannot use bicarbonate as a source of inorganic carbon, for example, the Chrysophyceae sensu latu (group of heterokont algae without CCM) have not been able to show the use of HCO_3^- (Saxby-Rouen *et al.,* 1998; Maberly *et al*., 2009; Bhatti & Colman, 2011; Maberly & Gontero, 2017). This list is far from exhaustive but generally, the local conditions appear to be an important factor for the presence of a CCM or not. However, despite the absence of direct evidence to characterise the mechanisms of inorganic carbon uptake potential hypotheses can be made. The present data show that a smaller cell size appeared to correlate with the absence of CCM. However it remains difficult to explain why *Chr. clathrata* has a small cell size. Does the small size reflect limited growth due to the absence of a CCM or is the small cell size a way to increase the concentration of the small Ci pool around Rubisco? It is important to remember that more anciently diverged organisms such as cyanobacteria also have small sizes but express a CCM (carboxysomes, Badger *et al*., 2002; Price *et al.,* 2008), therefore, small sizes cannot directly be linked to absence of CCM.

In addition, the thicker cell walls were consistent with some of the hypotheses developed in Lucas & Berry (1985) and could support the presence of either active HCO_3^- pump transport system or carbonic anhydrases directly in cell walls to avoid leakiness.

Finally, an important point to discuss here is the growth conditions of these strains. All the analyses performed in this study have been conducted at room temperature (20°C), however, most of these strains are snow algae, growing in environments with much lower temperatures. Low temperatures are usually associated with reduced photorespiration (Linke & Seidell, 1965; Kaye & Laby, 1973; Long, 1991). Decreasing temperatures increase solubility of CO_2 relative to O_2. With a CO_2 solubility in water at 0° and 5°C estimated at 24 and 21.8 µM (Figure 5.5 A), an O_2 solubility measured at 457 and 399 µM at 0° and 5°C respectively (Figure 5.5 C), a specificity of Rubisco for CO_2 relative to O_2 estimated at 250 and 220 µM at 0° and 5°C (Figure 5.5 B) and a selectivity for CO_2 over O_2 (W) describes as:

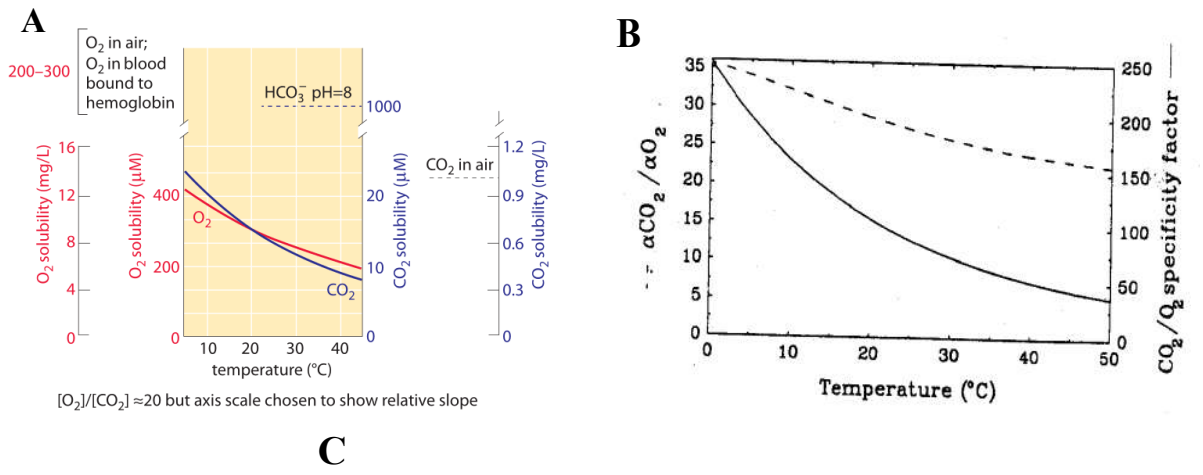
Equation 2

$$W = \frac{V_c.K_o}{V_o.K_c} . \frac{[CO_2]}{[O_2]}$$

Where V_c and V_o are the maximum velocities for carboxylation and oxygenation and K_c and K_o are the relative Michaelis-Menten constants for CO_2 and O_2 respectively and $[CO_2]$ and $[O_2]$ the concentrations of CO_2 and O_2 dissolved in water. The ratios between photosynthesis and photorespiration are therefore estimated at 13 at 0°C and 12 at 5°C. This means that there are 13 carboxylase for 1 oxygenase event at 0°C vs 12 carboxylase for 1 oxygenase event at 5°C relative to 3.5 at 20°C in absence of CCM. Therefore, the data collected in the *Chloromonas* strains are likely to vary in natural conditions. Such observations raises the question of the requirement for a CCM in low temperatures environments. If the temperature decreases the solubility of O_2 in water, it also reduces diffusion rate of CO_2 (Boudreau, 1997) but also slows down the turnover rate of Rubisco (Dutkiewicz *et al*., 2009. Young *et al.,* 2015). Therefore, lower temperatures cannot be directly associated to an absence of CCM. Work in diatoms (Kranz *et al*., 2015) showed that the lower temperatures were associated to a reduction of the energetic requirement of a CCM and not an absence of CCM expression.

5.3.2 $\delta^{13}C$ values reflect carbon accumulation capacity in *Chlamydomonas/Chloromonas* strains

The stable carbon isotope composition ($\delta^{13}C$) of organic matter is a good second proxy for CCM activity (Section 4.2.2; Meyer *et al.,* 2008) and once again the present data perfectly lined up with the oxygen evolution and morphological findings. Although, the $\delta^{13}C$ obtained in streptophyte algae are consistent with a CCM phenotype prone to leakiness (potential retro-diffusion of CO_2), the values obtained in the different *Chlamydomonas/Chloromonas* strains are likely to reflect carbon accumulation capacity. Generally, the absence of pyrenoids is correlated with more negative $\delta^{13}C$ value. However, this correlation is not absolute (Kevekordes *et al.*, 2006) as $\delta^{13}C$ reflects the trade-off between the expected diffusive CO_2 entry and the occurrence of some sort of CCM (Diaz & Maberly, 2009). In our study, the clear shift towards more negative $\delta^{13}C$ values is directly associated with the absence of pyrenoid. Interestingly, *Chr. serbinowii* (CCM+ pyrenoid-) showed a value intermediate between the *Chlamydomonas* strains and the two other *Chloromonas*, which surely reflects the presence of some sort of CCM, as the $K_{0.5}$ values also suggested. The absence of a similar $\delta^{13}C$ value in *Chr. rosae* is surprising, however, as, the ratio probably reflects the weak diffusive CO_2 entry.



C

Pressure abs	*mm Hg*	760		
	psi	14.7		
	bar	1		
	kPa	101.1		
Temperature				
°C	*°F*	*µMol*	*mg/l*	*ml/l*
0	32	457	14.6	10.2
5	41	399	12.8	9.1
10	50	353	11.3	8.2
15	59	315	10.1	7.5
20	68	284	9.1	6.8
25	77	258	8.3	6.3
30	86	236	7.6	5.9
35	95	218	7	5.5
40	104	202	6.5	5.2
45	113	189	6	4.9
50	122	177	5.6	4.6

Figure 5.5 Tables used to estimate the ratio between photosynthesis and photorespiration at 0° and 5°C **A** Oxygen and carbon dioxide solubility in water and their dependence on temperature under normal air composition. **B** Specificity of Rubisco for CO_2 relative to O_2 and the ration of solubilities (α) of CO_2 and of O_2 in water (pH=7) from Long (2011). **C** Oxygen solubility in water with temperature at 101 kPa.

5.4 Conclusions and future works

Chlamydomonas and *Chloromonas* are two genera phylogenetically close to each other (Buchheim *et al.,* 1997, Pröchold *et al.,* 2001; Nozaki *et al*., 2002), only differentiated by the absence of pyrenoid in *Chloromonas* (Figure 5.2 and 4). Carbon concentrating mechanisms are the result of a balance between local environmental pressures such as temperature and salinity, alongside Rubisco catalytic properties and traits specific to genus. The present data confirms the potential of these strains to help understand pyrenoid expression.

However, following the present experiments, further work needs be undertaken to fully characterize the physiological mechanisms of these five strains. Inorganic carbon uptake mechanisms in these related strains especially in *Chloromonas* remain unknown.

Thicker cell walls have been observed in species without pyrenoid, subsequent identification and characterisation of potential pumps or active transport systems in cell walls would need to be undertaken either in generating *Chloromonas* mutants or in looking for some specific genes in genome sequencing. In addition, as suggested in Chapter 4 for the streptophyte algae, characterising the main source of inorganic carbon is required. This would not only help to differentiate potential leakiness from weak whole cell affinity for CO_2 but to also further characterise the diversity of CCM in these related genera. Similar work in diatoms (Clement *et al.,* 2017) showed that *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* use two different sources inorganic carbon (HCO_3^- and CO_2 respectively). It is, therefore, possible to hypothesize that the different *Chloromonas* strains could have different inorganic carbon uptake mechanisms. Finally, such data would also help us, combined with cell measurements and Raven's equation (Raven, 1991) to calculate CO_2 diffusion through the external cell wall and therefore to fully characterise the physiological mechanisms of these five strains.

Chapter 6: Genomic comparison of strains lacking CCM and/or pyrenoid

6.1 Introduction

In Chapter 5, the physiology of five related strains of *Chlamydomonas* and *Chloromonas* were compared. The physiological analyses showed that the presence of a CCM is not necessarily linked to the presence of a pyrenoid. Similar to land plants, which show a range of physiological mechanisms, green algae species exhibit a range of CCM activity which may be associated with a pyrenoid. However, no study to date has attempted to explain the genetic basis to the expression of a CCM outside *Chl. reinhardtii*. In light of the development of new technologies, such as whole genome sequencing, the aim of this chapter was to investigate the genetic basis behind the observed presence/absence of a CCM with/without a pyrenoid. Specifically, this chapter will build upon the physiological measurements introduced in Chapter 5, and seek to answer three main questions: *i)* Based on whole genome sequencing can we infer the multiple, independent origins of the pyrenoid? *ii)* Do the interactions between LSUs and SSUs change between *Chlamydomonas* and *Chloromonas,* and could this affect the way EPYC1 binds to Rubisco? *iii)* What are the essential genes necessary for pyrenoid or CCM formation that are present in the *Chlamydomonas* strains and perhaps absent in *Chloromonas* (or vice-versa)?

6.1.1 Using Carbon Concentrating Mechanisms to improve crop production

Due to climate change and a growing population, the pressure to feed the world is increasing. World cereal production needs to be increased by at least 70% by 2050 to meet the demand for food (Covshoff & Hibberd, 2012; Freibauer *et al.,* 2011). One limitation on crop production is photosynthesis, due to the low efficiency of Rubisco, particularly in C_3 plants (Jordan & Ogren, 1981; Sage *et al*., 2002), which includes beans, rice, wheat and potatoes. Rubisco has a slow carboxylation rate and a low affinity for CO_2 and is also limited by photorespiration (see General Introduction). Although, CO_2 is directly fixed during the day in C_3 plants, 20-30% of potential CO_2 fixation is lost to photorespiration in moderate conditions (Andersson, 2008, Bauwe *et al.,* 2010; Zhu *et al*., 2010).

Different approaches have been attempted to overcome these limitations (Whitney *et al*., 2011; Parry *et al.,* 2012; Carmo-Silva *et al.,* 2015; Ort *et al.,* 2015). Firstly by placing C_4 photosynthesis into C_3 leaves in order to reduce the oxygenation activity by concentrating CO_2 around Rubisco (Hibberd *et al*., 2008) but also by optimising Rubisco catalytic properties to reduce oxygenation potential (Ellis, 2010). The photorespiratory pathway is often considered as a wasteful process (Betti *et al.,* 2016; Busch *et al*., 2018), therefore, engineering a photorespiratory bypass to limit photorespiration activity has also been the focus of research (Maier *et al.,* 2012; Carvalho *et al.,* 2011; Kebeish *et al.,* 2007; South *et al.,* 2019).

Introducing components from biochemical CCMs (carboxysomes or pyrenoid) in order to elevate CO_2 concentration in the chloroplast has also been considered (Von Caemmerer *et al.,* 2012; Price *et al.,* 2012; Meyer *et al.,* 2016). This approach is not only supported by modelling (Price *et al*., 2012; McGrath & Long, 2014) but also by preliminary work which showed that a complex Rubisco construct with *Chlamydomonas* SSU and EPYC1 could be stably expressed and localized in the chloroplast of *Arabidopsis thaliana* (Atkinson *et al.,* 2019).

Following the work of Meyer *et al.* (2012), Mackinder *et al*. (2016, 2017), transcriptome data from synchronised cells (Mitchell *et al.,* 2014; Zones *et al.,* 2015), has allowed analyses of gene regulatory networks and protein-protein interactions (Gita Yadav/Citu Gulia, University of Cambridge, unpublished data), to identify a total of 88 genes essential components for pyrenoid formation in *Chl. reinhardtii* (see Table 2.5; Chapter 2). However, at this stage how these proteins interact with each other and their expression profiles are not fully understood, particularly across closely related genera which differ in expression of a CCM with or without a pyrenoid (Chapter 5).

6.1.2 *Chloromonas*, a good study model to introduce a pyrenoid?

Despite some progress towards the introduction of the algal CCM in higher plants (Atkinson *et al*., 2019), this research is at an early stage and the molecular interactions/expressions between the different proteins are not fully understood even in *Chl. reinhardtii*. Therefore, characterising the expression of a pyrenoid in a closely related species might inform the work currently being undertaken in higher plants. Eukaryotic algae expressing a CCM but without a pyrenoid, such as some *Chloromonas* strains (Chapter 5), would be a good model to study gene expression associated with such compartments. This would facilitate the

understanding of the mechanisms of inorganic carbon uptake in organisms without a pyrenoid, compared to the a molecular definition established in *Chl. reinhardtii*.

6.1.3 Objectives of this study

The last twenty years have seen the rapid development of new technologies such as Whole Genome Sequencing (WGS). The relatively low cost of this technology has facilitated analysis and publication of whole-genome sequences from algal study models such as *Chl. reinhardtii* (Merchant *et al*., 2007), *Klebsormidium flaccidum* (Hori *et al.,* 2014) *Chara braunii* (Nishiyama *et al.,* 2018), and more recently from *Spirogloea muscicola* and *Mesotaenium endlicherianum* (Cheng *et al.,* 2019). Given the limited time available to process this data because close to the end the data collection period, the aim of this Chapter was to use WGS of the five *Chlamydomonas* and *Chloromonas* strains (introduced in the chapter 5; *Chlamydomonas augustae* UTEX LB 1969; *Chlamydomonas mutabilis* UTEX 578; *Chloromonas serbinowii* UTEX LB 492; *Chloromonas rosae* UTEX B 1337 and *Chloromonas clathrata* UTEX LB 1970) to address three specific questions:

1. Does the pyrenoid come from a common ancestor or does it have multiple and independent origins?
2. Secondly, are there any patterns of interactions between SSUs and LSUs in the different *Chlamydomonas* and *Chloromonas* strains, which could reflect the way EPYC1 binds to Rubisco (if present)?
3. Finally, based on the list of the 88 genes essential for the pyrenoid formation found in *Chl. reinhardtii*, which are present in the *Chlamydomonas* strains but absent in *Chloromonas* (and vice-versa)?

6.2 Results

6.2.1 The five strains have different genome size

The re-sequencing of the five strains using short reads (BGI) was generally successful with high quality reads and good coverage, whereas long read PacBio sequencing was very challenging with a very poor coverage. *Chr. rosae* had the largest genome size (141 Mb) but only 8.3 Mb of *Chr. serbinowii* was sequenced with the PacBio technology. However, reasonable genome sizes were obtained through a combination of both short and long reads (Table 6.1).

Interestingly, the five genomes exhibited different genome sizes (Table 6.1). *Chl. augustae* showed the smallest genome size (109 Mb) whereas *Chl. mutabilis* and *Chr. serbinowii* exhibited genome sizes of 151.1 and 144.1 Mb respectively. Finally, *Chr. rosae* and *Chr. clathrata* appeared to have the largest genomes (289 and 302 Mb). Overall, compared to the size of *Chl. reinhardtii,* the results suggested that the five new strains were successfully sequenced, but probably with quality variations between the different strains.

Table 6.1 Genome sizes of the five newly sequenced strains obtained with the two sequencing methods and of the hybrid assembly compared to *Chlamydomonas reinhardtii*

	Genome size (short reads/BGI)	**Genome size (long reads/PacBio)**	**Genome size (hybrid)**
Chlamydomonas reinhardtii			120 Mb (Merchant *et al*., 2007)
Chlamydomonas augustae	102.3 Mb	79.6 Mb	109 Mb
Chlamydomonas mutabilis	150.2 Mb	17.1 Mb	151.1Mb
Chloromonas serbinowii	144.1 Mb	8.3 Mb	144.1 Mb
Chloromonas rosae	180.9 Mb	141 Mb	289 Mb
Chloromonas clathrata	274 Mb	32.2 Mb	302 Mb

6.2.2 The pyrenoid has multiple and independent origins

Chloroplastic CDS were successfully extracted (Appendix 22) from the new *de novo* assemblies with BLAST and were included in the green algae chloroplastic phylogeny (Figure 6.1; black arrows). With the exception of *Palmophyllum crissum*, which was clustered with the prasinophytes, the 3 main clades (streptophyte algae, prasinophyte and chlorophyte) were reconstructed with generally good support, except for some nodes leading to branches which recently diversified, and with a tree topology generally consistent with the most recent green algae phylogenies (Lelieart *et al.,* 2012; Leebens-Mack *et al.,* 2019). The pyrenoid occurrence, when mapped on to this chloroplast phylogeny, appeared to be a widely distributed trait.

Of the 64 green algal species used to build this phylogeny, 15 were reported to be without a pyrenoid, which represent 23% of the species used in this phylogeny. Interestingly, none of the streptophyte algae were reported without pyrenoid. The absence of a cluster including all the species without pyrenoid suggested that the pyrenoid is not inherited from a common ancestor and can be lost multiple times during evolution. The loss of a pyrenoid would be

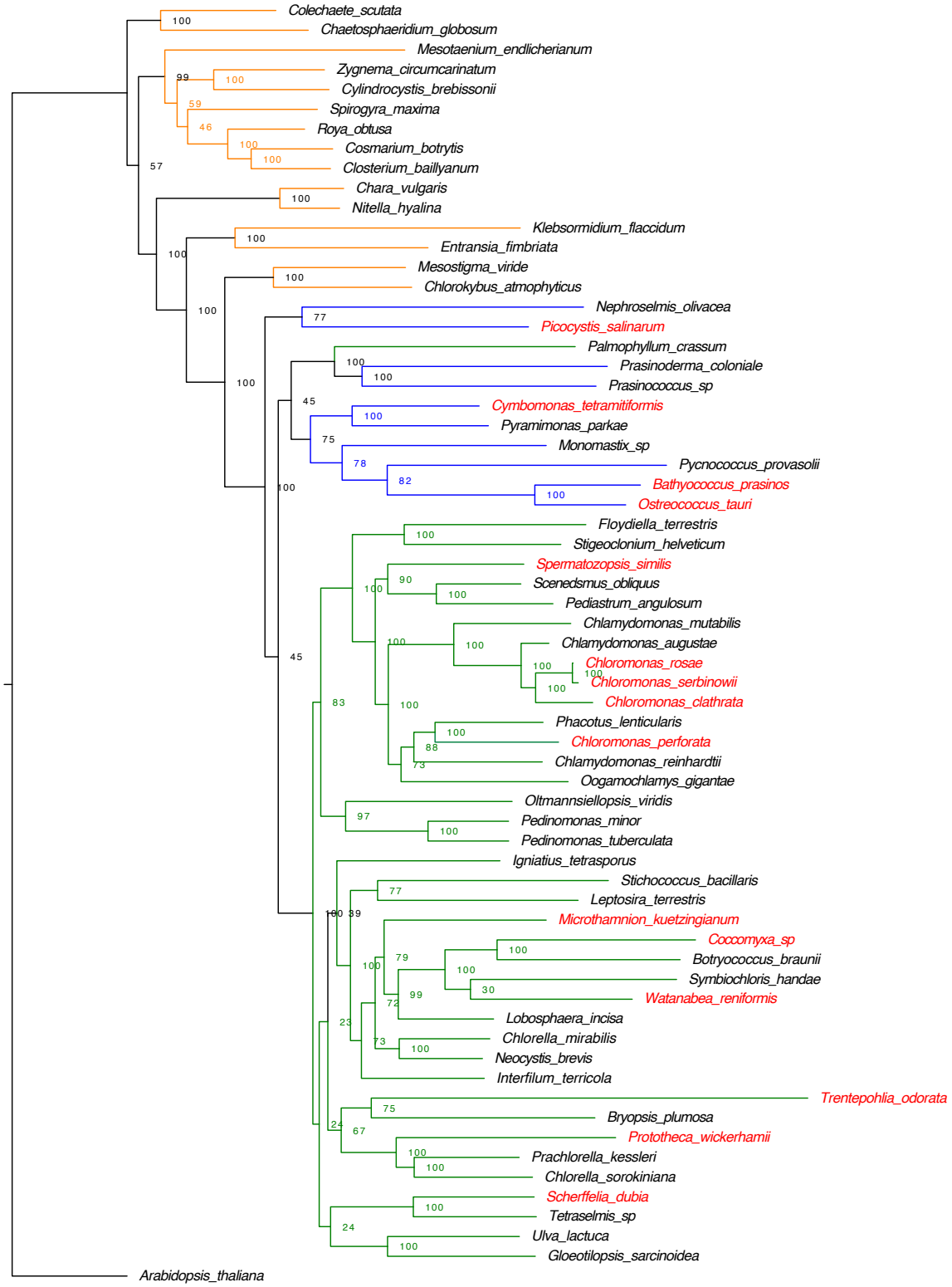


Figure 6.1 Phylogenetic tree of 64 green algae species, including the *Chlamydomonas* and *Chloromonas* strains used in Chapter 5 and Figure 6.2. This tree was built with RAxML (Stamatakis, 2014) based on the nucleotide alignment of 44 chloroplastic genes. *Arabidopsis thaliana* was used as an outgroup. Species without pyrenoid were highlighted in red and the new sequenced strains are indicated with a black arrow. Streptophyte algae were labelled in orange, prasinophytes in blue and chlorophytes in green. The pyrenoid appears to have been lost 12 times across this phylogeny of green algae. Node labels indicate bootstrap values (support values for 100 replicates).

characterised by the transition from a pyrenoid positive (pyr+) to pyrenoid negative (pyr-) state. The distribution of species without a pyrenoid across the phylogeny suggested that this structure was lost at least 10 times through the evolution of the green algae.

Within the prasinophytes, the pyrenoid was lost, for example, from *Pycnococcus provasolii* to *Bathyococcus* and *Ostreococcus tauri* which were clustered together with the transition from an external branch (pyr +) to two internal (pyr-) branches.

In addition, the five newly sequenced strains were clustered together, with high support. However, *Chl. reinhardtii* and *Chr. perforata* appeared to be in a sister clade to the five newly sequenced strains, also joined by *Phacotus lenticularis* and *Oogamochlamys gigantea.*

6.2.3 The genus *Chloromonas* is not monophyletic

The chloroplast phylogeny of the four *Chloromonas* and the three *Chlamydomonas* strains was successfully built with the 44 chloroplastic genes (Figure 6.2) previously extracted from the new genomes assemblies (Appendix 22). The aim of this phylogeny was to see whether the *Chloromonas* strains would form a monophyletic clade including all the strains without a pyrenoid. The two strains from GenBank (*Chl. reinhardtii*; NC005353 and *Chloromonas perforata*: KT625416) appeared to be clustered together whereas all the newly sequenced strains formed a separate clade (Figure 6.2). The two new sequenced *Chlamydomonas* strains (*augustae* and *mutabilis*) formed the basal branches of this cluster and the three *Chloromonas* strains were located in the most internal branches.

Following the explanation in the previous paragraph, the data in Figure 6.2 suggested two loss events in this restricted phylogenetic grouping. These losses were observed at the level of the *Chloromonas* strains. The first one occurred with *Chr. perforata* (pyr-) whereas the second event occurred in the ancestor of the three new sequenced *Chloromonas* strains (*serbinowii*, *rosae* and *clathrata*). Interestingly, *Chr. serbinowii* and *Chr. clathrata*, the two strains without a pyrenoid but with the presence of CCM activity, formed a monophyletic clade in the two most internal branches with a very high support value (100). Therefore, this multi-marker phylogeny did not support the monophyly of *Chloromonas*.

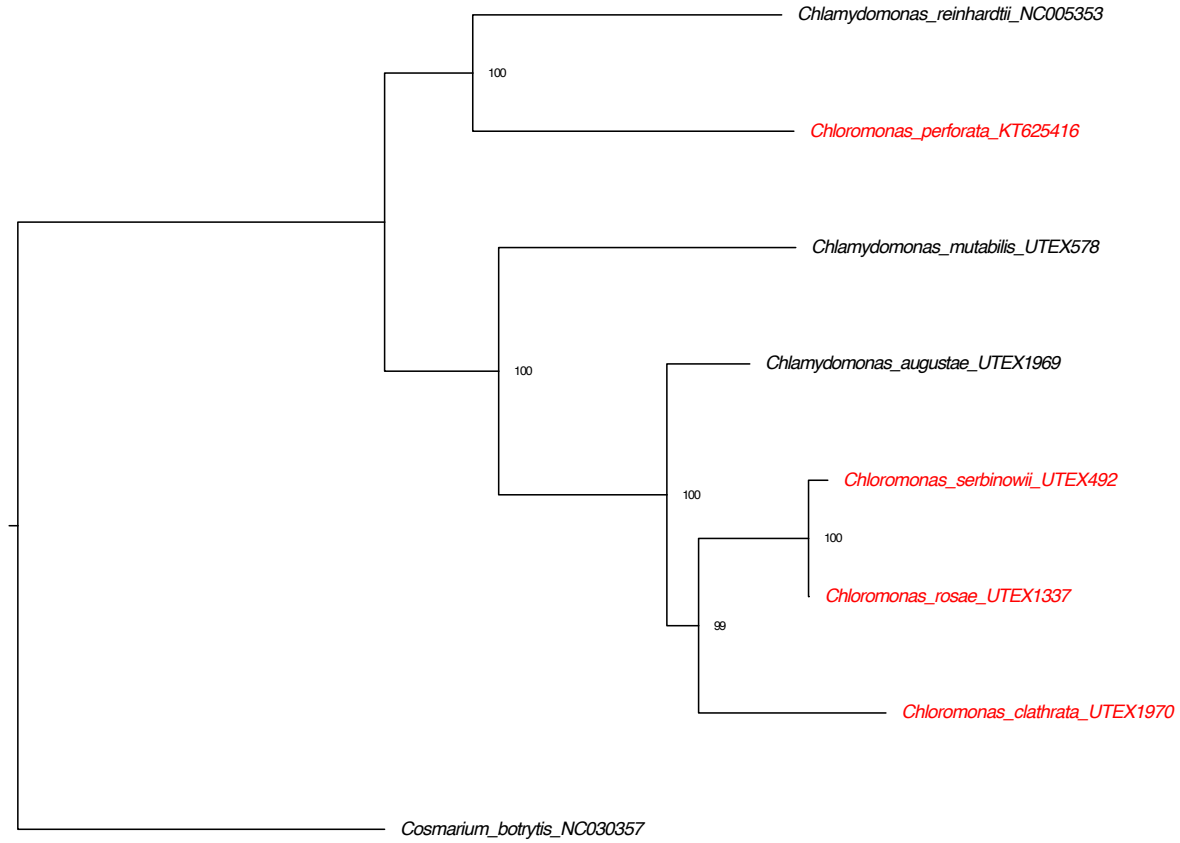


Figure 6.2 Phylogenetic tree of 4 *Chloromonas* and 3 *Chlamydomonas* strains based on the nucleotide alignment of 44 chloroplastic genes and built with RAxML (Stamatakis, 2014). *Cosmarium botrytis* was used as an outgroup. Strains without pyrenoid (See Chapter 5; Mackinder *et al*., 2016) are highlighted in red (all the *Chloromonas* strains). *Chloromonas* strains do not form a monophyletic clade. The pyrenoid appeared to have been lost with (*Chloromonas perforata*) and in the three *Chloromonas* strains' ancestor (*Chloromonas serbinowii, rosae* and *clathrata*). Node labels indicate bootstrap values (support values for 100 replicates).

6.2.4 The interactions between LSUs and SSU do not reflect the pyrenoid occurrence in *Chlamydomonas* and *Chloromonas*

The analysis of the Rubisco PDB structure of *Chl. reinhardtii* (1GK8; Taylor *et al.,* 2001) was successful and allowed the comparison of the interactions between LSUs and SSUs in this study model and the new sequenced strains. The *Chl. reinhardtii* Rubisco was characterised by the presence of 8 LSUs encoded by rbcL (475 amino acids) and 8 SSUs (140 amino acids) (Figure 6.3 and 6.5). Generally, one SSU appeared to interact with 3 LSUs confirming Spreitzer's observations (Spreitzer, 2003) with a total of 66 residues interacting with at least one of the 3 LSUs. Over these 66 residues, 4 were located in the first α-helix (A), 1 in the second α-helix (B) and 16 in the βA-βB loop (Figure 6.6).

Figure 6.3 Alignment of rbcL sequences used for Rubisco modelling. The *Chlamydomonas reinhardtii* sequence was used as template (PDB number: 1gk8) and the 4 other rbcL sequences were extracted from the whole genome sequencing.. * indicates positions which have a single fully conserved residues; «:» indicates a site belonging to group exhibiting strong similarity (strong score >0.5); «.» indicates sites belonging to a group from weak similarity (weak score =<0.5).

Rubisco modelling and the subsequent interaction analysis were only possible after extraction of *rbcL* and *RbcS* sequences from the new assemblies. In the newly sequenced strains, *rbcL* sequences were shown a great consistency between the different sequencing methods (Figure 6.3). Overall, the sequences were quite conserved across the five newly sequenced strains with more than 50% similarity (Figure 6.3; Table 6.2). However, two distinct patterns were observed between *Chlamydomonas* and *Chloromonas*. C*hlamydomonas* sequences were on average similar to each other with 90% of similarity. Comparisons between *Chloromonas* sequences showed similar values, whereas inter-genus comparison of *Chlamydomonas* and *Chloromonas* sequences showed only around 50% similarity (Figure 6.3; Table 6.2).

Analysis of *RbcS* sequences was more challenging. The *RbcS* sequences appeared to be very different from the *RbcS* sequences in *Chl. reinhardtii*, even in the new *Chlamydomonas* strains sequenced. Multiple copies were found and inconsistencies between the different sequencing methods were observed as well as the presence of a short βA-βB loop, which was not consistent with the observations in Chapter 3. In addition, the absence of full long read sequencing with the PacBio technology for some strains meant that only the short read BGI sequences were available for *Chl. mutabilis* and *Chr. serbinowii*. Furthermore, it was not possible to establish the functionality of these R*bcS* copies. Therefore, only the copies with relatively consistent α-helices sequences between the different methods were used for Rubisco modelling (Figure 6.4-5). In contrast to rbcL, the α-helices exhibited great variation between sequences and there was an absence of specific patterns between the two genera. The α-helices were never more than 67% similar (*Chr. serbinowii* and *Chr. clathrata*) and the least similar were *Chr. serbinowii* and *Chl. mutabilis* (30%). Among the 27 residues of the two α-helices, five amino acids were perfectly conserved across the five species: Q25, I26, Q29, Y32 in α-helix A and 92E in α-helix B (Figure 6.5, 7).

Following the extractions of *rbcL* and *RbcS* sequences in the new strains, Rubisco modelling of the new strains was successfully performed (using *Chl. reinhardtii* as template; Appendix 23) on 4 of the 5 strains: *Chl. augustae* and *mutabilis* (CCM+ pyr+), for *Chr. serbinowii* (CCM + pyr-) and *clathrata* (CCM- pyr-*). Chr. rosae* (CCM+ pyr-) was unused for this analysis due to a lack of consistency between the two sequencing methods, especially at the level of the second α-helix (Figure 6.4).

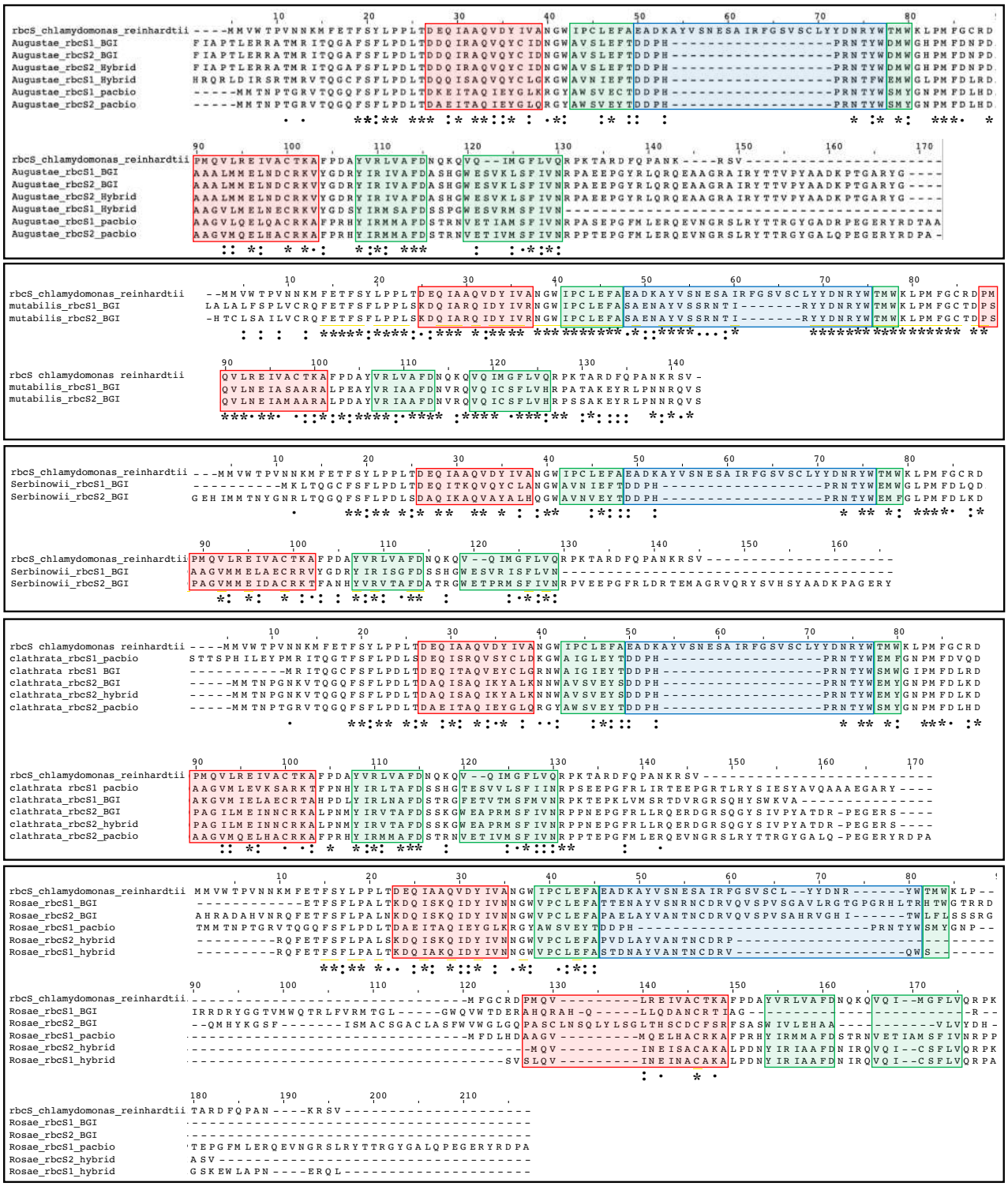


Figure 6.4 Alignments of the different *RbcS* sequences extracted from the different whole-genome sequencing (BGI, PacBio and hybrid). Sequences were aligned on the first copy of *RbcS* (*RbcS1*) in *Chlamydomonas reinhardtii*. α-helices were framed in red, β-strands in green and the βA-βB loop in dark blue. * indicates positions which have a single fully conserved residues; «:» indicates a site belonging to group exhibiting strong similarity (strong score >0.5); «.» indicates sites belonging to a group from weak similarity (weak score =<0.5)

Using the newly modelled Rubiscos, residues of the SSUs interacting with the LSUs were identified and compared to those found in *Chl. reinhardtii*. Overall, one SSU appeared to interact with 3 LSUs as described in *Chl. reinhardtii* in all the strains.

However, the number of interactions and where these interactions occurred varied between species (Table 6.6; Appendix 24) and as compared to *Chlamydomonas reinhardtii. Chl. augustae* and *Chl. mutabilis* (CCM+ pyr+) had generally more interactions between LSUs and SSU relative to the *Chloromonas* strains. With 67 residues of the SSU interacting with 3 LSUs, *Chl. augustae* was the strain with the most interactions, whereas *Chl. mutabilis* exhibited the same number of interactions as *Chl. reinhardtii* (66 interactions; Figure 6.6). *Chloromonas* strains had generally fewer interactions compared to their related *Chlamydomonas* strains. *Chr. serbinowii* (CCM+ pyr-) exhibited 63 interacting residues, whilst *Chr. clathrata* (CCM- pyr-) had the smallest number of interactions with 59 residues of the SSU interacting with the 3 LSUs.

In addition, the number of interactions occurring within the two SSU α-helices and within the βA-βB loop, our secondary structures of interest, varied between strains but also compared to *Chl. reinhardtii*. *Chl. augustae* (CCM+ pyr+) appeared to be the only new strain with residues within the second SSU α-helix interacting with a LSU whereas *Chl. mutabilis* (CCM+ pyr+), *Chr. serbinowii* (CCM+ pyr-) and *Chr. clathrata* (CCM- pyr-) did not show any residues of this helix interacting. Over the 28 residues forming βA-βB loop, most of them were identified as interacting residues in all the strains (Figure 6.6) without specific patterns differentiating strains with/without CCM or with/without pyrenoid. However, work is currently underway to confirm (via PCR) whether the extent that the *RbcS* βA-βB loop-lengths inferred from the sequencing are short or long.

Overall, the analyses of the interactions between LSUs and SSU did not show clear differences between the different strains. Therefore, the way subunits interact with each other did not reflect the different range of CCM activities.

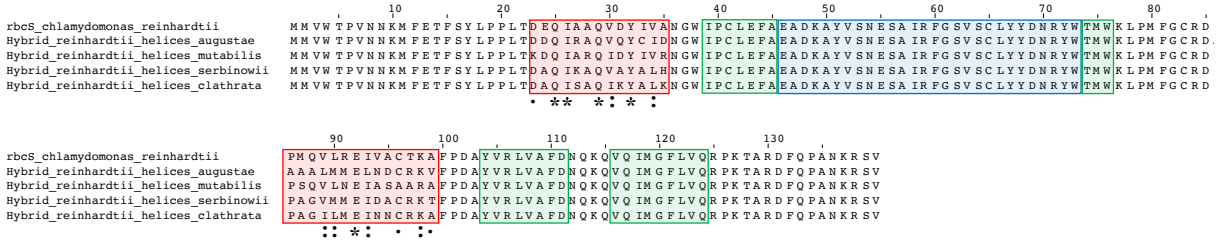
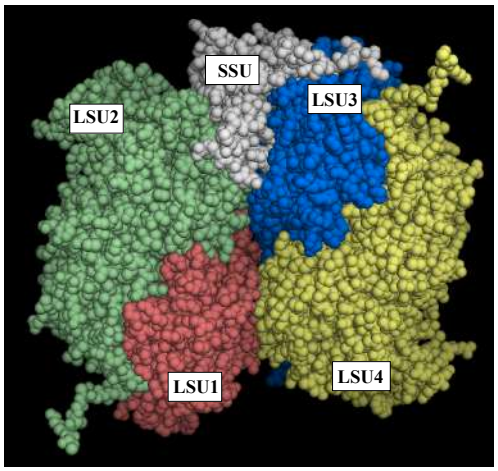


Figure 6.5 *RbcS* alignment of the sequences used for Rubisco modelling. *Chlamydomonas reinhardtii* was used as a template (PDB number : 1gk8). α-helices were framed in red, β-strands in green and the βA-βB loop in dark blue. Only the α-helices were annotated.* indicates positions which have a single fully conserved residues; «:» indicates a site belonging to group exhibiting strong similarity (strong score >0.5); «.» indicates sites belonging to a group from weak similarity (weak score =<0.5).

Table 6.2 rbcL and *RbcS* (only the α-helices) percentage of similarities between the different *Chlamydomonas* and *Chloromonas* strains.

rbcL	***Chlamydomonas reinhardtii***	***Chlamydomonas augustae***	***Chlamydomonas mutabilis***	***Chloromonas serbinowii***	***Chloromonas clathrata***	***Chloromonas rosae***
Chlamydomonas reinhardtii	100%					
Chlamydomonas augustae	91%	100%				
Chlamydomonas mutabilis	93%	95%	100%			
Chloromonas serbinowii	60%	55%	57%	100%		
Chloromonas clathrata	59%	56%	57%	93%	100%	
Chloromonas rosae	58%	54%	55%	86%	84%	100%

RbcS					
Chlamydomonas reinhardtii	100%				
Chlamydomonas augustae	37%	100%			
Chlamydomonas mutabilis	56%	22%	100%		
Chloromonas serbinowii	52%	52%	30%	100%	
Chloromonas clathrata	48%	48%	37%	67%	100%



Species name		Number of residues interacting with all the LSUs on SSU	Number of residues interacting on LSU1	Number of residues interacting on LSU2	Number of residues interacting on LSU3	Number of residues interacting on LSU4
Chlamydomonas reinhardtii	Total	66	14	31	50	0
	helix A	4				
	helix B	1				
	loop	16				
Chlamydomonas augustae	Total	67	11	31	49	0
	helix A	4				
	helix B	2				
	loop	21				
Chlamydomonas mutabilis	Total	66	13	29	51	0
	helix A	5				
	helix B	0				
	loop	18				
Chloromonas serbinowii	Total	63	10	27	50	0
	helix A	3				
	helix B	0				
	loop	21				
Chloromonas clathrata	Total	59	10	27	43	0
	helix A	6				
	helix B	0				
	loop	19				

Figure 6.6 Summary of the interactions between small subunit (SSU) and large subunit (LSU) of Rubisco after Rubisco modelling in 4 *Chlamydomonas* and *Chloromonas* strains compared to *Chlamydomonas reinhardtii*. One SSU interacts with 3 LSUs. The third column indicates the total number of residues interacting on the SSUs and how many of them are located within the 2 α-helices and in the βA-βB loop. The fourth, fifth, sixth and seventh columns indicate the number of residues interacting with the SSU on each of the four LSUs.

6.2.5 Comparison of the two α-helices and pyrenoid occurrence in *Chlamydomonas* and *Chloromonas* strains

As previously undertaken in Chapter 3 (Figure 3.2), a comparison of the amino acid composition of the two Rubisco SSU α-helices was undertaken in order to see if amino acid composition could explain CCM/pyrenoid occurrence. Except for the presence of 5 residues (see paragraph above) conserved across all the strains, none of the residues located in the α-helices showed any kind of pattern specific to *Chlamydomonas* or *Chloromonas* strains (Figure 6.7) or specific to a particular CCM activity, confirming the results previously found in Chapter 3. In addition, the comparison of the biochemical properties of the residues with a solvent-exposed side chain within the two α-helices between the new sequenced strains and with the α-helices of *Chl. reinhardtii*, confirmed the absence of specific residues which explain pyrenoid occurrence.

However, over the five amino acids conserved across the five species and in *Chl. reinhardtii*, four appeared to be buried residues, which would explain why they are less inclined to change and reflecting potential importance for Rubisco assembly. Overall, there is an absence of any common patterns which could discriminate species with or without a pyrenoid and with or without a CCM.

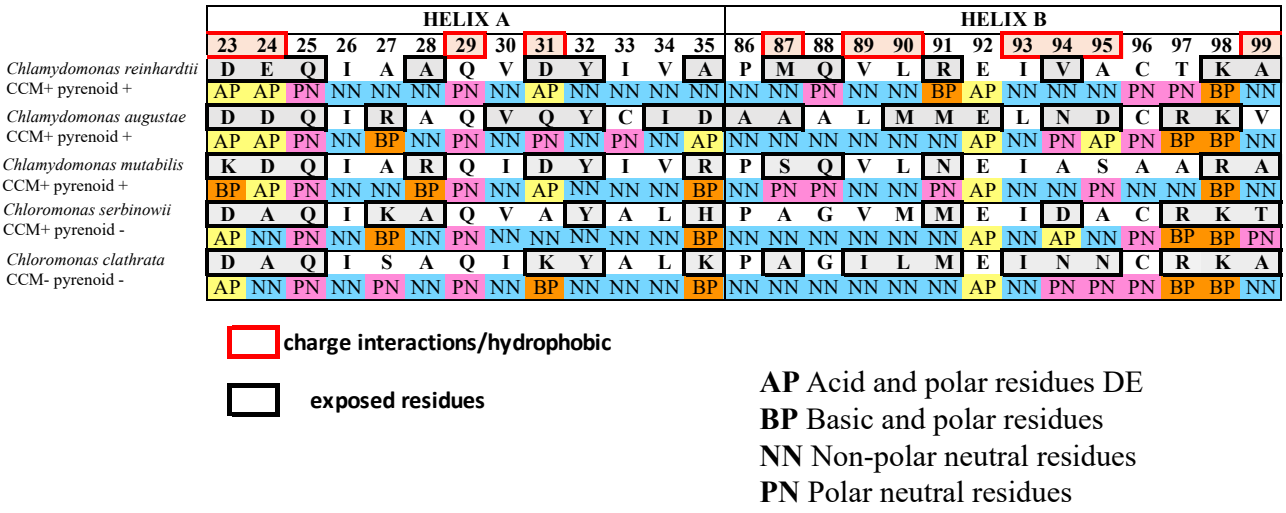


Figure 6.7 Comparison of the amino acids composition of the two Rubisco SSU α-helices for the 4 *Chlamydomonas* and *Chloromonas* strains compared to the study model *Chlamydomonas reinhardtii.* Residues which have potentially identified as binding sites for interactions with EPYC1) are framed in red in *Chlamydomonas reinhardtii*. Exposed residues are framed in black. Acid and polar residues are in yellow, basic and polar residues are in orange, non-polar neutral residues are in blue and polar neutral residues are in pink.

6.2.6 The 88 essential genes for pyrenoid formation are not found in the same proportions across the different *Chlamydomonas* and *Chloromonas* strains

Following the identification of 88 genes essential for pyrenoid formation characterised by the analysis of transcriptome data in synchronised cells, gene regulatory network analyses and protein-protein interactions (See Table 2.5; Chapter 2 and paragraph 6.1), BLAST analysis was used to identify what proportion of the 88 genes essential for pyrenoid formation in *Chl. reinhardtii* (See Chapter 2; Table 2.5; Appendix 25) were present in the five newly sequenced genomes (Figure 6.8). Based on the BLAST analysis, the five strains showed a great disparity in the results (Figure 6.8; Appendix 25). *Chl. augustae* (CCM+ pyr+) and *Chr. serbinowii* (CCM+ pyr-) were the two strains with the fewest matches (28 and 31 genes respectively did not match) whereas in *Chl. mutabilis* (CCM+ pyr+) more than 70% (64/88 genes) of the essential genes were present. Surprisingly *Chr. rosae* (CCM+ pyr-) and *Chr. clathrata* (CCM-pyr-) obtained more hits than *Chr. serbinowii* (CCM+ pyr-) and *Chl. augustae* (CCM+ pyr+) with 57 and 51 hits respectively, which is more than 57% of the 88 essential genes. However, the absence of a match does not necessarily mean absence of the gene in the assembly, as the gene may not have been sufficiently similar to match the reference genome. If the sequences from *Chl. reinhardtii* appeared to be too different from those in the new assemblies, the software might have not been able to detect them. Therefore, negative hits cannot be associated to the absence of these genes.

Of the 88 genes, 21 were found in all five strains (Table 6.3). Most of them were genes with important metabolic functions and it was therefore not surprising to find them. Along with *RbcS* and *rbcL*, they included genes such as kinases, a carbonic anhydrase, and an important protein in the Calvin cycle (Cre03.g185550.t1.2). Some of the genes identified in all five strains were very specific to CCM expression, such as BST1, a bestrophin protein thought to be located on the thylakoid membrane and potentially associated with chloroplast CCM components in *Chl. reinhardtii* (Mukherjee *et al*., 2019). In addition, the new BST4 protein, thought to be the missing thylakoid bicarbonate transporter of the *Chl. reinhardtii* CCM, was also found in all the five strains even though *Chr. clathrata* does not express CCM activity or a pyrenoid. Therefore, the results at this stage do not show evidence of particular trends in expression of the 88 essential genes between species with/without CCM and/or pyrenoid.

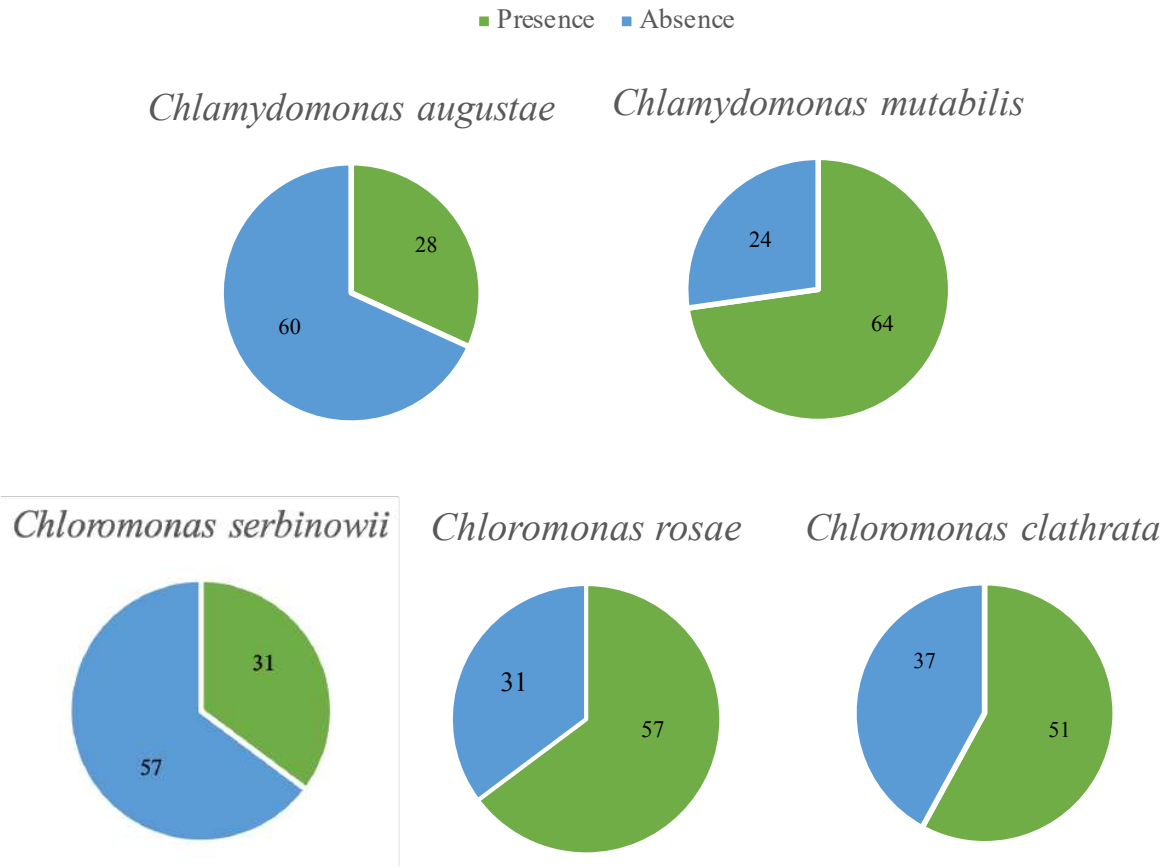


Figure 6.8 Proportion of the 88 genes essential for pyrenoid formation identified by transcriptome analysis in synchronised cells (Mitchell *et al.,* 2014; Zones *et al.,* 2015) but also by gene regulatory network analysis and protein-protein interactions (Gita Yadav, Citu Gulia, University of Cambridge, unpublished data) found with BLAST in the 2 *Chlamydomonas* and 3 *Chloromonas* strains. Genes found are labelled in green and the absence of genes are in blue. The two *Chlamydomonas* strains are CCM + pyrenoid +, *Chloromonas serbinowii* and *rosae* are CCM + pyrenoid -, finally *Chloromonas clathrata* is CCM- pyrenoid -.

Table 6.3 List of the genes common to the five new sequenced strains

Gene ID	Function
Cre01.g014350.t1.2	Peroxiredoxin type II
Cre01.g030900.t1.1	CoA ligase/ OSB CoA synthetase
Cre01.g045902.t1.1	Protein high chlorophyll fluorescence 101
Cre02.g111550.t1.1	Serine-threonine protein kinase
Cre02.g120100-150	RbcS
Cre03.g185550.t1.2	Sedoheptulose-1,7-bisphosphatase calvin cycle
Cre05.g248450.t1.	Mitochondrial carbonic anhydrase
Cre06.g259900.t1.2	ATP synthase gamma chain chloroplastic
Cre06.g261750	RBMP1; Bestrophin, RFP-TM, chloride channel (Bestrophin)
Cre06.g295450.t1.2	Hydroxypyruvate reductase
Cre06.g309000.t1.2	Anion transporter
Cre09.g396950.t1.1	Candidate Na+/HCO3- transporter from screens
Cre10.g444700.t1.1	SBE3 - Starch branching enzyme
Cre12.g484200.t1.2	GGPS1
Cre12.g494850.t1.2	ADK3 - Adenylate kinase 3
Cre13.g581850.t1.2	Kinase
Cre14.g626700.t1.	Fd/FDX1 - Ferredoxin
Cre16.g659050.t1.1	SEPHCHC synthase
Cre16.g662600.t1.2	BST1 bestrophin expressed low CO2
Cre16.g663450.t1.2	LCI11
Cre17.g721500.t1.2	STA2 - Starch synthase, chloroplastic/amyloplastic
	BST4
	rbcL

6.2.7 Multiple candidates for EPYC1-like protein found in the 5 new genomes.

The absence of any BLAST hits for EPYC1 in the five new genomes led to use of Mackinder's method (2016) to detect EPYC1-like candidates (See Chapter 2: Materials & Methods, paragraph *2.5.3.3*), and X-stream was performed successfully. The *Chl. augustae* genome contained the fewest repeats of the 40-80 amino acid lengths, with 26 repeats found (Table 6.4a). In contrast, *Chr. rosae* and *Chl. mutabilis* had 406 and 203 repeats respectively (Table 6.4a-b). Around 50% of the repeats exhibited pI superior to 8 (11 for *Chl. augustae,* 162 for *Chl. mutabilis*, 40 for *Chr. serbinowii,* 148 for *Chr. rosae* and 97 *Chr. clathrata*). However, only a few of these repeats with a pI>8 showed the typical oscillating disorder profile. In total, two repeats with physiochemical properties similar to *Chl. reinhardtii* EPYC1 were identified in *Chl. augustae*, *Chr. serbinowii* and *Chr. rosae*, 6 repeats in *Chl. mutabilis*, and 1 repeat was detected in *Chr. clathrata.*

Furthermore, any of the consensus repeats generated with Xstream looked similar to the general consensus repeat found in *Chl. reinhardtii,* which suggests that if any of these sequences were EPYC1, they would probably have different biochemical properties and consequently a different binding mechanism to Rubisco.

Table 6.4a Analysis of the five new strains for proteins with EPYC1-like physiochemical properties following Mackinder *et al*. (2016) method.

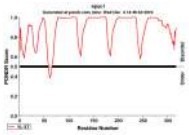
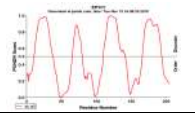
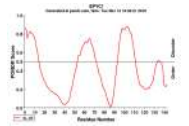
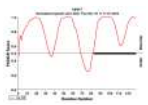
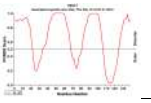
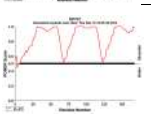
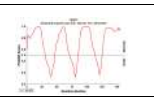
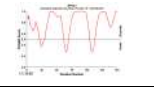
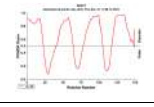
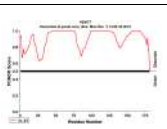
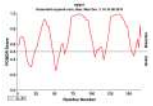
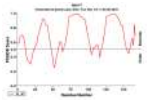
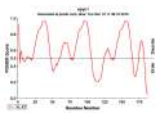
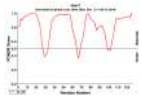
Species	CCM/ Pyrenoid	Number of proteins with …				repeat ID Xstream	Protein characteristics				Consensus repeat sequence from Xstream	Disorder profile
		...>=3 repeats with a 40-80aa repeat length	... And a pI>8	… and an oscillating disorder profile	… and no transmembrane domains		Length	Repeat length	Repeat copy #	pI		
Chlamydomonas reinhardtii	CCM+ pyrenoid+	18	8	1	1	Cre10.g436550	318	61	3.84	11.8	VTPSRSALP SNWKQELE SLRSSSPAP ASSAPAPAR SSSASWRD AAPASSAPA RSSSASKKA	
Chlamydomonas augustae	CCM+ pyrenoid+	26	11	2	2	repeat 16	205	50	4.12	9.13	TGCRCCSSP CSGPSRSRS TGCRWNWS PCSARSMSK STGSTRCCS SCSDPT	
						repeat 22	142	45	3.04	11.76	LQRGGHGL ERTGPGGE APFVARREI GLRCEQLLL MVQRMAA GLLPFL	
Chlamydomonas mutabilis	CCM+ pyrenoid+	203	162	6	6	repeat 171	128	40	3.23	12.04	RPDTGVPLV PPVPPARPS RPQQPQRH AGAARADA AAAERG	
						repeat 134	138	43	3.23	12.27	PWQWDGAP RQREGGAV AMWGAAA AGGGRRGN GVGCRGSG RGA	
						repeat 53	174	57	3.05	11.74	GCRSRGGC RCGRRHGH EPADLGGG RRYRHQGR RPGRCHGC RWRCRCSR AVCRGRG	
						repeat 160	156	41	3.8	11.68	AAGALRHA PQRAPARP QAAAACPA PRPHAPMPL CPLRCRHS	
						repeat162	149	41	3.63	11.71	ERRQRSGQ SGRGAYGR GVGHAAAA WGGAGARC GACRIAPAA	
						repeat162b	150	41	3.66	11.81	AERRQRRG QSGRGAYS RGAGHVAA ACGSGGVR CGACRSAP A	

Table 6.4b Analysis of the five new strains for proteins with EPYC1-like physiochemical properties following Mackinder *et al*. (2016) method.

Chloromonas serbinowii	CCM+ pyrenoid-	71	40	2	2	repeat 11	192	62	3.02	12.35	ARRCRAAR SRARRPACR ATSTTPRW CSTRTAPMP APCRARAA PRPARATRP APTATAA	
						repeat 15	166	53	3.15	11.53	LSPSFGTGS AATARATL ASTALSIAA ASGSAGTG AGRAAMAS STICLITAGS T	
Chloromonas rosae	CCM+ pyrenoid -	406	148	2	2	repeat 119	53	166	3.15	11.53	LSPSFGTGS AATARATL ASTALSIAA ASGSAGTG AGRAAMAS STICLITAGS T	
						repeat 52	63	199	3.17	9.87	ESNLHERIM TGCTSIRPE LRSMGIELV SFDNQKRK SKAPTERDG VMPNHCIRL YRNRTG	
Chloromonas clathrata	CCM- pyrenoid-	150	97	1	1	repeat 114	135	41	3.32	11.51	RTRKWPQP APSLSSCSA PPRQQCSGT PLAPCRLCH KTS	

6.3 Discussion

The aim of this chapter was to use whole genome sequences to investigate the molecular basis of different CCM activity in five closely related strains. Three biological questions were addressed: *i*) Can the multiple origins of the pyrenoid be inferred from sequencing? *ii*) Do the interactions between LSUs and SSUs reflect the different CCM activities and do the α-helices reflect the same interactions with EPYC1 found in *Chl. reinhardtii* ? *iii*) Can the 88 essential genes for CCM expression be found in the five new strains, which express different types of CCM activity?

6.3.1 Whole genome sequencing: a challenging but a powerful tool to answer biological questions.

Whole genome sequencing is an incredible source of information but generates a lot of data which can be difficult to handle. Specialist companies now provide high level sequencing expertise with relatively low costs. For our five green algal strains more detailed analysis of the whole genome sequences is ongoing, but initial results can be discussed. For the five strains the genome size estimation was based on the *Chl. reinhardtii* genome size (120 Mb) and overall, the final genome sizes obtained were similar to the initial estimation (slightly higher than *Chl. reinhardtii*). The short read BGI sequencing was very successful, with a coverage of x90 based on a genome size of 120 Mb. In contrast, the long read PacBio sequencing was very disappointing with only a third of each genome sequenced. The PacBio technology aims to generate very long reads of around 15 000 bp, however, the sequencing platform never managed to produce such reads on our strains probably due to problems with the polymerase, possibly due to the high GC content of these species? The Norwegian Centre is currently trying to understand why the sequencing failed.

Despite being closely related, the *Chl. reinhardtii* genome appeared to be quite different to the five new strains. Few previous attempts have been made to use *Chl. reinhardtii* as a reference to reconstruct new genomes, but on average, only 4% of the reads were mapped onto the *Chl. reinhardtii* genome using BWA (Li & Durbin, 2009). The same problem happened during the chloroplastic genome reconstruction with Geneious. However, the annotations of the full genomes are still in progress as well as the full chloroplastic reconstructions.

The genome sizes were surprising large in *Chr. rosae* and *Chr. clathrata*. However, the three other strains had genome sizes in the range of the other green algae recently sequenced

(*Mesotaenium endlicherianum*: 163 Mb; Cheng *et al*., 2019 or *Klebsormidium flaccidum*: 117 Mb; Hori *et al.*, 2014). Sadly, attempts to estimate genome size by flow cytometry of nuclei failed multiple times. Breaking the cell walls was generally a challenge during this PhD programme. However, such experiments remain important in the future in order to evaluate whether some parts of the genome were not sequenced or in case contamination by sequences from other organisms needs to be removed.

Overall, these sequencing approaches remain unique, as whole genome sequencing in green algae has mainly been aimed at understanding land colonisation (Liang *et al*., 2019; Leebens-Mack *et al.,* 2019), but these genome comparisons represent an unique opportunity for a deeper investigation of the genetic basis behind the different extent of CCM activity and pyrenoid occurrence.

6.3.2 The pyrenoid is the result of a convergent evolution in green algae

The multiple origins of the pyrenoid was a long standing hypotheses. The early studies on *Chloromonas* (Buchheim *et al.*, 1997; Nozaki *et al*., 2002) showed the gain and losses of the pyrenoid, but never across the whole green algae lineage and using only one or two markers. Later on, Meyer and Griffiths (2013) used a phylogeny based on ribosomal 18S that included more organisms. However, drawing strong and reliable conclusions from a phylogeny built with one marker remains difficult. Nozaki *et al.* (2002) showed that different tree topologies can be obtained when built with different markers (Figure 6.9). The phylogeny of *rbcL* published in Nozaki *et al.* (2002) placed *Chl. augustae* in the most internal branches whereas the chloroplast phylogeny here showed that the two *Chlamydomonas* strains are basal. The phylogeny of green algae in Lelieart *et al.* 2012 was one of the most complete trees, but only showed branches leading to the main lineages and not at the species level. More recently, the one thousand plant transcriptomes and the phylogenomics of green plants (Leebens-Mack *et al*., 2019) was released and this included dozens of green algae. However, the study focused mainly on streptophyte algae and did not provide a detailed phylogeny of the green algae at the species level. Chloroplast genomes are often used in plant phylogenetics because they possess properties common to Viridiplantae: the chloroplast genome is haploid because it is maternally inherited. Consequently, it is not recombined and highly conserved in gene structure, order and arrangement (Downie & Palmer, 1992; Olmstead & Palmer, 1994) making it a good marker for phylogenetic studies. In addition, a single *Chl. reinhardtii* cell contains on average 83 copies of the chloroplast genome (Gallaher *et al.,* 2018) which represents a substantial part

of the reads after sequencing. Reconstructing a phylogeny with species diagnosed with and without a pyrenoid, and based on multiple chloroplastic genes, generated results with good support and permits stronger conclusions to be drawn.

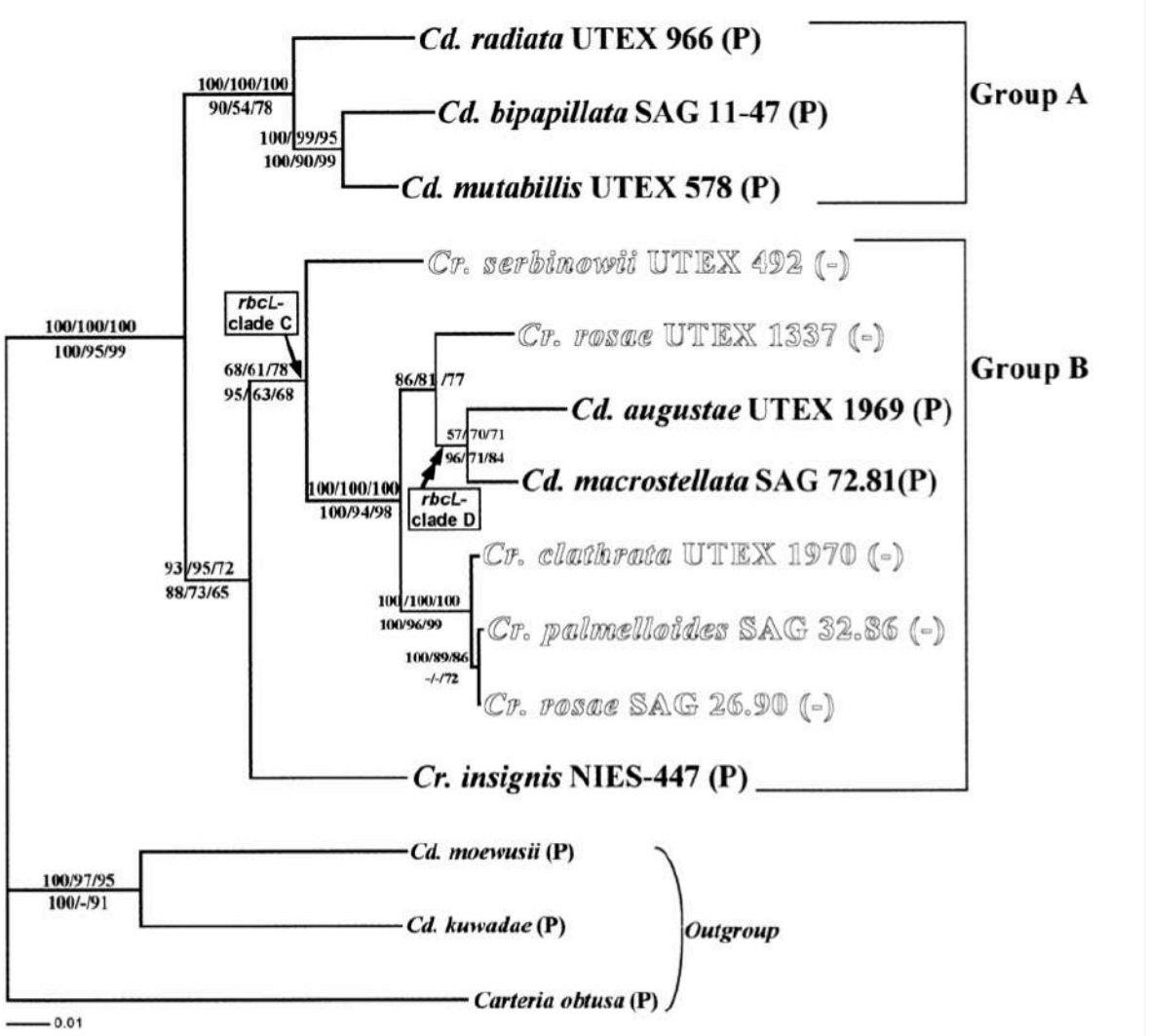


Figure 6.9 Maximum likelihood tree of thee 11 strains obtained in Nozaki *et al*. (2002), based on nucleotide sequences of rbcL.

The list of species in this phylogeny is far from exhaustive but provided evidence in support of multiple origins of the pyrenoid. Except for the newly sequenced *Chloromonas* and *Chlamydomonas*, for which the presence of CCM and/or pyrenoid have been clearly characterised, only the pyrenoid occurrence had previously been reported for the other species.

Therefore, the total absence of CCM activity in this species cannot be inferred without more physiological analyses of all of them. Even if this study has been the first to use multiple

markers to infer the multiple losses of the pyrenoid, the results were far from surprising. This analyses lines up with the other studies on CCM evolution in hornworts and C_4 metabolism (Sage *et al.,* 2011; Villareal & Renner, 2012) and confirmed our initial hypotheses. However, new results mean new questions arise. If the pyrenoid is not inherited from a common ancestor in green algae what was the state of the most common ancestor to the algae with a primary chloroplast? As explained in Raven *et al*. (2005), the single origin of the CCM in algae with a primary chloroplast and without horizontal gene transfer, would imply that the conditions at the time favoured firstly the origin but also the retention of a CCM over a long period of time [at least 1.2 billion years (Butterfield, 2000)]. The multiple origins of the CCM is therefore a stronger hypotheses, but it also does not rule out that the common ancestor did not have a CCM at all, and that it could have been lost when the environmental conditions would not require such a function.

6.3.3 Rubisco interactions and α-helices reflect potential different types of CCM

Interactions between LSUs and SSUs were investigated to see if any correlations could be observed between CCM activity and intensity of the interactions between the different Rubisco subunits. The higher number of interactions between LSUs and SSUs in *Chlamydomonas* strains was consistent with the presence of pyrenoid in this species. However, it remains difficult to understand the reason behind it. In addition, the present Rubisco modelling only took into account the α-helices and not the full sequences.

Generally, all the analyses described above suggested the existence of multiple forms of CCMs, which would be consistent with multiple independent origins and would explain the difficulty identifying commonalities between CCMs, such as the absence of an obvious EPYC1 in species other than *Chl. reinhardtii*. The potential EPYC1 candidates found in *Volvox carteri* (XP_002946604.1) and *Gonium pectoral* (KXZ46518.1) were very similar, but Mackinder *et al.* (2016) did not find any other similar sequences in other algal genomes. In addition, the results of this chapter showed that even in closely related species to *Chl. reinhardtii* (*Chl. augustae* and *mutabilis*) no similar sequences were identified, whilst a long list of potential candidates have been identified in the related genus *Chloromonas*. However, the presence of repeats with similar physiochemical properties to EPYC1 does not necessarily mean that one of these repeats is derived from EPYC1. Only further analyses will determine if any of these candidates are functionally equivalent to EPYC1, especially in the *Chloromonas* strains which do not have a pyrenoid. Interestingly, *Chloromonas*

strains showed the highest number of repeats in their genomes, but the consensus repeats were very different to each other, therefore if any of these repeats were to be EPYC1, the potential binding mechanisms are likely to be different compared to the one found in *Chl. reinhardtii.* The α-helices comparison confirmed the results found in Chapter 3, with not only an absence of similar biochemical properties which could differentiate species with a pyrenoid from species without, but also an absence of any pattern in the exposed residues. Therefore, in assuming that EPYC1 homologues are present in these strains, it remains difficult to see how the interactions between SSU and EPYC1 would remain similar to those in *Chl. reinhardtii*.

The existence of other types of interactions between Rubisco and a protein linker have been shown to exist in the diatom *Phaeodactylum tricornutum* (Oh Zhen Guo, PhD thesis). The Rubisco binding motifs were characterised and it was determined that the interactions take place with the surface exposed C-terminus residues of the Rubisco SSU.

The disorderly nature of EPYC1 makes this protein difficult to find. Intrinsically disordered proteins are known to be regulatory proteins (Launay *et al.,* 2018) that either do not have a unique stable structure or may have a significant propensity to fold into secondary structural elements (Marsh *et al.,* 2006; Sharma *et al*., 2019). In addition, this ability to fold depends on different factors [post-translational modifications, redox-conditions, pH or temperature (Yao *et al*., 2001; Csizmok *et al*., 2007; Cremers *et al*., 2010; Bah & Forman-Kay, 2016; Launay *et al.,* 2018)]. A large number of proteins have been found in *Chl. reinhardtii* (Zhang *et al*., 2018) with disordered protein properties. Among them, the C-terminus extension of Rubisco activase (RCA) is highly disordered, showing variations between chlorophytes and streptophyte algae. Streptophyte algae possess two regulatory cysteine residues, absent in the chlorophytes, which allow RCAs to be activated by light through a redox protein called thioredoxin (Carmo-Silva & Salvucci, 2013; Nagarajan & Gill, 2018). However, variations in this C-terminus extension also exists within the chlorophytes (Comparison between *Tetraselmis* sp and *Ostreococcus tauri*;), suggesting potential different mode of interactions between RCAs and Rubisco (Sena & Uversky, 2016). The absence of a C-terminus extension in chlorophytes suggests that RCAs could also act as an additional chaperone (Portis *et al.,* 2007; Mueller-Cajar *et al.,* 2014) and could potentially use different mechanisms to remodel inhibited Rubisco (Hauser *et al.,* 2015b). However, we can also hypothesize that in streptophytes, α-RCA with the disordered extension could also act as a linker protein. Disordered extremities enable proteins to increase the number of partners that

can directly affect their regulation. The interactions between the α-RCA and the A_2B_2-GAPDH have already shown to participate in photosynthesis regulation (Thieulin-Pardo *et al.,* 2015). This could support potential other interactions with other partners such as Rubisco.

The presence of potential EPYC1 candidates in our five new strains and those found in other algae supports the presence of EPYC1 in species other than just *Chl. reinhardtii*. However, the present study and those found in Chapter 3 suggest that the mechanisms of interactions with Rubisco are not ubiquitous across the green algae. However, if the final aim of fully understanding CCM expression in *Chl. reinhardtii* is to introduce it in C_3 crops, only a deep understanding of one model organism is perhaps necessary.

6.3.4 Genomes comparison

The genome comparison of the 88 genes essential for pyrenoid formation was somewhat interesting and few points need to be discussed here. The *Chlamydomonas* strains showed very different results. *Chl. mutabilis* turned out to be the strain with the most hits compared to *Chl. augustae*. Such results could be an effect of the genome size. *Chl. mutabilis* is 48 Mb larger than *Chl. augustae* but would it be enough to explain these differences? It is important to remember than the absence of hits does not mean absence of these genes and only the full annotations will give us the final presence or absence of certain genes. However, this initial screening gave us a first taste of the future results. The presence of BST1 and BST4 in the five strains suggest that either these genes are expressed at different levels, simply not expressed at all in the strains without a pyrenoid, or potentially expressed only in certain conditions. Based on the BLAST analyses, other similar results are expected to be found.

6.4 Conclusion and future works

This work represents only preliminary analyses of the complete genomes and reflects the potential for these large datasets. It appears clear from the data in this chapter, but also from Chapter 3, that the mechanistic interactions between the different subunits of Rubisco, and also with EPYC1 found in *Chl. reinhardti,* are not ubiquitous to all the green algal species.

Future work will mainly need to include 3 main analyses, firstly, genome size estimations will need to be undertaken using a different protocol. So far, the method used was optimised for *Arabidopsis thaliana* and provided by Sebastian Eves-van den Akker. The protocol

developed by Winck *et al*. (2011) will be tested and includes an optimized protocol from the CellLytic PN kit (Sigma-Aldrich, Steinheim, Germany) and liquid nitrogen. Secondly, structural and functional annotations of the five strains are currently ongoing. So far, only the structural annotation of *Chl. augustae* is complete. Such annotations will help us to definitively make conclusions on the occurrence of the 88 genes essential for pyrenoid formation, and will help us to determine the genetic basis behind the inorganic carbon uptake (Chapter 5.3.3). Finally, full chloroplast reconstruction is currently underway. Because of the lack of similarity between *Chl. reinhardtii* and the five new strains, only manual reconstruction is possible and therefore several methods are being tested.

Chapter 7: General Discussion

The starting point for this research project had been the observations that the small subunit (SSU) of Rubisco was directly involved in Rubisco aggregation mechanism which lead to pyrenoid formation and optimal photosynthetic carbon concentrating mechanism (CCM) expression in *Chl. reinhardtii.* The overall goal of the programme was to investigate the wider relationship between the Rubisco SSU sequence and CCM occurrence across the green algal lineages. Using bioinformatics, biochemistry and physiological approaches, the aim was to investigate whether any specifically conserved SSU residues might be related to the origins and functioning of the CCM, pyrenoid formation and associated Rubisco kinetic properties.

This General Discussion summarises the key findings of the four data Chapters and considers the results within a wider context, whilst also identifying lessons learned from the work and suggesting where future research should be focussed.

7.1 *RbcS*, a gene of interest to understanding pyrenoid formation

Following Meyer *et al*. (2012), the first step of this study was to use a phylogenetic approach to identify the residues that could explain pyrenoid occurrence in green algae based on early access to the incredible data collection arising from the 1kP programme (Leebens-Mack *et al*., 2019). Owing to their variable quality, only a small proportion of the sequences could be used, however the new phylogeny of *RbcS* (Figure 3.1) gave us more insight into the evolution of *RbcS* and Rubisco more generally. The analyses of *RbcS* showed that the structure of streptophyte algal Rubisco SSU is similar to land plant Rubisco SSU and that *RbcS* was not under any specific type of selection.

The main challenge of Chapter 3 came from the nature of *RbcS* itself. As part of a gene family and being nuclear encoded, *RbcS* is more inclined to variations compared to the chloroplast genes. More than two copies of *RbcS* were present in most of the species in the phylogeny. The two copies of *RbcS* in *Chl. reinhardtii* have been shown to have different timing of maximum gene transcript expression, with *RbcS1* highest at night and *RbcS2* that is expressed continuously in diurnal growth (Zones *et al.,* 2015). The manipulation of the α-helices by Meyer *et al.,* (2012) and associated changes in Rubisco aggregation and pyrenoid formation, were undertaken with *RbcS1*, which suggests this copy is the dominant

contributor to Rubisco synthesis following CCM induction. With four copies of *RbcS*, the newly sequenced *Prasinoderma coloniale* (Leebens-Mack *et al.,* 2019) is another recent example that highlights the difficulty in understanding the functional attributes of each copy of the gene for pyrenoid occurrence or growth under contrasting environmental conditions. The generation of an *Arabidopsis* mutant expressing only one of the four *AtRbcS* genes has provided a transformation platform for the complementary introduction of the *Chlamydomonas RbcS1* constructs, or specific alpha helices thereof, into Arabidopsis (Atkinson *et al.,* 2017; Meyer *et al*., 2012; Izumi *et al.,* 2012).

The *Chl. reinhardtii* CCM is, so far, the best described in the literature. No other green algae species has been characterised to the same depth, therefore it remains difficult to estimate both how widespread the *Chl. reinhardtii* CCM is across the green algal phylogeny, and also to what extent *RbcS* plays a similar central role. Overall, across the green algae, generally little is known about pyrenoid formation.

The shorter βA-βB loop length found in streptophytes algae is a major observation arising from work undertaken during the preparation of this thesis. This trait can now be used to differentiate streptophyte algae from chlorophytes, in addition to the other traits previously discussed (see General Introduction). However, loss of amino acids from the βA-βB loop raises again the importance of the central solvent channel and the impact that a change in channel diameter in the streptophytes might have on Rubisco photosynthetic efficiency, or capacity to adapt to changing CO_2 availability during life on land. It has been shown that CO_2 molecules were not only using the channel to access the active sites but could also diffuse through the Rubisco molecule itself, with the SSU potentially acting as a CO_2 reservoir (Van Lun *et al.,* 2014).

The short βA-βB loop arose between 750 and 1000 Mya, but in the absence of relaxed selection on *RbcS,* (Chapter 3) what can explain the loss of five amino acids? Chapter 4 showed that a shorter loop had no impact on the kinetic properties of Rubisco, so was not a selective advantage to improve Rubisco operating efficiency. It is difficult to associate the shorter loop with environmental changes as the split between streptophytes and chlorophytes probably occurred during the Boring billion, a period during which the environmental conditions were very stable. However, a one billion year old multicellular chlorophyte fossil has just been discovered (Tang *et al.,* 2020), pushing back the origin of the green algae. Therefore, we can also hypothesise an older split between chlorophytes and streptophytes during which, environmental conditions were different from the Boring billion. Another

possibility would be the presence of selection that was too weak to be detected by current methods with the small sample size available. However, with the increasing amount of data available in algae, positive selection may become apparent. To conclude, despite the absence of evidence which could explain the loss of amino acids, the shorter βA-βB loop in streptophyte algae is older than initially thought (Spreitzer, 2003).

7.2 Understanding photosynthesis during land colonisation

Following Chapter 3, the aim of Chapter 4 was to observe the extent to which CCM accumulation occurs across the different green algal lineages but also the extent to which Rubisco catalytic properties, and the physiological and molecular characteristics of CCMs are intrinsically linked to each other. Although the physiological data based upon O_2 exchange by cells grown under high and low CO_2 conditions were challenging to obtain with somewhat variable in $K_{0.5}$ values (Chapter 4), a consistent relationship was observed between organic material carbon isotope composition as an indicator of CCM engagement and Rubisco affinity for CO_2 (Km). The relationship between carbon isotope composition and CCM activity was confirmed in the subsequent study comparing *Chloromonas* with *Chlamydomonas* (Chapter 5).

The interest in streptophyte algae continues to increase with recent studies showing the extent that land plants and streptophyte algae share numerous genes (Cheng *et al*., 2019). Newly sequenced streptophyte genomes are regularly being published (Wang *et al.*, 2019). The insights in this thesis for the selective pressures on photosynthetic physiology during the transition between chlorophytes and land plants represent another major observation arising from this thesis, in demonstrating co-evolution between Rubisco catalytic properties and CCM activity. Whilst the morphological, reproductive and biochemical traits may show the evolutionary link within the streptophytes (between Charophytes and Embryophytes), this thesis has demonstrated the capacity for Rubisco to evolve and adapt to local CO_2 availability. Within the aerial environment, the improved diffusive supply of CO_2 has seemingly minimised selection pressure to retain a pyrenoid-based CCM, and explain the shift in Rubisco kinetic properties seen in today's C3 plants (Meyer & Griffiths, 2013), and perhaps the need for hornworts to reinvent a pyrenoid-based, biophysical CCM within the past 100 million years (Villarreal *et al.,* 2014).

However, some physiological observations are still missing. The source of inorganic carbon for the different streptophyte algae used in this study remain unknown, whether as

bicarbonate or free CO_2 (Lucas & Berry, 1985). A number of methods were attempted to analyse the carbon isotope composition of culture media for the streptophyte algae grown under high and low CO_2, with the intention of partitioning carbon sources (Raven *et al.,* 1982). Alternative methods need to be developed in which a detectable amount of inorganic carbon can be concentrated, purified and analysed with an isotope ratio mass spectrometer. Samples have been stored to allow these analyses to be undertaken.

Another question relates to the extent that algae did colonise terrestrial habitats directly. Recently, Cheng *et al.,* (2019) showed that early diverging Zygnematophyceae, the closest sister group to land plants (Embryophytes), share some subaerial/terrestrial habitats with the earliest terrestrial forms of plants, surviving desiccation and other biotic and abiotic stresses through horizontal gene transfer from soil bacteria. Measurement of Rubisco catalytic properties in this group of algae would provide additional understanding of the evolution of photosynthetic carbon acquisition. Unfortunately it was not possible to extract and purify Rubisco from *Chlorokybus atmophyticus* (Zygnematophyceae) during the procedures presented in Chapter 4, as it would have provided more insights for the evolution of Rubisco traits during the transition to land. The presence of algae on land raises the question of the expression of CCM in such conditions: Are the pyrenoids in these subaerial/terrestrial species expressed on land? Do the CCMs have reduced activities? If yes, would this influence the catalytic properties? Unfortunately, the only terrestrial species sampled for Rubisco kinetic properties failed, but other candidates could be measured. Other *Klebsormidium* species (*nitens* or *flaccidum*) are, for example, more adapted to subaerial/terrestrial environments and could now be compared under both environments.

7.3 Towards the use of new algal model organisms

Chapters 3 and 4 did not provide evidence of specific *RbcS* residues which could explain pyrenoid occurrence, but showed that the presence of a pyrenoid was widespread across green algal species. Therefore, Chapter 5 focused on five related species which could help us answer the questions which had arisen from Chapters 3 and 4. The *Chlamydomonas* and *Chloromonas* strains had previously been suggested to be closely related, and a comparison of rbcL sequences had suggested that losses and gains of the pyrenoid could be found within and between these two closely related genera (Nozaki *et al.* 2002). Earlier work by this group had suggested that the presence of a pyrenoid was not always linked to CCM activity in *Chloromonas* (Morita *et al.,* 1998, 199). These two genera confirmed their potential as study organisms not only because they showed physiological traits which could help us to

understand the mechanisms behind the absence of CCM/pyrenoid, but also proved to be good candidates to practice and to improve our knowledge on the determinants of pyrenoid and CCM co-expression, in parallel to the work currently ongoing in terms of land plants (Atkinson *et al.,* 2017). The aim was initially to define the responsiveness of any CCM in two species closely related to *Chl. reinhardtii*, as well as three species of *Chloromonas* which varied in their expression of a CCM but lacked a pyrenoid. This study presented initial results with a basic description of inorganic carbon uptake, although more experiments would be required to complete a detailed characterisation of Rubisco kinetic properties. As previously explained, the analyses of growth media would also be crucial to determine the source of inorganic carbon, although (as discussed above) the relationship between carbon isotope composition and $K_{0.5}$ provided compelling evidence. The SEM images showed that even in related species, pyrenoid morphologies were extremely different. *Chlamydomonas reinhardtii* exhibited very defined starch sheaths whereas *Chl. augustae* had multiple granules. Such observations raise the question of potential different forms of inorganic carbon uptake, as well as different mechanisms of CCM expression or different protein-protein interactions. Additional work would be required to determine that the activity of any CCM was related to the capacity of the starch sheath to minimise leakiness, or alternatively whether the coupling of inorganic carbon pumps and carbonic anhydrase activities alter the efficiency of the uptake mechanisms. The relative size of cells and the temperature for growth under natural conditions could also indicate that direct uptake of CO_2 (Raven, 1991), or some of CO_2 -bicarbonate interconversion in the periplasmic region (Lucas *et al.,* 1985), could support Rubisco operating in a low temperature, low photorespiratory environment. In addition, the multiple and independent origins of other metabolic pathways, such as C_4 and CAM, support the above hypothesis and consequently multiple type of CCM expression or other mechanisms which could be specific and adapted to local environments.

7.4 Using new technologies to understand CCM expression

The aim of Chapter 6 was to extend the physiological experiments on the *Chlamydomonas* and *Chloromonas* strains in order to start investigating the genetic basis behind the differential occurrence of CCM and pyrenoid, and potential different inorganic carbon uptake mechanisms in these species.

The absence of full genome annotations only allowed preliminary analyses to be completed within the timescale of this study. Indeed, the sequencing analysis presented represented a significant quantity of data for analysis. The genetic distance between *Chl. reinhardtii* and

the five newly-sequenced species slowed down data processing, as everything had to be processed manually and individually adapted to the different species. However, the multiple gain or loss events leading to pyrenoid occurrence were confirmed in such closely related species, and was confirmed by a wider analysis of chloroplast genome phylogeny for the green algae with and without a pyrenoid (Nozaki *et al*., 2002; Morita *et al*., 1998, 1999). In addition, the interactions between small and large subunits were interpreted in terms of how densely packaged the Rubisco molecules are in the chloroplasts. However, the analyses of the α-helices confirmed the results found in Chapter 3, with the absence of any residues (or common chemical properties) shared by all the green algae which could help to differentiate species with or without a pyrenoid. No direct EPYC1 homologues were found in the different genome assemblies, which could either suggest either that EPYC1 is specific to *Chlamydomonas reinhardtii* and the limited number of species in which EPYC1 has been identified (Mackinder *et al*., 2016; 2017), or that EPYC1/similar linker proteins do exist in other forms but have not yet been identified. Finally, the first screening of the 88 genes essential for pyrenoid formation showed various results, and some promising new CCM-associated candidates, but no consistent trends were identified.

Overall, this chapter did not identify common rules to explain pyrenoid occurrence, suggesting a broad CCM diversity. The SEM images of the species used in this research study, not only in *Chlamydomonas* but also in the streptophyte algae showed that pyrenoid morphologies could vary between species, particularly in terms of the starch sheath structures. The absence of starch sheaths in some species (naked pyrenoid; *Klebsormidium subtile*) highlights the difficulty of understanding how the proteins could interact the same way as they do in *Chl. reinhardtii,* suggesting the presence of other mechanisms or interactions are required to bring about Rubisco packaging and pyrenoid tubule formation in other systems (Meyer *et al.*, 2017.)

In the light of the results found in Chapter 6, major experiments/analysis will need to be undertaken. First of all, genome sizes need to be confirmed by cytometry, but a different methods will need to be developed for lysing the cells, possibly following the protocol in Winck *et al.* (2011). In addition, the present phylogeny of green algae could be dated in order to see when the different species lost and/or regained their pyrenoid. Dated phylogenies of green algae already exist (Laurin-Lemay *et al*., 2012; Becker, 2013; Del Cortona *et al.,* 2020) and therefore could be used to calibrate the current tree. Due to the importance of *RbcS* in this dissertation, *RbcS* will need to be re-amplified and specifically

re-sequenced for the 5 species of interest. Primers have been designed, based on the current sequence analyses, and attempts to amplify each specific RbcS should reveal the detailed structure. This will help us to identify variations between the multiple copies, and, more importantly, will hopefully confirm that the short βA-βB loop observed in some of these strains are just sequencing issues. The long βA-βB loop in chlorophytes is supported by the general phylogeny of *RbcS* built in Chapter 3, which was based on thousand sequences of *RbcS* across green algae but also supported by the only other *RbcS* sequence deposited on GenBank (*Chloromonas sp*; AAD00448.1). The presence of EPYC1 could be experimentally tested for in the *Chlamydomonas* strains using Zhan *et al.* (2018) protocol developed on *Chl. reinhardtii* or by pull-down assays. Pull-down assays could also be tested in the *Chloromonas* strains to identify potential interactions between Rubisco and EPYC1 or other linker proteins (Mackinder *et al.*, 2017). Based on the current observations two hypotheses are possible: either EPYC1/another linker protein do not exist in species without pyrenoid, or it exists but does not have the same level of expression.

Finally, structural and functional annotations of the new sequences will need to be completed, including genome polishing and genome comparisons.

7.5 Overall conclusion

The algal pyrenoid has been the subject of many studies over the years. This research project has built on recent studies and tested whether the local mechanisms observed in particular species could be extended across the green algae more generally. The dissertation showed that the algal CCM is no exception in exhibiting the diversity that Nature has generated over millions of years at all different levels of life. Structures, functions and mechanisms are convergently coupled to the selection pressure of limited CO_2 in the aquatic environment, and the algal pyrenoid reflects the fragile interactions between the contrasting components of an algal CCM, and more complex selective pressures dependent on external environmental conditions. In addition, analyses focussed around a single gene (*RbcS*) have proved that more can continue to be learned from the algal CCM and aquatic photosynthesis more generally from comparisons across the green algal phylogeny. Also, the study has revealed that the occurrence of a CCM can lead to relatively rapid changes in Rubisco kinetic properties and operating efficiency, which can also be inferred from the more recent development of C_4 and CAM from terrestrial C_3 photosynthesis. In addition, the whole genome sequencing used in this research study has only just started to reveal its potential.

Future work should focus on the remaining Rubisco catalytic properties in species found in subaerial/terrestrial habitats, in conjunction with the more detailed assembly of the 5 new genomes of *Chlamydomonas* and *Chloromonas.*

References

Abascal F, Zardoya R. & Posada D. (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*, 21(9), 2104-2105.

Acton E. (1909). *Coccomyxa subellipsoidea*, a new member of the Palmellaceae. *Annals of Botany*, 4, 573-578.

Allen JF. & Raven JA. (1996). Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. *Journal of Molecular Evolution*, 42(5), 482-492.

Allen JF. (2017). The CoRR hypothesis for genes in organelles. *Journal of theoretical biology*, 434, 50-57.

Amoroso G, Sültemeyer D, Thyssen C. & Fock HP. (1998). Uptake of HCO_3^- and CO_2 in cells and chloroplasts from the *microalgae Chlamydomonas reinhardtii* and *Dunaliella tertiolecta. Plant Physiology*, 116(1), 193-201.

Andersson I. (2008). Catalysis and regulation in Rubisco. *Journal of experimental botany*, 59(7), 1555-1568.

Andersson I. & Backlund A. (2008). Structure and function of Rubisco. *Plant Physiology and Biochemistry*, 46(3), 275-291.

Andrews TJ. (1988). Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large subunits in the complete absence of small subunits. *Journal of Biological Chemistry*, 263(25), 12213-12219.

Aquino CAN, Bueno NC, Servat LC. & Bortolini JC. (2017). The genus *Euastrum* Ehrenberg ex Ralfs (Desmidiaceae) in a subtropical stream adjacent to the Parque Nacional do Iguaçu, Paraná State, Brazil. *Hoehnea*, 44, 1-9.

Archibald JM. (2009). The puzzle of plastid evolution. *Current Biology*, 19(2), R81-R88.

Asamizu E, Miura K, Kucho K, Inoue Y, Fukuzawa H, Ohyama K, Nakamura Y. & Tabata S. (2000). Generation of expressed sequence tags from low-CO_2 and high-CO_2 adapted cells of *Chlamydomonas reinhardtii*. *DNA research*, 7(5), 305-307.

Atkinson N, Leitão N, Orr DJ, Meyer MT, Carmo-Silva E, Griffiths H, Smith AM. & McCormick AJ. (2017). Rubisco small subunits from the unicellular green alga *Chlamydomonas* complement Rubisco-deficient mutants of *Arabidopsis. New Phytologist*, 214(2), 655-667.

Atkinson N, Velanis CN, Wunder T, Clarke DJ, Mueller-Cajar O. & McCormick AJ. (2019). The pyrenoidal linker protein EPYC1 phase separates with hybrid Arabidopsis–*Chlamydomonas* Rubisco through interactions with the algal Rubisco small subunit. *Journal of experimental botany*, 70(19), 5271-5285.

Axelsson L, Ryberg H. & Beer S. (1995). Two modes of bicarbonate utilization in the marine green macroalga Ulva lactuca. *Plant, Cell & Environment*, 18(4), 439-445.

Badger MR, Kaplan A. & Berry JA. (1980). Internal inorganic carbon pool of *Chlamydomonas reinhardtii*: evidence for a carbon dioxide-concentrating mechanism. *Plant Physiology*, 66(3), 407-413.

Badger MR. (1987). The CO_2-concentrating mechanism in aquatic phototrophs. In *Photosynthesis* (pp. 219-274). Academic press.

Badger MR. & Gallagher A. (1987). Adaptation of Photosynthetic CO_2 and HCO_3^- Accumulation by the Cyanobacterium *Synechococcus* PCC6301 to Growth at Different Inorganic Carbon Concentrations. *Functional Plant Biology*, 14(2), 189-201.

Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W. & Price GD. (1998). The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. *Canadian Journal of Botany*, 76(6), 1052-1071.

Badger MR, Hanson D. & Price GD. (2002). Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Functional Plant Biology*, 29(3), 161-173.

Badger MR. & Price GD. (2003). CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *Journal of experimental botany*, 54(383), 609-622.

Bah A. & Forman-Kay JD. (2016). Modulation of intrinsically disordered protein function by post-translational modifications. *Journal of Biological Chemistry*, 291(13), 6696-6705.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA. & Pevzner P. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*, 19(5), 455-477.

Bar-On YM. & Milo R. (2019). The global mass and average rate of rubisco. *Proceedings of the National Academy of Sciences*, 116(10), 4738-4743.

Bathellier C, Tcherkez G, Lorimer GH. & Farquhar GD. (2018). Rubisco is not really so bad. *Plant, cell & environment*, 41(4), 705-716.

Bauwe H, Hagemann M. & Fernie AR. (2010). Photorespiration: players, partners and origin. *Trends in plant science*, 15(6), 330-336.

Beardall J, Griffiths H. & Raven JA. (1982). Carbon isotope discrimination and the CO_2 accumulating mechanism in *Chlorella emersonii*. *Journal of Experimental Botany*, 33(4), 729-737.

Beardall J. (1991). Effects of photon flux density on the CO_2-concentrating mechanism of the cyanobacterium *Anabaena variabilis*. *Journal of Plankton Research*, 13(1), 133-141.

Beardall J. & Giordano M. (2002). Ecological implications of microalgal and cyanobacterial CO_2 concentrating mechanisms, and their regulation. *Functional Plant Biology*, 29(3), 335-347.

Becker B. (2013). Snow ball earth and the split of Streptophyta and Chlorophyta. *Trends in Plant science*, 18(4), 180-183.

Bedoshvili YD, Popkova TP. & Likhoshway YV. (2009). Chloroplast structure of diatoms of different classes. *Cell and Tissue Biology*, 3(3), 297-310.

Beerling DJ, Osborne CP. & Chaloner WG. (2001). Evolution of leaf-form in land plants linked to atmospheric CO_2 decline in the Late Palaeozoic era. *Nature*, 410(6826), 352.

Behrenfeld MJ, Randerson JT, McClain CR, Feldman GC, Los SO, Tucker CJ, Falkowski PG, Field CB, Frouin R, Esaias WE, Kolber DD. & Pollack NH. (2001). Biospheric primary production during an ENSO transition. *Science*, 291, 2594-2597.

Belcher JH. & Swale EMF. (1961). Some new and uncommon British Volvocales. *British Phycological Bulletin*, 2(2), 56-62.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J. & Sayers EW. (2012). GenBank. *Nucleic acids research*, 41(D1), D36-D42.

Bernacchi CJ, Singsaas EL, Pimentel C, Portis Jr AR. & Long SP. (2001). Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell & Environment*, 24(2), 253-259.

Berner RA. (2001). Modelling atmospheric O_2 over Phanerozoic time. *Geochimica et Cosmochimica Acta*, 65(5), 685-694.

Berner RA. (2003). The long-term carbon cycle, fossil fuels and atmospheric composition. *Nature*, 426(6964), 323.

Berry J, Boynton J, Kaplan A. & Badger M. (1976). Growth and photosynthesis of *Chlamydomonas reinhardtii* as a function of CO_2 concentration. *Annual report of Food and Agriculture Organization of the United Nations*. 7: 423-432.

Betti M, Bauwe H, Busch FA, Fernie AR, Keech O, Levey M, Ort DR, Parry MA, Sage R, Timm S, Walker B. & Walker B. (2016). Manipulating photorespiration to increase plant productivity: recent advances and perspectives for crop improvement. *Journal of Experimental Botany*, 67(10), 2977-2988.

Bhattacharya D. & Ehlting J. (1995). Actin coding regions: gene family evolution and use as a phylogenetic marker. *Archiv für Protistenkunde*, 145(3-4), 155-164.

Bhatti S. & Colman B. (2011). Evidence for the occurrence of photorespiration in synurophyte algae. *Photosynthesis research*, 109(1-3), 251-256.

Blankenship RE. & Hartman H. (1998). The origin and evolution of oxygenic photosynthesis. *Trends in biochemical sciences*, 23(3), 94-97.

Booton GC, Floyd GL. & Fuerst PA. (1998). Polyphyly of tetrasporalean green algae inferred from nuclear small-subunit ribosomal DNA. *Journal of Phycology*, 34, 306-311.

Borges AV. & Frankignoulle M. (2002). Distribution and air-water exchange of carbon dioxide in the Scheldt plume off the Belgian coast. *Biogeochemistry*, 59(1-2), 41-67.

Boudreau BP. (1997). *Diagenetic models and their implementation* (Vol. 505). Berlin: Springer.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A. & Drummond AJ. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology*, 10(4), e1003537.

Bozdogan H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3), 345-370.

Brook AJ. (1992). Spirotaenia alpina (Zygnemaphyceae: Mesotaeniaceae), a saccoderm desmid new to the British Isles: Revised description and observations on cell division. *British Phycological Journal*, 27(1), 29-38.

Buchheim MA, Buchheim JA. & Chapman RL. (1997). Phylogeny of *Chloromonas* (Chlorophyceae): a study of 185 Ribosomal RNA gene sequences. *Journal of Phycology*, 33(2), 286-293.

Buchmann K. & Becker B. (2009). The system of contractile vacuoles in the green alga *Mesostigma viride* (Streptophyta). *Protist*, 160(3), 427-443.

Burkhardt S, Amoroso G, Riebesell U. & Sültemeyer D. (2001). CO_2 and HCO_3^- uptake in marine diatoms acclimated to different CO_2 concentrations. *Limnology and Oceanography*, 46(6), 1378-1391.

Burrows EM. (1991). *Seaweeds of the British Isles: chlorophyta* (Vol. 2). Natural History Museum Publications.

Butterfield NJ. (2000). *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology*, 26(3), 386-404.

Busch FA, Sage RF. & Farquhar GD. (2018). Plants increase CO_2 uptake by assimilating nitrogen via the photorespiratory pathway. *Nature plants*, 4(1), 46.

Butcher RW. (1952). Contributions to our knowledge of the smaller marine algae. *Journal of the Marine Biological Association of the United Kingdom*, 31, 175-191.

Carmo-Silva AE. & Salvucci ME. (2013). The regulatory properties of Rubisco activase differ among species and affect photosynthetic induction during light transitions. *Plant physiology*, 161(4), 1645-1655.

Carmo-Silva E, Scales JC, Madgwick PJ. & Parry MA. (2015). Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant, Cell & Environment*, 38(9), 1817-1832.

Carpenter EJ, Matasci N, Ayyampalayam S, Wu S, Sun J, Yu J, Vieira FRJ, Bowler C, Dorrell RG, Gitzendanner MA, Li L, Du W, Ullrich KK, Wickett NJ, Barkmann TJ, Barker MS, Leebens-Mack JH. & Wong GKS. (2019). Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). *GigaScience*, 8(10), giz126.

Carvalho JFC, Madgwick PJ, Powers SJ, Keys AJ, Lea PJ. & Parry MA. (2011). An engineered pathway for glyoxylate metabolism in tobacco plants aimed to avoid the release of ammonia in photorespiration. *BMC biotechnology*, 11(1), 111.

Chan KX. (2018). Morphological and physiological studies of the carbon concentrating mechanism in *Chlamydomonas reinhardtii*. PhD thesis, University of Cambridge, UK.

Chapman RL. (1981). Ultrastructure of *Cephaleuros virescens* (Chroolepidaceae; Chlorophyta). III. Zoospores. *American Journal of Botany*, 68, 544-556.

Chapman RL. (2013). Algae: the world's most important "plants"—an introduction. *Mitigation and Adaptation Strategies for Global Change*, 18(1), 5-12.

Chappell DF, O'Kelly CJ. & Floyd GL. (1991). Flagellar apparatus of the biflagellate zoospores of the enigmatic marine green alga *Blastophysa Rhizopus*. *Journal of phycology*, 27, 423-428.

Cheng S, Xian W, Fu Y, Marin B, Keller J, Wu T, Sun W, Li X, Xu Y, Zhang Y, Wittek S, Reder T, Günther G, Gontcharov A, Wang S, Li L, Liu X, Wang J, Yang H, Xu X, Delaux PM, Melkonian B, Wong GK. & Melkonian M. (2019). Genomes of Subaerial Zygnematophyceae Provide Insights into Land Plant Evolution. *Cell*, 179(5), 1057-1067.

Chengwu Z. & Hongjun H. (2018). Taxonomy and ultrastructure of five species of *Tetraselmis (*Prasinophyceae) isolated from China seas. *海洋学报 (中文版)*, (4), 557-579.

Chihara M, Inouye I. & Takahata N. (1986). *Oltmannsiellopsis*, a new genus of marine flagellate (Dunaliellaceae, Chlorophyceae). *Archiv für Protistenkunde*, 132(4), 313-324.

Clement R, Jensen E, Prioretti L, Maberly SC. & Gontero B. (2017). Diversity of CO_2-concentrating mechanisms and responses to CO_2 concentration in marine and freshwater diatoms. *Journal of experimental botany*, 68(14), 3925-3935.

Coat G, Dion P, Noailles MC, De Reviers B, Fontaine JM, Berger-Perrot Y. & Loiseaux-De Goër S. (1998). *Ulva armoricana* (Ulvales, Chlorophyta) from the coasts of Brittany (France). II. Nuclear rDNA ITS sequence analysis. *European journal of phycology*, 33(1), 81-86.

Coesel PF. & Van Geest A. (2016). New or otherwise interesting desmid taxa from the Bangweulu region (Zambia). 2. Genera *Staurodesmus, Staurastrum* and *Xanthidium* (Desmidiales). *Plant Ecology and Evolution*, 149, 101-111.

Colman B, Miller AG. & Grodzinski B. (1974). A study of the control of glycolate excretion in *Chlorella*. *Plant physiology*, 53(3), 395-397.

Colman B, Huertas IE, Bhatti S. & Dason JS. (2002). The diversity of inorganic carbon acquisition mechanisms in eukaryotic microalgae. *Functional Plant Biology*, 29(3), 261-270.

Colman B. & Balkos KD. (2005). Mechanisms of inorganic carbon acquisition in two *Euglena* species. *Canadian journal of botany*, 83(7), 865-871.

Conesa MÀ, Muir CD, Molins A. & Galmés J. (2019). Stomatal anatomy coordinates leaf size with Rubisco kinetics in the Balearic Limonium. *AoB PLANTS*.

Cook ME. (2004). Structure and asexual reproduction of the enigmatic Charophycean green alga *Entransia fimbriata* (Klebsormidiales, Charophyceae). *Journal of Phycology*, 40(2), 424-431.

Covshoff S. & Hibberd JM. (2012). Integrating C_4 photosynthesis into C_3 crops to increase yield potential. *Current Opinion in Biotechnology*, 23(2), 209-214.

Cremers CM, Reichmann D, Hausmann J, Ilbert M. & Jakob U. (2010). Unfolding of metastable linker region is at the core of Hsp33 activation as a redox-regulated chaperone. *Journal of Biological Chemistry*, 285(15), 11243-11251.

Crespo C, Rodríguez H, Segade P, Iglesias R. & García-Estévez JM. (2009). *Coccomyxa sp.* (Chlorophyta: Chlorococcales), a new pathogen in mussels (*Mytilus galloprovincialis*) of Vigo estuary (Galicia, NW Spain). *Journal of invertebrate pathology*, 102(3), 214-219.

Croasdale H. & Grönblad R. (1964). Desmids of Labrador 1. Desmids of the southeastern coastal area. *Transactions of the American Microscopical Society*, 83(2), 142-212.

Csizmók V, Dosztányi Z, Simon I. & Tompa P. (2007). Towards proteomic approaches for the identification of structural disorder. *Current Protein and Peptide Science*, 8(2), 173-179.

Cummins PL, Kannappan B. & Gready JE. (2018). Directions for optimization of photosynthetic carbon fixation: Rubisco's efficiency may not be so constrained after all. *Frontiers in plant science*, 9, 183.

Darden JR. & William H. (1966). Sexual differentiation in Volvox aureus. *The Journal of protozoology* 13: 239-255.

Dean C, Pichersky E. & Dunsmuir P. (1989). Structure, evolution, and regulation of *RbcS* genes in higher plants. *Annual review of plant biology*, 40(1), 415-439.

DeLano WL. (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography*, 40(1), 82-92.

Del Cortona A, Jackson CJ, Bucchini F, Van Bel M, D'hondt S, Škaloud P, Delwiche CF, Knoll AH, Raven JA, Verbruggen H, Vandepoele K., De Clerck O. & Leliaert F. (2020). Neoproterozoic origin and multiple transitions to macroscopic growth in green seaweeds. *Proceedings of the National Academy of Sciences*, https://doi.org/10.1073/pnas.1910060117

Delwiche CF. & Palmer JD. (1997). The origin of plastids and their spread via secondary symbiosis. In *Origins of algae and their plastids* (pp. 53-86). Springer, Vienna.

Dempsey GP, Lawrence D. & Cassie V. (1980). The ultrastructure of *Chlorella minutissima* Fott et Nováková (Chlorophyceae, Chlorococcales). *Phycologia*, 19(1), 13-19.

De Queiroz Mendes MC, González AAC, Moreno MLV, Figueira CP. & de Castro Nunes JM. (2012). Morphological and ultrastructural features of a strain of *Botryococcus terribilis* (Trebouxiophyceae) from Brazil. *Journal of Phycology*, 48, 1099-1106.

Diaz MM. & Maberly SC. (2009). Carbon-concentrating mechanisms in acidophilic algae. *Phycologia*, 48(2), 77-85.

Dion P, De Reviers B. & Coat G. (1998). Ulva armoricana sp. nov.(Ulvales, Chlorophyta) from the coasts of Brittany (France). I. Morphological identification. *European Journal of Phycology*, 33(1), 73-80.

Domozych DS, Elliott L, Kiemle SN. & Gretz MR. (2007). *Pleurotaenium trabecula*, a desmid of wetland biofilms: the extracellular matrix and adhesion mechanisms. *Journal of phycology* 43, 1022-1038.

Doolittle WF. (1998). You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends in Genetics*, 14(8), 307-311.

Downie SR. & Palmer JD. (1992). Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In *Molecular systematics of plants* (pp. 14-35). Springer, Boston, MA.

Drummond AJ. & Rambaut A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*, 7(1), 214.

Du YC, Hong S. & Spreitzer RJ. (2000). *RbcS* suppressor mutations improve the thermal stability and CO_2/O_2 specificity of rbcL-mutant ribulose-1, 5-bisphosphate carboxylase/oxygenase. *Proceedings of the National Academy of Sciences*, 97(26), 14206-14211.

Duanmu D, Miller AR, Horken KM, Weeks DP. & Spalding MH. (2009). Knockdown of limiting-CO_2–induced gene HLA3 decreases HCO_3^- transport and photosynthetic Ci affinity in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*, pnas-0812885106.

Dutkiewicz S, Follows MJ. & Bragg JG. (2009). Modeling the coupling of ocean ecology and biogeochemistry. *Global Biogeochemical Cycles*, 23(4).

Edwards EJ. (2019). Evolutionary trajectories, accessibility and other metaphors: the case of C_4 and CAM photosynthesis. *New Phytologist*, 223(4), 1742-1755.

Eikrem W. & Throndsen J. (1990). The ultrastructure of *Bathycoccus* gen. nov. and *B. prasinos* sp. nov., a non-motile picoplanktonic alga (Chlorophyta, Prasinophyceae) from the Mediterranean and Atlantic. *Phycologia*, 29(3), 344-350.

Ellis RJ. (1979). The most abundant protein in the world. *Trends in biochemical sciences*, 4(11), 241-244.

Ellis RJ. (2010). Biochemistry: Tackling unintelligent design. *Nature*, 463(7278), 164.

Engel BD, Schaffer M, Cuellar LK, Villa E, Plitzko JM. & Baumeister W. (2015). Native architecture of the *Chlamydomonas* chloroplast revealed by in situ cryo-electron tomography. *Elife*, *4*, e04889.

Erb TJ. & Zarzycki J. (2018). A short history of RubisCO: the rise and fall (?) of Nature's predominant CO_2 fixing enzyme. *Current opinion in biotechnology*, 49, 100-107.

Espie GS. & Kimber MS. (2011). Carboxysomes: cyanobacterial Rubisco comes in small packages. *Photosynthesis research*, 109(1-3), 7-20

Esquivel MG, Genkov T, Nogueira AS, Salvucci ME. & Spreitzer RJ. (2013). Substitutions at the opening of the Rubisco central solvent channel affect holoenzyme stability and CO_2/O_2 specificity but not activation by Rubisco activase. *Photosynthesis research*, 118(3), 209-218.

Ettl H. (1967). Die Feinstruktur vonChloromonas rosae Ettl. *Protoplasma*, 64(2), 134-146.

Ettl H. (1970). Die Gattung Chloromonas Gobi emend. Wille.

Ettl H. (1976). Über den Teilungsverlauf des Chloroplasten bei Chlamydomonas. *Protoplasma*, 88, 75-84.

Ettl H. (1983). Süßwasserflora von Mitteleuropa. Band 9. Chlorophyta I: Phytomonadina. *G. Fischer, Stuttgart, New York*.

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O. & Taylor FJR. (2004). The evolution of modern eukaryotic phytoplankton. *science*, 305(5682), 354-360.

Falkowski PG. & Raven JA. (2007). Photosynthesis and primary production in nature. *Aquatic photosynthesis 2nd ed*. Princeton University Press, Princeton.

Falkowski PG. & Raven JA. (2013). *Aquatic photosynthesis*. Princeton University Press.

Fawley MW, Dean ML, Dimmer SK. & Fawley KP. (2006). Evaluating the morphospecies concept in the selenastraceae (chlorophyceae, chlorophyta). *Journal of phycology,* 42, 142-154.

Field CB, Behrenfeld MJ, Randerson JT. & Falkowski P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *science*, 281(5374), 237-240.

Finet C, Timme RE, Delwiche CF. & Marlétaz F. (2010). Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. *Current Biology*, 20(24), 2217-2222.

Flachmann R. & Bohnert HJ. (1992). Replacement of a conserved arginine in the assembly domain of ribulose-1, 5-bisphosphate carboxylase/oxygenase small subunit interferes with holoenzyme formation. *Journal of Biological Chemistry*, 267(15), 10576-10582.

Franks PJ. & Beerling DJ. (2009). Maximum leaf conductance driven by CO_2 effects on stomatal size and density over geologic time. *Proceedings of the National Academy of Sciences*, 106(25), 10343-10347.

Freibauer A, Mathijs E, Brunori G, Damianova Z, Faroult E, Girona IGJ, O'Brien L. & Treyer S. (2011). Sustainable Food Consumption and Production in a Resource-Constrained World; the 3rd SCAR Foresight Exercise. *European Commission: Brussels, Belgium*.

Friedl T. (1989). Comparative ultrastructure of pyrenoids in *Trebouxia* (Microthamniales, Chlorophyta). *Plant Systematics and Evolution*, 164, 145-159.

Friedl T. (1997). The evolution of the green algae. *Plant Systematics and Evolution*, 11, 87-101. *In* D Bhattacharya, ed, Origins of algae and Their Plastids. Springer-Verlag, Wien, Germany.

Fritsch FE. & West GS. (1927). A treatise on the British Freshwater Algae. *By the late GS West. New and revised ed*.

Fulton AB. (1978). Colonial development in *Pandorina morum*: I. Structure and composition of the extracellular matrix. *Developmental biology*, 64, 224-235.

Gallaher SD, Fitz-Gibbon ST, Strenkert D, Purvine SO, Pellegrini M. & Merchant SS. (2018). High-throughput sequencing of the chloroplast and mitochondrion of *Chlamydomonas reinhardtii* to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. *The Plant Journal*, 93(3), 545-565.

Galmes J, Kapralov MV, Andralojc PJ, Conesa MÀ, Keys AJ, Parry MA. & Flexas J. (2014). Expanding knowledge of the R ubisco kinetics variability in plant species: environmental and evolutionary trends. *Plant, Cell & Environment*, 37(9), 1989-2001.

Galmes J, Kapralov MV, Copolovici LO, Hermida-Carrera C. & Niinemets Ü. (2015). Temperature responses of the Rubisco maximum carboxylase activity across domains of

life: phylogenetic signals, trade-offs, and importance for carbon gain. *Photosynthesis research*, 123(2), 183-201.

Galmés J, Hermida-Carrera C, Laanisto L. & Niinemets Ü. (2016). A compendium of temperature responses of Rubisco kinetic traits: variability among and within photosynthetic groups and impacts on photosynthesis modeling. *Journal of Experimental Botany*, 67(17), 5067-5091.

Galmés J, Capó-Bauçà S, Niinemets Ü. & Iñiguez C. (2019). Potential improvement of photosynthetic CO_2 assimilation in crops by exploiting the natural variation in the temperature response of Rubisco catalytic traits. *Current opinion in plant biology*, 49, 60-67.

Garbayo I, Torronteras R, Forján E, Cuaresma M, Casal C, Mogedas B, Ruiz-Domínguez MC, Márquez C. Vaquero I, Fuentes-Cordera JL, Fuentes R, González-del-Valle M. & Vílechez C. (2012). Identification and physiological aspects of a novel carotenoid-enriched, metal-resistant microalga isolated from an acidic river in huelva (Spain). *Journal of phycology*, 48, 607-614.

Gehl KA, Colman B, Sposato LM. (1990). Mechanism of inorganic carbon uptake in *Chlorella saccharophila*: the lack of involvement of carbonic anhydrase. *Journal of experimental botany*, 41(11), 1385-1391.

Genkov T, Du YC. & Spreitzer RJ. (2006). Small-subunit cysteine-65 substitutions can suppress or induce alterations in the large-subunit catalytic efficiency and holoenzyme thermal stability of ribulose-1, 5-bisphosphate carboxylase/oxygenase. *Archives of biochemistry and biophysics*, 451(2), 167-174.

Genkov T. & Spreitzer RJ. (2009). Highly conserved small subunit residues influence rubisco large subunit catalysis. *Journal of Biological Chemistry*, 284(44), 30105-30112.

Genkov T, Meyer M, Griffiths H. & Spreitzer RJ. (2010). Functional hybrid rubisco enzymes with plant small subunits and algal large subunits engineered rbcS cDNA for expression in *Chlamydomonas*. *Journal of Biological Chemistry*, 285(26), 19833-19841.

Gensel PG. (2008). The earliest land plants. *Annual Review of Ecology, Evolution, and Systematics*, 39, 459-477.

Gerloff J. (1940). Beitrage zur kenntnis der Variabilitat und Systematik der Gattung Chlamydomonas. *Arch. Protistenkunde*, 94, 311-502.

Gerrath JF. (1968). *Studies on the ultrastructure of desmids and its relation to their taxonomy.* PhD thesis, University of British Columbia.

Gerrath JF. (2003) Conjugating green algae and desmids. – In: Wehr JD, Sheath RG. (eds), Fresh-water algae of North America: 363–365. – San Diego

Goldberg WM, Makemson JC. & Colley SB. (1984). *Entocladia endozoica* sp. nov., a pathogenic chlorophyte: structure, life history, physiology, and effect on its coral host. *The Biological Bulletin*, 166, 368-383.

Giordano M, Beardall J. & Raven JA. (2005). CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu. Rev. Plant Biol.*, 56, 99-131.

Goldschmidt-Clermont M. & Rahire M. (1986). Sequence, evolution and differential expression of the two genes encoding variant small subunits of ribulose bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *Journal of molecular biology*, 191(3), 421-432.

Gontcharov AA. & Melkonian M. (2010). Molecular phylogeny and revision of the genus *Netrium* (Zygnematophyceae, Streptophyta): *Nucleotaenium* gen. nov. *Journal of Phycology*, 46, 346-362.

Goodenough UW. (1970). Chloroplast division and pyrenoid formation in *Chlamydomonas reinhardtii*. *Journal of Phycology*, 6(1), 1-6.

Goodenough UW. & Levine R. (1970). Chloroplast structure and function in ac-20, a mutant strain of *Chlamydomonas reinhardtii*: III. Chloroplast, Ribosomes and Membranes Organization. *The Journal of cell biology*, 44(3), 547-562.

Gottlieb B. & Goldstein ME. (1977). Colony development in *Eudorina elegans* (Chlorophyta, Volvocales). *Journal of Phycology*, 13, 358-364.

Gradstein FM, Ogg JG, Smith AG, Bleeker W. & Lourens LJ. (2004). A new geologic time scale, with special reference to Precambrian and Neogene. *Episodes*, 27(2), 83-100.

Graham LE, Arancibia-Avila P, Taylor WA, Strother PK. & Cook ME. (2012). Aeroterrestrial Coleochaete (Streptophyta, Coleochaetales) models early plant adaptation to land. *American Journal of Botany*, 99(1), 130-144.

Granick S. (1965). Evolution of heme and chlorophyll. In *Evolving genes and proteins* (pp. 67-88). Academic Press.

Griffiths H, Robe WE, Girnus J. & Maxwell K. (2008). Leaf succulence determines the interplay between carboxylase systems and light use during Crassulacean acid metabolism in Kalanchoë species. *Journal of Experimental Botany*, 59(7), 1851-1861.

Griffiths H, Meyer MT. & Rickaby RE. (2017). Overcoming adversity through diversity: aquatic carbon concentrating mechanisms. *Journal of Experimental Botany*, 68(14), 3689-3695.

Guillard RR, Bold HC. & MacEntee FJ. (1975). Four new unicellular chlorophycean algae from mixohaline habitats. *Phycologia*, 14, 13-24.

Guillard RR, Keller MD, O'Kelly CJ. & Floyd GL. (1991). *Pycnococcus provasolii* gen. et sp. Nov., a coccoid prasinoxanthin-containing phytoplankter from the western north Atlantic and Gulf of Mexico. *Journal of Phycology*, 27(1), 39-47.

Guillou L, Eikrem W, Chrétiennot-Dinet MJ, Le Gall F, Massana R, Romari K, Pedrós-Alió C. & Vaulot D. (2004). Diversity of picoplanktonic prasinophytes assessed by

direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist*, 155(2), 193-214.

Gunn LH, Valegård K. & Andersson I. (2017). A unique structural domain in Methanococcoides burtonii ribulose-1, 5-bisphosphate carboxylase/oxygenase (Rubisco) acts as a small subunit mimic. *Journal of Biological Chemistry*, 292(16), 6838-6850.

Guo OZ. (2019). Towards a biochemical characterization of the diatom pyrenoid. Nanyang Technological University, Singapore

Gurevich A, Saveliev V, Vyahhi N. & Tesler G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075.

Halder N. & Halder N. (2015). Taxonomy and ecology of *Coleochaete irregularis* Pringsheim and *Coleochaete orbicularis* Pringsheim, West Bengal, India. *J. Algal Biomass Utln*,. 6, 47-49.

Hanagata N, Karube I, Chihara M. & Silva PC. (1998). Reconsideration of the taxonomy of ellipsoidal species of Chlorella (Trebouxiophyceae, Chlorophyta), with establishment of *Watanabea* sen. nov. *Phycological Research*, 46(4), 221-229.

Hanson TE. & Tabita FR. (2001). A ribulose-1, 5-bisphosphate carboxylase/oxygenase (RubisCO)-like protein from *Chlorobium tepidum* that is involved with sulfur metabolism and the response to oxidative stress. *Proceedings of the National Academy of Sciences*, 98(8), 4397-4402.

Harholt J, Moestrup Ø. & Ulvskov P. (2016). Why plants were terrestrial from the beginning. *Trends in Plant Science*, 21(2), 96-101.

Hasegawa T, Miyashita H, Kawachi M, Ikemoto H, Kurano N, Miyachi S. & Chihara M. (1996). *Prasinoderma coloniale* gen. et sp. nov., a new pelagic coccoid prasinophyte from the western Pacific ocean. *Phycologia*, 35(2), 170-176.

Hauser T, Bhat JY, Miličić G, Wendler P, Hartl FU, Bracher A. & Hayer-Hartl M. (2015). Structure and mechanism of the Rubisco-assembly chaperone Raf1. *Nature structural & molecular biology*, 22(9), 720-728.

Haworth M, Elliott-Kingston C. & McElwain JC. (2011). Stomatal control as a driver of plant evolution. *Journal of Experimental Botany*, 62(8), 2419-2423.

Hazen TE. (1922). New British and American species of *Lobomonas*: a study in morphogenesis of motile algae. *Bulletin of the Torrey Botanical Club* 123-140.

Hedges SB, Blair JE, Venturi ML. & Shoe JL. (2004). A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC evolutionary biology*, 4(1), 2.

Hepperle D. & Krienitz L. (1996). The extracellular calcification of zoospores of *Phacotus lenticularis (*Chlorophyta, Chlamydomonadales). *European Journal of Phycology*, 31(1), 11-21.

Hermida-Carrera C, Kapralov MV. & Galmés J. (2016). Rubisco catalytic properties and temperature response in crops. *Plant Physiology*, 171(4), 2549-2561.

Heureux AM, Young JN, Whitney SM, Eason-Hubbard MR, Lee RB, Sharwood RE. & Rickaby RE. (2017). The role of Rubisco kinetics and pyrenoid morphology in shaping the CCM of haptophyte microalgae. *Journal of experimental botany*, 68(14), 3959-3969.

Hibberd JM, Sheehy JE. & Langdale JA. (2008). Using C_4 photosynthesis to increase the yield of rice rationale and feasibility. *Current opinion in plant biology*, 11(2), 228-231.

Hoffman LR. (1968). Observations on the fine structure of *Oedogonium* IV. The mature pyrenoid of *Oe. cardiacum*. *Transactions of the American Microscopical Society* 178-185.

Hoffmann L. (1989). Algae of terrestrial habitats. *The botanical review*, 55(2), 77-105.

Hoham RW. (1975). The life history and ecology of the snow alga *Chloromonas pichinchae* (Chlorophyta, Volvocales). *Phycologia*, 14(4), 213-226.

Hoham RW. & Mullet JE. (1977). The life history and ecology of the snow alga *Chloromonas cryophila* sp. nov.(Chlorophyta, Volvocales). *Phycologia*, 16(1), 53-68.

Hoham RW, Roemer SC. & Mullet JE. (1979). The life history and ecology of the snow *alga Chloromonas brevispina* comb. nov.(Chlorophyta, Volvocales). *Phycologia*, 18(1), 55-70.

Holdsworth RH. (1971). The isolation and partial characterization of the pyrenoid protein of *Eremosphaera viridis*. *The Journal of cell biology*, 51, 499-513.

Holzinger A, Karsten U, Lütz C. & Wiencke C. (2006). Ultrastructure and photosynthesis in the supralittoral green macroalga *Prasiola crispa* from Spitsbergen (Norway) under UV exposure. *Phycologia*, 45, 168-177.

Holzinger A, Albert A, Aigner S, Uhl J, Schmitt-Kopplin P, Trumhová K. & Pichrtová M. (2018). Arctic, Antarctic, and temperate green algae *Zygnema spp*. under UV-B stress: vegetative cells perform better than pre-akinetes. *Protoplasma*, 255(4), 1239-1252.

Honegger R. (2018). Fossil lichens from the Lower Devonian and their bacterial and fungal epi-and endobionts. *Biosyst. Ecol. Ser*, 34, 547-563.

Hori T, Norris RE. & Chihara M. (1986). Studies on the Ultrastructure and Taxonomy of the Genus Tetraselmis (Prasinophyceae)-3-Subgenus Parviselmis. *Botanical Magazine, Tokyo*, 99(1053), p123-135.

Hori K, Maruyama F., Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, Moriyama T, Ikeuchi M, Watanabe M, Wada H, Kobayashi K, Saito M, Masuda T, Sasaki-Sekimoto Y, Mashiguchi K, Awai K, Shimojima M, Masuda S, Iwai M, Nobusawa T, Narise T, Kondo S, Saito H, Sato R, Murakawa M, Ihara Y, Oshima-Yamada Y, Ohtaka K, Satoh M, Sonobe K, Ishii M, Ohtani R, Kanamori-Sato M, Honoki R, Miyazaki D, Mochizuki H, Umetsu J, Higashi K, Shibata D, Kamiya Y, Sato N, Nakamura Y, Tabata S, Ida S, Kurokawa K. & Ohta H. (2014).

Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nature communications*, 5, 3978.

Hutner SH, Provasoli L, Schatz A. & Haskins CP. (1950). Some approaches to the study of the role of metals in the metabolism of microorganisms. *Proceedings of the American Philosophical Society*, 94(2), 152-170.

Iyengar MOP. & Desikachary TV. (1981). Volvocales (green algae). Indian Council of Agricultural Research, New Dehli.

Izumi M, Tsunoda H, Suzuki Y, Makino A. & Ishida H. (2012). RBCS1A and RBCS3B, two major members within the Arabidopsis RBCS multigene family, function to yield sufficient Rubisco content for leaf photosynthetic capacity. *Journal of experimental botany*, 63(5), 2159-2170.

Janson G, Zhang C, Prado MG. & Paiardini A. (2017). PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL. *Bioinformatics*, 33(3), 444-446.

Johnston AM, Maberly SC. & Raven JA. (1992). The acquisition of inorganic carbon by four red macroalgae. *Oecologia*, 92(3), 317-326.

Jordan DB. & Ogren WL. (1981). Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase. *Nature*, 291(5815), 513.

Joshi KR, Gavin JB. & Wheeler EE. (1975). The ultrastructure of *Prototheca wickerhamii. Mycopathologia*, 56(1), 9-13.

Juárez ÁB, Vélez CG, Iñiguez AR, Martínez DE, Rodríguez MC, Vigna MS. & del Carmen Ríos de Molina M. (2011). A *Parachlorella kessleri* (Trebouxiophyceae, Chlorophyta) strain from an extremely acidic geothermal pond in Argentina. *Phycologia*, 50(4), 413-421.

Junkins EN, Stamps BW, Corsetti FA, Oremland RS, Spear JR. & Stevenson BS. (2019). Draft Genome Sequence of *Picocystis sp*. Strain ML, Cultivated from Mono Lake, California. *Microbiol Resour Announc*, 8(4), e01353-18.

Kanevski I, Maliga P, Rhoades DF. & Gutteridge S. (1999). Plastome engineering of ribulose-1, 5-bisphosphate carboxylase/oxygenase in tobacco to form a sunflower large subunit and tobacco small subunit hybrid. *Plant Physiology*, 119(1), 133-142.

Kaplan A. & Reinhold L. (1999). CO_2 concentrating mechanisms in photosynthetic microorganisms. *Annual review of plant biology*, 50(1), 539-570.

Kapralov MV. & Filatov DA. (2007). Widespread positive selection in the photosynthetic Rubisco enzyme. *BMC evolutionary biology*, 7(1), 73.

Kapralov MV, Kubien DS, Andersson I. & Filatov DA. (2010). Changes in Rubisco kinetics during the evolution of C_4 photosynthesis in Flaveria (Asteraceae) are associated

with positive selection on genes encoding the enzyme. *Molecular Biology and Evolution*, 28(4), 1491-1503.

Karkehabadi S, Peddi SR, Anwaruzzaman M, Taylor TC, Cederlund A, Genkov T, Andersson I. & Spreitzer RJ. (2005). Chimeric small subunits influence catalysis without causing global conformational changes in the crystal structure of ribulose-1, 5-bisphosphate carboxylase/oxygenase. *Biochemistry*, 44(29), 9851-9861.

Katoh K, Kuma KI, Toh H. & Miyata T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic acids research*, 33(2), 511-518.

Kaye GWC. & Laby TH. (1973) *Tables of Physical and Ciiemical Constants,* 14th edn. Longman, London.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P. & Drummond A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647-1649.

Kebeish R, Niessen M, Thiruveedhi K, Bari R, Hirsch HJ, Rosenkranz R, Stäbler N, Schönfeld B, Kreuzaler F. & Peterhänsel C. (2007). Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. *Nature biotechnology*, 25(5), 593.

Keeling PJ. (2004). Diversity and evolutionary history of plastids and their hosts. *American journal of botany*, 91(10), 1481-1493.

Keeling PJ. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1541), 729-748.

Kenrick P, Wellman CH, Schneider H. & Edgecombe GD. (2012). A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1588), 519-536.

Kent WJ. (2002). BLAT—the BLAST-like alignment tool. *Genome research*, 12(4), 656-664.

Kevekordes K, Holland D, Häubner N, Jenkins S, Koss R, Roberts S, Raven JA, Scrimgeour CM, Shelly K, Stojkovic S. & Beardall J. (2006). Inorganic carbon acquisition by eight species of *Caulerpa* (Caulerpaceae, Chlorophyta). *Phycologia*, 45(4), 442-449.

Kochert G. & Olson LW. (1970). Ultrastructure of *Volvox carteri*. *Archiv für Mikrobiologie*, 74, 19-30.

Komárek J, Fott B. & Huber-Pestalozzi G. (1983). *Das Phytoplankton des Süßwassers*. Systematik und Biologie-Teil 7, 1. Hälfte.

Kostov RV, Small CL. & McFadden BA. (1997). Mutations in a sequence near the N-terminus of the small subunit alter the CO_2/O_2 specificity factor for ribulose bisphosphate carboxylase/oxygenase. *Photosynthesis research*, 54(2), 127-134.

Kranz HD. & Huss VA. (1996). Molecular evolution of pteridophytes and their relationship to seed plants: evidence from complete 18S rRNA gene sequences. *Plant Systematics and Evolution*, 202(1-2), 1-11.

Kranz SA, Young JN, Hopkinson BM, Goldman JA, Tortell PD. & Morel FM. (2015). Low temperature reduces the energetic requirement for the CO_2 concentrating mechanism in diatoms. *New Phytologist*, 205(1), 192-201.

Krogh A, Larsson B, Von Heijne G. & Sonnhammer EL. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology*, 305(3), 567-580.

Kubien DS, Whitney SM, Moore PV. & Jesson LK. (2008). The biochemistry of Rubisco in Flaveria. *Journal of Experimental Botany*, 59(7), 1767-1777.

Lacoste-Royal G. & Gibbs SP. (1987). Immunocytochemical localization of ribulose-1, 5-bisphosphate carboxylase in the pyrenoid and thylakoid region of the chloroplast of *Chlamydomonas reinhardtii*. *Plant Physiology*, 83(3), 602-606.

Launay H, Barré, P, Puppo C, Zhang Y, Maneville S, Gontero B. & Receveur-Bréchot V. (2018). Cryptic disorder out of disorder: Encounter between conditionally disordered CP12 and glyceraldehyde-3-phosphate dehydrogenase. *Journal of molecular biology*, 430(8), 1218-1234.

Laurin-Lemay S, Brinkmann H. & Philippe H. (2012). Origin of land plants revisited in the light of sequence contamination and missing data. *Current Biology*, 22(15), R593-R594.

Le SQ. & Gascuel O. (2008). An improved general amino acid replacement matrix. *Molecular biology and evolution*, 25(7), 1307-1320.

Lee B. & Tabita FR. (1990). Purification of recombinant ribulose-1, 5-bisphosphate carboxylase/oxygenase large subunits suitable for reconstitution and assembly of active L_8S_8 enzyme. *Biochemistry*, 29(40), 9352-9357.

Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, Grosse I, Li Z, Melkonian M, Mirarab S, Porsh M, Quint M, Rensing SA, Soltis PS, Stevenson DW, Ullrich KK, Wickett NJ, DeGironima L, Edger PP, Jordon-Thaden IE, Joya S, Liu T, Melkonian B, Miles NW, Pokorny L, Quigley C, Thomas P, Villareal JC, Augustin MM, Barett MD, Baucom RS, Beerling DJ, Benstein RB, Biffin E, Brockington SF, Burge DO, Burris KP, Burtet-Sarramegna V, Caicedo AL, Cannon SB, Çebi Z, Chang Y, Chater C, Cheeseman JM, Chen T, Clarke ND, Clayton H, Covshoff S, Crandall-Stotler BJ, Cross H, DePamphilis CW, Der JP, Determann R, Dickson RC, Di Stilio VS, Ellis S, Fast E, Feja N, Field KJ, Filatov DA, Finnegan PM, Floyd SK, Fogliani B, Garcia N, Gâteblé G, Godden GT, Goh F, Greiner S, Harkess A, Heaney JM, Helliwell KE, Heyduk K, Hibberd JM, Hodel RGJ, Hollingsworth PM, Johnson MTJ, Jost R, Joyce B, Kapralov MV, Kazamia E, Kellogg

EA, Koch MA, Von Konrat M, Könyves K, Kutchan TM, Lam V, Larsson A, Leitch AR, Lentz R, Li FW, Lowe AJ, Ludwig M, Manos PS, Mavrodiev E, McCormick MK, McKain M, McLellan T, McNeal JR, Miller RE, Nelson MN, Peng Y, Ralph P, Real D, Riggins CW, Ruhsam M, Sage RF, Sakai AK, Scascitella M, Schilling EE, Schlösser EM, Sederoff H, Servick S, Sessa EB, Shaw AJ, Shaw SW, Sigel EM, Skema C, Smith AG, Smithson A, Stewart jr N, Stinchcombe JR, Szövényi P, Tate JA, Helga T, Trapnell D, Villegente M, Wang CN, Weller SG, Wenzel M, Weststrand S, Westwood JH, Whigham DF, Wu S, Wulff AS, Yang Y, Zhu D, Zhuang C, Zuidof J, Chase MW, Pires JC, Rothfels CJ, Yu J, Chen C, Chen L, Cheng S, Li J, Li R, Li X, Lu H, Ou Y, Sun X, Tan X, Tang J, Tian Z, Wang F, Wang J, Wei X, Xu X, Yan Z, Yang F, Zhong X, Zhou F, Zhu Y, Zhang Y, Ayyampalayam S, Barkman TJ, Nguyen NP, Matasci N, Nelson DR, Sayyari E, Wafula EK, Walls RL, Warnow T, An H, Arrigo N, Baniaga AE, Galuska S, Jorgensen SA, Kidder TI, Kong H, Lu-Irving P, Marx HE, Qi X, Reardon CR, Sutherland BL, Tiley GP, Welles SR, Yu R, Zhan S, Gramzow L, Theiβen G. & Wong GKS. (2019).One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, 574, 679-685

Leghari SM. (2001). Fresh Water Algae of Sindh. V. The Desmids from the lakes and ponds of Sindh, Pakistan. *OnLine Journal of Biological Sciences*, 1(6), 456-460.

Leliaert F, Verbruggen H. & Zechman FW. (2011). Into the deep: new discoveries at the base of the green plant phylogeny. *Bioessays*, 33(9), 683-692.

Leliaert F, Smith DR, Moreau H, Herron MD, Verbruggen H, Delwiche CF. & De Clerck O. (2012). Phylogeny and molecular evolution of the green algae. *Critical reviews in plant sciences*, 31(1), 1-46.

Lembi CA. & Lang NJ. (1965). Electron microscopy of *Carteria* and *Chlamydomonas*. *American Journal of Botany*, 52, 464-477.

Lenton TM, Dahl TW, Daines SJ, Mills BJ, Ozaki K, Saltzman MR. & Porada P. (2016). Earliest land plants created modern levels of atmospheric oxygen. *Proceedings of the National Academy of Sciences*, 113(35), 9704-9709.

Lewin RA, Krienitz L, Goericke R, Takeda H. & Hepperle D. (2000). Picocystis salinarum gen. et sp. nov.(Chlorophyta)–a new picoplanktonic green alga. *Phycologia*, 39(6), 560-565.

Leyon H. (1954). The structure of chloroplasts: IV. The development and structure of the Aspidistra chloroplast. *Experimental cell research*, 7(1), 265-273.

Li WH, Wu CI. & Luo CC. (1985). A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Molecular biology and evolution*, 2(2), 150-174.

Li H. & Durbin R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, 25(14), 1754-1760.

Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, Yue R, Robertson JM, Lefebvre PA, Fitz-Gibbon ST, Grossman AR. & Jonikas MC. (2016). An indexed,

mapped mutant library enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. *The Plant Cell*, 28(2), 367-387.

Liang Z, Geng Y, Ji C, Du H, Wong CE, Zhang Q, Zhang Q, Zhang Y, Zhang P, Riaz A, Chachar S, Ding Y, Wen J, Wu Y, Whang M, Zheng H, Wu Y, Demko V, Shen L, Han X, Zhang P, Gu X. & Yu H. (2019). *Mesostigma viride* Genome and Transcriptome Provide Insights into the Origin and Evolution of Streptophyta. *Advanced Science*.

Linke WF. & Seidell A. (1965). Solubilities, Inorganic and Metal–Organic Compounds, 4th. *Washington DC: American Chemical Society*.

Ljunggren B. & Oja T. (1961). *The Uppsala Spectral Classification: Intrinsic Colours and Absolute Magnitudes* (Vol. 18, No. 1). Almqvist & Wiksell.

Lloyd ND, Canvin DT. & Culver DA. (1977). Photosynthesis and photorespiration in algae. *Plant physiology*, 59(5), 936-940.

Long SP. (1991). Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO_2 concentrations: has its importance been underestimated?. *Plant, Cell & Environment*, 14(8), 729-739.

Long BM, Rae BD, Badger MR. & Price GD. (2011). Over-expression of the β-carboxysomal CcmM protein in Synechococcus PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. *Photosynthesis research*, 109(1-3), 33-45.

Lopez-Bautista JM, Rindi F. & Guiry MD. (2006). Molecular systematics of the subaerial green algal order Trentepohliales: an assessment based on morphological and molecular data. *International journal of systematic and evolutionary microbiology*, 56(7), 1709-1715.

Losh JL, Young JN. & Morel FM. (2013). Rubisco is a small fraction of total protein in marine phytoplankton. *New Phytologist*, 198(1), 52-58.

Lucas WJ, Keifer DW. & Sanders D. (1983). Bicarbonate transport in *Chara corallina*: Evidence for cotransport of HCO_3^- and H^+. *Journal of membrane biology*, 73, 263-274.

Lucas WJ. & Berry JA. (1985). Inorganic carbon transport in aquatic photosynthetic organisms. *Physiologia plantarum*, 65(4), 539-543.

Maberly SC. (1990). Exogenous sources of inorganic carbon for photosynthesis by marine macroalgae. *Journal of Phycology*, 26(3), 439-449.

Maberly SC, Ball LA, Raven JA. & Sültemeyer D. (2009). Inorganic carbon acquisition by Chrysophytes. *Journal of phycology*, 45(5), 1052-1061.

Maberly SC. & Gontero B. (2017). Ecological imperatives for aquatic CO_2-concentrating mechanisms. *Journal of experimental botany*, 68(14), 3797-3814.

Mackinder LC, Meyer MT, Mettler-Altmann T, Chen VK, Mitchell MC, Caspari O, Freeman Rosenzweig ES, Pallesen L, Reeves G, Itakura A, Roth R, Sommer F, Geimer

S, Mühlhaus T, Schroda M, Goodenough U, Stitt M, Griffiths H. & Jonikas MC. (2016). A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proceedings of the National Academy of Sciences*, 113(21), 5958-5963.

Mackinder LC, Chen C, Leib RD, Patena W, Blum SR, Rodman M, Ramundo S, Adams CM. & Jonikas MC. (2017). A Spatial Interactome Reveals the Protein Organization of the Algal CO_2-Concentrating Mechanism. *Cell*, 171(1), 133-147.

MacFarlane JJ. & Raven JA. (1985). External and internal CO_2 transport in Lemanea: interactions with the kinetics of ribulose bisphosphate carboxylase. *Journal of Experimental Botany*, 36(4), 610-622.

MacFarlane JJ. & Raven JA. (1989). Quantitative determination of the unstirred layer permeability and kinetic parameters of RUBISCO in *Lemanea mamillosa*. *Journal of Experimental Botany*, 321-327.

MacFarlane JJ. & Raven JA. (1990). C, N and P nutrition of *Lemanea mamillosa* Kütz.(Batrachospermales, Rhodophyta) in the Dighty Burn, Angus, UK. *Plant, Cell & Environment*, 13(1), 1-13.

Maeda N, Kitano K, Fukui T, Ezaki S, Atomi H, Miki K. & Imanaka T. (1999). Ribulose bisphosphate carboxylase/oxygenase from the hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1 is composed solely of large subunits and forms a pentagonal structure. *Journal of molecular biology*, 293(1), 57-66.

Maier A, Fahnenstich H, Von Caemmerer S, Engqvist MK, Weber AP, Flügge UI. & Maurino VG. (2012). Transgenic introduction of a glycolate oxidative cycle into A. thaliana chloroplasts leads to growth improvement. *Frontiers in plant science*, 3, 38.

Manhart JR. (1994). Phylogenetic analysis of green plant rbcL sequences. *Molecular Phylogenetics and Evolution*, 3, 114–127

Manton I. & Parke M. (1960). Further observations on small green flagellates with special reference to possible relatives of *Chromulina pusilla* Butcher. *Journal of the Marine Biological Association of the United Kingdom*, 39(2), 275-298.

Martin W. & Herrmann RG. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant physiology*, 118(1), 9-17.

Marsh JA, Singh VK, Jia Z. & Forman-Kay JD. (2006). Sensitivity of secondary structure propensities to sequence differences between α-and γ-synuclein: Implications for fibrillation. *Protein science*, 15(12), 2795-2804.

Massalski A, Mrozinska T. & Olech M. (2001). Ultrastructural observations on five pioneer soil algae from ice denuded areas (King George Island, West Antarctica). *Polar bioscience*, 14: 61-70

McCourt RM, Delwiche CF. & Karol KG. (2004). Charophyte algae and land plant origins. *Trends in Ecology & Evolution*, 19(12), 661-666.

McKay RML. & Gibbs SP. (1991). Composition and function of pyrenoids: cytochemical and immunocytochemical approaches. *Canadian Journal of Botany*, 69(5), 1040-1052.

McKay RML, Gibbs SP. & Vaughn KC. (1991). RuBisCo activase is present in the pyrenoid of green algae. *Protoplasma*, 162(1), 38-45.

McGrath JM. & Long SP. (2014). Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant Physiology*, 164(4), 2247-2261.

Melkonian M. (1975). The fine structure of the zoospores of *Fritschiella tuberosa* Iyeng.(Chaetophorineae, Chlorophyceae) with special reference to the flagellar apparatus. *Protoplasma*, 86, 391-404.

Melkonian M. (1984). Flagellar root-mediated interactions between the flagellar apparatus and cell organelles in green algae. In *Compartments in algal cells and their interaction* (pp. 96-108). Springer, Berlin, Heidelberg.

Melkonian M. & Preisig HR. (1986). A light and electron microscopic study of Scherffelia dubia, a new member of the scaly green flagellates (Prasinophyceae). *Nordic journal of botany*, 6(2), 235-256.

Melkonian M, McFadden GI, Reize IB. & Preisig HR. (1987). A light and electron microscopic study of the quadriflagellate green alga *Spermatozopsis exsultans*. *Plant systematics and evolution*, 158, 47-61.

Mercado JM. & Niell FX. (1999). Carbonic anhydrase activity and use of HCO_3^- in Bostrychia scorpioides (Ceramiales, Rhodophyceae). *European Journal of Phycology*, 34(1), 13-19.

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Maréchal-Drouard L, Marshall WF, Qu L-H, Nelson DR, Sanderfoot AA, Spalding MH, Kapitonov VV, Ren Q, Ferris P, Lindquist E, Shapiro H, Lucas SM, Grimwood J, Schmutz J, *Chlamydomonas* Annotation Team, JGI Annotation Team, Grigoriev IV, Rokhsar DS. & Grossman AR. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*, 318(5848), 245-250.

Meyer M, Seibt U. & Griffiths H. (2008). To concentrate or ventilate? Carbon acquisition, isotope discrimination and physiological ecology of early land plant life forms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1504), 2767-2778.

Meyer MT, Genkov T, Skepper JN, Jouhet J, Mitchell MC, Spreitzer RJ. & Griffiths H. (2012). RuBisCO small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. *Proceedings of the National Academy of Sciences, USA,* 109, 19474-19479.

Meyer M. & Griffiths H. (2013). Origins and diversity of eukaryotic CO_2-concentrating mechanisms: lessons for the future. *Journal of experimental botany*, 64(3), 769-786.

Meyer M. & Griffiths H. (2015). Photosynthesis: The internal plumbing of algal chloroplasts. *Elife*, *4*, e05983.

Meyer MT, McCormick AJ. & Griffiths H. (2016). Will an algal CO_2-concentrating mechanism work in higher plants?. *Current Opinion in Plant Biology*, 31, 181-188.

Meyer MT, Whittaker C. & Griffiths H. (2017). The algal pyrenoid: key unanswered questions. *Journal of experimental botany*, 68(14), 3739-3749.

Mikhailyuk TI, Sluiman HJ, Massalski A, Mudimu O, Demchenko EM, Kondratyuk SY. & Friedl T. (2008). New streptophyte green algae from terrestrial habitats and an assessment of the genus *Interfilum* (Klebsormidiophyceae, streptophyta). *Journal of Phycology*, 44(6), 1586-1603.

Mikhailyuk T, Holzinger A, Massalski A. & Karsten U. (2014). Morphology and ultrastructure of *Interfilum* and *Klebsormidium* (Klebsormidiales, Streptophyta) with special reference to cell division and thallus formation. *European journal of phycology*, 49(4), 395-412.

Mikhailyuk T, Glaser K, Holzinger A. & Karsten U. (2015). Biodiversity of *Klebsormidium* (Streptophyta) from alpine biological soil crusts (Alps, Tyrol, Austria, and Italy). *Journal of phycology*, 51(4), 750-767.

Mimura T, Müller R, Kaiser WM, Shimmen T. & Dietz KJ. (1993). ATP-dependent carbon transport in perfused *Chara* cells. *Plant, Cell and Environment*. 16, 653-661.

Mitchell M. (2014). Regulation of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. University of Cambridge.

Mitchell MC, Meyer MT. & Griffiths H. (2014). Dynamics of carbon-concentrating mechanism induction and protein relocalization during the dark-to-light transition in synchronized Chlamydomonas reinhardtii. *Plant Physiology*, 166(2), 1073-1082.

Miyachi S, Tsuzuki M, Maruyama I, Gantar M, Miyachi S. & Matsushima H. (1986). Effects of CO_2 concentration during growth on the intracellular structure of *Chlorella* and *Scenedesmus* (Chlorophyta). *Journal of phycology*, 22(3), 313-319.

Miyata T, Miyazawa S. & Yasunaga T. (1979). Two types of amino acid substitutions in protein evolution. *Journal of Molecular Evolution*, 12(3), 219-236.

Moestrup Ø. (1974). Ultrastructure of the scale-covered zoospores of the green alga *Chaetosphaeridium*, a possible ancestor of the higher plants and bryophytes. *Biological Journal of the Linnean Society*, 6(2), 111-125.

Moestrup Ø. (1991).Further studies of presumedly primitive green algae, including the description of Pedinophyceae class. nov. and resultor gen. nov. *Journal of Phycology*, 27(1), 119-133.

Moestrup Ø, Inouye I. & Hori T. (2003). Ultrastructural studies on *Cymbomonas tetramitiformis* (Prasinophyceae). I. General structure, scale microstructure, and ontogeny. *Canadian Journal of Botany*, 81(7), 657-671.

Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K. & Nozaki H. (1998). Presence of the CO_2-concentrating mechanism in some species of the pyrenoid-less free-living algal genus *Chloromonas* (Volvocales, Chlorophyta). *Planta*, 204(3), 269-276.

Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K. & Nozaki H. (1999). Role of pyrenoids in the CO_2-concentrating mechanism: comparative morphology, physiology and molecular phylogenetic analysis of closely related strains of *Chlamydomonas* and *Chloromonas* (Volvocales). *Planta*, 208(3), 365-372.

Moroney JV. & Tolbert NE. (1985). Inorganic carbon uptake by *Chlamydomonas reinhardtii. Plant physiology*, 77(2), 253-258.

Moroney JV. & Chen ZY. (1998). The role of the chloroplast in inorganic carbon uptake by eukaryotic algae. *Canadian Journal of Botany*, 76(6), 1025-1034.

Moroney JV. & Ynalvez RA. (2007). Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*. *Eukaryotic cell*, 6(8), 1251-1259.

Morse D, Salois P, Markovic P. & Hastings JW. (1995). A nuclear-encoded form II RuBisCO in dinoflagellates. *Science*, 268(5217), 1622-1624.

Mueller-Cajar O, Stotz M. & Bracher A. (2014). Maintaining photosynthetic CO_2 fixation via protein remodelling: the Rubisco activase. *Photosynthesis research*, 119(1-2), 191-201.

Mukherjee A, Lau CS, Walker CE, Rai AK, Prejean CI, Yates G, Emrich-Mills T, Lemoine SG, Vinyard DJ, Mackinder LCM. & Moroney JV. (2019). Thylakoid localized bestrophin-like proteins are essential for the CO_2 concentrating mechanism of *Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences*, 116(34), 16915-16920.

Mulkidjanian AY. & Junge W. (1997). On the origin of photosynthesis as inferred from sequence analysis. *Photosynthesis Research*, 51(1), 27-42.

Müller OF. (1882) Flora Danica, Band V, Heft 15.

Murru M. & Sandgren CD. (2004). Habitat matters for inorganic carbon acquisition in 38 species of red macroalgae (Rhodophya) from Puget sound, Wahsington, USA. *Journal of Phycology*, 40(5), 837-845.

Nagarajan R. & Gill KS. (2018). Evolution of Rubisco activase gene in plants. *Plant molecular biology*, 96(1-2), 69-87.

Nakada T, Suda S. & Nozaki H. (2007). A taxonomic study of *Hafniomonas* (Chlorophyceae) based on a comparative examination of cultured material. *Journal of phycology*, 43, 397-411.

Newman AM. & Cooper JB. (2007). XSTREAM: a practical algorithm for identification and architecture modeling of tandem repeats in protein sequences. *BMC bioinformatics*, 8(1), 382.

Neyman J. & Pearson ES. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference: Part II. *Biometrika*, 263-294.

Nisbet EG. (1995). Origins of photosynthesis. *Nature*, 373, 479-480.

Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK, Haas FB, Vanderstraeten L, Becker D, Lang D, Vosolsobe S, Rombauts S, Wilhelmsson PKI, Janitza P, Kern R, Heyl A, Rümpler F, Calderón Villalobos LIA, Clay JM, Skokan R, Toyoda A, Suzuki Y, Kagoshima H, Schijlen E, Tajeshwar N, Catarino B, Hetherington AJ, Saltykova A, Bonnot C, Breuninger H, Symeonidi A, Radhakrishnan GV, Van Nieuwerburgh F, Deforce D, Chang C, Karol KG, Hedrich R, Ulvskov P, Glöckner G, Delwiche CF, Petrásek J, Van de Peer Y. & Friml J. (2018). The Chara genome: secondary complexity and implications for plant terrestrialization. *Cell*, 174(2), 448-464.

Nozaki H, Onishi K. & Morita E. (2002). Differences in pyrenoid morphology are correlated with differences in the rbcL genes of members of the *Chloromonas* lineage (Volvocales, Chlorophyceae). *Journal of Molecular Evolution*, 55(4), 414-430.

Ogawa S. (1988). Disappearance of chloroplast nucleoids during male gamete formation in *Bryopsis plumosa* (Hudson) C. Ag.(Chlorophyceae). *Botanical Gazette*, 149(1), 25-29.

Ohad I, Siekevitz P. & Palade GE. (1967). Biogenesis of chloroplast membranes: I. Plastid Dedifferentiation in a Dark-Grown Algal Mutant (*Chlamydomonas reinhardtii*). *The Journal of cell biology*, 35(3), 521-552.

Ohnishi N, Mukherjee B, Tsujikawa T, Yanase M, Nakano H, Moroney JV. & Fukuzawa H. (2010). Expression of a low CO_2-inducible protein, LCI1, increases inorganic carbon uptake in the green alga *Chlamydomonas reinhardtii*. *Plant Cell*, 22, 3105–3117.

O'Kelly CJ, Bellows WK. & Wysor B. (2004). Phylogenetic position of *Bolbocoleon piliferum* (Ulvophyceae, Chlorophyta): evidence from reproduction, zoospore and gamete ultrastructure, and small subunit rRNA gene sequences. *Journal of Phycology*, 40, 209-222.

O'Kelly CJ. (2007). The origin and early evolution of green plants. In: Evolution of Primary Producers in the Sea, pp. 287–309. Falkowski PG. and Knoll AH, Eds., Elsevier Academic, Burlington, MA.

O'Leary MH. (1988). Carbon isotopes in photosynthesis. *Bioscience*, 38(5), 328-336.

Olmstead RG. & Palmer JD. (1994). Chloroplast DNA systematics: a review of methods and data analysis. *American journal of botany*, 81(9), 1205-1224.

Oltrogge LM, Chaijarasphong T, Chen AW, Bolin ER, Marqusee S. & Savage DF. (2019). α-carboxysome formation is mediated by the multivalent and disordered protein CsoS2. *bioRxiv*, 708164.

Öpik H. & Flynn KJ. (1989). The digestive process of the dinoflagellate, *Oxyrrhis marina* Dujardin, feeding on the chlorophyte, *Dunaliella primolecta* Butcher: a combined study of ultrastructure and free amino acids. *New phytologist*, 113, 143-151.

Orr DJ, Alcântara A, Kapralov MV, Andralojc PJ, Carmo-Silva E. & Parry MA. (2016). Surveying Rubisco diversity and temperature response to improve crop photosynthetic efficiency. *Plant Physiology*, 172(2), 707-717.

Orr DJ. & Carmo-Silva E. (2018). Extraction of Rubisco to determine catalytic constants. In *Photosynthesis* (pp. 229-238). Humana Press, New York, NY.

Ort DR, Merchant SS, Alric J, Barkan A, Blankenship RE, Bock R, Croce R, Hanson MR, Hibberd JM, Long SP, Moore TA, Moroney J, Niyogi KK, Parry MA, Peralta-Yahya PP, Prince RC, Redding KE, Spalding MH, Van Wijk KJ, Vermaas WFJ, Von Caemmerer S, Weber APM, Yeates TO, Yuan JS. & Zhu XG. (2015). Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proceedings of the national academy of sciences*, 112(28), 8529-8536.

Osterhout WJV. (1945). Water relations in the cell: The chloroplasts of *Nitella* and of *Spirogyra. The Journal of general physiology*, 29(2), 73.

Palmer JD. (1997). Organelle Genomes--Going, Going, Gone. *Science*, 275(5301), 790-790.

Palmqvist K, Sjöberg S. & Samuelsson G. (1988). Induction of inorganic carbon accumulation in the unicellular green algae *Scenedesmus obliquus* and *Chlamydomonas reinhardtii. Plant physiology*, 87(2), 437-442.

Palmqvist K, Ögren E. & Lernmark U. (1994a). The CO_2-concentrating mechanism is absent in the green alga Coccomyxa: a comparative study of photosynthetic CO_2 and light responses of *Coccomyxa, Chlamydomonas reinhardtii* and barley protoplasts. *Plant, Cell & Environment*, 17(1), 65-72.

Palmqvist K, Máguas C, Badger MR. & Griffiths H. (1994b). Assimilation, accumulation and isotope discrimination of inorganic carbon in lichens: further evidence for the operation of a CO_2 concentrating mechanism in cyanobacterial lichens. *Cryptogamic Botany*, 4, 218-226.

Palmqvist K, Sültemeyer D, Baldet P, Andrews TJ. & Badger MR. (1995). Characterisation of inorganic carbon fluxes, carbonic anhydrase (s) and ribulose-1, 5-biphosphate carboxylase-oxygenase in the green unicellular alga *Coccomyxa*. *Planta*, 197(2), 352-361.

Palmqvist K, de Los Rios A, Ascaso C. & Samuelsson G. (1997). Photosynthetic carbon acquisition in the lichen photobionts *Coccomyxa* and *Trebouxia* (Chlorophyta). *Physiologia plantarum*, 101(1), 67-76.

Parry MA, Keys AJ, Madgwick PJ, Carmo-Silva AE. & Andralojc PJ. (2008). Rubisco regulation: a role for inhibitors. *Journal of experimental botany*, 59(7), 1569-1580.

Parry MA, Andralojc PJ, Scales JC, Salvucci ME, Carmo-Silva AE, Alonso H. & Whitney SM. (2012). Rubisco activity and regulation as targets for crop improvement. *Journal of experimental botany*, 64(3), 717-730.

Pearson BR. & Norris RE. (1975). Fine structure of cell division in *Pyramimonas parkae* Norris and Pearson (Chlorophyta, Prasinophyceae). *Journal of Phycology*, 11(1), 113-124.

Pei ZY, Mu GL, Pan J. & Zhang DM. (2013). Codon usage and coevolution of the large and small subunits of ribulose-1, 5-bisphosphate carboxylase/oxygenase. *Journal of Systematics and Evolution*, 51(5), 511-521.

Pedersen O, Colmer TD. & Sand-Jensen KAJ. (2013). Underwater photosynthesis of submerged plants–recent advances and methods. *Frontiers in Plant Science*, 4, 140.

Peterfi LS. & Manton I. (1968). Observations with the electron microscope on *Asteromonas gracilis* Artari emend.(Stephanoptera gracilis (Artari) Wisl.), with some comparative observations on *Dunaliella sp*. *British Phycological Bulletin***,** 3, 423-440.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC. & Ferrin TE. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13), 1605-1612.

Petersen J, Teich R, Becker B, Cerff R. & Brinkmann H. (2006). The GapA/B gene duplication marks the origin of Streptophyta (charophytes and land plants). *Molecular biology and evolution*, 23(6), 1109-1118.

Phadwahl K. & Singh PK. (2003). Isolation and characterization of an indigenous isolate of Dunaliella sp. for β-carotene and glycerol production from a hypersaline lake in India. *Journal of Basic Microbiology: An International Journal on Biochemistry, Physiology, Genetics, Morphology, and Ecology of Microorganisms*, 43(5), 423-429.

Phillips R. & Milo R. (2009). A feeling for the numbers in biology. *Proceedings of the National Academy of Sciences*, 106(51), 21465-21471.

Pichrtova M, Vesela J, Holzinger A. & Hajek T. (2013). Diversity and desiccation tolerance of *Zygnema* (Zygnematophyceae, Streptophyta) on Svalbard (high arctic). 229. *Phycologia*, *52*(4).

Pie MR. (2006). The influence of phylogenetic uncertainty on the detection of positive Darwinian selection. *Molecular biology and evolution*, 23(12), 2274-2278.

Pierrehumbert RT, Abbot DS, Voigt A. & Koll D. (2011). Climate of the Neoproterozoic. *Annual Review of Earth and Planetary Sciences*, 39, 417-460.

Pond SLK. & Muse SV. (2005). HyPhy: hypothesis testing using phylogenies. In *Statistical methods in molecular evolution* (pp. 125-181). Springer, New York, NY.

Porra RJ, Thompson WA. & Kriedemann PE. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b

extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 975(3), 384-394.

Portis Jr AR, Li C, Wang D. & Salvucci ME. (2007). Regulation of Rubisco activase and its interaction with Rubisco. *Journal of experimental botany*, 59(7), 1597-1604.

Posada D. (2008). jModelTest: phylogenetic model averaging. *Molecular biology and evolution*, 25(7), 1253-1256.

Prasanna R, Ratha SK, Rojas C. & Bruns MA. (2011). Algal diversity in flowing waters at an acidic mine drainage "barrens" in central Pennsylvania, USA. *Folia microbiologica*, 56, 491-496.

Preisig HR. & Melkonian M. (1984). A light and electron microscopical study of the green flagellate *Spermatozopsis similis* spec. nova. *Plant systematics and evolution*, 146(1-2), 57-74.

Price GD, Howitt SM, Harrison K. & Badger MR. (1993). Analysis of a genomic DNA region from the cyanobacterium Synechococcus sp. strain PCC7942 involved in carboxysome assembly and function. *Journal of bacteriology*, 175(10), 2871-2879.

Price GD, Badger MR, Woodger FJ. & Long BM. (2008). Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *Journal of experimental botany*, 59(7), 1441-1461.

Price GD. (2011). Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism. *Photosynthesis Research*, 109(1-3), 47-57.

Price GD, Pengelly JJ, Forster B, Du J, Whitney SM, von Caemmerer S, Badger MR, Howitt SM. & Evans JR. (2012). The cyanobacterial CCM as a source of genes for improving photosynthetic CO_2 fixation in crop species. *Journal of experimental botany*, 64(3), 753-768.

Prins A, Orr DJ, Andralojc PJ, Reynolds MP, Carmo-Silva E. & Parry MA. (2016). Rubisco catalytic properties of wild and domesticated relatives provide scope for improving wheat photosynthesis. *Journal of Experimental Botany*, 67(6), 1827-1838.

Pröschold T, Marin B, Schlösser UG. & Melkonian M. (2001). Molecular phylogeny and taxonomic revision of Chlamydomonas (Chlorophyta). I. Emendation of *Chlamydomonas* Ehrenberg and *Chloromonas* Gobi, and description of *Oogamochlamys* gen. nov. and Lobochlamys gen. nov. *Protist*, 152(4), 265-300.

Race HL, Herrmann RG. & Martin W. (1999). Why have organelles retained genomes? *Trends in Genetics*, 15(9), 364-370.

Raimundo SC, Sørensen I, Tinaz B, Ritter E, Rose JK. & Domozych DS. (2018). Isolation and manipulation of protoplasts from the unicellular green alga *Penium margaritaceum*. *Plant methods*, 14(1), 18.

Rambaut A. (2007). FigTree, a graphical viewer of phylogenetic trees.

Rasmussen B, Fletcher IR, Brocks JJ. & Kilburn MR. (2008). Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature*, 455(7216), 1101.

Raven JA. & Beardall J. (1981). Carbon dioxide as the exogenous inorganic carbon source for Batrachospermum and Lemanea. *British Phycological Journal*, 16(2), 165-175.

Raven J, Beardall J. & Griffiths H. (1982). Inorganic C-sources for *Lemanea*, *Cladophora* and *Ranunculus* in a fast-flowing stream: measurements of gas exchange and of carbon isotope ratio and their ecological implications. *Oecologia*, 53(1), 68-78.

Raven JA, Osborne BA. & Johnston AM. (1985). Uptake of CO_2 by aquatic vegetation. *Plant, Cell & Environment*, 8(6), 417-425.

Raven JA. & Richardson K. (1986). Marine environments. *Photosynthesis in Contrasting Environments. Elsevier Science Publishing*, 337-396.

Raven JA. (1991). Implications of inorganic carbon utilization: ecology, evolution, and geochemistry. *Canadian Journal of Botany*, 69(5), 908-924.

Raven JA, Ball LA, Beardall J, Giordano M. & Maberly SC. (2005). Algae lacking carbon-concentrating mechanisms. *Canadian Journal of Botany*, 83(7), 879-890.

Raven JA. (2010). Inorganic carbon acquisition by eukaryotic algae: four current questions. *Photosynthesis research*, 106(1-2), 123-134.

Raven JA, Giordano M, Beardall J. & Maberly SC. (2011). Algal and aquatic plant carbon concentrating mechanisms in relation to environmental change. *Photosynthesis Research*, 109(1-3), 281-296.

Raven JA, Giordano M, Beardall J. & Maberly SC. (2012). Algal evolution in relation to atmospheric CO_2: carboxylases, carbon-concentrating mechanisms and carbon oxidation cycles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1588), 493-507.

Raven JA, Beardall J. & Sánchez-Baracaldo P. (2017). The possible evolution and future of CO_2-concentrating mechanisms. *Journal of Experimental Botany*, 68(14), 3701-3716.

Rawat M, Henk MC, Lavigne LL. & Moroney JV. (1996). *Chlamydomonas reinhardtii* mutants without ribulose-1, 5-bisphosphate carboxylase-oxygenase lack a detectable pyrenoid. *Planta*, 198(2), 263-270.

Ray S, Klenell M, Choo KS, Pedersén M. & Snoeijs P. (2003). Carbon acquisition mechanisms in *Chara tomentosa*. Aquatic botany, 76, 141-154.

Read BA. & Tabita FR. (1992). A hybrid ribulose bisphosphate carboxylase/oxygenase enzyme exhibiting a substantial increase in substrate specificity factor. *Biochemistry*, 31(24), 5553-5560.

Renberg L, Johansson AI, Shutova T, Stenlund H, Aksmann A, Raven JA. & Samuelsson G. (2010). A metabolomic approach to study major metabolite changes during acclimation to limiting CO_2 in *Chlamydomonas reinhardtii*. *Plant physiology*, 154(1), 187-196.

Rice P, Longden I. & Bleasby A. (2000). EMBOSS: the European molecular biology open software suite.

Rickaby RE. & Hubbard ME. (2019). Upper ocean oxygenation, evolution of Rubisco and the Phanerozoic succession of phytoplankton. *Free Radical Biology and Medicine*, 140, 295-304.

Rindi F, Mikhailyuk TI, Sluiman HJ, Friedl T. & López-Bautista JM. (2011). Phylogenetic relationships in *Interfilum* and *Klebsormidium* (Klebsormidiophyceae, Streptophyta). *Molecular phylogenetics and evolution*, 58(2), 218-231.

Romero P, Obradovic Z, Li X, Garner EC, Brown CJ. & Dunker AK. (2001). Sequence complexity of disordered protein. *Proteins: Structure, Function, and Bioinformatics*, 42(1), 38-48.

Rosenzweig ESF, Xu B, Cuellar LK, Martinez-Sanchez A, Schaffer M, Strauss M, Cartwright HN, Ronceray P, Plitzko JM, Förster F, Wingreen NS, Engel BD, Mackinder LCM. & Jonikas MC. (2017). The eukaryotic CO_2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell*, 171(1), 148-162.

Rotatore C. & Colman B. (1991). The acquisition and accumulation of inorganic carbon by the unicellular green alga *Chlorella ellipsoidea*. *Plant, Cell & Environment*, 14(4), 377-382.

Ruan J. & Li H. (2019). Fast and accurate long-read assembly with wtdbg2. *BioRxiv*, 530972.

Sáez PL, Bravo LA, Cavieres LA, Vallejos V, Sanhueza C, Font-Carrascosa M, Gil-Pelegrin E, Peguero-Pina JJ. & Galmés J. (2017). Photosynthetic limitations in two Antarctic vascular plants: importance of leaf anatomical traits and Rubisco kinetic parameters. *Journal of experimental botany*, 68(11), 2871-2883.

Sage RF. (2002). Variation in the k_{cat} of Rubisco in C_3 and C_4 plants and some implications for photosynthetic performance at high and low temperature. *Journal of Experimental Botany*, 53(369), 609-620.

Sage RF, Christin PA. & Edwards EJ. (2011). The C_4 plant lineages of planet Earth. *Journal of experimental botany*, 62(9), 3155-3169.

Salvucci ME. & Ogren WL. (1996). The mechanism of Rubisco activase: insights from studies of the properties and structure of the enzyme. *Photosynthesis Research*, 47(1), 1-11.

Sasanuma T. (2001). Characterization of the *rbcS* multigene family in wheat: subfamily classification, determination of chromosomal location and evolutionary analysis. *Molecular genetics and genomics*, 265(1), 161-171.

Satagopan S. & Spreitzer RJ. (2008). Plant-like substitutions in the large-subunit carboxy terminus of *Chlamydomonas* Rubisco increase CO_2/O_2 Specificity. *BMC plant biology*, 8(1), 85.

Saxby-Rouen KJ, Leadbeater BSC. & Reynolds CS. (1998). The relationship between the growth of Synura petersenii (Synurophyceae) and components of the dissolved inorganic carbon system. *Phycologia*, *37*(6), 467-477.

Savir Y, Noor E, Milo R. & Tlusty T. (2010). Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proceedings of the National Academy of Sciences, USA*, 107: 3475–3480.

Sena L. & Uversky VN. (2016). Comparison of the intrinsic disorder propensities of the RuBisCO activase enzyme from the motile and non-motile oceanic green microalgae. *Intrinsically disordered proteins*, 4(1), e1253526.

Schopf JW. (1992). Proterozoic prokaryotes: affinities, geologic distribution, and evolutionary trends. *The Proterozoic Biosphere*, 195-218.

Schopf JW. (1993). Microfossils of the Early Archean Apex chert: new evidence of the antiquity of life. *Science*, 260(5108), 640-646.

Schrödinger LLC. (2010). The PyMOL molecular graphics system, version 1.3 r1.

Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F. & Serrano L. (2005). The FoldX web server: an online force field. *Nucleic acids research*, 33(2), W382-W388.

Sharwood RE, Sonawane BV, Ghannoum O. & Whitney SM. (2016). Improved analysis of C_4 and C_3 photosynthesis via refined in vitro assays of their carbon fixation biochemistry. *Journal of Experimental Botany*, 67(10), 3137-3148.

Sharma R, Sharma A, Raicar G, Tsunoda T. & Patil A. (2019). OPAL+: Length-Specific MoRF Prediction in Intrinsically Disordered Protein Sequences. *Proteomics*, 19(6), 1800058.

Shih PM, Occhialini A, Cameron JC, Andralojc PJ, Parry MA. & Kerfeld CA. (2016). Biochemical characterization of predicted Precambrian RuBisCO. *Nature communications*, *7*, 10382.

Shiraiwa Y. & Miyachi S. (1983). Factors controlling induction of carbonic anhydrase and efficiency of photosynthesis in *Chlorella vulgaris* llh cells. *Plant and cell physiology*, 24(5), 919-923.

Shiraiwa Y. & Miyachi S. (1985). Effects of temperature and CO_2 concentration on induction of carbonic anhydrase and changes in efficiency of photosynthesis in *Chlorella vulgaris* 11h. *Plant and Cell Physiology*, 26(3), 543-549.

Shiratori T, Fujita S, Shimizu T, Nakayama T. & Ishida KI. (2017). *Viridiuvalis adhaerens* gen. et sp. nov., a novel colony-forming chlorarachniophyte. *Journal of plant research*, 130(6), 999-1012.

Schmitz F. (1882). Die Chromatophoren der Algen. Vergleichende untersuchungen über Bau und Entwicklung der Chlorophyllkörper und der analogen Farbstoffkörper der Algen. M. Cohen & Sohn (F. Cohen), Bonn, Germany.

Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD. & Higgins DG. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, 7(1).

Škaloud P, Nemjová K, Veselá J, Černá K. & Neustupa J. (2011). A multilocus phylogeny of the desmid genus *Micrasterias* (Streptophyta): Evidence for the accelerated rate of morphological evolution in protists. *Molecular Phylogenetics and Evolution*, 61, 933-943.

Skuja H. (1943). Ein Fall von fakultativer Symbiose zwischen operculatem Discomycet und einer Chlamydomonade. *Arch. Protistenkunde*, 96, 365-376.

Sluiman HJ. (1985). Mitosis and cell division in *Cylindrocapsa geminella* (Chlorophyceae). *Journal of phycology*, 21, 523-532.

Soo RM, Hemp J, Parks DH, Fischer WW. & Hugenholtz P. (2017). On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science*, 355(6332), 1436-1440.

South PF, Cavanagh AP, Liu HW. & Ort DR. (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science*, 363(6422), eaat9077.

Spalding MH, Spreitzer RJ. & Ogren WL. (1983). Carbonic anhydrase-deficient mutant of *Chlamydomonas reinhardtii* requires elevated carbon dioxide concentration for photoautotrophic growth. *Plant Physiology*, 73(2), 268-272.

Spreitzer RJ, Esquivel MG, Du YC. & McLaughlin PD. (2001). Alanine-Scanning Mutagenesis of the Small-Subunit βA−βB Loop of Chloroplast Ribulose-1, 5-Bisphosphate Carboxylase/Oxygenase: Substitution at Arg-71 Affects Thermal Stability and CO_2/O_2 Specificity. *Biochemistry*, 40(19), 5615-5621.

Spreitzer RJ. & Salvucci ME. (2002). Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annual review of plant biology*, 53.

Spreitzer RJ. (2003). Role of the small subunit in ribulose-1, 5-bisphosphate carboxylase/oxygenase. *Archives of biochemistry and biophysics*, 414(2), 141-149.

Spreitzer RJ, Peddi SR. & Satagopan S. (2005). Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco. *Proceedings of the National Academy of Sciences*, 102(47), 17225-17230.

Stabenau H. & Winkler U. (2005). Glycolate metabolism in green algae. *Physiologia Plantarum*, 123(3), 235-245.

Stamatakis A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312-1313.

Steel M. & McKenzie A. (2001). Properties of phylogenetic trees generated by Yule-type speciation models. *Mathematical biosciences*, 170(1), 91-112.

Stewart KD, Mattox KR. & Floyd GL. (1973). Mitosis, cytokinesis, the distribution of Plasmodesmata and other cytological characteristics in the Ulotrichales, Ulvales and Chaetophorales: Phylogenetic and taxonomic considerations. *Journal of Phycology*, 9(2), 128-141.

Student. (1908). The probable error of a mean. *Biometrika*, 1-25.

Studer RA, Christin PA, Williams MA. & Orengo CA. (2014). Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. *Proceedings of the National Academy of Sciences*, 111(6), 2223-2228.

Suda S, Watanabe MM. & Inouye I. (2004). Electron microscopy of sexual reproduction in *Nephroselmis olivacea* (Prasinophyceae, Chlorophyta). *Phycological Research*, 52(3), 273-283.

Sültemeyer DF, Miller AG, Espie GS, Fock HP. & Canvin DT. (1989). Active CO_2 transport by the green alga *Chlamydomonas reinhardtii. Plant Physiology*, 89(4), 1213-1219.

Sültemeyer DF, Fock HP. & Canvin DT. (1991). Active uptake of inorganic carbon by *Chlamydomonas reinhardtii*: evidence for simultaneous transport of HCO_3^- and CO_2 and characterization of active CO_2 transport. *Canadian Journal of Botany*, 69(5), 995-1002.

Süss KH, Prokhorenko I. & Adler K. (1995). In situ association of Calvin cycle enzymes, ribulose-1, 5-bisphosphate carboxylase/oxygenase activase, ferredoxin-NADP+ reductase, and nitrite reductase with thylakoid and pyrenoid membranes of *Chlamydomonas reinhardtii* chloroplasts as revealed by immunoelectron microscopy. *Plant Physiology*, 107(4), 1387-1397.

Tabita FR. (1999). Microbial ribulose 1, 5-bisphosphate carboxylase/oxygenase: a different perspective. *Photosynthesis Research*, 60(1), 1-28.

Tabita FR, Hanson TE, Li H, Satagopan S, Singh J. & Chan S. (2007). Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.*, 71(4), 576-599.

Tabita FR, Satagopan S, Hanson TE, Kreel NE. & Scott SS. (2008). Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *Journal of experimental botany*, 59(7), 1515-1524.

Tabita FR, Hanson TE, Satagopan S, Witte BH. & Kreel NE. (2008). Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1504), 2629-2640.

Taiz L. & Zeiger E. (1998). Plant Physiology, second ed. Sinauer Associates, Sunderland.

Tang Q, Pang K, Yuan X. & Xiao S. (2020). A one-billion-year-old multicellular chlorophyte. *Nature Ecology & Evolution*, 1-7.

Taniguchi GM, Peres AC, Senna PAC. & Compère P. (2003). The desmid genera Cosmarium, Actinotaenium and Cosmocladium from an oxbow lake, Jataí Ecological Station (Southeastern Brazil). *Systematics and Geography of Plants*, 133-159.

Tavaré S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on mathematics in the life sciences*, 17(2), 57-86.

Tawfik DS. (2014). Accuracy-rate tradeoffs: how do enzymes meet demands of selectivity and catalytic efficiency?. *Current opinion in chemical biology*, 21, 73-80.

Taylor TC. & Andersson I. (1997). The structure of the complex between rubisco and its natural substrate ribulose 1, 5-bisphosphate. *Journal of molecular biology*, 265(4), 432-444.

Taylor TC, Backlund A, Bjorhall K, Spreitzer RJ. & Andersson I. (2001). First crystal structure of Rubisco from a green alga, *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry*, 276(51), 48159-48164.

Tcherkez G, Farquhar GD. & Andrews TJ. (2006). Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proceedings of the National Academy of Sciences, USA*, 103, 7246–7251

Tcherkez G. (2013). Modelling the reaction mechanism of ribulose-1,5-bisphosphate carboxylase/oxygenase and consequences for kinetic parameters. *Plant, Cell & Environment*, 36(9), 1586-1596.

Thieulin-Pardo G, Avilan L, Kojadinovic M. & Gontero B. (2015). Fairy “tails”: flexibility and function of intrinsically disordered extensions in the photosynthetic world. *Frontiers in molecular biosciences*, 2, 23.

Throndsen J. & Zingone A. (1997). *Dolichomastix tenuilepis* sp. nov., a first insight into the microanatomy of the genus *Dolichomastix* (Mamiellales, Prasinophyceae, Chlorophyta). *Phycologia*, 36, 244-254.

Tortell PD. (2000). Evolutionary and ecological perspectives on carbon acquisition in phytoplankton. *Limnology and Oceanography*, 45(3), 744-750.

Tsuzuki M, Gantar M, Aizawa K. & Miyachi S. (1986). Ultrastructure of *Dunaliella tertiolecta* cells grown under low and high CO_2 concentrations. *Plant and cell physiology*, 27, 737-739.

Valegård K, Andralojc PJ, Haslam RP, Pearce FG, Eriksen GK, Madgwick PJ, Kristoffersen AK, Van Lun M, Klein U, Eilertsen HC, Parry MA. & Andersson I. (2018). Structural and functional analyses of Rubisco from arctic diatom species reveal unusual posttranslational modifications. *Journal of Biological Chemistry*, 293(34), 13033-13043.

Van Baren MJ, Bachy C, Reistetter EN, Purvine SO, Grimwood J, Sudek S. & Grigoriev IV. (2016). Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC genomics*, 17, 267.

Van Lun M, van der Spoel D. & Andersson I. (2011). Subunit interface dynamics in hexadecameric rubisco. *Journal of molecular biology*, 411(5), 1083-1098.

Van Lun M, Hub JS, van der Spoel D. & Andersson I. (2014). CO_2 and O_2 distribution in Rubisco suggests the small subunit functions as a CO_2 reservoir. *Journal of the American Chemical Society*, 136(8), 3165-3171.

Vance P. & Spalding MH. (2005). Growth, photosynthesis, and gene expression in *Chlamydomonas* over a range of CO_2 concentrations and CO_2/O_2 ratios: CO_2 regulates multiple acclimation states. *Canadian Journal of Botany*, 83(7), 796-809.

Vaughn KC, Campbell EO, Hasegawa J, Owen HA. & Renzaglia KS. (1990). The pyrenoid is the site of ribulose 1, 5-bisphosphate carboxylase/oxygenase accumulation in the hornwort (Bryophyta: Anthocerotae) chloroplast. *Protoplasma*, 156(3), 117-129.

Villarreal JC. & Renner SS. (2012). Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proceedings of the National Academy of Sciences*, 109(46), 18873-18878.

Villarreal JC. & Renner SS. (2014). A review of molecular-clock calibrations and substitution rates in liverworts, mosses, and hornworts, and a timeframe for a taxonomically cleaned-up genus Nothoceros. *Molecular phylogenetics and evolution*, *78*, 25-35.

Von Caemmerer S. & Quick WP. (2000). Rubisco: physiology in vivo. In *Photosynthesis* (pp. 85-113). Springer, Dordrecht.

Von Caemmerer S, Quick WP. & Furbank RT. (2012). The development of C_4 rice: current progress and future challenges. *Science*, 336(6089), 1671-1672.

Wagner G. & Klein K. (1981). Mechanism of chloroplast movement in Mougeotia. *Protoplasma*, 109, 169-185.

Wang Y. & Spalding MH. (2006). An inorganic carbon transport system responsible for acclimation specific to air levels of CO_2 in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences*, 103(26), 10110-10115.

Wang Y, Stessman DJ. & Spalding MH. (2015). The CO_2 concentrating mechanism and photosynthetic carbon assimilation in limiting CO_2: how *Chlamydomonas* works against the gradient. *Plant J*, 82, 429–448.

Wang H, Yan X, Aigner H, Bracher A, Nguyen ND, Hee WY, Long BM, Price GD, Hartl FU. & Hayer-Hartl M. (2019). Rubisco condensate formation by CcmM in β-carboxysome biogenesis. *Nature*, 566(7742), 131.

Wang L, Yamano T, Kajikawa M, Hirono M. & Fukuzawa H. (2014). Isolation and characterization of novel high-CO_2-requiring mutants of *Chlamydomonas reinhardtii*. *Photosynthesis research*, 121(2-3), 175-184.

Warren SD, Clair LLS, Stark LR, Lewis LA, Pombubpa N, Kurbessoian T, Stajich JE. & Aanderud ZT. (2019). Reproduction and dispersal of biological soil crust organisms. *Frontiers In Ecology Evolution.* 7, 344.

Wasmann CC, Ramage RT, Bohnert HJ. & Ostrem JA. (1989). Identification of an assembly domain in the small subunit of ribulose-1, 5-bisphosphate carboxylase. *Proceedings of the National Academy of Sciences*, 86(4), 1198-1202.

Watanabe S, Himizu A, Lewis LA, Floyd GL. & Fuerst PA. (2000). *Pseudoneochloris marina* (Chlorophyta), a new coccoid ulvophycean alga, and its phylogenetic position inferred from morphological and molecular data. *Journal of Phycology*, 36, 596-604.

Watanabe S. & Nakayama T. (2007). Ultrastructure and phylogenetic relationships of the unicellular green algae *Ignatius tetrasporus* and *Pseudocharacium americanum* (Chlorophyta). *Phycological Research*, 55(1), 1-16.

Watson MW. (1975). Flagellar apparatus, eyesport and behavior of *Microthamnion kuetzingianum* (Chlorophyceae) zoospores. *Journal of Phycology*, 11(4), 439-448.

Webster-Smith NK, Chapman GB. & Sze P. (1983). The ultrastructure of *Helicodictyon planctonicum* (Chlorophyceae). *Phycologia,* 22, 295-301.

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL. & Scheffler K. (2014). RELAX: detecting relaxed selection in a phylogenetic framework. *Molecular biology and evolution*, 32(3), 820-832.

West GS. (1904). *A treatise on the British freshwater algae*. University Press.

Whitney SM. & Andrews TJ. (1998). The CO_2/O_2 specificity of single-subunit ribulose-bisphosphate carboxylase from the dinoflagellate, *Amphidinium carterae*. *Functional Plant Biology*, 25(2), 131-138.

Whitney SM, Houtz RL. & Alonso H. (2011). Advancing our understanding and capacity to engineer nature's CO_2-sequestering enzyme, Rubisco. *Plant Physiology*, 155(1), 27-35.

Wille N. (1903). Algologische Notizen IX-XIV. *Nyt Magazin for Naturvidenskaberne* 41: 89-185.

Wilcox LW. & Floyd GL. (1988). Ultrastructure of the gamete of *Pediastrum duplex* (Chlorophyceae). *Journal of Phycology*, 24(2), 140-146.

Wilson RH. & Hayer-Hartl M. (2018). Complex chaperone dependence of Rubisco biogenesis. *Biochemistry*, 57(23), 3210-3216.

Winck FV, Kwasniewski M, Wienkoop S. & Mueller-Roeber B. (2011). An optimized method for the isolation of nuclei from *Chlamydomonas reinhardtii* (Chlorophyceae). *Journal of phycology*, 47(2), 333-340.

Wolf FR. & Cox ER. (1981). Ultrastructure of active and resting colonies of *Botryococcus braunii* (Chlorophyceae) 1. *Journal of Phycology*, 17(4), 395-405.

Wujek DE. & Thompson RH. (1999). The Algal Genera Chaetopeltis, Oligochaetophora, and Polychaetophora (Chaetopeltidales, Chlorophyta). *Transactions of the Kansas Academy of Science*, 40-46.

Wunder T, Cheng SLH, Lai SK, Li HY. & Mueller-Cajar O. (2018). The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. *Nature communications*, 9(1), 5076.

Wygasch J. (1964). Elektronenmikroskopische Untersuchungen über die Bauweise der Pyrenoide vonHaematococcus pluvialis. *Protoplasma*, 59, 266-276.

Yamada K, Davydov II, Besnard G. & Salamin N. (2019). Duplication history and molecular evolution of the *rbcS* multigene family in angiosperms. *Journal of experimental botany*, 70(21), 6127-6139.

Yamano T, Sato E, Iguchi H, Fukuda Y. & Fukuzawa H. (2015). Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga *Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences*, 112(23), 7315-7320.

Yang Z. (1998). Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular biology and evolution*, 15(5), 568-573.

Yang Z. & Nielsen R. (1998). Synonymous and non synonymous rate variation in nuclear genes of mammals. *Journal of molecular evolution*, 46(4), 409-418.

Yang Z, Wong WS. & Nielsen R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular biology and evolution*, 22(4), 1107-1118.

Yang Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution*, 24(8), 1586-1591.

Yang Z. & Rannala B. (2012). Molecular phylogenetics: principles and practice. *Nature reviews genetics*, 13(5), 303.

Yao J, Chung J, Eliezer D, Wright PE. & Dyson HJ. (2001). NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry*, 40(12), 3561-3571.

Yokota A, Harada A. & Kitaoka S. (1989). Characterization of ribulose 1, 5-bisphosphate carboxylase/oxygenase from *Euglena gracilis* Z. *The Journal of Biochemistry*, 105(3), 400-405.

Yoon HS, Hackett JD, Pinto G. & Bhattacharya D. (2002). The single, ancient origin of chromist plastids. *Journal of phycology*, 38, 40-40.

Yoon HS, Hackett JD, Ciniglia C, Pinto G. & Bhattacharya D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Molecular biology and evolution*, 21(5), 809-818.

York PV. & Johnson LR. (2002). *The freshwater algal flora of the British Isles: an identification guide to freshwater and terrestrial algae*. Cambridge University Press.

Young EB. & Beardall J. (2005). Modulation of photosynthesis and inorganic carbon acquisition in a marine microalga by nitrogen, iron, and light availability. *Canadian Journal of Botany*, 83(7), 917-928.

Young JN, Rickaby REM, Kapralov MV. & Filatov DA. (2012). Adaptive signals in algal Rubisco reveal a history of ancient atmospheric carbon dioxide. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1588), 483-492.

Young JN, Goldman JA, Kranz SA, Tortell PD. & Morel FM. (2015). Slow carboxylation of Rubisco constrains the rate of carbon fixation during Antarctic phytoplankton blooms. *New Phytologist*, 205(1), 172-181.

Young JN, Heureux AM, Sharwood RE, Rickaby RE, Morel FM. & Whitney SM. (2016). Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *Journal of Experimental Botany*, 67(11), 3445-3456.

Zettler LAA, Gómez F, Zettler E, Keenan BG, Amils R. & Sogin ML. (2002). Microbiology: eukaryotic diversity in Spain's River of Fire. *Nature*, 417(6885), 137.

Zhan Y, Marchand CH, Maes A, Mauries A, Sun Y, Dhaliwal JS, Uniacke J, Arragain S, Jiang H, Gold ND, Martin VJJ, Lemaire SD. & Zerges W. (2018). Pyrenoid functions revealed by proteomics in *Chlamydomonas reinhardtii*. *PloS one*, 13(2), e0185039.

Zhang Y, Launay H, Schramm A, Lebrun R. & Gontero B. (2018). Exploring intrinsically disordered proteins in *Chlamydomonas reinhardtii*. *Scientific reports*, 8(1), 6805.

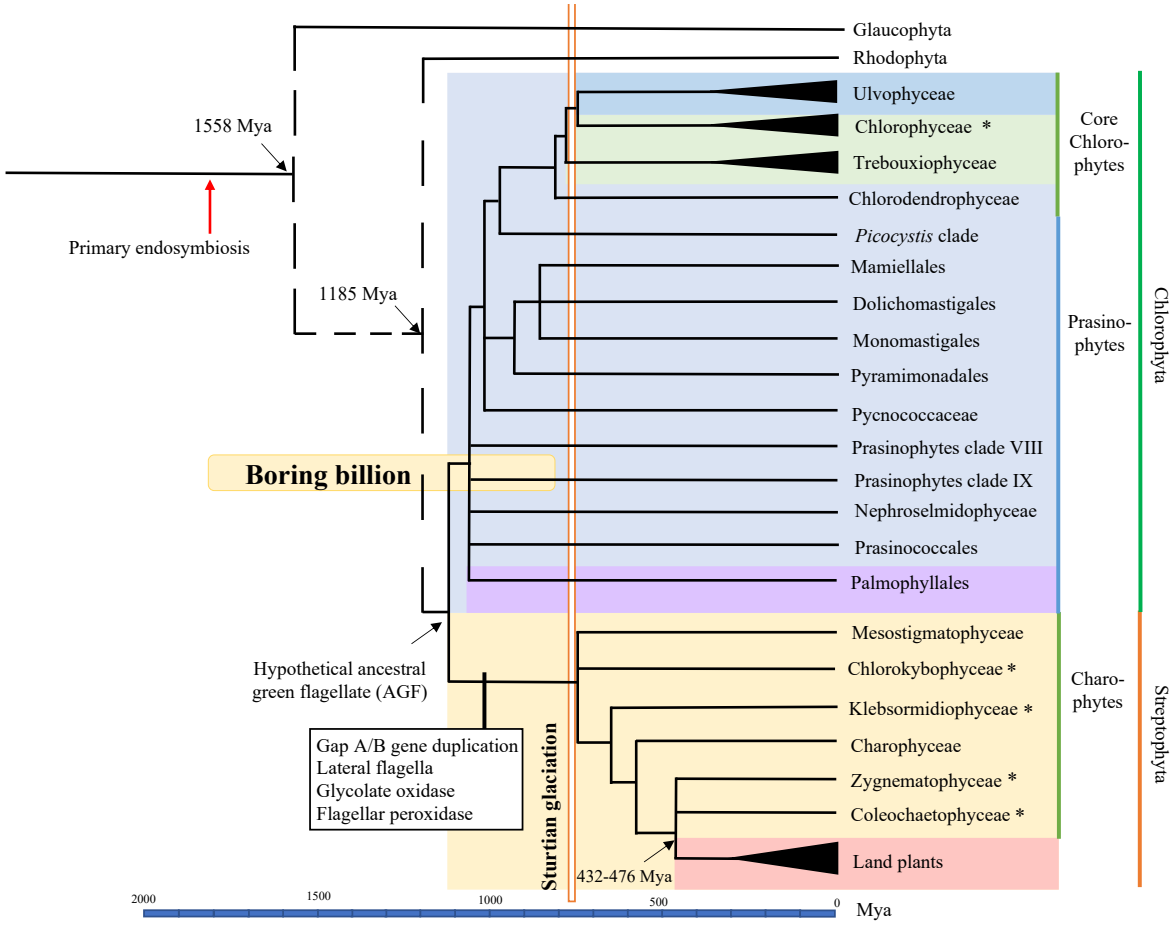
Zhao ZJ, Zhu H, Liu GX. & Hu ZU. (2016). *Rhizoclonium ramosum* sp. nov.(Cladophorales, Chlorophyta), a new fresh-water algal species from China. *Fottea, Olomouc*, 16(1), 12-21.

Zhu XG, Long SP. & Ort DR. (2010). Improving photosynthetic efficiency for greater yield. *Annual review of plant biology*, 61, 235-261.

Zones JM, Blaby IK, Merchant SS. & Umen JG. (2015). High-resolution profiling of a synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. *The Plant Cell*, 27(10), 2743-2769.

Appendices

Appendix 1 Evolutionary relationship of algae issued of the primary endosymbiosis and the major glaciation events which occurred during the diversification of the green algae lineages modified from Leliaert *et al*. (2012) and Becker (2013). Evolutionary hypotheses of the streptophyta (morphological and molecular characters) are indicated in the black box. Asterisks indicate lineages to which the sampling representatives belong. Primary symbiosis is indicated by a red arrow and dashed lines represent uncertain relationships.



Appendix 2 Medium recipe used to grow *Cosmarium subtumidum*, *Klebsormidium subtile, Chlorokybus atmophyticus* and *Coleochaete scutata.*

SAG Sammlung von Algenkulturen Göttingen
Culture Collection of Algae

Medium Recipe
Version 03.2007

1. Basal Medium (= ES "Erddekokt + Salze")

	stock solution [g/100 ml]	nutrient solution [ml]
KNO_3	1	20
K_2HPO_4	0.1	20
$MgSO_4 . 7H_2O$	0.1	20
soil extract *		30
micronutrient solution **		5
de-ionized or distilled water		905

* Preparation of soil extract: Fill a 6 litre flask one third with garden or leaf soil of medium, but not too great humus content which does not contain fertilizers or plant protective agents. Success of soil extract depends on selection of suitable soils. Those with high clay content are usually less satisfactory. Add de-ionized water until it stands 5 cm above the soil and sterilize by heating in a steamer for one hour twice in a 24 h interval. Separate the decanted extract from particles by centrifugation. Fill into small containers of stock solution each of a size appropriate to making a batch of media, autoclave for 20 min at 121°C and store in the refrigerator.

** Preparation of the micronutrient solution:

	stock solution [g/100 ml]	applied solution
$ZnSO_4 . 7H_2O$	0.1	1 ml
$MnSO_4 . 4H_2O$	0.1	2 ml
H_3BO_3	0.2	5 ml
$Co(NO_3)2 . 6H_2O$	0.02	5 ml
$Na_2 MoO_4 .2H_2O$	0.02	5 ml
$CuSO_4 . 5H_2O$	0.0005	1 ml
de-ionized or distilled water		981 ml
$FeSO_4 . 7H_2O$		0.7 g
EDTA (Titriplex III, Merck)		0.8 g

Autoclave the components separately in two solutions which are united after cooling.

Solution I: 881 ml distilled water + stock solutions of salts without $FeSO_4$ + 0.4 g EDTA
Solution II: 100 ml distilled water + 0.7 g $FeSO_4$ + 0.4 g EDTA
The following modifications of the Basal Medium proved suitable for many strains:

a) Basal Medium with Beef Extract (= ESFl. "Erddekokt + Salze + Fleisch"): Basal Medium (Medium 1) with 0.1 % beef extract.
b) Basal Medium with Peptone (= ESP "Erddekokt + Salze +Peptone"):
Basal Medium (Medium 1) with 0.1% proteose-peptone.
c) Basal Medium with 10 % Euglena Medium and Vitamins (= +V "Erddekokt + Salze + Euglena gracilis Medium + Vitamine"):
Basal Medium (Medium 1) plus 10 % Euglena Medium (medium 9) and the vitamins B_1 (5 x 10-4 g/l) and and B_{12} (5 x 10-6 g/l), added in sterile solution after autoclaving.
d) Acidified Basal Medium (= ES + H_2SO_4):
Basal Medium (Medium 1) plus 1% conc. H_2SO_4.

Appendix 3 Medium recipe used to grow *Onychonema laeve* and *Spirogyra sp.*

SAG Sammlung von Algenkulturen Göttingen
Culture Collection of Algae

Medium Recipe
Version 11.2008

7. Desmidiacean Medium (= MiEB$_{12}$ "Micrasterias + Erddekokt + Vitamin B$_{12}$")

	stock solution [g/100 ml]	nutrient solution [ml]
KNO_3	1	10
$(NH4)_2HPO_4$	0.2	5
$MgSO_4 . 7H_2O$	0.1	10
$CaSO_4$	saturated solution	10
soil extract *		20
peat extract (prepared like soil extract)		10
micronutrient solution **		5
de-ionized or distilled water		930

* Preparation of soil extract (as in medium 1):
Fill a 6 litre flask one third with garden or leaf soil of medium, but not too great humus content which does not contain fertilizers or plant protective agents. Success of soil extract depends on selection of suitable soils. Those with high clay content are usually less satisfactory. Add de-ionized water until it stands 5 cm above the soil and sterilize by heating in a steamer for one hour twice in a 24 h interval. Separate the decanted extract from particles by centrifugation. Fill into small containers of stock solution each of a size appropriate to making a batch of media, autoclave for 20 min at 121°C and store in the refrigerator.

** Preparation of the micronutrient solution (as in medium 1):

	stock solution [g/100 ml]	applied solution
$ZnSO_4 . 7H_2O$	0.1	1 ml
$MnSO_4 . 4H_2O$	0.1	2 ml
H_3BO_3	0.2	5 ml
$Co(NO_3)2 . 6H_2O$	0.02	5 ml
$Na_2 MoO_4 .2H_2O$	0.02	5 ml
$CuSO_4 . 5H_2O$	0.0005	1 ml
de-ionized or distilled water		981 ml
$FeSO_4 . 7H_2O$		0.7 g
EDTA (Titriplex III, Merck)		0.8 g

Add vitamin B_{12} (5×10^{-6} g/l) in sterile solution after cooling.

a) Desmidiacean Medium with Vitamin B$_1$ (= MiEB$_1$B$_{12}$)
Add vitamin B_1 (5×10^{-4} g/l) in sterile solution to medium 7 after cooling.

Appendix 4 Medium recipe used to grow the *Chlamydomonas* and *Chloromonas* strains .

Bold 3N Medium Recipe

Directions

Modification of Bold's recipe. General purpose freshwater medium used for xenic cultures, especially blue-greens and reds.

For 1 L Total
1. To approximately 850 mL of dH_2O, add each of the components in the order specified (except vitamins) while stirring continuously.
2. Bring the total volume to 1 L with dH_2O.
*For 1.5% agar medium add 15 g of agar into the flask; do not mix.
3. Cover and autoclave medium.
4. When cooled add Vitamin B_{12}.
*For agar medium add vitamin, mix, and dispense before agar solidifies.
5. Store at refrigerator temperature.

#	Component	Amount	Stock Solution Concentration	Final Concentration
1	$NaNO_3$ (Fisher BP360-500)	30 mL/L	10 g/400mL dH2O	8.82 mM
2	$CaCl_2 \cdot 2H_2O$ (Sigma C-3881)	10 mL/L	1 g/400mL dH2O	0.17 mM
3	$MgSO_4 \cdot 7H_2O$ (Sigma 230391)	10 mL/L	3 g/400mL dH2O	0.3 mM
4	K_2HPO_4 (Sigma P 3786)	10 mL/L	3 g/400mL dH2O	0.43 mM
5	KH_2PO_4 (Sigma P 0662)	10 mL/L	7 g/400mL dH2O	1.29 mM
6	NaCl (Fisher S271-500)	10 mL/L	1 g/400mL dH2O	0.43 mM
7	P-IV Metal Solution	6 mL/L		
8	Soilwater: GR+ Medium	40 mL/L		
9	Vitamin B_{12}	1 mL/L		

Appendix 5 Codes and parameters used to fit the Michaelis-Menten kinetics equation to the curves of external inorganic carbon versus photosynthetic rate.

```
jsv5D.,
;umol O2 sec-1 per unit chl=
;(water O2 conc(at 25C= 253uM)* volume*
;(fcond-icond))/(water cond*time*chl wt)

;col(4)= rate of change of voltage
;col(4)= (fcond-icond)/time
;col(7)= rate of change of o2 conc
;            wrt time for a conc of HCO3-

col(7)=col(4)*253/0.8

; this value is not normalized
; for chl amt

col(8)=col(7)*30
;col(8)=oxygen evolved in 30 sec

;col(9)=cumulative HCO3 released
cell(9,1)=0
If cell(8,1)> 0 then cell(9,2)=cell(8,1)
else cell(9,2)=cell(9,1)
end if
If cell(8,2)> 0 then cell(9,3)=cell(9,2)+cell(8,2)
else cell(9,3)=cell(9,2)
end if
If cell(8,3)> 0 then cell(9,4)=cell(9,3)+cell(8,3)
else cell(9,4)=cell(9,3)
end if
If cell(8,4)> 0 then cell(9,5)=cell(9,4)+cell(8,4)
else cell(9,5)=cell(9,4)
end if
If cell(8,5)> 0 then cell(9,6)=cell(9,5)+cell(8,5)
else cell(9,6)=cell(9,5)
end if
If cell(8,6)> 0 then cell(9,7)=cell(9,6)+cell(8,6)
else cell(9,7)=cell(9,6)
end if
If cell(8,7)> 0 then cell(9,8)=cell(9,7)+cell(8,7)
else cell(9,8)=cell(9,7)
end if
If cell(8,8)> 0 then cell(9,9)=cell(9,8)+cell(8,8)
else cell(9,9)=cell(9,8)
end if
If cell(8,9)> 0 then cell(9,10)=cell(9,9)+cell(8,9)
else cell(9,10)=cell(9,9)
end if

;col(6)=cumulative HCO3 added
;col(10)= unused HCO3 t each addition
col(10)=col(6)-col(9)

;cell(11,1) - chl in ug/mL
;col(12) normln of dP/dT wrt chl amt

col(12)=col(7)/cell(11,1)
col(13)=col(10)
; col 14-y and col15-x
col(14)=col(12)
col(15)=col(13)
```

```
Equation
f=(Vmax*x)/(Km+x)
fit f to y
''fit f to y with weight reciprocal_ysquare
''fit f to y with weight reciprocal_y

Initial parameters
Vmax = 25 ' {{previous: 96.5429}}
Km = 25 ' {{previous: 6.84101}}

Variables
x=col(13)
y=col(12)

Constraints
Km>0
```

Appendix 6 R codes used to statistically test $K_{0.5}$ values between streptophyte algae for cells grew under low CO_2 conditions and to statistically compare $K_{0.5}$ values obtained under low and high CO_2 conditions in the *Chloromonas* and *Chlamydomonas* strains.

R CODE FOR COMPARISON $K_{0.5}$ VALUES FOR CELLS GREW UNDER LOW CO_2 CONDITION BETWEEN STREPTOPHYTE ALGAE

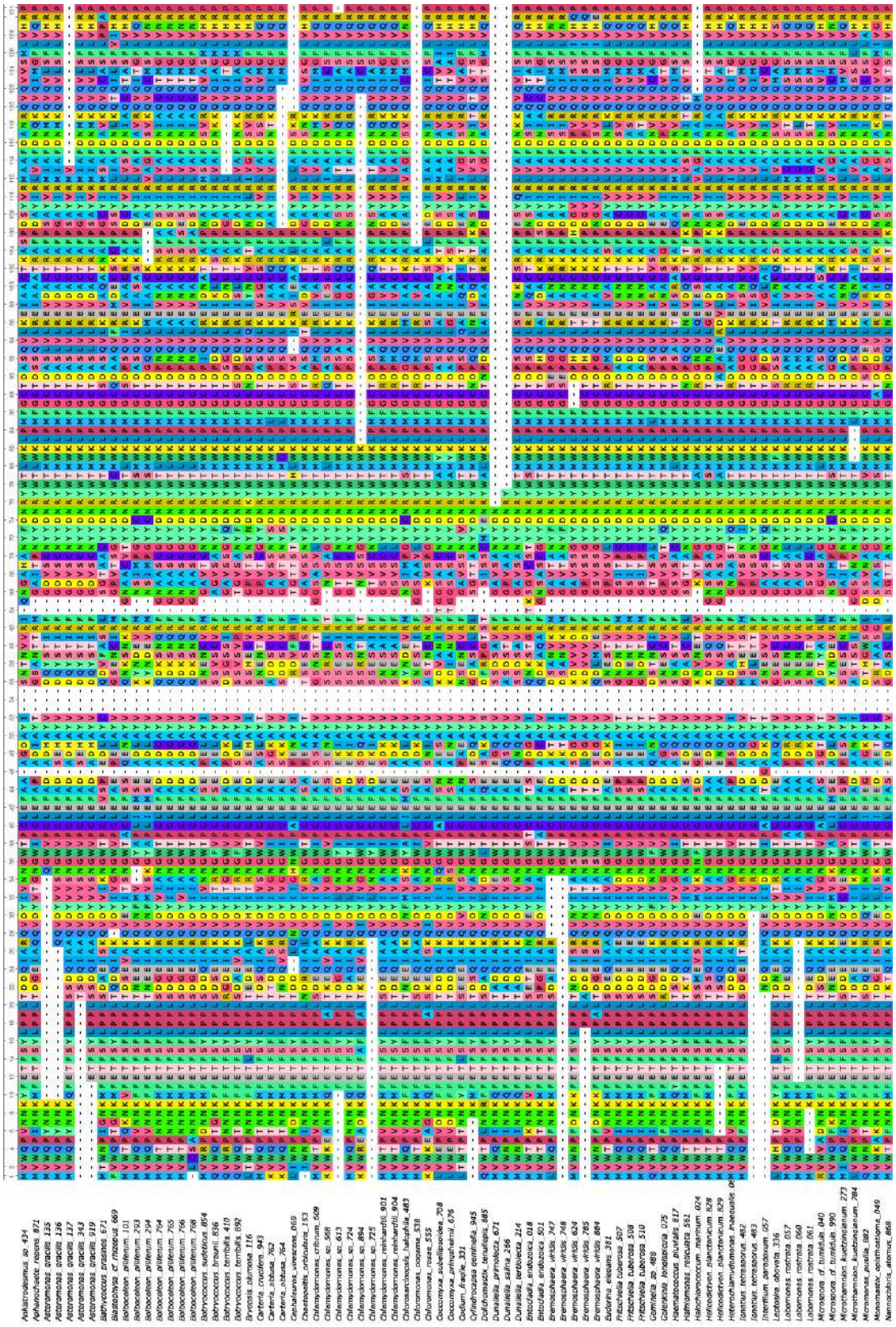
```
CA_LCK05<-c(rnorm(3, mean= 62))
KS_LCK05<-c(83.2, 45.2,49.7)
Cosm_LCK05<-c(rnorm(3, mean = 64))
OL_LCK05<-c(rnorm(3, mean= 62))
Spi_LCK05<-c(rnorm(3, mean = 48))
Coscu_LCK05<-c(rnorm(3, mean = 45))
Rein_LCK05<-c(53, 54, 55)

#1CAvsKS
CAvsKS=t.test(CA_LCK05,KS_LCK05, var.equal = TRUE)
CAvsKS
#CAvsSpi
CAvsSpi=t.test(CA_LCK05,Spi_LCK05, var.equal = TRUE)
CAvsSpi
#CAvsCoscu
CAvsCoscu=t.test(CA_LCK05,Coscu_LCK05, var.equal = TRUE)
CAvsCoscu
#CAvsRein
CAvsRein=t.test(CA_LCK05,Rein_LCK05, var.equal = TRUE)
CAvsRein

#2CosmvsKS
CosmvsKS=t.test(Cosm_LCK05,KS_LCK05, var.equal = TRUE)
CosmvsKS
#CosmvsSpi
CosmvsSpi=t.test(Cosm_LCK05,Spi_LCK05, var.equal = TRUE)
CosmvsSpi
#CosmvsCoscu
CosmvsCoscu=t.test(Cosm_LCK05,Coscu_LCK05, var.equal = TRUE)
CosmvsCoscu
#CosmvsRein
CosmvsRein=t.test(Cosm_LCK05,Rein_LCK05, var.equal = TRUE)
CosmvsRein

#3OLvsKS
OLvsKS=t.test(OL_LCK05,KS_LCK05, var.equal = TRUE)
OLvsKS
#OLvsSpi
OLvsSpi=t.test(OL_LCK05,Spi_LCK05, var.equal = TRUE)
OLvsSpi
#OLvsCoscu
OLvsCoscu=t.test(OL_LCK05,Coscu_LCK05, var.equal = TRUE)
OLvsCoscu
#OLvsRein
OLvsRein=t.test(OL_LCK05,Rein_LCK05, var.equal = TRUE)
OLvsRein
```

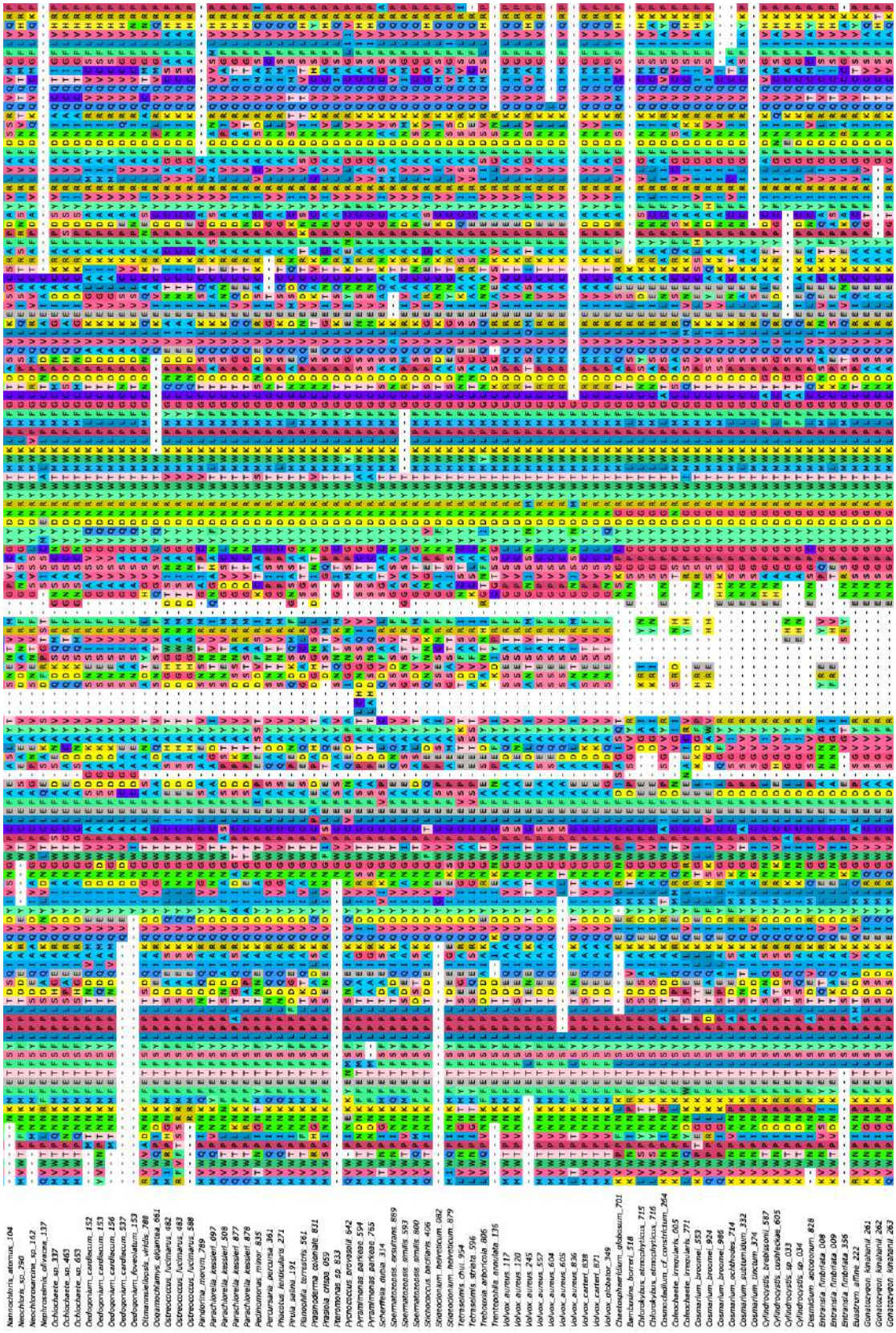
R CODE COMPARISON $K_{0.5}$ VALUES LOW VS HIGH CO_2 CONDITIONS

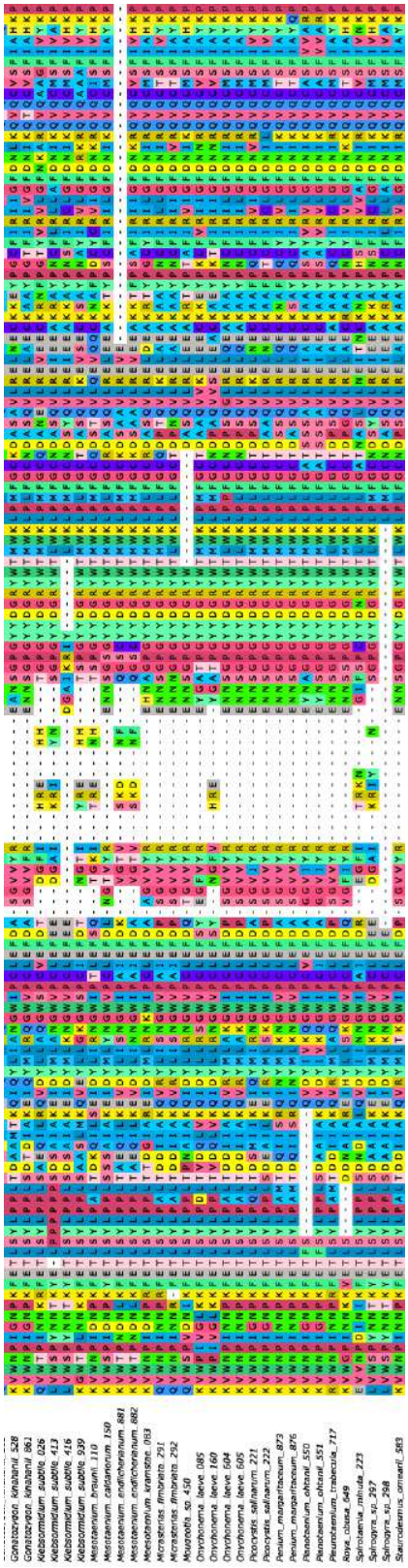
```
Aug_LC<-c(71.3, 35.34,60.34)
Muta_LC<-c(77.19, 81.13,93.35)
Cla_LC<-c(155.39, 226.66, 313.49)
Rosae_LC<-c(233.02, 338.19, 340.97)
Serbi_LC<-c(209.86, 222.30, 250.06, 264.86)
Rein_LC<-c(53, 54, 55)

Aug_HC<-c(444.45,525.58)
Muta_HC<-c(461.93, 1134.02, 600.52)
Cla_HC<-c(716.12, 598.51, 149.181)
Rosae_HC<-c(1049.66, 783.87, 743.65, 926.66)
Serbi_HC<-c(823, 821, 822)
Rein_HC<-c(148, 149, 150)

#NORMALITY TESTS
shapiro.test(Aug_LC)#p-value = 0.5766
```

```
shapiro.test(Muta_LC)#p-value = 0.4507
shapiro.test(Cla_LC)#p-value = 0.8916
shapiro.test(Rosae_LC)#p-value = 0.04314
shapiro.test(Serbi_LC)#p-value = 0.6996
shapiro.test(Rein_LC)#p-value =

shapiro.test(Aug_HC)#
shapiro.test(Muta_HC)#p-value = 0.3754
shapiro.test(Cla_HC)#p-value = 0.3778
shapiro.test(Rosae_HC)# p-value = 0.6112
shapiro.test(Serbi_HC)#p-value = 1
shapiro.test(Rein_HC)#p-value =

#T-TESTS
testAug=t.test(Aug_LC, Aug_HC, var.equal=TRUE)
testAug#p-value = 0.000999

testMuta=t.test(Muta_LC, Muta_HC, var.equal=TRUE)
testMuta#p-value = 0.03408

testCla=t.test(Cla_LC, Cla_HC, var.equal=TRUE)
testCla#p-value = 0.2251

testRosae=t.test(Rosae_LC, Rosae_HC, var.equal=TRUE)
testRosae#p-value = 0.001284

testSerbi=t.test(Serbi_LC, Serbi_HC, var.equal=TRUE)
testSerbi#p-value = 2.012e-07

testRein=t.test(Rein_LC, Rein_HC, var.equal=TRUE)
testRein#p-value=3.272e-08
```

Appendix 7 *RbcS* protein alignment used to reconstruct the phylogeny of *RbcS* in green algae




Namnochloris_atomus_104
Neochloris_sp_760
Neochloris_sp_162
Nephroselmis_olivacea_137
Ochlochaete_sp_337
Ochlochaete_sp_463
Ochlochaete_sp_653
Oedogonium_cardiacum_152
Oedogonium_cardiacum_153
Oedogonium_cardiacum_156
Oedogonium_cardiacum_537
Oedogonium_foveolatum_153
Ostreococcus_lucimarinus_388
Pandorina_morum_789
Parachlorella_kessleri_697
Parachlorella_kessleri_877
Parachlorella_kessleri_878
Pedinomonas_minor_835
Percursaria_percursa_361
Phacotus_lenticularis_271
Picocystis_salinarum_191
Planophila_terrestris_561
Prasinoderma_coloniale_831
Prasiola_crispa_059
Pteromonas_sp_833
Pychnococcus_provasoli_642
Pyramimonas_parkeae_594
Pyramimonas_parkeae_765
Scherffelia_dubia_314
Spermatozoopsis_exsultans_869
Spermatozoopsis_similis_593
Spermatozoopsis_similis_800
Stichococcus_bacillaris_406
Stigeoclonium_helveticum_082
Stigeoclonium_helveticum_879
Tetraselmis_chui_954
Tetraselmis_striata_596
Trebouxia_arboricola_806
Trentepohlia_annulata_136
Volvox_aureus_117
Volvox_aureus_120
Volvox_aureus_245
Volvox_aureus_557
Volvox_aureus_604
Volvox_aureus_605
Volvox_aureus_836
Volvox_carteri_838
Volvox_carteri_871
Volvox_globator_349
Cheetomorpha_sp_701
Bambusina_borreri_518
Chlorokybus_atmophyticus_715
Chlorokybus_atmophyticus_716
Cosmocladium_cf_constrictum_264
Cylindrocystis_brebissonii_587
Cylindrocystis_cushleckae_605
Cylindrocystis_sp_033
Cylindrocystis_sp_034
Desmidium_aeruginosum_828
Entransia_fimbriata_008
Entransia_fimbriata_009
Entransia_fimbriata_356
Euastrum_affine_222
Gonatozygon_kinahanii_261
Gonatozygon_kinahanii_262
Gonatozygon_kinahanii_263


Gonatozygon_kinahanii_528
Gonatozygon_kinahanii_861
Klebsormidium_subtile_026
Klebsormidium_subtile_413
Klebsormidium_subtile_416
Klebsormidium_subtile_599
Mesotaenium_braunii_110
Mesotaenium_caldariorum_150
Mesotaenium_endlicherianum_881
Mesotaenium_endlicherianum_882
Molodanium_krameshe_053
Micrasterias_fimbriata_251
Micrasterias_fimbriata_292
Mougeotia_sp_450
Onychonema_laeve_085
Onychonema_laeve_160
Onychonema_laeve_604
Onychonema_laeve_605
Picocystis_salinarum_221
Picocystis_salinarum_222
Penium_margaritaceum_873
Penium_margaritaceum_875
Planotaenium_ohtanii_550
Planotaenium_ohtanii_551
Pleurotaenium_trabecula_717
Roya_obtusa_649
Spirotaenia_minuta_223
Spirogyra_sp_297
Spirogyra_sp_298
Staurodesmus_omearii_583

Appendix 8 Pyrenoid diagnostic for all the species present in the phylogeny of *RbcS* and the associated references. Species without pyrenoid are highlighted in light grey.

Species name	presence or absence of the pyrenoid	Charophytes/ Chlorophytes	References
Bambusina borreri	Presence	Charophytes	Algaebase
Chaetosphaeridium globosum	Presence	Charophytes	Moestrup (1974)
Chlorokybus atmophyticus	Presence	Charophytes	SEM image
Cosmocladium cf. constrictum	Presence	Charophytes	Leghari (2001)
Coleochaete irregularis	Presence	Charophytes	Halder & Halder (2015)
Cosmarium broomei	Presence	Charophytes	Gerrath (1968)
Cosmarium ochthodes	Presence	Charophytes	Gerrath (1968)
Cosmarium subtumidum	Presence	Charophytes	SEM image
Cosmarium tinctum	Presence	Charophytes	Gerrath (2003)
Cylindrocystis brebissonii	Presence	Charophytes	Croasdale & Grönblad (1964)
Cylindrocystis cushleckae	Presence	Charophytes	York *et al*. (2002)
Cylindrocystis sp.	Presence	Charophytes	Prasanna *et al*. (2011)
Desmidium aptogonum	Presence	Charophytes	Algaebase
Entransia fimbriata	Presence	Charophytes	Cook (2004)
Euastrum affine	Presence	Charophytes	Aquino *et al*. (2017)
Gonatozygon kinahanii	Presence	Charophytes	West (1904)
Klebsormidium subtile	Presence	Charophytes	SEM image
Mesotaenium braunii	Presence	Charophytes	York *et al*. (2002)
Mesotaenium caldariorum	Presence	Charophytes	West (1904)
Mesotaenium endlicherianum	Presence	Charophytes	West (1904)
Mesotaenium kramstae	Presence	Charophytes	York *et al*. (2002)
Micrasterias fimbriata	Presence	Charophytes	Škaloud *et al*. (2011)
Mougeotia sp.	Presence	Charophytes	Wagner & Klein (1981)
Nucleotaenium eifelense	Presence	Charophytes	Gontcharov & Melkonian (2010)
Onychonema laeve	Presence	Charophytes	SEM image
Penium margaritaceum	Presence	Charophytes	Gerrath (1968)
Planotaenium ohtanii	Presence	Charophytes	Gontcharov & Melkonian (2010)
Pleurotaenium trabecula	Presence	Charophytes	Domozych *et al*., (2007)
Roya obtusa	Presence	Charophytes	Ljunggren & Oja (1961)
Spirotaenia minuta	Presence	Charophytes	Brook (1992)
Spirogyra sp.	Presence	Charophytes	SEM image
Staurodesmus omearii	Presence	Charophytes	Coesel & Van Geest (2016)
Microspora cf. tumidula	Absence	Chlorophytes	Stewart *et al*. (1973)
Ankistrodesmus sp.	Absence	Chlorophytes	Fawley *et al*. (2006)
Aphanochaete repens	Presence	Chlorophytes	Algaebase

Asteromonas gracilis	Presence	Chlorophytes	Peterfi & Manton (1968)
Blastophysa cf. rhizopus	Presence	Chlorophytes	Chappell *et al*. (1991)
Bolbocoleon piliferum	Presence	Chlorophytes	O'Kelly *et al*. (2004)
Botryococcus sudeticus	Presence	Chlorophytes	Komarek *et al*. (1983)
Botryococcus braunii	Presence	Chlorophytes	Wolf & Cox (1981)
Botryococcus terribilis	Presence	Chlorophytes	Queiroz Mendes *et al*. (2012)
Bryopsis plumosa	Presence	Chlorophytes	Ogawa (1988)
Carteria crucifera	Presence	Chlorophytes	Lembi & Lang (1965)
Carteria obtusa	Presence	Chlorophytes	York *et al*. (2002)
Cephaleuros virescens	Absence	Chlorophytes	Chapman *et al*. (1981)
Chaetopeltis orbicularis	Presence	Chlorophytes	Wujek & Thompson (1999)
Chlamydomonas cribrum	Presence	Chlorophytes	Ettl (1976)
Chlamydomonas sp.	Presence	Chlorophytes	Lembi & Lang (1965)
Chlamydomonas reinhardtii	Presence	Chlorophytes	Mackinder *et al*. (2016)
Chlorosarcinopsis halophila	Presence	Chlorophytes	Guillard *et al*. (1975)
Chloromonas oogama	Absence	Chlorophytes	Buckheim *et al*. (1997)
Chloromonas rosae	Absence	Chlorophytes	Ettl (1976)
Coccomyxa pringsheimii	Presence	Chlorophytes	Garbayo *et al*. (2012)
Coccomyxa subellipsoidea	Presence	Chlorophytes	Acton (1909)
Codium fragile	na	Chlorophytes	na
Cylindrocapsa geminella	Presence	Chlorophytes	Sluiman (1985)
Dunaliella primolecta	Presence	Chlorophytes	Öpik & Flynn (1989)
Dunaliella salina	Presence	Chlorophytes	Melkonian & Preisig (1986)
Dunaliella tertiolecta	Presence	Chlorophytes	Tsuzuki *et al*. (1986)
Entocladia endozoica	Presence	Chlorophytes	Goldberg *et al*. (1984)
Eremosphaera viridis	Presence	Chlorophytes	Holdsworth (1971)
Eudorina elegans	Presence	Chlorophytes	Gottlieb & Goldstein (1977)
Fritschiella tuberosa	Presence	Chlorophytes	Melkonian (1975)
Geminella sp.	Presence	Chlorophytes	York *et al*. (2002)
Golenkinia longispicula	Presence	Chlorophytes	York *et al*. (2002)
Haematococcus pluvialis	Presence	Chlorophytes	Wygash (1964)
Hafniomonas reticulata	Presence	Chlorophytes	Nakada *et al*. (2007)
Halochlorococcum marinum	Presence	Chlorophytes	Komarek *et al*. (1983)
Helicodictyon planctonicum	Presence	Chlorophytes	Webster-Smith *et al*. (1983)
Heterochlamydomonas inaequalis	Presence	Chlorophytes	Algaebase
Ignatius tetrasporus	Presence	Chlorophytes	Watanabe & Nakayama (2007)
Interfilum paradoxum	Presence	Chlorophytes	Mikhailyuk *et al*. (2008)
Leptosira obovata	Presence	Chlorophytes	Stewart *et al*. (1973)
Lobomonas rostrata	Presence	Chlorophytes	Hazen (1922)
Microthamnion kuetzingianum	Absence	Chlorophytes	Watson (1975)
Nannochloris atomus	Absence	Chlorophytes	Butcher (1952)

Neochloris sp.	Presence	Chlorophytes	Watanabe *et al*. (2000)
Neochlorosarcina sp.	Presence	Chlorophytes	Watanabe & Nakayama (2007)
Ochlochaete sp.	Presence	Chlorophytes	Burrows (1991)
Oedogonium cardiacum	Presence	Chlorophytes	Hoffman (1968)
Oedogonium foveolatum	Presence	Chlorophytes	Hoffman (1968)
Oltmannsiellopsis viridis	Presence	Chlorophytes	Chihara *et al*. (1986)
Oogamochlamys gigantea	Presence	Chlorophytes	Pröschold *et al*. (2001)
Pandorina morum	Presence	Chlorophytes	Fulton (1978)
Parachlorella kessleri	Presence	Chlorophytes	Juárez *et al*. (2011)
Pedinomonas minor	Presence	Chlorophytes	Moestrup (1991)
Percursaria percursa	Presence	Chlorophytes	Stewart *et al*. (1973)
Phacotus lenticularis	Presence	Chlorophytes	Hepperle & Krienitz (1996)
Pirula salina	Presence	Chlorophytes	Burrows (1991)
Planophila terrestris	Presence	Chlorophytes	Booton *et al*. (1998)
Prasiola crispa	Presence	Chlorophytes	Holzinger *et al*. (2006)
Pteromonas sp.	Presence	Chlorophytes	West (1904)
Scherffelia dubia	Absence	Chlorophytes	Melkonian & Preisig (1986)
Spermatozopsis exsultans	Absence	Chlorophytes	Melkonian *et al*. (1987)
Spermatozopsis similis	Absence	Chlorophytes	Preisig & Melkonian (1984)
Stichococcus bacillaris	Presence	Chlorophytes	Massalski *et al*. (2001)
Stigeoclonium helveticum	Presence	Chlorophytes	Stewart *et al*. (1973)
Tetraselmis chui	Presence	Chlorophytes	Chengwu & Hongjun (2018)
Tetraselmis striata	Presence	Chlorophytes	Hori *et al*. (1986)
Trebouxia arboricola	Presence	Chlorophytes	Friedl (1989)
Trentepohlia annulata	Absence	Chlorophytes	Algaebase
Volvox aureus	Presence	Chlorophytes	Darden (1966)
Volvox carteri	Presence	Chlorophytes	Kochert & Olsen (1970)
Volvox globator	Presence	Chlorophytes	Darden & William (1966)
Picocystis salinarum	Absence	*Incertae sedis*	Lewin *et al*. (2000)
Bathycoccus prasinos	Absence	Prasinophytes	Eikrem & Throndsen (1990)
Dolichomastix tenuilepis	Presence	Prasinophytes	Throndsen & Zingone (1997)
Micromonas pusilla	Presence	Prasinophytes	Van Baren *et al*. (2016)
Monomastix opisthostigma	Presence	Prasinophytes	Belcher & Swale (1961)
Nephroselmis olivacea	Presence	Prasinophytes	Suda *et al*. (2004)
Ostreococcus lucimarus	Absence	Prasinophytes	Meyer & Griffiths (2013)
Prasinococcus capsulatus	Presence	Prasinophytes	Guillou *et al*. (2004)
Prasinoderma coloniale	Presence	Prasinophytes	Hasegawa *et al*. (1996)
Pycnococcus provasolii	Presence	Prasinophytes	Guillard *et al*. (1991)
Pyramimonas parkeae	Presence	Prasinophytes	Pearson & Norris (1975)
coccoid-prasinophyte	na	Prasinophytes	na

Appendix 9 Script used with PyMOL (Schrödinger, 2010) to identify interacting residues in new modelled Rubisco.

```
from pymol import stored

def interfaceResidues(cmpx, cA='c. A', cB='c. B', cutoff=0.5, selName="interface"):
	"""
	interfaceResidues -- finds 'interface' residues between two chains in a complex.

	PARAMS
		cmpx
			The complex containing cA and cB

		cA
			The first chain in which we search for residues at an interface
			with cB

		cB
			The second chain in which we search for residues at an interface
			with cA

		cutoff
			The difference in area OVER which residues are considered
			interface residues.  Residues whose dASA from the complex to
			a single chain is greater than this cutoff are kept.  Zero
			keeps all residues.

		selName
			The name of the selection to return.

	RETURNS
		* A selection of interface residues is created and named
			depending on what you passed into selName
		* An array of values is returned where each value is:
			( modelName, residueNumber, dASA )

	NOTES
		If you have two chains that are not from the same PDB that you want
		to complex together, use the create command like:
			create myComplex, pdb1WithChainA or pdb2withChainX
		then pass myComplex to this script like:
			interfaceResidues myComlpex, c. A, c. X

		This script calculates the area of the complex as a whole.  Then,
		it separates the two chains that you pass in through the arguments
		cA and cB, alone.  Once it has this, it calculates the difference
		and any residues ABOVE the cutoff are called interface residues.

	AUTHOR:
		Jason Vertrees, 2009.
	"""
	# Save user's settings, before setting dot_solvent
	oldDS = cmd.get("dot_solvent")
	cmd.set("dot_solvent", 1)

	# set some string names for temporary objects/selections
	tempC, selName1 = "tempComplex", selName+"1"
	chA, chB = "chA", "chB"

	# operate on a new object & turn off the original
	cmd.create(tempC, cmpx)
	cmd.disable(cmpx)

	# remove cruft and inrrelevant chains
	cmd.remove(tempC + " and not (polymer and (%s or %s))" % (cA, cB))



# get the area of the complete complex
	cmd.get_area(tempC, load_b=1)
	# copy the areas from the loaded b to the q, field.
	cmd.alter(tempC, 'q=b')
```

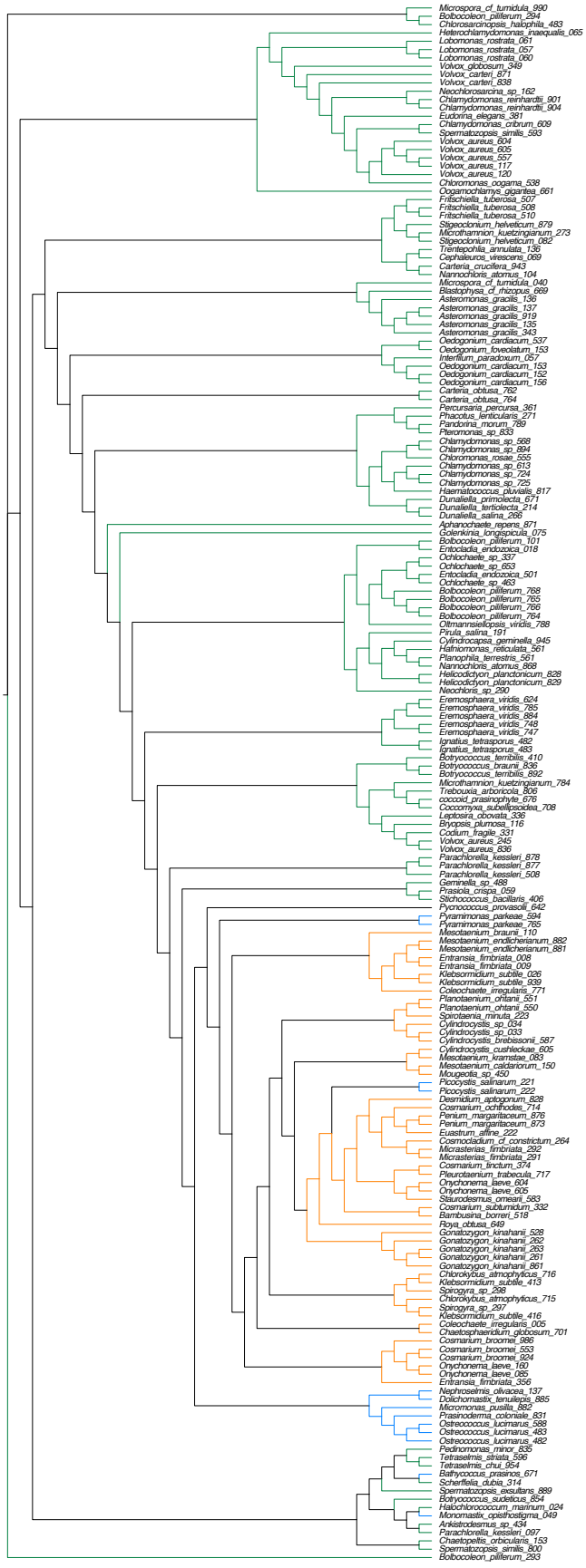
```
        # extract the two chains and calc. the new area
        # note: the q fields are copied to the new objects
        # chA and chB
        cmd.extract(chA, tempC + " and (" + cA + ")")
        cmd.extract(chB, tempC + " and (" + cB + ")")
        cmd.get_area(chA, load_b=1)
        cmd.get_area(chB, load_b=1)

        # update the chain-only objects w/the difference
        cmd.alter( "%s or %s" % (chA,chB), "b=b-q" )

        # The calculations are done.  Now, all we need to
        # do is to determine which residues are over the cutoff
        # and save them.
        stored.r, rVal, seen = [], [], []
        cmd.iterate('%s or %s' % (chA, chB), 'stored.r.append((model,resi,b))')

        cmd.enable(cmpx)
        cmd.select(selName1, None)
        for (model,resi,diff) in stored.r:
                key=resi+"-"+model
                if abs(diff)>=float(cutoff):
                        if key in seen: continue
                        else: seen.append(key)
                        rVal.append( (model,resi,diff) )
                        # expand the selection here; I chose to iterate over stored.r instead of
                        # creating one large selection b/c if there are too many residues PyMOL
                        # might crash on a very large selection.  This is pretty much guaranteed
                        # not to kill PyMOL; but, it might take a little longer to run.
                        cmd.select( selName1, selName1 + " or (%s and i. %s)" % (model,resi))

        # this is how you transfer a selection to another object.
        cmd.select(selName, cmpx + " in " + selName1)
        # clean up after ourselves
        cmd.delete(selName1)
        cmd.delete(chA)
        cmd.delete(chB)
        cmd.delete(tempC)
        # show the selection
        cmd.enable(selName)

        # reset users settings
        cmd.set("dot_solvent", oldDS)

        return rVal

cmd.extend("interfaceResidues", interfaceResidues)
```

Appendix 10 Scripts used with PyMOL (Schrödinger, 2010) to identify accessible residues in new modelled Rubisco.

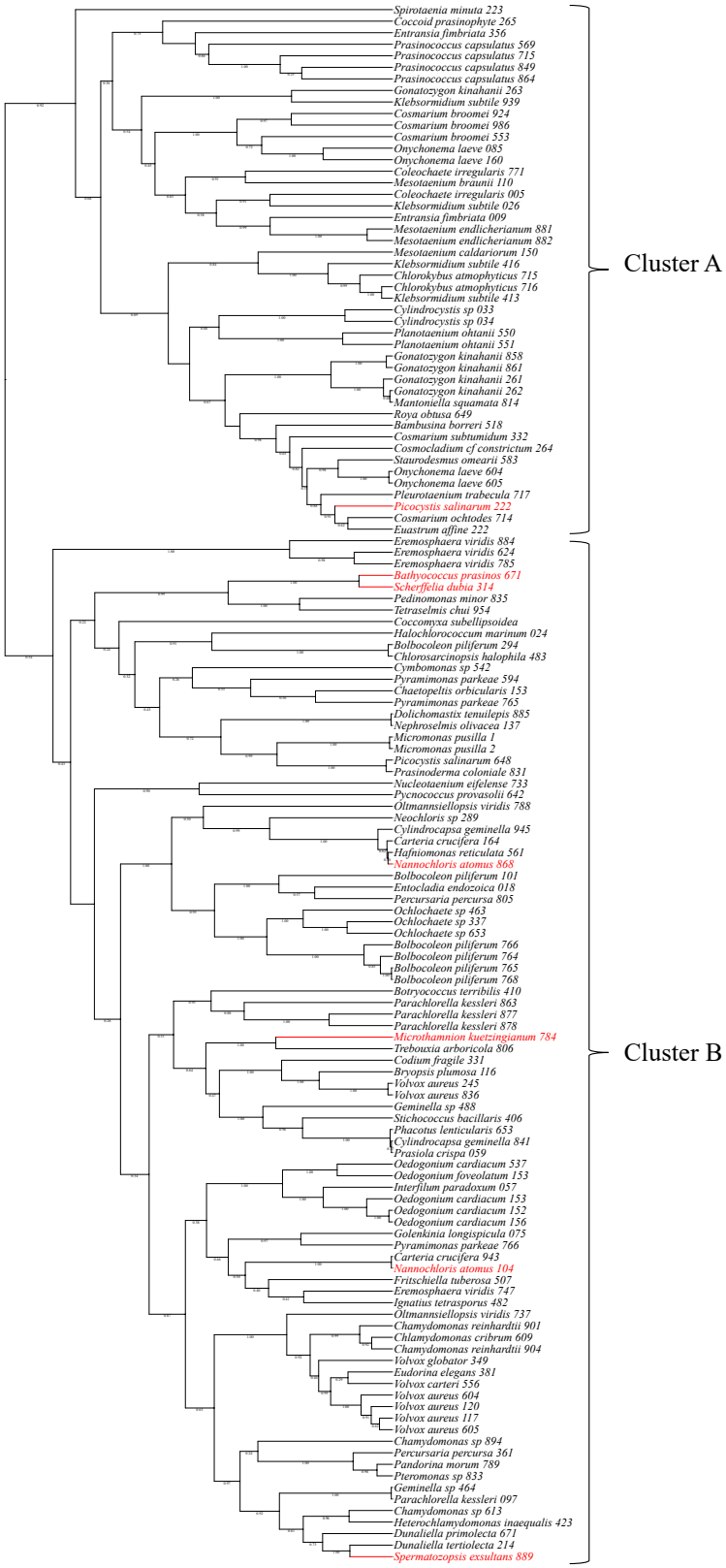
```
http://pymolwiki.org/index.php/FindSurfaceResidues
'''

from __future__ import print_function
from pymol import cmd

def findSurfaceAtoms(selection="all", cutoff=2.5, quiet=1):
    """
DESCRIPTION

    Finds those atoms on the surface of a protein
    that have at least 'cutoff' exposed A**2 surface area.

USAGE

    findSurfaceAtoms [ selection, [ cutoff ]]

SEE ALSO

    findSurfaceResidues
    """
    cutoff, quiet = float(cutoff), int(quiet)

    tmpObj = cmd.get_unused_name("_tmp")
    cmd.create(tmpObj, "(" + selection + ") and polymer", zoom=0)

    cmd.set("dot_solvent", 1, tmpObj)
    cmd.get_area(selection=tmpObj, load_b=1)

    # threshold on what one considers an "exposed" atom (in A**2):
    cmd.remove(tmpObj + " and b < " + str(cutoff))

    selName = cmd.get_unused_name("exposed_atm_")
    cmd.select(selName, "(" + selection + ") in " + tmpObj)

    cmd.delete(tmpObj)

    if not quiet:
        print("Exposed atoms are selected in: " + selName)

    return selName


def findSurfaceResidues(selection="all", cutoff=2.5, doShow=0, quiet=1):
    """
DESCRIPTION

    Finds those residues on the surface of a protein
    that have at least 'cutoff' exposed A**2 surface area.

USAGE

    findSurfaceResidues [ selection, [ cutoff, [ doShow ]]]

ARGUMENTS

    selection = string: object or selection in which to find exposed
    residues {default: all}

    cutoff = float: cutoff of what is exposed or not {default: 2.5 Ang**2}

RETURNS

    (list: (chain, resv ) )
        A Python list of residue numbers corresponding


        to those residues w/more exposure than the cutoff.

    """
    cutoff, doShow, quiet = float(cutoff), int(doShow), int(quiet)
```

```
    selName = findSurfaceAtoms(selection, cutoff, quiet)

    exposed = set()
    cmd.iterate(selName, "exposed.add((chain,resv))", space=locals())

    selNameRes = cmd.get_unused_name("exposed_res_")
    cmd.select(selNameRes, "byres " + selName)

    if not quiet:
        print("Exposed residues are selected in: " + selNameRes)

    if doShow:
        cmd.show_as("spheres", "(" + selection + ") and polymer")
        cmd.color("white", selection)
        cmd.color("yellow", selNameRes)
        cmd.color("red", selName)

    return sorted(exposed)

cmd.extend("findSurfaceAtoms", findSurfaceAtoms)
cmd.extend("findSurfaceResidues", findSurfaceResidues)
```

Appendix 11 Phylogeny of *RbcS* built with RAxML (Stamatakis, 2014) without the bA-bB loop. Streptophyte algae were labelled in orange, prasinophytes in blue and chlorophytes in green. The pyrenoid appears to have been lost 12 times across this phylogeny of green algae.



Appendix 12 DNA phylogeny of *RbcS* used for the PAML analysis and built with BEAST v2.3.1. As observed with the protein phylogeny, all the core chlorophytes are clustered together (Cluster B) with some of the prasinophytes. The top cluster includes all the streptophyte algae with the remaining prasinophytes. Species without pyrenoid are labelled in red. Foreground branches used for the branch model (codeml) in Paml are also labelled in red.



Appendix 13 Summary of the different parameters obtained after the ten different tests for RELAX selection (Wertheim *et al*., 2014).

All the streptophyte algae labelled	**K**	**p**	**LR**	
Test 1	1.25	0.065	3.4	intensification not significant
Test 2	1.26	1	-276.86	intensification not significant
Test 3	1.31	0.024	5.12	intensification significant
Test 4	1.36	1	-6.78	intensification not significant
Test 5	1.84	1	-135.5	intensification not significant
Test 6	1.87	1	-199.2	intensification not significant
Test 7	1.81	1	-171.75	intensification not significant
Test 8	1.29	1	-0.09	intensification not significant
Test 9	1.31	0.024	5.1	intensification significant
Test 10	1.82	1	-207.25	intensification not significant
				RELAXATION NOT SIGNIFICANT

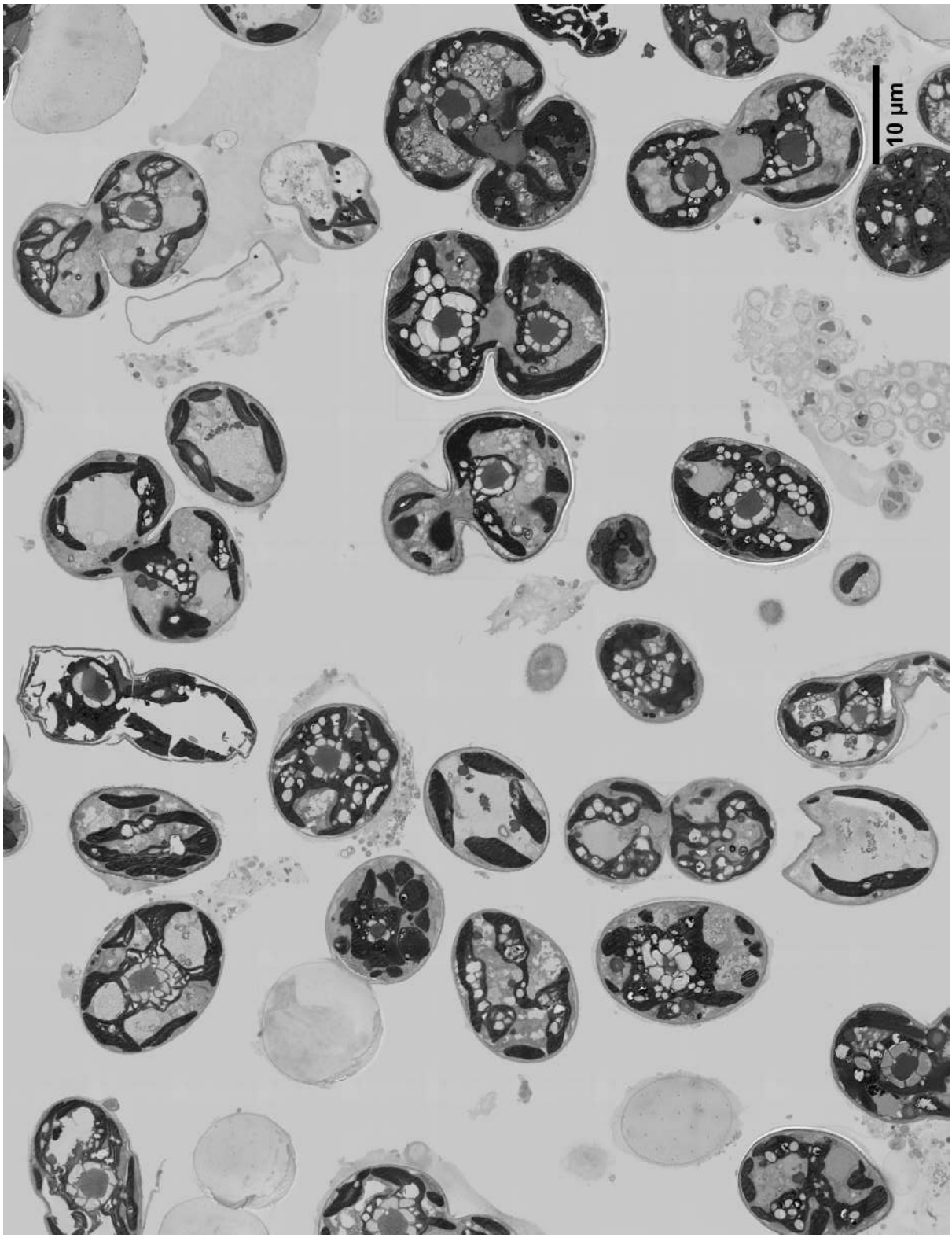
Basal branch of the streptophyte algae labelled	**K**	**p**	**LR**	
Test 1	1	1	-11.92	relaxation not significant
Test 2	0.28	0.002	9.3	relaxation significant
Test 3	1	1	-71.11	relaxation not significant
Test 4	1	1	-0.51	relaxation not significant
Test 5	1	1	0	relaxation not significant
Test 6	1	1	-7.3	relaxation not significant
Test 7	1	1	-0.01	relaxation not significant
Test 8	1	1	-3.31	relaxation not significant
Test 9	1	1	-7.11	relaxation not significant
Test 10	1	1	0	relaxation not significant
				RELAXATION NOT SIGNIFICANT

All the chlorophytes labelled	**K**	**p**	**LR**	
Test 1	0.82	0.022	5.27	relaxation significant
Test 2	1.02	1	-0.28	intensification not significant
Test 3	0.74	1	-109.2	relaxation not significant
Test 4	1.31	1	-60.28	intensification not significant
Test 5	1.09	1	-13.36	intensification not significant
Test 6	1.11	1	-49.27	intensification not significant
Test 7	1.07	1	-78.17	intensification not significant
Test 8	0.8	1	-5.57	relaxation not significant
Test 9	1.08	1	-5.34	intensification not significant
Test 10	0.79	0.001	11	relaxation significant
				RELAXATION NOT SIGNIFICANT

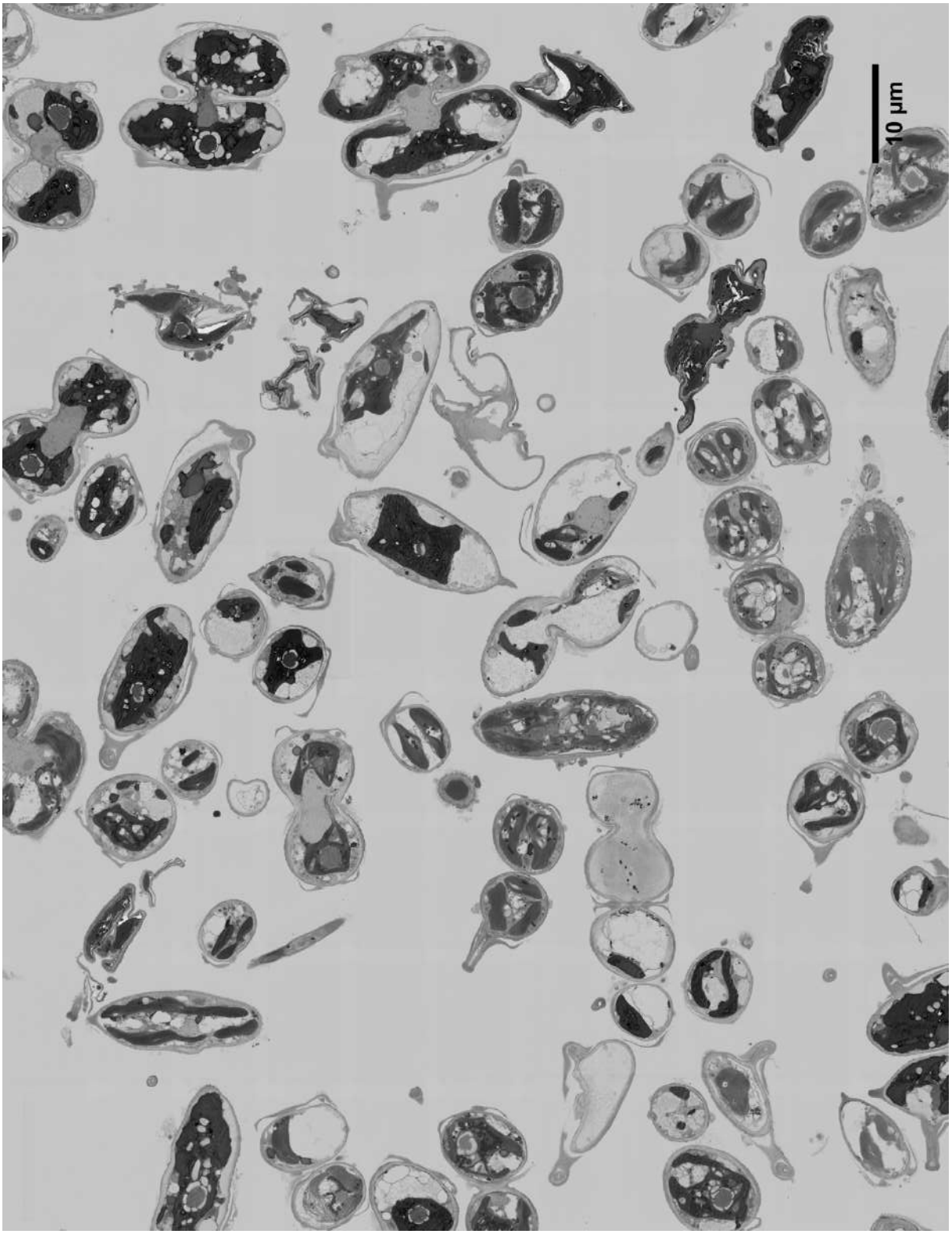
Basal branch of the chlorophyte labelled	K	p	LR	
Test 1	0.28	0.032	4.6	relaxation significant
Test 2				
Test 3	1	1	-0.09	relaxation not significant
Test 4	1	1	-17.45	relaxation not significant
Test 5	0.29	1	-22.03	relaxation not significant
Test 6	1	1	-30.39	relaxation not significant
Test 7	1	1	-2.38	relaxation not significant
Test 8	1	1	-0.03	relaxation not significant
Test 9	0.28	0.017	5.65	relaxation significant
Test 10	1	1	-0.03	relaxation not significant
				RELAXATION NOT SIGNIFICANT

Basal branches of the tree labelled	K	p	LR	
Test 1	0.28	0.009	6.88	relaxation significant
Test 2	0.35	0.003	8.77	relaxation significant
Test 3	1	1	-27.8	relaxation not significant
Test 4	1	1	-0.01	relaxation not significant
Test 5				
Test 6	1	1	-37.85	relaxation not significant
Test 7	1	1	-0.4	relaxation not significant
Test 8	1	1	-0.03	relaxation not significant
Test 9	1	1	-4.49	relaxation not significant
Test 10	1	1	-1.83	relaxation not significant
				RELAXATION NOT SIGNIFICANT

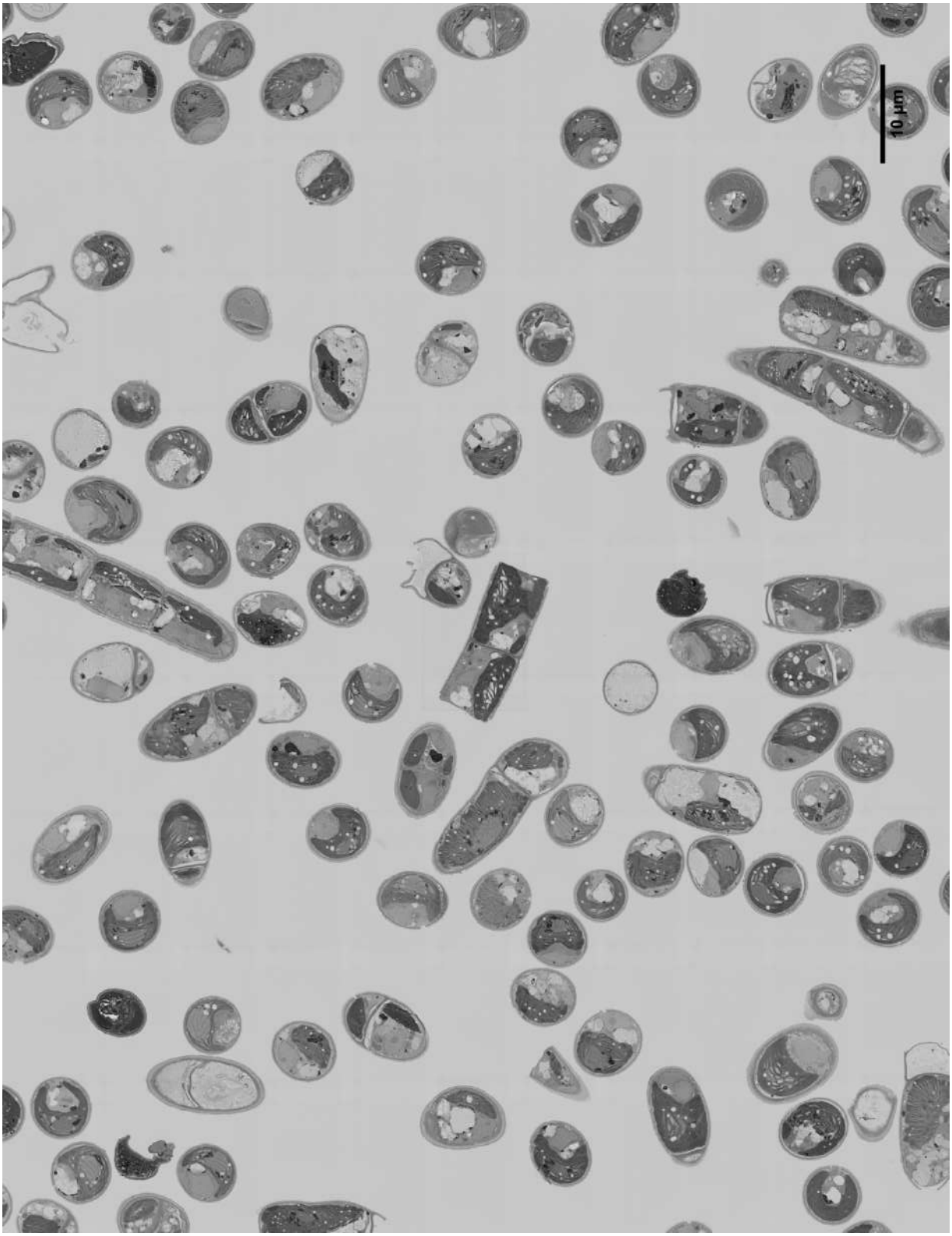
Appendix 14 Scanning Electronic Microscopy image of *Cosmarium subtumidum.*



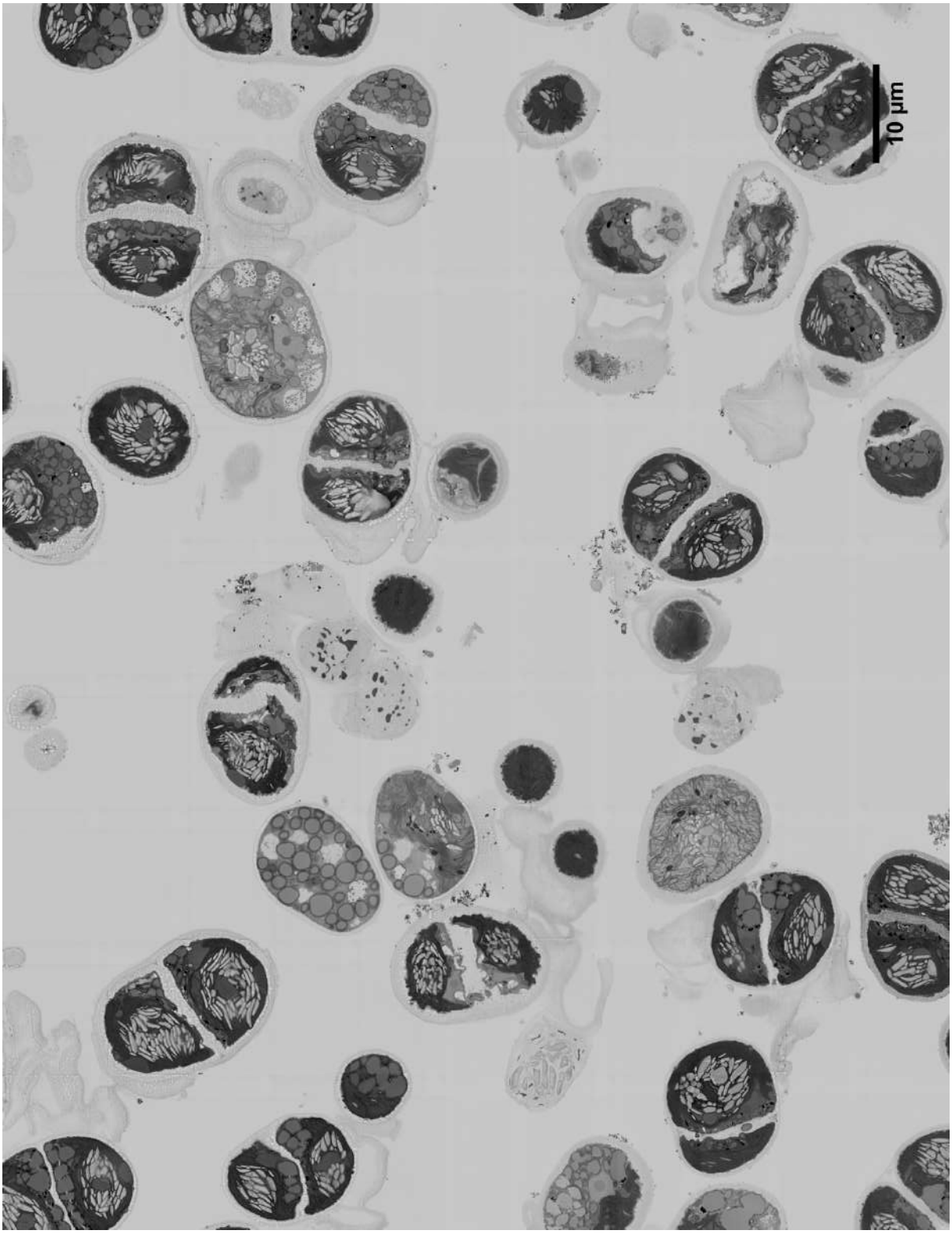
Appendix 15 Scanning Electronic Microscopy image of *Onychonema laeve.*



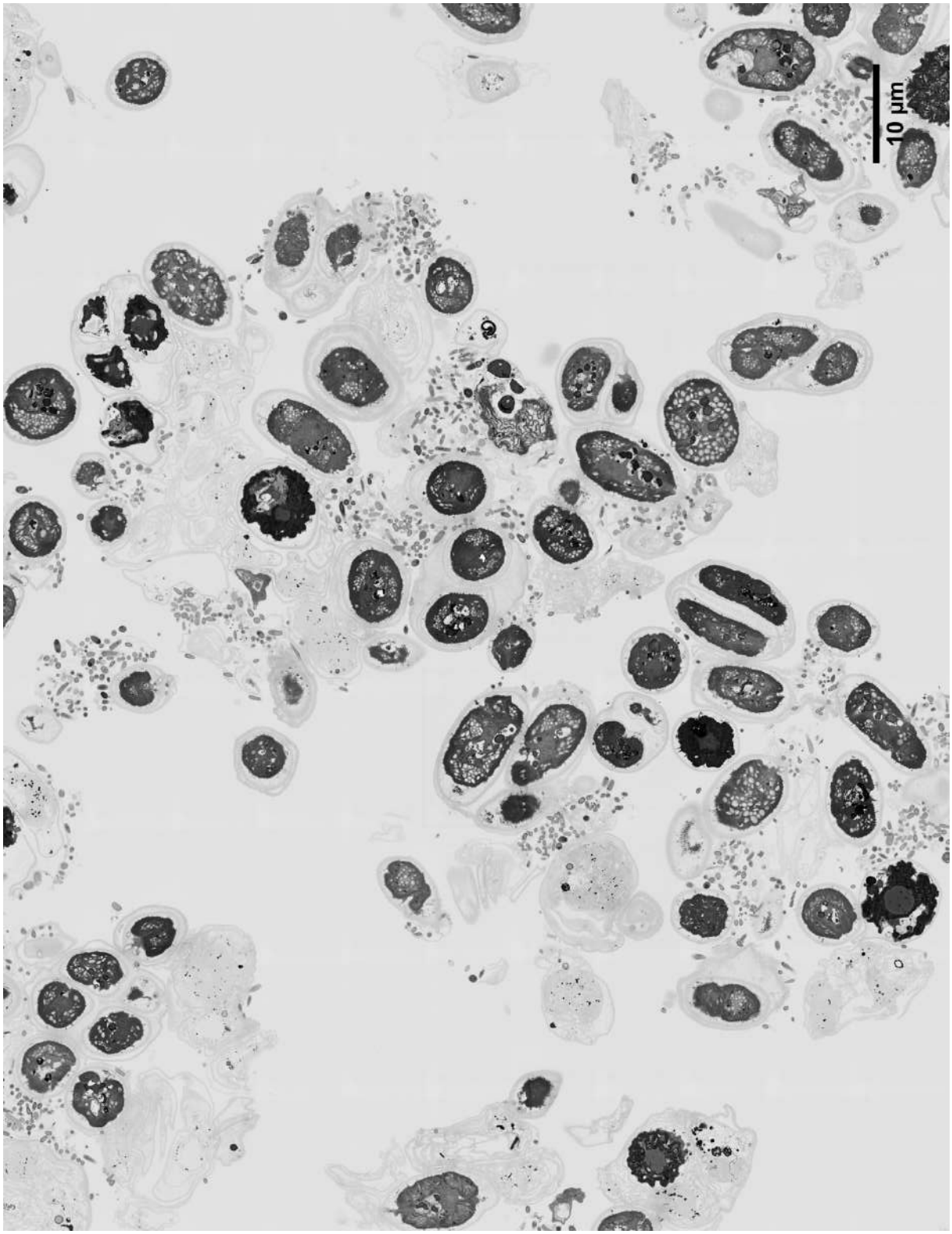
Appendix 16 Scanning Electronic Microscopy image of *Klebsormidium subtile.*



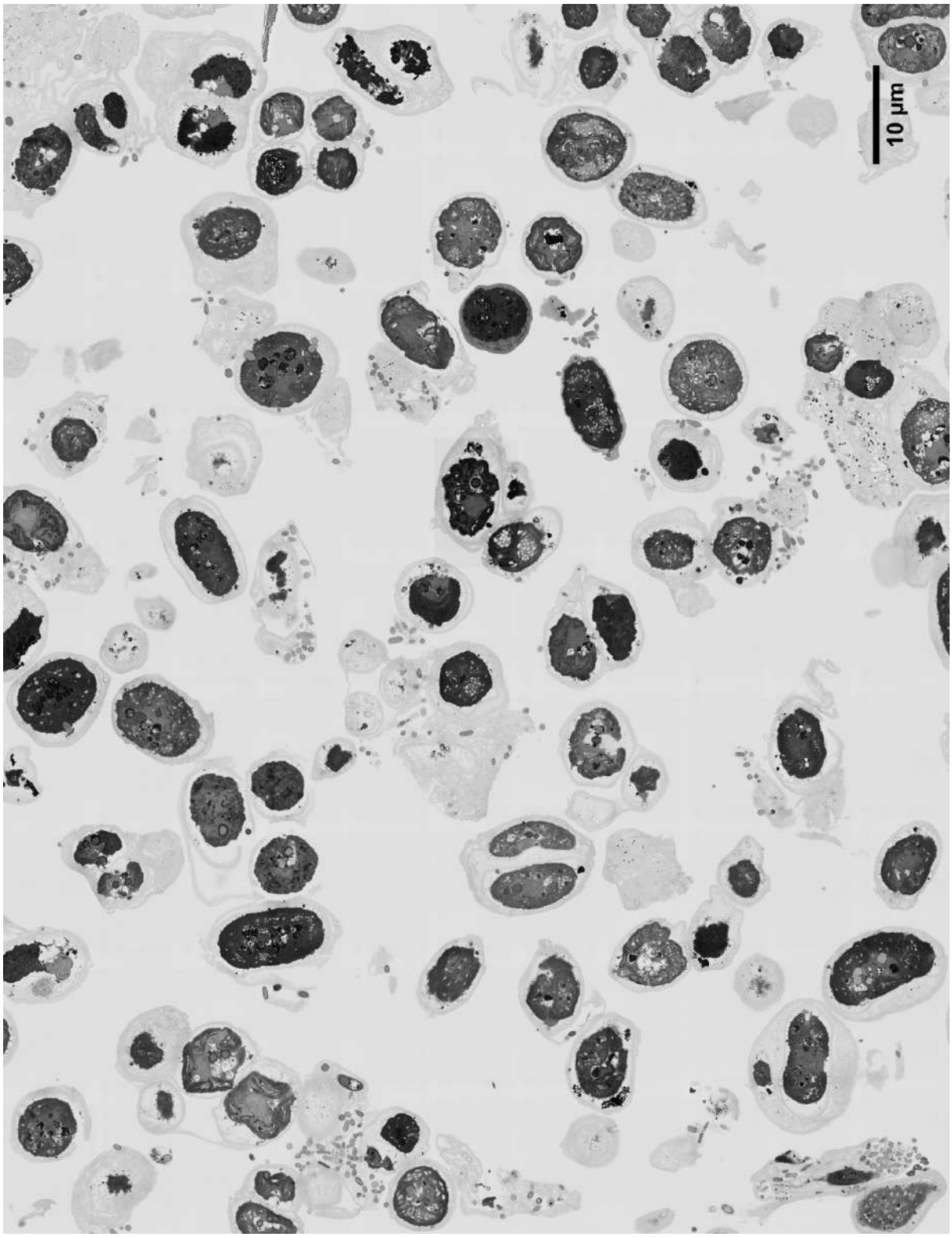
Appendix 17 Scanning Electronic Microscopy image of *Chlorokybus atmophyticus*.



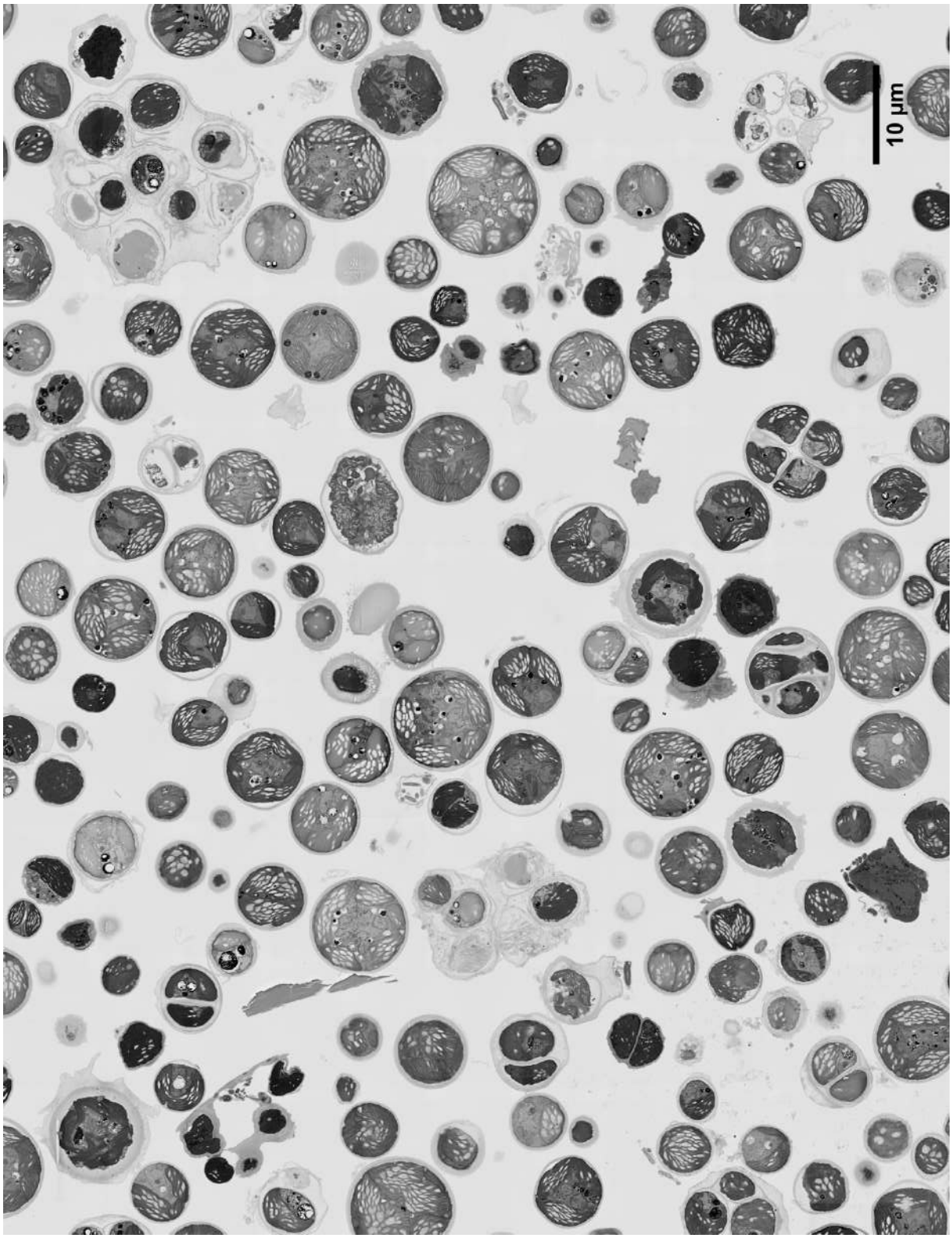
Appendix 18 Scanning Electron Microscopy of *Chlamydomonas mutabilis.*



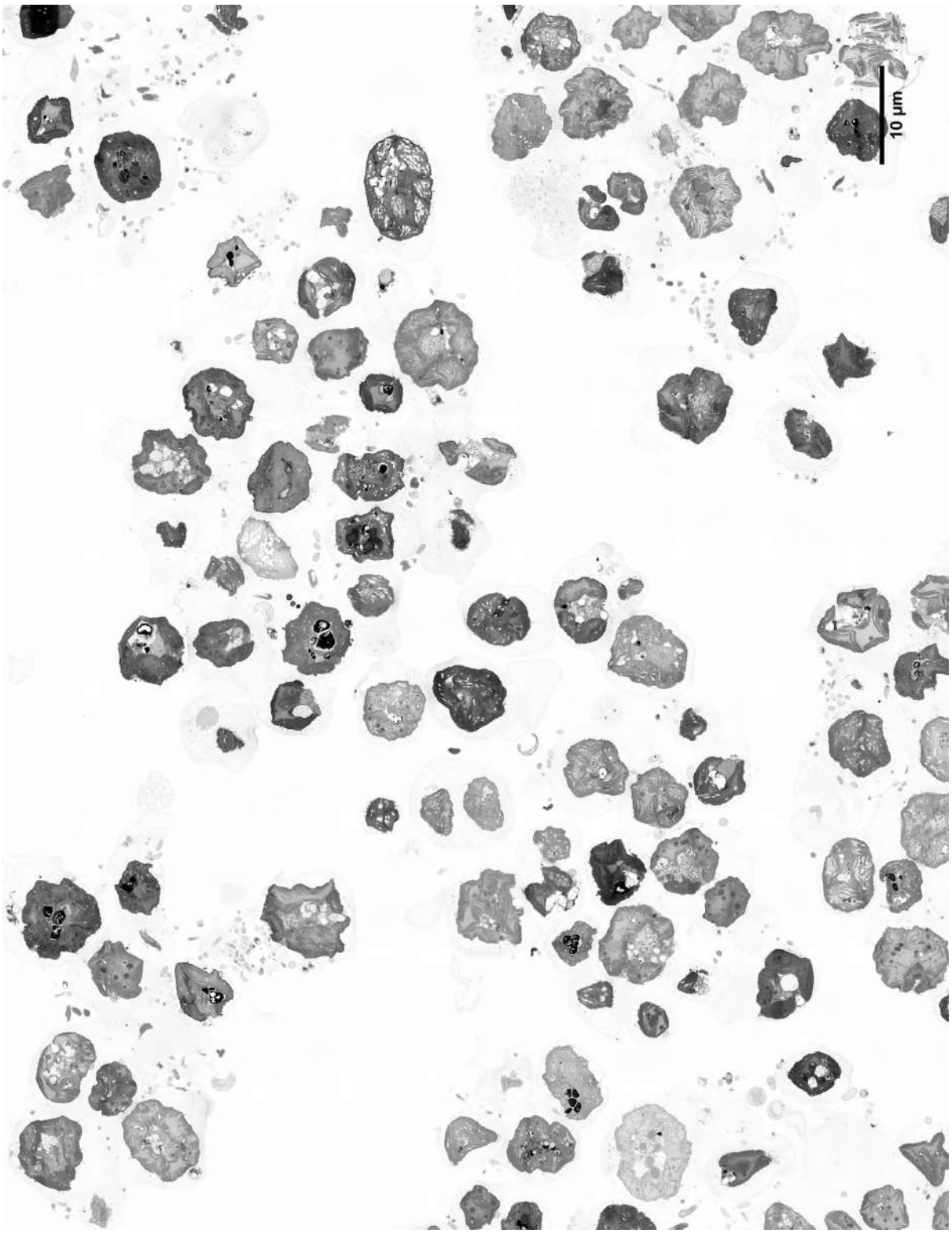
Appendix 19 Scanning Electron Microscopy of *Chloromonas rosae.*



Appendix 20 Scanning Electron Microscopy of *Chloromonas serbinowii.*



Appendix 21 Scanning Electron Microscopy of *Chloromonas clathrata.*



Appendix 22 List of the all the 44 CDS chloroplastic genes used for phylogeny reconstruction and extracted from the new five whole genome sequencing.

>Chlamydomonas_augustae-atpA
atggcaatgcgtactccagaagaattaagtaatttaattaaagacctaattgagcaatatactccggaagtaaaaatggttgattttggtattgtgtttcaagtaggggatggtattgctcgtgtgtatggctta
gaaaaagcaatgtctggagaacttttagaatttgaagatgggactcttggtatcgctttaaacttagaaactaataacgttggagctgttttacttggtgacggcattaaaataacagaaggaagtcgagttc
gttgtacaggaaaaatcgctgaaatcccagttggtgatggctatttaggtcgtgttgtagatgctttagcgcgtccagttgatggaaaaggcgcaatttcaacaaaagaaacaagagctattgagtctgttg
cacctggtattattacacgtcgttctgtatatgaaccttttagctactggactagtttctattgatgctatgattccagttggtcgtggtcaacgtgaattaattattggtgaccgtcaaacaggaaaaacatcaatt
gcaattgatacaatttaaaccaaaaaggaaaaggtgttgtttgtgtttatgttgcaattggacaaaaagcatcatctgttgcgcaagttttaaatactttaaaagaacgtggtgcattagattatactattattgtt
atggcaaacgctaatgaaccatcaactttacaatatttagctccttatacaggtgcaacattagctgaatactttatgtatacaggacgtgctacattagttatttatgatgatttatctaaacaagctcaagctta
ccgtgaaatgtcattattacttcgccgtccaccaggacgtgaagcatacccaggtgacgttttttatcttcactcacgtctttagaaagagcagctaaattaagtacagctttaggagaaggaagtatgaca
gcacttccaattattgaaacacaagaaggtgatgtatcagcttatattccaacaaatgttatttcaattacggatgggcaaatatttttagccggcgacttgtttaatgctggtattcgtcctgctattaatgttggt
atttctgtttctcgtgtaggttcagctgcgcaacctaaagcaatgaaacaagttgctggatctcttaaattatctttagcacaatttgccgaattagaagcttttttctcaatttgcttcagactagaccaagcaac
gcaaaatcaattagcacgtggtgtacgtttacgtgaaattcttaaacaagctcaatctgctccttttatctttagaagatcaagttttaacaatttattcaggtactcaaggttttcttgataaacttgaagtaaacca
agtacgagcttttgttattggtttacgaagttttatagcttcaaaatatcctaaatacagcgaaattataaaagctactttagcatttactcctgaagctgaaggaatcgtaaaacaaggaattaaagaatatttaa
gtgagtttcaagtgacaaactctccaaaagctagtagt

>Chlamydomonas_augustae-atpB
atgagcgattctgtagaaacaaaaaatattggacgtgttgtacaaattatcggtccagttttagacattgtttttttctaaaggtcaagtaccaaatatttacaatgctttagttattcgttctaaaaacgcagcagg
atcagaggtaagcgttacggtggaagttcaacaactgcttggtgataattgtgtacgcgcagtatcaatgaatccatcagatggtttaacacgcggtattgaagttatcgatacaggcaaaccattaactgtt
cctgtaggaaaagcaactttaggtcgtatttttaacgttcttggtgaaccggtagacaatttaggtcctgtaaaagctgatacagcattaccaattcaccgtacagcaccggctttttgttgacttagatacacgt
ctttcaatttttgaaacaggaataaaagtagttgatcttttagcgccatatcgtcgtggcggaaaaatcggtttatttggtggtgcgggtgtaggaaaaactgtattgattatggaattaatcaataatattgcaa
aagctcatggaggggtttcagtttttgctggagtaggtgaaagaacacgtgaaggaaatgaccttttatacagaaatgaaagaatctggtgtaattgtagaaaaaagtctttctgattcaaaagtagcacttgt
ttatggtcaaatgaatgaaccaccaggagctcgtatgcgtgttgcattaacagctttaacaatggccgaatattttagagatttcaataaacaagacgttcttttctttattgataatattttccgatttgttcaagc
cggtgctgaagtttcagctttattaggtcgtatgccttctgctgtagggtatcaaccaacgttagcaacagaaatgggtactttacaagaacgtattacatctactaaagaaggttctattacatcaattcaagc
agtatatgtacctgctgatgaccttactgacccggcaccagctgttacttttacgcatttagatgctactactgtattatcaagaggtttagcaagtaaaggtatttatccagctgttgatcctttagattcaacat
caactatgttacaaccttggattgttggtgaaaaacattatggcgtagcacaagacgttaaaaaaacgcttcaaagatacaaagagcttactgatattattgctatttttaggtttagacgaattatctgaagaag
atagattagttgttgctcgtgcacgcaaaatcgaaagatttcttagccaaccgtttttcgttgctgaagttttcacaggttcgcctggaaaatatgttagtttaacagagtcaatggaaggttttggtaaaatttttt
actggggaattagacagtttaccagaacaagcgttctatttagtaggaaatattaatgaagtaattgcaaaagcagctacatta
aaa

>Chlamydomonas_augustae-atpE
Atgagtttacaaatttcaatttttaacaccagaacgccctttttggaatggtcaagcagaagaaataattcttcctactgaaacaggagaaatgggtgttttaaaaaaccacgctccacttattaccggtttaga
tgttggagcaatgttgattcgttctaaagatgagtggaattcatatgcaattatgggaggatttgctttagttaaacaaaatcaagtaacaattttagcaaacgaagctgaatcagctgaaaatattgatccaga
agaagcaaaaactagttttgaaactgctaaagctaatttagaaaaagctgaaggtgtaaaagaaaaagtagaggcaaatttgcctacaaacgttcaaaagctcgttttcaaacagtaaaattacgcttacgt
t

>Chlamydomonas_augustae-atpF
Atggaatgcgtaacatttattaacggagttagtataggacatggtggttttggttttaatagcaatatttttgaaacaaatattattaacttagctgcagtggtgggcattgttgtaacatttgttggaagcaatctt
actgcattattagaagaccgtaaaaaaacgattttaaataatttacaagaagcaaatcaaagagctattgaagcgcaagaaaaattaagccaagcacgtgcacaattagaatcagcaaaaaaaaagctc
aagaaattcgtgaagaaggaatatttaagagcaactcaagaaattaataatttaaaatttaaccatgatattagattagcaagattacaagagtttaaacaagaaactcttcaacaagctgaacaaaaagctttt
aaacaagcttatatgtatttaattgcaaaaatctcaaaaagagttcgtgaaagattaaatactggattagattcaacttatcacgttgtagttaataacttttatgtatcgcgttttactgattttaaa

>Chlamydomonas_augustae-atpH
Atggcaattgatagtttaattggtgcagctagtgttttagctgctgggatcgctgttggtgttggaagtattggccctggcactgggcaaggaactgctgcaggatatgcagtagaaggtatcgctcgtca
accagaagctgaaggtaaaatccgtggtgctcttttacttttgcttttatggaatctctaacgatttatggtttagttgttgctttagctctactatttgctaacccttttttaggctaa

>Chlamydomonas_augustae-atpI
Atgattaatcctttattagagatcggtgaggtgtctgttggtcaacatttctattggaacatcgcaggttacagtgtccacggacaggttcttctaacttcatggttcgtattagcaataattggaatttttaagtttt
ttagggaactccaatttaaaggccacgcctgaagggacacaaaatttcactgaacttgttaccgaattcatccgtgatttatcaaaaacacaaattggtgaatcacatggtgactatctgcgctgggttccttt
tctagggactatttttttattcatttttgtatctaattggtctggagctttaattccttggaaacttattgaattaccaaatggtgaattagctgctccaacaaatgacattaatacaacagttgcattagcgcttttaac
ttcaatctcttatttttacgcgggtattaataaaaaaggcttaggctactttaatagatatgttcaaccggcggcttttttattaccaattaacgttttggaggattttaccaaaccgttatcattatctttccgtctttt
gggaatattcttgctgatgaactagttgttggagttcttgttgctcttgtacccccttataatcccaattcctcttatgcttcttggtcttttacgagtggaattcaagctctagtatttgcaacacttgcaggcgctta
tattggtgaagctcttgaagatcatcattaa

>Chlamydomonas_augustae-ccsA
atggcttctagcttaatttccgtattactaagcaactctgttacaaatttctttaatttgacttttttacacccttttagccgctttaggtggaactactggattagagaatgaaagcgctagcttttataacaacat
atatgcgtttgcttctaactttagttacactaatcctacgaatgcaggagcgctcccattaatagaaactaactttagttatggttctgctctaattattaacggagttttaagaaattgttcttttttagtctctttcttt
caatgatttttttattggattcaaactgcttttagttcgccacctgttgcttactatcggaacgatgtaatagtagccccttccttactccatccaataactgtatcttcaaattctaatagattagacgctccagtctttt
ttaatttctttttttggaaaaagtaaaacaagttccaaattttgccggcacagctttaatgattatttctaatttttattaatgtgtttattgatttttcgatggaaagaatctggacattttccattaagtaacttgtatga
atctttaatgttttttatcttggagttgtacgtttattcatcttgttttagagttaaaaacagaactaattagatctgatgcaatacttttaaagcaatgcttgggttctattacttctccaatagctttatttacaaatgcttt
tgcaaccttaatttgccagcagaaatgcaaaaagcagctcctttagttccggcacttcaatcaaattggttaatgatgcatgttactgttatgattatcagttactcagctttaattcttggctcattattatcgatt
gcttttctaatttaataagttcttttttcttttttcttttttgattttagttttttaaaacaaaaggggggcaaaaaatggcaatcaaacttctaccaaaactccctttaatactccatcctcttcttccaagcaactgacgt
cttacttctcttctcaaaaagcgacaaaaaaaaatcagccctcgttccttcaaaagcaaaaatgacttttagcgtggaatttgataatttaagttatcgagttttaggaataggatttccttttttaactataggcat
tttatcgggcgcggtctgggctaatgaagcatgggggtcgtattggagttgggatccaaaagaaacatgggcattattaacatggttaatttttgctatctatttacatgccagaataacaaaaggttggcaa
ggcaaaaaaccagctattatagcttcatttggttttttagctatttggatttgtttttaggagtaaatttaattggtgaaggtttacatagttatggttggttttcaaattcg

>Chlamydomonas_augustae-cemA
Atgtttcctttccaagcaagcagaagcttgcccatcctatccttgccagcatgttacacatgctactgctccgcaagaccgaagcacagaaagagcctcccagctactaactttagttacactaacaccttt
ggtggaaaggatggggttgctgcttacttttagttacactaatcctatccttcctttctctttatataaataaaaagggatgggaggatgctaagcatacaccaaaggtgtgctacgcagggaaaggagcagg
gggtccct

ctttgcttctaaatttgcttctaaatttatagaatatacgtaccaaaagtctgcttcgcataaaaaagaaaggggcttatttgcttgtaaagcttctattcatagattgttattaaaagtatatgcaaacaaacagc
aagctttcgcttacccaacctatttaaaagataaacaagtagtttctattacatatgaagaaatagggtttatttccaagatcatttagccgtgtatttgatcgatttattaaacaattgttttctgatgtcgaaaattta
gttattcaagaatatcgtttttatcggtatttatttttaacaacagttaaatgttttttattcttcttttgttccttttttaataaacgtagctagtaaaaattatttaatcagaccgctcacagaatattgctggaatcaaa
aacagggtgaaatttttttaaatcaagataaacaaaaacatgcttttgctgaattacaagagtttgaagaaaaagtttattttgaatcgctagtttcctttccaccaaaggtggatgggggggaaagggccccta
ccgaaaaccgaaattggtaaaactgataatggcaaacaaatttttacaagaaaaaacaattcaattggctattgattataataatgctagcattgaagctattagttgtatgtttgctgacttaattagcctcagc
gtgtttggatggttacttatttttatggaggtgcaaataagcgtaagccgtttatttatacttgaagtctttttgggcttgatgacactaaaaaatctcttattattttattagtcaccgatcttttagtaggatatcatt
ccgttggcccatggtcaattttatttgaatctttatttaatcgctatggcttgccaaatagccaagctgcgatttatttactcacaggaagcttaccagtcataatggatgttttttttaaatacttgattttcagacatt
taaaccgagcttctccagcttcagttgcgacttatcatgcaatgattgaataa

>Chlamydomonas_augustae-ChlB
atgaaattagcgtattggatgtatgccgggccagctcatattggaactttacgtgtggcaagttcatttaaaaacgtacatgctattatgcatgctccattaggtgatgattattttaatgtaatgcgttcgatgtt
agaaagagaaaagagattttacgccggttacagctagtattgtagatcgtcatgtgttagcacgtggatcacaagaaaaagttgtagaaaatattacacgaaaagacaaagaagaacaacctgatttaatc
gttcttacaccgacatgtacaagttcaattttacaagaggatttacagaattttgtagatcgagctgctatggaatctcaaagtgatgtccttttagctgacgtaaatcattatagagttaatgaactacaagcag
ctgacagaactctagaacaaattgttcggttttatatcgaaaaagcaagaaaacaaaataatttacagactaacaaaacagaaaaaccttctgccaatattataggtatttttacttttaggttttcacaatcagc
acgattgtcgcgaattaagacgtttattaaatgatttaggcattgaagttaatgaggttttaccagagggtggctcagttaataacttaaaaaatttacctaaagcgtggtttaattttatcccttatcgcgaagtg
ggtttaatgtctgcgatttatctagaaaaagaatttaatatgccttatgttgcaattacccctatgggtgttgttgacacagcagcatgtattcgagaaattggcactattgttactaaaatagatccatctttaaca
atgattaaaaatgagtatgacaattatattgataaacaaacacgttttgtatctcaagcggcgtggttttcacgttctattgattgtcagaatttaactggcaaaaaagctgtagtttttggagatgcgactcatgc
tgcttctatgactaaaatttagcacgtgaaatggggattcgtgttgtttgtgctggaacttattgtaaacatgatgcagattggtttagggagcaagtcgttggtttttgtgatcaagttttaattactgatgacca
cactttagttggagacattattgctcgaatggaaccagcagcgatttttggaacacaaatggaacgccacgttggcaaacgattagatattccttgcggtgtgatttccgctcctatacacattcaaaattttc
cacttggttatcgaccttttttaggttatgaaggaaccaatcaaatagctgatttagtttataactcttttgttcttggaatggaagatcacttgttagaaatttttcggtggacatgataataaagaagtcattacaa
aatcattgtccactgattcagatttaagctggtcatctgatggtttagctgaattaaataaaatacctggttttgttcgtggtaaagttaaacgtaatacagaaaaatttgcacgtcaaaaaaatattggagttata
acagttgaagttatgtttgctgctaaagaagcagctggtgcataa

>Chlamydomonas_augustae-ChlL
Atgaaattagcagtttatgggaaaggtggtattggtaaatcaacaacaagttgtaacatttcaattgctttagctagacgtggaaaaaaagttttacaaatcggttgtgatccaaaacatgatagcacttttac
ccttacaggttttttaattcccacaattatagatactcttcaacaaaaagattaccattatgaagatgtttggccagaagatgttatttatcaaggttatggaggtgttgacagtgtcgaagctggaggtccacc
agctggtgctggctgtggtggatatgttgttggtgaaacagtaaagttattaaaagaattaaatgcttttatgaatatgatattatttttatttgatgttttaggagatgttgtatgtggtggatttgcagcaccttaa
actatgcagattattgtattattgtaacggataatggttttgatgcattatttgcagcaaatcgtattgctgcttcagttcgcgaaaaagctcgtacacatccattaagattagcgggcttaatcggaaatcgaac
agctaaaagagatttaattgataaatatgttgaagcttgtccaatgccagttcttgaagttttacctctgattgaagaaatcagagtttcccgtgtcaaaggtaaaacattatttgaaatggttgaatctgaacca
gctcttcaatatatttgtgatttttatttaaacattgcagatcaactattaacagaaccagaaggggttgttccaagagaattgtctgaccgagaattatttagtttactatcagatttctatttaaatcctgtcgatcc
gacaaaaaaaacagaatctgttgaaacattagacttttattagtt

>Chlamydomonas_augustae-ChlN
atgaagcttaatttaaatattagtaaagttttatatcaaataaaaatgtcaagcaatatgcttatgtctaataaatcgactactttaaacaattcaggctccggtagttttagtgaatatgctacgcatgcaaacac
agataatatagcagtaactacgcagactaatgatgactcgttaacatttgaatgtgaaacaggtaattaccatacttttgtcctattagttgtgtcgcatggctttatcaaaaaattgaagatagctttttttggt
aattggaactaagacatgtggttacttttacaaaacgctttaggagttatgatttttgctgagccgcgctatgctatggcagagttagaggaaagtgatatttcagcgcaattaaatgattataaagaattaaa
aagactgtgtttacaaatataaacaagatagaaatccaagtgtggtagtttggatagggacttgtactacagaaattatcaaaatggatttagagggtatggctcctcgtttagaaacagaaattggaatacca
attgtcgttgctcgagctaatggactagattatgcgtttacacaaggcgaagacactgttctagctgcaatggctcaaagatgtcctcaagctttaagttctgaaactaaaaaaaatataacaatagagcta
ctttacattctactgcttctttctcttctgcggagctcatccccattccagcgtttgggaaccaagcaacaaaagtggatggggaaggggcatgcgcggcaagtactgctacgcagggggatggcgctgc
gcattcccaaataagaagggaaaaaagcactttaaaaaatactaaacaactcgttttatttggttctttacctagcacagtaacttctcaattatcaatggaatttaaaaaaccaaggcattttagtttcaggttgg
ttaccttcgcaacgttataatgatttacccgttttaggtgaagatgtttatgtttgtggtgttaatcctttcctaagtcgaactgcaacgactttgatgcgccgtcgaaaatgtaaactgataggagctccatttcc
aatcgggcctgatggaactcgtgcttgggtagaaaaaatttgtagtatttttgatattgttccacaaggtttagaagaacgtgaatcaaaaatatggcaaagtttagaagattatttacaattagttcgtggtaa
atctgtatttttttatgggagattcgctactagaagttcccttagcgagatttttaacacgttgtggaatgactgtttttgagattgggattccctatatggataaacgattccaagcagcggaacttgctttattaga
aaaaacttgtaaagatatgaaagttccaatgccgcgaattgtcgaaaaaccagataattattatcaacttcaacgaattcgtgaacttttacctgatcttgtaattactggaatggcccatgcaaatcctttaga
agctcgaggcattagtactaagtggtcagtagaatttacatttgctcaaattcatggttttgccaatgctcgagatattttagaactagtaacgagaccattaagaagaaatcaaagtctagaagctttaggttg
gatgaaattagtaaaataa

>Chlamydomonas_augustae-ClpP
atgccaattggagtaccaagaattatttattgttggggtgaagaacttcctccacaatggactgatatctataatttttatttttcgtcgacgaatggtttttttaatgcaatatttagatgatgaacttttgtaatcaaat
ttgtggattattaattaatattcatatggaagatcgatcaaaagaacttgaaaaaaaagaaatcgaaaaaagtggcatttttaaaagtgctaataaaagtggtaaagaatctagtccaaattcttctgtcggag
gaccgtcgcaagctttatttgatgctaaaaaagacaaacaatttttagagctaaaaaataacgaagcaataaatgatggtaaaaagagaaaaaaagataaatcaatggaagatcttttaagttcttatgcaa
gctctgcttatgaagaagatttagctattgatgaaaattatactttagaacagtatactttacaaaaaataacattagaatggttaaattggaatgctcaatttttgactattctgatgaaccgtatttattttatttag
ctgaaattttatctaaagattttacaaaagatgatacaggtcaattattttataatttttacttacgctaaaacaaatcaaaaataataaatttggtttcgtgcttccatccccatctaccgttggtgtgcgaagcaagg
ggatggagccgcaggaggggggaaagattcgaagccactcctcgtttgcttgctccgtttcctgcagaaactagggaccccaagcatcctacagatggacaacaaacgcagaaggcattagaaaattta
gcagcggagcttataaaaccgacagttgacactacaaaatttagcagcataaaaaagaatgcctttcaagagtactggccgtttcgatcaattttaaatttagcaaatttaatttctaaagatcagaattttact
ataaataatttattacctcaacttagtaataacaaagttattaatcaagctctgctggttaagaatggttctttgcatcctttataatagcaagcaaagcttacccaatctcttacagaccaaaggaggaggggag
ttattaaatcaatttaatcaattaaaccaaaaacctggtttaactaaattacttggagttttggcttcaaaagaatttgtcaataaaaatacccaactctcaaaagcaacaatctgtagcaacaaaagttttaggac
aaagaaatttaaaaaaagagtatgctttagataataatttttaataattatggctccttatccagtccagcaatcttatccgtagctccgaccccatcctcgctacgcataggtaggatgcaaggacaggaagg
ggatacacctttggtgaatgttctatatagtgcacaaaaacacaacaaaaaagaaaagagctttaaattatttagatttgattctgctggatctattccctttggtggtgctacgaagcatccatcatattccaac
cctacctcttccccctaataagcgcagcagtccttcctgtgcgaagcaaggataggaacgggatggggatggatggagcgaagcaaataaataaaataaaaaaaacttcagtgctgctacacgaaa
acaaacaaaaagagctcttcaagaagaagaatctaaaaaagtctttgttatcattaattcttttggtggctcagttgggaatggtattactgttcatgatgctttacagtttataaaagctggctctcttactttagg
tcttggagttgcggccagtgcagcatctttagtattagctggcggaactatttctgagcgttatgtaactgaagggtgtcacactatgatacatcagcccgaaggaggccttaatggtcaagcttctgatatc
tggattgatagtcaagaaattatgaaaattcgattagatgtagcagaaatttattcattatctgcacatcgacctcgtcataaaatcttacgtgatttagatcgagatttctatttaactgcaacagaaactatcca
ttatggtttagcagatgaaatcgctactaacgatgtaatgcaccagataatcgaaatgactagcaaagtttgggattatcatgatagcaaacaacaacgtttacttgaaacacgaaatagcgcaaatagtgtt
gatacccaaactcaaagctaa

>Chlamydomonas_augustae-petA
Atgtctaaaaattcggcatacacctttggtgtaagcgtaacgagaaatttaatagcaaatgctgtattagatgggagttttgcaaaaacatttttttgttcaatctttttttggaatttttctttttctaaattgtgcagtt
tctgactacgctatcgcttatccagtttttgcgcaacaaaactatgaaaatccgcgtgaagcgaatggacgtattgtttgtgcaaattgtcacttagctcaaaaaccagctgaaattgaaattcctcaagctgt
attaccagatactgtttttgaagctgtagttcaaattccttatgataaacaagtacaacaagtacaagctaatggtaaaaaaggcgatttaaatgtaggtatggttttaattctgcctgaaggatttgaacttgcg
ccgccagatcgcattccagaagaattaaaaaaaaaaagttggaaatatttattatcaaccgtatagtcctgaaaaaaaaaatattttagtagcaggtccgcttccaggtaaaaaatatagtgacatgataattc
caattctttctccagatcctgctaaaaattcaaaagttgcgtatttaaaatatcctatttatttaggtgctaatagaggtcgtggtcaagtttatccagatggttctaaatcaaataatactgtttataacgcatctgtt
agtggtaaaattgtatcaat

taatccgggtgaaaaaaaaggatctattaacattaccattgaaaaagcaaatggagaaaatgttattgaaaaaattcctgcaggacctgaagttttagtacaggaaggacaattaatccaagctgatcaac
cgttaacaaataatccaaatattggtggttttggacaaaaagatgtcgaaatcgttcttcaaaattcatcgagaataacgggattattaattttctttgtatttattcttatagcacaagttctttttagttcttaaaaag
aaacaatttgaaaaagttcaattagctgaaatgaatttt

>Chlamydomonas_augustae-petB
Atgagtaaagtttacgactggtttgaagaacgtttagaaattcaatcaattgctgatgatattagtagtaaatatgttccaccgcatgttaatatttttttattgtttaggtggcattacttttacgtgtttttttagttcaa
gttgcaacaggctttgctatgacttttttattatagaccaaccgtagcagaagcttttgcatctgttcaatatataatgactgatgttaatttttggttggttaattcgatctattcatcgttggtcagcaagcatgatgg
ttttaatgatgattttacatgttttccgtgtttacttaactggtggttttaaaaaaccacgtgagttaacatgggtaagtggggttattatggctgtatgtacggtttctttttggggtaacaggttattcattaccttgg
gaccaaattggatattgggctgttaaaattgttacaggtgttcctgaagctattcctgttattggtgggcctttagtagaactattaagaggtggtgttggcgttggccaaagtacattaacgcgttttatagtct
tcatacttttgtattaccattagcaactgctgcttttatgttaatgcatttcttaatgatcagaaaacaaggtatttcaggaccactataa

>Chlamydomonas_augustae-petD
Atgtcagtaacaaaaaaacctgatctaactgatccagttttaaaagctaaattagcaaaaggtatgggtcacaatagttatggtgaaccagcatggcctaatgatttacttttatatttttccagttgttatttttgg
gacgtttgcttgtgttgttggtttagctgtattagatcctgctgctataggtgaacctgctaatccatttgcaacaccgttagaaatttttaccagaatggtatttttatccagtattccaacttttacgaacagttccc
aataaactattaggtgttttgttaatggccgcagttcctttcggtctaggtttagtaccgtttatcgaaaatatcaataaatttcaaaacccatatcgtagaccaattgcaacaatttttatttctttttagggactgttg
ctgctatttggttaggtattggtgcaacatttccaattgatatctcattaacgtttggtttattt

>Chlamydomonas_augustae-psaC
Atggctcatttagttaaaatatacgatacttgtattggttgtacacaatgtgttcgtgcatgtcctttagacgtattagaaatggtaccttgggatggttgtaaagcgaatcaaatggcttcagctcctcgtactg
aagactgtgtaggttgtaaacgttgtgaaacagcttgtcctactgacttttaagtgttagagtctatttaggctcagaaagtacacgaagtatgggattagcttattaa

>Chlamydomonas_augustae-psaJ
atgaaagattttacaacttatttatcaactgctcctgtggtaagcttagcatggttagtattaactgcagtattattaattggttttaacaaagtattccctgatcctcttgtttttacttttttaa

>Chlamydomonas_augustae-psbA
Atgactgctgttatcaacaaacaagatgtttcaactagcttatgggctcgcttctgcgaatgggttacgtctacagaaaaccgtatctacgtaggatggtttggtacaattatgttcccaactctattaactgca
acttcagtttatattattgctttcgttgcggctcctcctgtagatatcgatggtatccgtgaaccagtttctggttcattactttacggaaacaacatcatttctggtgctgtaatccctactagtaacgctatcggtc
ttcacttctaccctatctgggaagcagcttctcttgacgaatggttatacaacggtggtccttaccaaatgattgtttgccacttcttcatcggtatctgtgcgtacatgggtagagaatgggagttatcttaccg
tttaggcatgagaccatggatctgtgtagcttattcagcaccagttgctgcagctactgctgtattcatcatctatcctatcggacaaggtagtttctcagacggtatgccttttaggaatttcaggaacgttcaa
cttcatgatcgttttccaagcagagcacaatatccttatgcacccttccacatgttaggcgttgctggtgtatttggtggttctttattctcagctatgcatggttcgttagtaacttcatctttaatccgtgaaact
actgaaaacgaatcagctaacgctggttacaaattcggtcaagaagaagaaacttacaacatcgttgctgctcatggatattttggtcgtttaatcttccaatatgcttcattcaacaactcacgttcattacact
tcttcttagctgcatggccagttatcggaatttggttcactgctttaggtatttctacaatggcgtttaaccttaacggtttcaacttcaaccaatctgtagtagactctcaaggtcgtgtattaaacacttgggctg
acatcatcaaccgtgctaacttaggtatggaagttatgcacgaacgcaacgcgcacaatttccctcttgacttagcttctgttgaagctcctagcgtaaacgct

>Chlamydomonas_augustae-psbB
atgggattaccatggtatcgtgtacatactgtagttattaatgacccgggccgcttaatttcagtgcatttaatgcatacagctcttgtagctggttgggcaggttcgatgacacttttttgaaattgctgtttttgat
ccatcagatccagttttaaatcctatgtggcgtcaaggaatgtttgtacttccttttatgacacgtttagggattacacaatcttggggtggttggacaattagtggagaaacagcatcaaatccaggcatttgg
agctatgaaggtgtagctgcttctcatatcgttctttcagggcttctttttttagcttctgtttggcattgggtgtattgggaccttgaattattccgtgatccaagaacaggtaaaacagcattagatttaccaaaa
atttttggaattcatttattcttatcaggtcttctttgttttggttttggtgcttttcatgtaactggtgtttttggtcctggtatttgggtttctgatccttatggattaacaggaagcgtacaacccgtttctccttcatgg
ggtgcggatggctttgacccttataaccctggaggtattgctgcacatcacattgctgcaggaattttaggtgtacttgctggtcttttttcacctttgtgttcgtccatcaattcgtctatatttcgggctttcaatg
ggtagtattgaaactgtattatcaagtagtattgcagctgttttctgggcagctttttgttgttgcaggcacaatgtggtatggttcagcagctactccaattgaacttttacggtccaacacgttatcaatgggattt
aggctttttccaacaagaaatccaaaaacgtgttcaaactagtttaagtgaaggttcttctttacctgctgcttgggctaaaattcctgaaaaattagctttctatgattacatcggtaataacccagcaaaaggt
ggtcttttccgtacaggagctatgaacagtggcgatggtattgctgttggatggttaggtcatgctgtatttaaagatcaagatggacgtgacttatttgttcgtcgtatgccgacattctttgaaacattccctg
ttattttaattgacaaagatggggttgttcgtgctgacgttcctttccgtaaagcagaatctaaatatagtattgaacaagtaggtgtttctgttactttctatggtggtgaattagatggtttaacttttaatgatcca
gctactgttaaaaaaatatgctcgtaaagctcaattaggtgaaatttttgaatttgatcgttcaacttttacaatctgatggggttttccgtagtagcccacgcggatggtttacttttggtcatgtttgttttgctttatt
attcttctttggtcatatttggcatggtgcaagaactattttcagagatgtttttgccggaattgatgacgatttaaatgaacaattagaatttggtaaatacaaaaaacttggagatacttcgtctcttcgtgaagc
tttt

>Chlamydomonas_augustae-psbC
Atgacaaatattttaggttatcaccttattatgttaggtcttggtgcttggcttcttgtaattaaagcaatgtttttaggaggggtttatgatacttgggctccaggtggtggagacgttagaattatttcaaatcca
acaacaaatgctgctattatttttggttatttattaaaatcaccatttggtggtgacggttggatcgtaagcgttgataatatggaagatattattggcggacacatttggattggtacacttttactattcggtgga
atttggcacatttatacaacgccgtggccatgggctagacgtgcttttatttggtctggtgaagcttatttatcatatagtttagctgctatatcattaatgggttttatttcttgttgtatgtcttggttcaataatact
gcttatccaagcgagttttacgggccgacgggtcctgaagcatctcaatcacaaacattcacattcttagtaagagaccaacgtttaggtgctaatgttgcatctgctcaaggtccaacaggtttaggaaaa
tatttaatgcgttctcctactggagaaattatctttggtggtgaaacaatgcgtttctgggatttccgtggtccttggttagaacctttaagaggtccgaacggtttagatttaaataaactaaaaaatgatattca
accatggcaagaacgtcgtgcagctgaatacatgacgcatgctcctcttggatcattaaattcagtaggtggagttgcaacagaaattaacgctgtaaacttttgttagtccacgttcatggcttgccacttca
catttctgtctaggtttcttcttctttgttggtcatttatggcattctggccgtgctcgtgctgctgcagctggttttgaaaaaggaattgacagatttgatgaacctgttctttctatgaaaccattagattaa

>Chlamydomonas_augustae-psbD
Atgactatagcgattggaacatatcaagaaaaacgtacttggtttgatgatgctgatgactggcttcgccaagaccgttttgtttttattggatggtcaggtcttttattattaccttgtgcatatttagcacttggt
ggttggtttacaggaactacttttgtaacttcttggtatacacacggattagcaacatcatacttagaaggttgtaatttcttaacagctgctgtgtctacacctgctaacagtatgggtcactctttactatttgttt
ggggtccagaagctcaaggagattttactcgttggtgccaacttggtggtttatggacttttgtagctctacatggcgcatttggtttaattggatttatgctacgtcagtttgaaattgctcgttctgtaaatcta
cgcccatataacgctattgctttctctgcgccaattgcagtatttgtttctgttttcctaatttatccattaggtcaatcaggttggttttttgcgccaagcttcggtgtagctgcaattttccgattcatattattctttc
aaggattccataactggacattaaaccctttccatatgatgggtgttgcaggtgtccttggtgctgctcttctttgtgctattcatggtgctactgtagaaaatactcttttttgaagatggtgatggtgcaaatact
ttccgtgcatttaaccctactcaatcagaagaaacatactctatggttactgccaatagattttggtctcaaatctttggtgtagcattctctaacaaacgttggttgcacttttttcatgcttttcgtgccagtgactg
gtttatggatgagcgctatcggtgttgttggtcttgctttcaacttaagagcatacgatttcgtatctcaagaaatcagagctgctgaagatcctgaatttgaaacattctacactaaaaatattcttttaaacgag
ggtattcgtgcttggatggctgcgcaagatcaaccacatgaaagacttgttttcccagaagaagtcttacctcgtggtaatgctctt

>Chlamydomonas_augustae-psbF
Atgtcaacaaaagctgaaactattacatatcctatttttactgtacgttggctgtctattcatgctttagcagtgccaacagttttcttttttaggtgctataacggcaatgcaattcattcaacgttaa

>Chlamydomonas_augustae-psbH
Atggcaacaggaacaacttctaaagtaaacttaaataaaactgataattcaaattttcaagaaccagggatttctactcctttaggtactttattacgcccattaaattcagaagctggtaaagttttaccagga
tggggtacgactgtcttaatggctgtttttatgctacttttgcagttttttactaattattttagaaatttataatagttcacttatcttagacggggttgttaataattgggattattctgctaaatag

>Chlamydomonas_augustae-psbI
Atgttaacactaaaaatttttgtttatactgttgttacatttttttgtatgtttatttctttttcggatttctttctaatgaccctgcacgtaacccaggaaaaggtaat

>Chlamydomonas_augustae-psbK
atgtcagctttttctattttacttgcaaaacttccagaagcttatgcacctttttgctccaatcgttgatgttatgccaattattcctgttttatttatttttattagcctttgtttggcaagcttcagtaagttttagataa

>Chlamydomonas_augustae-psbL
atggctagaccaaatccaaataaacaaagcgttgaacttaatcgtacaagtctatattggggattattattaatttttgtattagctgtattatttagcagctacatttttaactaa

>Chlamydomonas_augustae-psbM
atggaagtaaacatttttggattaacagcaactgctttatttatttttaattccaacttctttttctattaattttatatgtaaaaacagcttcaactcctgag

>Chlamydomonas_augustae-psbN
atgggaagctctgctttttttctttaccttttttttatggtttctgctgttaagcgtaacaggttattcagtatatataagttttggtcctccttcaaaaaaattaagagatccttttgaagaacatgaagattaa

>Chlamydomonas_augustae-psbZ
Atgacatctattcttcaactgactttatttgcattaattttagtttcttttggattagttgtgggtggtgcgactccaaatggttggacagaaaataaaggttttgttttttcaggtctgagcttatgggcagttcttgt
cttcactgttggtgtttttaattcctttgttgtttaa

>Chlamydomonas_augustae-rbcL
atggttcctcaaacacaaactagggttggtgcaggttttaaagctggtgttaaagattatcgtttaacatattatactcctgattacgttgtaagagaaacagatattcttgctgcattccgtatgactccacaag
ctggtgttcctatcgaagaagcaggtgctgctgtagctgctgaatcttcaacaggtacatggacaactgtatggactgatggtttaacaagtcttgaccgttataaaggtcgttgttatgatatcgaaccagtt
gctggtgaagataatcaatatatcgcttacgttgcataccctatcgaattatttgaagaaggttctgtaactagtttattaacatctattgttggtaacgtttttggtttcaaagctcttcgtgctctacgtcttgaag
atttacgtatttctactgcatattgtaaatcattccaaggacctcctcacggtattcaagtagaacgtgacaaattaaacaaatatggtcgtggtcttttaggttgtactatcaaacctaaattaggtctttcagcta
aaaactacggacgtgctgtttatgaatgtttacgtggtggacttgacttcacgaaagatgacgaaaacgttacttctcaatcatttatgcgttggagagaccgttttattttcgttgctgaagctctttacaaatct
caagctgaaactggtaaacttgaaattaaaggtcactatttaaacgctacttcaggaactgctgaagaaatgttaaaacgtgcagaagttgctaaagacttaggtgtacctatcatcatgcatgactatttaac
aggtggtttaacagctaatacatcattagcacactactgtcgtgataatggtttattattacacattcacagagctatgcacgcggttattgaccgtcaaagaaaccatggtatgcacttccgtgttttagctaa
agctctacgtctatcaggtggtgaccaccttcactcaggtactgttgtaggtaaacttgaaggtgaacgtgaagtaactttaggtttcgttgatttaatgcgtgatgaatacattgaaaaagaccgtagccgt
ggtatttacttcactcaagactggtgtggtttaggtggtgttatgccagttgcatctggtggtatccacgtatggcatatgcctgctttagttgaaatcttcggtgatgacgcttgtcttcaattcggtggtggtac
tttaggtcacccttggggtaacgctcctggtgctgcagctaaccgtgtagctctagaagcttgtacacaagctcgtaatgaaggacgtgatttagcacgtgaaggtggagatgtaatccgctcagcttgta
aatggagtcctgaattagctgctgcttgtgaagtttggaaagaaatcaaatttgaatttgatactattgacaaacta

>Chlamydomonas_augustae-rpl2
atgggaattcgttttcttcaagcatttacaccaggaacaagaaatcgttcagtttctgattttagtgaattaacaacaactaaacctgagagttcgttaacatacaatttacaaagagcaaaaggccgaaatca
ccgaggtgttattacttcacgtcatcgtggaggggacataaacgtctttatagactaatagattttcgtcgtgacaaaattggaatggaagcaaaagttattacaattgaatatgatcctaatcgaaatgcac
gaatcgctctccttcgttatgaagatggagaaaaaagatatatattacatccacgtggactaaatattggttcaacggtagttggagaaataaatgctccaattattattggaaattcacttcccttacgtaatat
tccgttaggtgctgaaattcataacgttgaatttcaaccaggttctggtggccaaattgcccgggcagctggagctgtagtagaaatttagcgaaagaaggcaattttgtaacttacgtttaccctctaaag
aaatccggttagtttcaaaaaattgttgggcaactataggtcaagtaggaaatattgaagcttataatttaacgttaggcaaagctggtcgaacacgttggttaggaattagaccaactgtaagaggttcggt
tatgaaccctgttgatcacccacatggtgggggagaaggccgtactccaattgggcatagtcgtcccttaacgccttggggcaaacctgctttaggtgttttaactcgaacacctaaaaaatatagtaatca
atttattattcgtaaaagaaaacaa

>Chlamydomonas_augustae-rpl5
atgacacaaagactcaaaacatattatacagaaaccataattccaaaattacaaaaacaatttcaatacgagaatattcaccaagtgcctaaaatagaaaaaattgttattaatagaggtattggagcggctt
ctcaaaatcaaaaaattgtagattcttctttaaaagaattagctataattgctggacaaaaaggaatcataacacgctcaaaaaaagcgattgcaggctttaaagtaagagaaaaaatgccagttggtattgt
agttagcttaagaggcgatcggatgtacagttttctagatcgactaataaacttagctttgcctcgtgttcgggattttcaaggaattaatccaaaaagttttgataaaaatggcaattatagtttaggtttagaa
gaacaattaatgtttcctgaaattgaatatgacaagattgatcaagttcgaggtatggacatttcaattgttacgacggcaacaaaacaagccgaaggtttagctcttttaaaagaatttggtttaccatttaaa
gcttaa

>Chlamydomonas_augustae-rpl14
atgattaaacctcaatcttatcttaatgttgcagacaatagcggagcacgaaaattaatgtgtattcgtgttttaggtggaagtcaacgtgaagctggaaatattggagatattattatcggagttgttaaagatt
ctattccgaatatgccattaaaaaaagctgatgttgttcgagctgtcattgtccgaacgagtaaaggattaaaacgtcaaaacggaatttcaattcgttttgatgataacgctgctgtcattataaataaagaag
gaaatcctagaggcacacgagtttttggtccaatagctcgagaattaagagatcgtaatttcacaaaaatagtttcattagctcctgaagttatttaa

>Chlamydomonas_augustae-rpl16
atgcttagcccgaaaagaacaaaatttcgtaaacaacatcgtggtagattaaatggaaaagcaactcgtggtaataaaatttctttggtgattttgctttacaagctttagaaccgtgttggataacttcgaga
caaatcgaagccggaagacgcgttttaactcgttatgttcgtagaggtggaaaattgtggataagaatttttcccgacaaacctattacacttcatccagctggaactcgtatgggctctggaaagggatat
cctgaatattgggtcgccgtagtccatcctggaaaaatcatttatgaaatgaaaggtgtttctgaaattattgcaaaacaagcatttcgtattgctgggcataaaatgccagttaaaactaaatttttaacaaatc
aaattatt

>Chlamydomonas_augustae-rpl20
atgactcgtgttaaacgtggtaatgtatctcgaaaacgtcataaaaaagtattaaatatgtctaaaggttttcgcggcgctggctctgttttatttagaacagcaaatcaacagaaatatgaaagcattacgatat
tcgtatcgaaatcgccgtcaaaaaaaacgtgatttagacgtctttggattgcacgtttaaatgctgctgttcggtgttatggtcttaattataatgagtttctgaattatttaaaaatacgtagtattaaatttaaatc
gaaaaattctagctcaattagcaacagctgatactgaagcctttatgcaattacttttattttaa

>Chlamydomonas_augustae-rpl23
atgattgatttaataaaatatccaattattacagaaaaatcttatttaacgttatttaaaaatcaacaatatacatttgatgtagatttacgattaagtaaacctcaaattaaaaaattatttgaaaatttatttaacgta
agtgtaattgcagtaaatactcatataccgccacgtaaaaatcttcgtgttggcacaactaaaggatatcgagctcgttataaacgagctatattaaccttaaaaaaaggtcaatcattaaaatttgctatacct
ttaacattt

>Chlamydomonas_augustae-rpl36
atgaaagtccgtgcatcagtaaaagctatttgtgataaatgtcgtgttattcgtcgaaaaggtacagtaatggttatttgttcaaatccaaaacataagcaacgtcaagga

>Chlamydomonas_augustae-rpoA

atgattaaatttgtaaaaaatgatttttttatttattgtaaagaagcccgtattgaaagtcatcgaagtttttatgggtcttttagtcttggccctttgaagctggacaaagtattacaatagcaaatgcattaagac
gaactttgttatcggaattaaaaggtttagcaattacttccgctgaaattgaaggtgcgtttcatgaatatgccacacttccaggcgttcgagattctatattagatatttttattaaatttaaaagatattgtcttaaa
atcaattaattatgggaataaaaaaaaaggaacttcatatacagcaacttttaaaaacgaacaacgtggatatttaaaaagcagaggacctggaattgttacagcatcagatttaaaactccctcctggaata
cagtgtgtcgatccagatcaatatattgcaactctgtcggaagaaggaattcttaatataaaatttaatataaatgaaggagcaagcgcttatggtacccatgaaagttcttcttttaataaacaaaaagcttct
gctgtaggagggacttgggacccaagtgcagcaagtatagaggggtccttcggaggacagcaacaagcaagtagcaaaggactggagcaagaagcgcccttcttatttgagaatggaaaaggcgg
gcaaagcccatattacggctttacagctagcgttacacctcctctattacctgcctacgtcgcagggcaaaaccttccaagcgttttcctacagttccagtagatagaaaaacttctaaaaaagctttgctta
taaaagccaatcaagttcctccaaagtatgcttcgcatacaaaggacggactcacaaaaacgctaaataaattatttttttctcctcgaccttcacttctatcgtctaattttagatccgctagcctcaaaaccct
tgctatgcatccatccacatccacctttgttgcgtttgggtaccaagcaagaatgggtaacaaaattaacccatgggttaatttaaaagtgctccgcactttagaagttaaagggacgggtgtttcgctctttc
ccctccttccttgctacgttttcttaggaaaggaggaagaacccctactacttgctgtccttccttctttgctacgctttcgtaagaaaggagggacagggagctcctatgggccaagggacgggatgag
atgggacaaagagcaaactcagcaaaaaaaaaagggcctttggccttggttatcccagccttcccttttgctcctttcccatcctacgcaccttcctatgcatatagaatatatacccccagcctcgttccaaca
aagggagatggggaaggatcagcgtcacgaagttaagcggataggagcaactaaaagggagggaaaacaagctacgcttgaacaaaaaaaaaataagcttctctactcaaagttttggcagtataaaa
cgtttatggttaaaaaaaataaaagcaattcaagctagctatgtacataaaattaatagaaaaaaaaatgcttttagtctaaataaaattaataaaaactgttttaaaagagcttggttacttaataaaaaaaaaaa
attcagatttgaaaaacaaaattccttataatgaagccataatagctttatctactccttagtaattgatcctgtttttatgcctgtgaataaagttaattatattattgaaataaatgaacaatcagttattgcaaaa
gatagtttaggtaacgttaataataatatataggaaataaccctataaatttccgtttcaagttattagcaaaattaacaaaattcaccaagtatccttccctttagcaaagctaacagggctactctccggccctt
ccccttgcctcctcctactaactttagttacactaatcctatggatgcctttggtggaaaggatgtgaaacagagcacaggtagtaggggccctccccccatacacctttggtgtatgccactagcttatact
aatttattaacccaccaatcaattaagcacgcaattaacccaggttccctcaagctacgcttaaataaatatgcttctaacagtattattaataatacttttttttcaattaaaaaaaaacaagcatggcaaatcga
accaaaaacgctcttaaaaaaaagccatattgttttttttaaagcaagtttttaagaacttttctactaattatgcattcaacctttctaaattaacaaatcaaattccttttttggcgctctaaatacgaataaaaa
caggtgtcttaaagaacaaaaatgagaatataaaaaaaaatttttaaagccttctaaatttgggactgcttcgcttattatgctaggggcccctaccccttccaccaaaggtgttagtgtaactaaagttagtagc
agcagtagaacctctaatttaagctctgcttggtcctgtttgtttcattcccatccattatttttccagtaaggatagtaagcttcgcttggaaaaggcaactgttagctacactaatagaaaaacaagttcaaatt
caattgttttttttaaacctttatagcactacgttccgcagtttggccattatcaattagtgccactaaaataaattacatgatagctcttaaaaaatttagctatttttatggagcccctaagacatttaaacaaag
cttaaaattatttaatgtcatagaaaaatcattatatggggcaactatagctcttgctaatttacgtttgcacagcttccacgtccaccaaaggtggaaggggggaaggaccccctacaccaaaggtggaaaag
gtcaagcgtacgcttgcccatggataggggcctctatttgcaacactcctcttatagctcatcctttccaccaacgctacgcacacatccaccaaaggtggaattggtggatgtgttgctacgcaaaagag
caggagatggatgggaaaagaatgcaacagccccctaatttagaagacaaaaacaaacttttagggcccaatttttatataaggggccctaaaaaaactaatagaaaatttagtagaggaactttgctctct
aactatcctcagataaaatttccctttaagtggaatacaaacatctttaaaaaaataatattattttagaaatttggacaaatggaagtattcatccacgccaagcactttatgatgcatgtaaaaatttattaagtatt
ttttcaaaattagaaaaagctcgctttactagagcgatttttcaatcagaaaaaacctatgaaaaattaggtttctcactcatctctagtgaaactacggctatggcacaaaatcctcagccttggtggaaatca
aatttaattaagcaaagcttaataaaaaatgaatggcaaacttcgaatggctctttaccttttcagtagtcaacaaattgaatcagacccgttttctacgattatggggaaaaacgtataataataatacttaca
gccattcccaaaaaatgctccctagcaacccccttaataagcgcatcagtctgggtaccaagcaatgctacgcttacacatccaccaaaggtgtgctacgcagaggtgtatgggatggggttggagaat
gctacgcacccattcatcctaaaacaaaaggcagagtaccaagcactagggaagtcaatagctatagttcaactccattagaaatagcaggttaccctaatgttttaacaaaaaaaaacacgaataaaga
aatggcttttaaacataactctacggtgcatcctagctacgcacaccaaaggtggatggaaggggcccctacttcatctatccccttaacttctaaagtgcggataaaaggcatatttcatatgccttttttaa
aagggttgcgaagcaacccttttttaacttttaaaattacccattcccttcctggcgactgctgcgcttattaaggggggttgaccctaaacaagaggaaggatccagtaggaggaaggaacgggcagagc
ctttaattaatgagagtaaattttcttttgttcctatttggggagttatgtttaacaaaaataaacaatttttttttagctatgccaaaaaatcaccaaatagctttttttcaaaatcaaacaagaaaaaaaaattttttaat
gcaaagcatagccattaagggctgttataaaaatggatttttatttatcaaaattagctaatacaaaagctaacgctaaagctagtttcactcctgttagcttaaataacctagcagtaaatacttctaaatcaact
aacgctttaaaagcaaatcaagttcctttatggctagaagctaaaatgaatcaagtattgggggcatcctacggagcccctgaaaaagggcctctaataattgaaaacttaaaaaaaattgatgttttttttaaa
acaaaaacaaataaattggaatataaaaatgtctctaacagattataatctttttaaaaacaaaaccagattttgttgatattggaaaacttaatatttcacttcgtccgtatacgtgtttaaaaagagctaatattaa
tactttaggagagcttctaaagaaatcaaaagaagaattattagcattaaaaaattttggaaaacaatcattaaaagagcttgaagaaaatttaaaacagcttggactttctttaagatct

>Chlamydomonas_augustae-rpoC2
atgttgttaaaaaaaaaaattactggccgctatcatgtatattcatatagtggatggtttggtgctctgtttcactggcaagggccccctagccataatagacgaattttggtacgctcaagcatactttccataa
ttgattctacaacccacgttttaaatcaagcgcgtcttgacttaaaaatcaagtccaataagttaaaattaaaaaaaagtgttacaatttttattattaagacaaagcactcaccttgtagcttttggctctaatagc
gaagttagaaaacctaagttaaaaatttataattacacaaataaaatgatttcttcttttaattcaaattttaaaacaacaaaagagcaccttaaatgcttcgcatttaaaaaaccaaatattaatggaaaaaataa
aaaccaaatatctccactacctttaaaaatccgttttggaatcaatgttttaaccgaggtagtttaaaaaatttttgttttatggttttagtaacgtatggagaatatcgtacagtagaattaattgaacaattaaaa
aatgttggttttgaatatgctactctagcaggaatttctttaggaattgatgatttaaaaattcctccgaaaaaatcttttttaatttttgatgcagaaaaaaatacaaatttagctcttaatcaatataatcgcggtga
aattactggagttgaacgttttcaaaaaatgattgatacatggcatagaacaagtgaaatttaaaacaagaagttattgactattttgaagtaacagatactttaaatcctgtttatatgatggcatttcaggag
ctcggggtaatatttcccaagttcgtcaattagtaggtatgcgaggcttaatggctaatccccaaggtcaaattttagattttccaattcgaagtaattttcgtgaaggattaagtttaacagaatatattatttcca
catatggagcacgaaaaggaattgtggatacagcgctaagaacagcaaatgctggctatttaacaagacgattagttgatgttgctcagcacgtaattgtttcaaattttgactgtgggactcaaagaggc
atttttttaaatgatatgaaagagggtaataaagttttatattcttgcaaaatcgattaattggtcgagttttaggaaaagatgttttaataatcaagaagcttttaaaaatcattatggatctcctgttatcggcga
agcctatcctcttaacgaagttactcgaatcaaaagtcataaaaatttatttttttgaggggattccttttttaacaaacaataaatctaataacaaaataactaattggactactaaaagtgggaaagaggttaca
aatttacaatctcaacacgaagtttcgctaaatgcctctatcccggcttcgattatagaaactaaattaggggcttactctccttttatttttagaaatagtgaaatttcaactgatttagcaagtaaacttagtaaa
ctctctaataaaatttttgttcgttcaactttaacttgtgaaactcaaaaatcagtttgtcaattatgttatggttggagtttagctcaaggaaatcttgtttcaatcggtgaagctgtcggcgttgtcgccgctcaat
caataggtgagccaggtacccagttaacaatgagaacatttcatactggtggagtattttcaggtgatataacagatcaaataaaagctctttataatggtattgttaaatatttagaacctatagcagggactt
tgatacgaacacctgaaggtaaaatttccttttttaacaaaaacagaaggatttttagaagtcatttcgacgggtacagaaaaagaattcgtaaaaaaaaaaatctatataaattccttctttttacttttactatatgtt
cggcatggcgaaaatgtttttgaaaaacaagttattgctcaaatttccaatatttcacgacaacaaaataaacgcgatgatgcagaattcactataaattctgaaatggaaggccaatttcgtattatgacttta
aacttattagaaaaacaagtcggtcctaaacctactttagcagtcaaaaaaataggtcccaatcaagcttcgtttaaaaacgttaacgttaaagaatctcgttttattataaatggcattaattccccccttaata
atgttaatgaaattaaattagctgtaaattccacaatggaaacattgcttgaagcaagcgattggggttatgcttggattttagctggaaaaatatatcaacttcctatttcatctaatatattacctcaatctggtg
attatgttaataaaaaaagtcttgtaaatgccgtagtactaaaaccgtcacttcaaggagtgaattctgggtttttagaaacaagattaaacaataatttgctgttaaataaacctattttatttttagagatcaata
aaatgttttatacaaaaattggatatttttttaaattacctgaatatccgttttttaatgcaaatattttaccctctttgtttatcatgggtgctcccatccgtttcaaattgctactaaaatttctttatgctgcgcattaa
gaaatttttaatttttcttcatgcaagcataaagaaaaatttaaccgcttcacaggaccacaggataagcgtttcacagttgttataagtagtaagggaacgtatgcatcattcgataatttacataaaaatgatttt
gttatttacacctctagcttcactctatttaacttctttttctaacttttttgggtcgcctatttgcatagccaaaaaaaatgtaaacactatttctttttttgcaacgagtcagaatatgagcaaagactatcgaaattt
attaactaagcttctaaaagggcaaataaaaggtatatttcatatgccttttaaaaaaggtatatttcatatgccttttttaccttttttaaagtgcaaagcacttttagcttaaataaaacaaacctctttctaggg
aacctaaaaacaaaacgggcattaaattcaaatagattgtttaatggttctaaaattataggtaaaaatagttctgcaaattggagtaaaaatttagttttatttttagaatggtttccacattataaaaaagttaca
agtggtgcattacttttgtgattctaaaactttacaaagtagttcttatctaccttcaatgtggccaaaaaaaatcactttctttaatcttttttcaaaaatcttttttctttttattatttacatcttctctagagttttttcttaat
acttcatctttatttttctgagcctttaacaatttttaaaaaataaaaaactttatcaacaattagctgtctaccattctctagcctcttgcccatccatcccctttaacttctaaagtgcggataaaaggcatatttcatat
gccttttttaaaaggttgcgaagcaaccctttttttaacttttaaattaacccatgggttaattttgttacccttctttctatccttgctgcgaaaagaaaggatgctaggggcccctaccccttccactgctacg
cacaccaaaggtgtgctacgccaactctgcctcctcctactaactttagttacactaatcctatggatgcctttggtggaaaggatgcgaagcagagctggcagaggtggatgcataggagcttattgtta
aaaaaaaagagtttaaaaaaacttcttgctttaactatgcaaaaagttcttaaaaataaaatgctacgcaaaatgccgttaatgacttctattttaacaaatgaaaagtctaatagcaataacttatcttcttttgtt
ggcaagcagagcttatatggagaatggaccaaacgactgaaagaaggacagattccttatttttataaatagagtaaaaagaataaaagggtatactaaattagtgaatttttttaaatcttattagttctgtaaa
gcttcgttgcttaaacagcaaaaacattttattataattattcaaaattaattgcaaaaaaagaacaaatttaaatgaaacttttatatgaaaaatttttatatttcaagacaatttttgaaattccacaagaacg
tttaactatctctaacttttttttttaaagataagatgtcttataatatttacatgacactttcgaaaaaaaatcgacaaggaataaaaaaagcacttttttcttaatgcaactaacactagtggga

aatttggtcaaacgactcttcacatagtgaaaaaagtcaataattattctaatgcaagctacacttataaaaatctgaatcatcaagctagtttaattagaaattcaagcaatgtaacttttaacaatcaactacta
tttcttatttttccaaacgtaaaagaaaaagttagtaaagaaaaggcagacaaagttacacaatttgaaactagttggaaggacacaaacttttttagtttcatttaagcgtagcttatttaagctatgcttaaaag
ctcatgtattaaacgattcggccttatggcaaagccatacttgttttcaggttccctattctgcttctaactttagttacactaatcctatacttgccagctctgcttctaactttagttacactaatcctatccttgcca
gctctgctacgcagcctttccaccaaaggtggatggatgcagagcacaggaaggatgcagagcacaggaaggatgcagagcacaggaaggattcgcttcttaaggagggagaagggaaaagata
ccaagaaaagtttacttttttagggaaggctaggaacccactctgggtccaaccctgccagctctgcctcctacggatgcgaagcagagcatcctagctactaacttttagttacactaacaccaaaggtg
gatggatgcagagcacggctaaagtgacaaatcaattactaaagtttaacttaaaaaaaaaaaaattttttaaagaatggacaataaaacaatttatttttgtaaaaacgtcttgtaattcagttatacaagtca

ccccgtttccaattaaaagaaaaagaggtattccgaaaaataaagtaaggcttttacaagctagagttaaaaatattaaagtagtagggctttgccctgtttctttagattatttaaaattgtttaaaattattaaa
actaaagttttagctggtaattctaaaaaattgaaaaaagatctaaaattaaaaaagaaaattggtcttaagttattaacttttattgatggcttgctacgcaggtccctatcttaagccaataacaaagttaaacta
tggctgttctgtacccttgggtacgctttatggggtcccaactacgttaatacgatgatagaagccataaagagagttaaaaaaagtttgagtggtacgtttctaaaatttaatgaaattaaattgactcaaaa
tataggtgaagcctataaaaagaaattaaatattaaaattctctaaaaacctaaacaatatcgctcttttcaactacgttaaagcaccctaattataaaacaagccttttttggcttaaaaaaaccagttttaa
atttgcagcaaaacttttaaaaagaaaatagttatttatcctcatcatttaaaaaaattaccaaataatttttttgaatttaaaagaaaataatttaagatttaatggttttcctataattaataaggacgcttatagtaa
cttattattaatggcccattatttaaaacggggtccatttaaagaacatcaaaaactcatttcattatctgttcttgaccgtcatgcacaaaaacaagcaacaactttaatccaaaaaaaattcgtttatgaaaaag
aaacttttttttaactacacaaaacaggccatataatataatactaaaagtagattccagaacctgttttagtagctttgctgctcgtgtaactccttcttcgcaacgtttttttcacccatctcctccgccaatctcaa
ataatgctttgcaacttgtttcccatccatcccctagctacgcacaccaaaggtggatgggtaacaaagttaaagagaaggagcgtagcaaaattaacggcagctacgcatccgaagaaaaggaatgg
ataagaaagattagtgaaactaaagttaaaagcgggagaaaggagctaaagaaaagaaaagaaccaaaaaaaggaaagtttgcggaacccataataactaacccatataaaagtacaagttctaagtt
aattaatcaaagtcctggattaattactatttttacaaaaaaaaggttgggttttttttccttctatttctattaaaaaaaaaaggagtatttcgaaatctcactcttatacgaagtctaaaatcggtgctttccctaagc
aaattcattatcgactaaaaaattcgcatttaataagtccaggacaaatggtgttgccttatttaatgtttgataatcataatatttacattgaatggtttaataatttttagttctacaaaatatatacataagcaccgg
ttatttccctttttatggttcattattaaaaatgtgggccctaaaaaaggttattaattttaaaacagctaaaaaaacagttagaattaatgttatataaaaaaaaacaatagaatggcaaaaccatcaccaagcttttag
tagcaaaaaaaaattgttaaaaaattcttattcttcggatccattatttggtgctcacgtttatgggcttatgtctatataaaccctttagtttgctctgtaccgatccattctcttgcaactgaacaagtaagagctt
gcaaaggagctattgaggaacaggctttctttaaaaaaaaaagaagaaaaaaaaataagctttgcttgaccattaatttatctaaagattgctccgcagcagagcttaactacctaaataacggaggtctttt
aggttctaaagccttttcaatatcttcttcttcgcttaatcaagctataggctttgcctatacaggtcccctgaatataacaacaagtgattatggttgggtacaaaataaacattgtcgttatcgtccaaatcattc
agctagtgctattgatgatttaactagtaagtgtgcagcattaaaacctattttaataattcaaaaagctttagaatttactattgctaaaccaattacctataaaaaagaaatcgcagaattatatgcctcaaata
gggatttagttattactggaactaaggatttttttaatttcaggtttagtcggaaaaacgcttcaatcaccatctaaacatgttttttttgagaaaacgaacggttctaaaataatgcgtagcataaataacaatctg
agttcttgtttatctagacttatgtttaaaaaaaatagtataataaacttaaataaaactcaatttagaggcacgcatctgccaaaagcagtgagcgagagacctttaaattatacaacaaaaaaagaaagtgc
tttacaaagtttctatagtagatacagagaaagctgttatgcttatagcattaaatacaagaggttaaagttctttatttggcctatagtaagcagcgaagctgtcaggttatcagttagaggatgggtaacaaa
gtttgctacgctccttgctacgcacaccaaaggtggatggatggattagggaacctgaaaaaaaaaagacaaaactcatagttaacaatatttttatttaaggtaaataactttttaacgaagggttct
atttctttatttgctgcctcgcaccaaaataaaattagtaaggtaagagttccttttttaggcaaagtttatacttttaagcccgaagggcttaaaaaacgtaatataagttttttttctaaaaaaatttatctttggca
atcaacaattaaaggaacgttaaatttagtagaaggcattaagggtttaagtattaaaaagtttaacttgtctcaatatattgtgcttaaacctataagtgatagggaaaataatagatgggtacgaaatagga
agggaaaagcagtagcattactcatgactaacaacaaattgaaaaataatcctttaactaaagtaactatttttcaattacgtaaaaataaatcttataattcaatattagttagaaagcagttttttggatctttac
taaataaacaaaataaaacgctagaaggtacagcgaatacaatgccttatttttgatttttaaagtcaaacataatgaattttacactaaaatggtttcattgttaaaattaaaaggaaacagtgtagtaaaagtttt
agcgaagctatctaaaagtcctcttttttaaattaaaattattaatgctaccatggctaaaaaacgcgaaaatcaataatcgttttaaaaaaggtttagttaggggcccctaccccatgtcttctaataggcacct
atccgtaggtaggagcccattaaaaaagccattattaaaaaaaaaaaattatgtacagaagggagtaggggccctgcccctacacctgagcttttaccaaaaagttttagcacaacaaaagctaatactt
tagttttaaaaaaccaattaatctaactcctctattaataggttataatatgccttatggtttgatttagattttaatgtgcctaacttttaattctaatattgcaagttgtaaacttcttggttctgaaacttatct
acatttcaatttttaacaaaaaaacaaaaattgaatttttataaagtcaattaatgaagctaaaattaaagtaactacgttaaatagaacgagcttgcgagttaatttacggaaataaaagttcttggttttttatttg
ctcaattacaagcacttaaaattccgtctttatttgcacaaccgtgttttttcttttttatgagcaaaaaattaacaaatttttataaaacgttttaaacaaagcttaaaatgggttcaagaatctgatccttttgcttcgt
taactaataataatttttcttatcaaataaaaaggagtattgggataagaaaagctaattcttcatttattaatttaacttcatttttaaatttaaaaaaaacaggggggaaaaaagtctgcattttgtgaaaagttacg
ctggaagctttttttgacttttattcttatggagcatctttatacaaaaataaattgcttttttctagtgcaagcaatgcttgttcttttttcgactccttttcttttctcagttttttcctgtccttcctatacgaggctcttatc
ccattttacagcagctaaaagctaatagtaaaacgaaacgaagagaaaaaagagtgcaaggtatcaaacaaagcttaatgaaattttcatttttaagttctttgaaggtgaatttgtgaaaacaatcaatagt
tatgaatctaaactgcgtttaactaatggtttaaatttcaataattatataagttagatttagtaaagcaacaatggacaaaacctatagtgggttcttttccttggcaatttgatagtctttctcttctatcttatcctt
gccagctctgcttcgctggcagagtacaaataagtcttctgtcctttcatcctctccttgctgcgtataagaaggatggggtaagaaccttagtttctgtaaaaaaaaaaggaaggcttagtacccattcta
ctctctgcttctaactttagttacactaatcctactgcttcgtacaaaaaaggcacagcaagtttggtgaagaaaggagtgaagctgctctaaatttaaacggaaatagagccttatggccgttgctaggctt
tatgcgtaaggacaaaaatagaaaaacaaaagtcaatttttaaaattaatcagttttcaggtattaacaagtcagctaaagttgggactagtttcgtttggtcagttcaaacaaccagtcaaattcaactgaa
aaaacatttaagatggtttgggcaaatgctacgcataaaaacaaatcagccaaaaaatctcctatccgtaggtaggagccctttaaataaaaataagacttatttaaataatttttattggaccttcaagtcttatt
ttaaccaaagctaatttaattagttatttttttaagtttaaccaaaaaaagcttaattcgaaaactaaatataacacatattcttataaaatcgctcgagaaaatctttcgaaggaattaatgaacttttgctagatatt
ataaatataataatatttttaattaaagttagttcgagtctttttgaggttgtttatggcgaaactattggtcaaagtaataaaaattacccttttgggtgctttgcatgaactcaaaaagttttatttctcccaatcttgcc
actaaattgttagctccgctcccagttccctctgaagaaacagcgggaatgggaaggaaggatgagcaagggaacttttaataatccaggttcccttatggctttgccatacctgaacaaggtgttactaga
taatactttagccaatcattaccaagaatctactttaaacggccaaactaagaaaaatgaaaagttttcttcttacaaaaaaatttgcaacaaaaaatttttttgcaggtgagaaggaaattggtttaaaagttcctt
tagcaagcaaaattaacaaagtgaattttaaagttaaaaaaggtatatttcatatgccttttattcgcactttagaagcttataaagaatctcgttttaaaaatcaatatatagtactacaaaatttaattaatgtagc
taaagctaaaaacattttatttgctactggaaaaaacgcagcttgcctacttccttttaatttaagcacacaagctccgaaaatagatattgggggtaggaacttatacttttagtgacaattctgatattaatgtca
gtcattacactgaagctataagccgagcaaaaatttttaaaattaatggattaattgcaggtttacctaaacaacccggtatttcattatttaaaaacaaattatctttaggggaatttttagtctatggcgataaa
ataatgtctataaaaacgaaaagttttggaattccttttgcaggtcaaattattcatattaatcaatacaaaattacgttaagacgtgcacaaccgatttttatttcacctaaaagtattttacattttcatgatggaa
gttttgtgcgtaataatgatcccgttatgaccttaacatatcaacgattaaaaactggtgatattgttcaaggtataccgaaagtggaacaatattttgaagctcggataactaaacaaggaagactatttcgtg
attgtttgcctaatttattaaaaggcttatttaaaagatacagagctcaattaccattagaaaaagctgttagacaaagttttttataaaattcaacaaatcatagtagatggtgttttacgtgtttatcgatctcaag
gagttacaattgcagataaacatttagaagttattgtaagacaaatgacttctaaagtcaaaattataaatggagcgcaaactggtttctttccaggtgaaataattgatttacaattcgttgaaagagttaatttta
tcattaatgaaaaaaatacaatatgaaccattaattttaggtataacacgagcagctttagaagtggatagctttttatcagcagcaagtttccaacaaactaccagagttttaagtagtgctgctattgaaaaa
agaaaagatttttttaaaaggtttaaaagaaaatattcttttaggaaatttaatgcctgctggaactggttatttagtatcttttgaagatcttaaaaaaaaaact

>Chlamydomonas_augustae-rps2
atggagttaaaaatcaatttaaaatatttaaaaaaaacattaactggttatcttttattaaataagaaaataaatagcgaattagaacttattaaaacaaaatatttatcggaacaacaaattattcgtgttttaaga
agaaaattaaaacgtcatgaagctgaaaaaaaattgttacaatttttaccaaaattacggtatttacctactcctaaaaatcaaatttctgtcgctattcaattattaatgaaacgatttgttgatcctaaaatgaaa
tatccgttagatcttatttatgatcaagcttcgcttagaaaaaaatcaaaaaaagtagctgcagctcgaaaaaaaaaatggcaacgattagaaacatatttttggtggtattgcaaacatgactaaattaaagaa
aacacaaatcgctaaaaatgttgctataataataggccaacaagaagaaattaatgctattcgtgaatgtaaaaaactgggtattaaaatgtttcatattgttgatactaattgtaatcctggtttagcggatcat
tttattccggcaaatgatgattctagaaattcaataaaatatttattatcaaaaatggtaacacgtattagattagctcaaaaaattcgtttacgactaaaaaaaaaaaatgtctagctcttctgaaattaactcttta
gataaaggcaatactgtccataatggcaaaaccttacttgaagatggaaaagaaaaaataaaggttggtgatattttaacagttaaattaaataaaattgcaattcaaacagataattcttctattggtatagttt
atacaacgcatgaagcactaaaatacaattcataattcctaacgtagctgagactaacgttggtcaagaaatcaaaattcaagtaactaaaatgaaattaaaatatgggtttgcttttaattaaatctagcata
aataaatcagatcaagtaaagcttccaggttcccttatggctttgccatacctgaaaagcttgcaaacctcaaattataagcaaaacttattaaatgagcccatttttaggtttacaaaaaggatttgcgaattta
ccaaaaccttttatttctttaagctcacgtccaagtgaagcttatgtttttggaaactctcttccattgcatcctacggattctgtacataagagcaccggctttccaagtctatacattccgaaagtcaataattat
ggtcaaaaaattactcctaataatagttcttttttcgtctttgatttctgggagccttgcacaagcaaagcttgctaaaaaaatcaattacaaaccttttaacatagcgaatatgaccaaaggaagtaaaataactt
taacattacctaaaacagcaaaaaaaatatgctaaatatttagtaattaacatgctaacgcataaaaatattgataaaaatttagttattgagcaaccatgaaaatgatggaaaagctaaccataattcttatatttt
atttataaaaaatttattaggagcaaaactaggtgataaagtacggttaaaatttacaaaaatcggttcttccttgctattggtgaaattatttgtatttcacccaattcatcattaacaaaattagctcgtgttcgt
aaaaatttaaatgaaatgcttataaatggaatgcattatggagaaaaagttgtaaaatgccaagctaaaatgcgcagcttcgcttggaatagaaaaaaaatctccgtaattagtaattcggcaactgaggga
acctttgctttaaattctaataattataatgcagttaaagactcagaaacttct

ggagcaagcgtagcttttgcttttgctaaaaaaagctactctttcggttatgaacctattagctttgctaattcccaagcttctaacgttctaaaagggcaaagccctttttagttacactaaaaacctttttatcctc
agctgctttcccatccgcaaaggggatgcaaccaaaggcgaagcctagaagccaaggatcgggcaagaggggtggggaaaggggcaaaaagcaagtagcgaagctaactaaccaaagggtaga
tagggagcaaccattagaagcttctaaaaatctatctttaaacaaaaacgttgctttaataaaaaaaggtagacatttaattaatttattaaaaactcgtcgttgttttaaatatagtaggacttaatcaattaacaaaat
atgcagctaaaggaagaacttttttatttgttggaacaaaaaaaccagcagcaggcttaatagctcgcgcagctttatttagtaaaacatcgtttttgtaaatacaagatggttaggtggtatgcttacgaatt
ggaaaacaattataaaatctatttccaaaatccgtcctattcttaaagaaaaacaaaaagtcatttccaatattcttcaaaaacgttcaaaaattcaacatcgtttactaaaaaaagtgttttcttaagaaaaaaa
agtaaactcttaattaaaaaaggtaaacaacttattgctaaagttaaaatagctaatggaactgcttcttcctctcttattaacgttaacgcaggtgcgaaggaaggattagcaaataacataagtgaagcatc
taatacaaaaaacagttttaatgaaaactatttaaagctaaatcagattaatataaaacgaaaagaacttttgaaaaaggtcaaagtttattagaaaaacgccaaaatcttttaaataaacaaaaagaattaaa
agaaaaaactttaaatcttaaagaaaatggttttttaattaccaataaatataaagttttattaacgcaattacgaactacgcatactaaacttcatgatttaagacttcttttttattaaataaagaacttcgaaaaa

taaaaaaattagctaaagaccaaggaaaaaataaatttattctttcttatggtaaatttaaacaattaactacacttgctacagaaaaaaacaatcaaggtacgcttttagcctctagttctaatgactctgccatt
attaatggaaataaaaacagagacagcagaaatttaatactgactttattaattctaaaaaatttggcttattccaaatccgcctcaagaaatttaaataaactgcttagcaaaataaatcaagctacgctta
aaacttccattcccttccagccaatgcgtagcaaggactcaagcaaagcagctaaagctgagcttacaaaacaagatataagtgataaaacgaaacaaaaacaagcaaaacttaatcaagctagcttaa
cttctttttaaattctatggcttggaataaaggtccttataccttagctccagaccaaaatcattttagtattaaagaaaaaaatactacaagtattttaaatgaaagtaaattaattatttttaactaaacttttaagta
aatttagtcgctttttaccgtatattaaaaattgtattaatttaataaaaaatttaatgcagcatatagaaaaaaaaatt

>Chlamydomonas_augustae-rps3
atgggacaaaaagtccatccaataggatttcgcgttggtattactaaaagacatcaatcacagtggtttgcaaggtttacaaaatatgcttatgctcaaagtttattagaagatcgtatgctacgaaatacatta
atcagtgtagctaatgaagctttaaattcaaattcaataaagaaacgtgactctgcttttcaacaaaacgtaataacacctaaagtaacacatataaaaattgaacgaggattaattccgtatgaaatcggaat
tcaaattcatgctcaagctccgttaaaaaaattgaaatcatcattaaataatttaaaaattaacagagaattattgtgtaaattacaaaaaactcgtcggtatttaatgagtttaggaaaacagcaaagttctaaa
ccaaaaagtgaaatgcttttaaatatacatgctaagcatagcttgaactcctcttcctctcctcttactgcttctaataggagtgacactaatcctacagatggaaagaacagcaagaaccaaacggaaaaa
gattgctacgcaggaacggtttctaaagatgacacattgcttgttggaaaaaaaaataatttttaaaaaagttaatatccttaaaaaaaaactttttttgttttagttcagcgttagcgactagtcaaggtaaacaaa
acttaaaagtcaacattccttttttctacaagcaaagcaggcacacagggaatgactaaaagtgggcccaaacaaagtttagaagctaaagtaggcaaagcctataactttaaaaaacgtagacgtcctaa
ttttttaagtctagtctacaaaaaaaaaaaaaattattataaaacgtttacgaaaacgtcaagaaatacgtcaacgttatagaagattaatgttaaaaggcttatttgttcaaaaaaaaggaaaacaaattagttgt
gttgcttctttatttttttagtaaaaagcaaagaaaaaaagctagtgcaattaaaaaaaataaaagtaataaatttgctaaaaaatatgaacagaaaatcagtaaaatcttctttaacaaaaaatactagtttactaa
aaaaaaaaagtttaggtgggtccttgctacgttctttccctgctgctacgcttcctacgctccatccatccacctttggtggaaaggagcgtagcaacaccaaaggtggatggaatagtgtaactaaagtta
gtcagaaggatgctaacaggagtgacactaagaaaggtccattggagaaacaagaaggactaaaacaaaataataatattaaacaatcaaaaactgataatttttttagctgcccctataaaatttaataaa
agacaaagtagcaatttaaccagcggagctacgcagacttcattattaacaaattctaaatctaataaaaatataaaagaaaacttatactttagcaaaattacgtaggagacaaagttcccctcggacaa
ctacgttgttagcttcgccgttacacacaacaatcatgggcaaaacccatccttcctacagataggatggaatgttggttatagaataaaaaaaaaaatttgtttcgttatttttaaataaaattaacaaacgctttta
aaagaattaaaagagcaaatgactcaatggaataaatctcttaaactacatcgacaagaacaaatcaaattatacggtacgcttcgttttgctcctcttggttataatcaaaaatggagtttaaaaagattaaat
tttttgaaaaataaaccgttatttaagcttttaaaattaattaaaattattgaaagaaaatctttaataaagttagaaacttaagaaaagattttatagcatttggcacaatttctaaaagtgaagctttaagttattat
caaataatcaatttttttaaaaaatttaaaaattgctgctttaaaattaaaaaaagaacaaaaacgtctgttagttaaaaatttaacaaaagattattctaaaatgtcagtattaacagcatcaaatcaagaaacac
atagtctaggtgagaataaagatgctactacgtttaaacgtggcgataaatcaatcaagcttcgcttaatgaaaacatgggcaacgaatcaatcaatagtatctttagatagaaaatctagcttcgcaactttg
ttgacttcttctgttttaagcggaactaactatcaaagcaagccattaatagaagctattaataatgaatgccgtaagattaatttttattgaatatttaaaagaaatggttaaaaaacatagaacaaataatttgttt
tattatttgtcaacaatttccgatgcccgtaaaaatttaaaaaaaattaaacaatttactaaagttcattcaaaattcctatttgggtttgagtttaaccaaccaacagccatgttagctcccgacggtcgctccgc
ttactcaaaaaatacagatggcaaagatggattagtttttaataaaataaaaactgtcattgatttagcttctaaaaaatctgattttgaaaaaggattaccagaaatttttcttagaacaactggaaaagcaacgt
aatatgtataaacaacttgttaaactacgtgctaaaatttcaattaaatttattcagtaaaatcagagaacgtaaaaataaaagcagcaataatatcggattctgttattgatgctttagaaaaacgaaaagcttt
cagaaaagttatataaagatgctaaagaaaatttaatgcgaaattccaaagtaaaaggtgttaaaatacaagtatctgggcgtttaaacggagctgaaattgctcgaacagaatgggtgagaagtggtcga
gtgccattacaaacttaagagccaatttagattattcgtataaaacagctaaaactatttatggaattattggtgttaaagtatggatatttaaaggttatactaaactagtttaa

>Chlamydomonas_augustae-rps4
atgtcacgttatcttggtccacgattaagaattattcgaagaattggcaaattaagaggatttacacgaaaaaaaccttttcgcagagtctttaaagggcgaggtgctttacgtggtaaagttattccacccgg
tcaacatggaatggtcaaattatttaaaacacgcccttatgactcttctgagtctgattatctaattcgtttaaaagtaaaacagagattacgttttaattatggtttaactgaacgtcaattagtaaacgtagttag
aaaagctaaaaaaaatataagaatctactggtggtatttttattacaattattagaaatgcgtttagataatatagtttttcgtttaaatatggccccaacaattgttgcagctcgacaattaatttgccatggtcatatt
aaagtaaataacaaaaaagttaatataggcagctacttgtgtaaacaaaaagatgttatttcggtttcaatgaaagaaaaatcattaaaacttattacaaacaatttacaaaattactatcaacgcgtacgcttc
tataaagaccgtttgaagaaaactttgtcttttattcttctgaaatcgcaaatagtaacaaatatggcagctgctattagtcttattcaaagcggaaattgtaaaataaataatcaagctacacggaaacctaatt
acatttgtggcgaacgtgatgtcattacaatagttacaaaaacaggaattcgtcaagtcaaacaaacc

>Chlamydomonas_augustae-rps7
atgcctcgtcgccctataacaaaaaaacgttctcttttgccagatcctatctataatagtatttctgttcatatgttagttaaccgtgttttaaaaaatggaaaaaaatctattgcttatcggattgtttataatgcctt
aaaagaagtaggggacataactaaaaaaaatccggttgaaatttcgaaaaagctttagataatgtcacaccacgtgtagaagttaaacctcgtcgcagagcgggaacggttcaacttgttccacgcgtt
cttcgtactggtgacaaggctcgcgcatcagcccttagatggatttagaagcttgtcaaaaacgatcaggtcaatcaatgattgcaaagttaaaaaatgaaattgttgaagcttataaaaaaaccgggtttg
ctgttcgaaaaaaagatgagcttcataaaattgcaattaataatgcgatgtatgctcgaaaaccgcaaacagtaattaatgctataaatcaagaaaaaactaat

>Chlamydomonas_augustae-rps8
atggttaatgatactattagcgatatgttaactcgtattcgaaatgcaagtttagcaaaaaaatcaacagtttgtattccttatacacagttaaatcagcaaattgctcaaatttttagaaaaagaaggatatatatt
aacggtccaaatgtcattagattcaaaaaacttaattgtccgtcttaagtatagatcaaaaaaaatttataacgggaaaactaaagagtcatgtttaacaaatttaagacgaattagtcgtccttctctgagaatt
tatacaaattctaaagaaattccaagagttttaggggaacaggaattataattctttcaacaccaagtggacttttaactgatcgtgaagctcgttctcgtggtattggtggtgaaatttttatgctctatatggta
a

>Chlamydomonas_augustae-rps9
atggatattttagcaagagcggttggacgtcgaaaagaagctgtagctcaaattaagcttattcgtggaaatggaaaatttttaattaatgataaacctgccgaagtttatctgcaaaataattcttgttctcttat
ttgtattaaatcaccattagaagcagcattaagtctttcttctaataatttttcaaatatacacaggttccctgatcaagactctgctctaattcaagctgttaatggaaaattagaagaacaaagtactatcttacg
gatgcaaactgaaaaaagtgaaatttttaactaacaaagaaggtggactagaagccaacgtagcgagtagttctgttaacactaacgactctcctccttctttccccaactctgctccagttccttcttctttagg
ggaaggaacccattacctttggtaccaagcaaagtgagaagcaaggagcagaaggaatggagcttctttctgtacaagcgaacgcgccaagcgaggcacgtagcatgaacgaagggagaagcttt
gagctaaatgtttctggcgacgagacacaaaatttaaatgattctacacataagctcaacttgctcccgttaaataatttagtttcgttagatgaaattgacgtactcataaaagtaaaaggggggaggactgat
tggtcaagcagaagcaatcaaattgggaattgcacgagctgtttgtttactgcaagctgcgtttaacgagggtattttcaaaaaaatttttaaaaactaaaggttatttaactcaagattctcgtgttaaagaacg
tagaaaatatggtctaaaaaaagcacgtaaagcttcgcaatatcataaacgttaa

>Chlamydomonas_augustae-rps11
atggcaagacaaacacgaaaggttgcaccgaataaagcaaaaaaaaaaatttatcggggagttgttcatatccaagcaggttaccataatacaattattacaataactaacgtaagaggagacgttctttg
ttggagttcagcaggggcttgtggatttaaaggaaaacgaaaatcaacaagttttgcagcaaaaaaagcagcagaaacagctgcgcgcaaatcaaaagatgctgctatgagagaagctaaagttttagt
aactggtcctggtcaaggtcgcgaaagtgcgattagagaaattttcaaagcaggtataaaagttaatgttattcgcgaaaaaacaggtattccccataatggatgtcgtccaccaaaaaaacgaagtgttta
a

>Chlamydomonas_augustae-rps12
atgccaactattcaacaattaattcgttcagcacgaaaaaaactaacaaataaactgaaagctcctgctctaaaatcttgtcctcaacgaagaggtatttgtcttagagtttatactattactccaaaaaagccc
aactccgctcttcgaaaagtagctagagttcgtttaacctcaggttacgaagtcacggcgtatattcctggaattggtcacaatttacaagaacatgctgttgttttagttagaggtggccgagtaaaagatct
acctggagttcgttatcatatcgttcgtggcacattggataccgctggagtgaaaaaccgtgtacaaagtcgctcaaaatatggcgtgaaaatagcttctaaaacagcagcgaaaaccgcatctaaaaaat
aa

>Chlamydomonas_augustae-rps14
atggcaaaaaaaagtatgattcaacgcgagttaaaacgacaaaatttagtaatgaaatatgctgaaaaacgagcttctttaaaagaacagattaaacaaacatctttttttaaaagaaaaattaacgttacatc
gtaagttacaacaacttccgcgtaatagcgctgctgttagactgcataatcgttgtatgattactggtcgtcctaaaggatattatagagattttggattatcacgacatgttttacgtgaaatggctcatgaag
gtcttttaccaggggtacaaaaatcaagttggtaa

>Chlamydomonas_augustae-rps18
atgaatcagcccgctccgcgtttccaaaataatataaatggaacatttattagtccgagtactcctaaaaaaaatttttttctaattctccaaaataagaatttattaaaagaaaatacttttaaaggtactaaatta
aattataagcatagcttaacaacaaaaacaggcaatattaagcaaagcttaaacaaaggaaatacatctgctcaatcaaaaggtaaattaaaaaaaaatctaaaaattaaacgtgttttatcattatctcaaat
tcttggtcggattagaatacaacgtaaaaaaaaagtagagcaacaaaaacgtcaaaaaccaataaaaccaataattccacccaaatcattaattattatttttgaaagataaaccagaaaaagcggtttataa
tcgtagaattattgattataagcattgcggtttattacaaagatatataggtttaggtggtaaaattttgccaagacgacaaacgcgattaactgcaaaacaacaacgctatgtagcaaaaacaattaaaagtg
ctagaataatgggattattacctttttgtaagtaaagaacgaggctttttttcgataa

>Chlamydomonas_augustae-rps19
atgccacgttcgattaaaaaaggtccgtttgttgctgatcacttattaaaaaaaattgaaaaattaaatgctcaaggtcaaaaaaaagttttaacaacatggtctcgttcttccatgattttacctccaatgatag
gccatacaattggagtatataatggtcgtgaacatattcctgtttttataactgatcaaatggttggccataaattaggggaattttctcctactcgcacttaccgcggtcatggtaaaacagataaaaaatcaa
aacgttaa

>Chlamydomonas_augustae-tufA
atggcacgtgctaaatttgaacgtaaaaaacctcatgtaaatattggaacaatcggtcatgttgatcatggaaaaacaacattaacagctgcgattacaatgactttagctgctcgcggaggcggtgcagg
aaaacgctacgacgaaattgactctgctccagaagaacgagcacgtggtattactattaatactgcacatgttgaatatgaaacagaaaaccgtcattatgcccatgttgattgcccaggtcatgctgacta
tgttaaaaatatgattacaggagcagctcaaatggatggcgctattttagtagtttcaggcgcagatgggccaatgccccaaacaaaagaacatatcttattagcaaaacaagttggtgttccaaatattgtt
gttttttaaataaagaagatcaagtagatgatgcagaacttttagaattagtagaattagaagttcgtgaaacttagataaatatgaatatcccggcgatgaaattccaattgttcgtggttccgctttattagc
tcttgaagcattagtagaaaatccaaaaattcaacgaggtgaacataaatgggttgataaaatttatgaattaatggcagcagtagatagttatattccaacacctgaacgtcaaatcgataagccttttttatt
agcagttgaatcaactgtgtcaattacgggtcgcggaactgtagctacaggccgtgttgaaagaggtactgttaaaatagggcgatactgtggaaattgttggtttaaaagaaactaaaaatactatcgtaac
agggcttgagatgtttaaaaaaacgttagaagaaagtattgctggtgataacgtaggagttttactgcgtggtattcaaaaaaaagatatcgaacgtggtatggtaattgctaagccgaaatcgatccttcct
catcacaaatttgaagcacaagtttatatccttacaaaagaagaaggtgggcgacattctgctttttagcaggttatcaacctcaatttttcgttcgaacaacggacgttactggaaaagttgttagttttagtc
atattcaaatgcgtaatccttcttctgttgcagaagaacattctaataaaatggcaatgcctggtgatcgaatcagtatgttagttgaactaatgtcaccgattgcaatcgaaaaaggagttagattcgcgattc
gtgaaggtggccgtaccgtaggggctggtgtggtaactgcaattattgaatcaaaa

>Chlamydomonas_augustae-ycf3
atgccaagaactcaaaaaaatgataatttattgacaaaacatttacagtaatcgcagacattttactgaaagttttaccaacctcacaacgagaaaaacaagctttttcatattatcgaaatgggatgtctgcc
caagctgaaggtgaatatgcagaagctcttcaaaattattatgaagccatgcgtttagaaattgatgcttatgatcgcagttatattctatataatattggtttaattcatacaagtaatggagaacatggtagag
cattagaatactattatcaagctttagaaagaaatccttctttacctagtgcattaaataatattgctgtgatttatcattatagaggcgaacaagctattcaagataatcagccagaaatctgtcaacttttatttg
aaaaagcagcggattattggaaagaagctattcgtttagcgcctacaaactatattgaagctcaaaattggttaaaaatgactggtcgagaataa

>Chlamydomonas_augustae-ycf4
atgaacaattccttttatcacaaaagagttcttctaaggttgtcgctttagagacaaattcaaaacaaacagaactaaatcgtcgttattttattgtcggagaacgtagactgagcaattattggtgggcttctg
taatttttttaggcggatttggtttttattaacgggaatttcgtcttatttgaactataatattttagcaaatgcttttaaactatttaacgtgactatgttaatgcaagctttgtctgatggggaagctataaaagcga
atttgactagcaatttaactccgttaaacgaaaattcagttattgctttttttcctcaaggtttactaatgtgttttacggcagcttaggtgttctattaagtatatattggtggcttctaatttttgggatgttggtgg
cggttttaatgaatttaataaaaaaagaaggctttatgagaatttttcgctggggatatccaggaaaaaatcgtcgaattgatatctcttacccattacaagatatcgaagctattcgagttgaatttaaacaagc
acaaggtttattagcttctgaacaaacaatttttgttcgtttgctttcaaccccgaacaacaacgttgttactgagcaaagcaataagatgggaagggagggaaaaagaaaagaagaaggaagtagcaag
aatgggaaagagttaaaagaaaaacgtgaaattcctttaggtgggattggccaactattaacttttaaaagaaatagaaaaacaagcatctgaattagcaaatttttacaagtagaactggaaggtctataa

>Chlamydomonas_mutabilis-atpA
atggcaatgcgcactccagaagaattaagcaatttaattaaagacctaattgaacaatatactccagaagttaaaatggttgattttggtattgtgtttcaagtaggggacggtattgctcgtatttatggtctag
aaaaggcaatgtctggagaacttttagaatttgaagatggaactcttggtattgctcttaacttagaagctaataacgtaggagcagttttattaggtgatggtgtaaaaataacagaaggaagtagagttcg
ttgtacaggcaaaattgctgaaattccggtaggcgaaggatatttaggccgcgtcgtggatgccttagcgcgtccagtagacggaaaaggtgctgtatcaacaagtgatacaagagctatcgaatctatg
gctcctggtatcatctctagacgctcggtatacgagcctttagcaacaggcttagtatcgattgatgcaatgattccagttggccgtggtcaacgtgagctgattattggtgaccgccaaacaggaaaaact
gcaattgctattgatacaattttaaatcaaaagggtaaaggcgtagtttgtgtttacgttgcaattggtcaaaaagcgtcatcagtagctcaagttttaaatactttaaaagagcgcggagcattagactatacg
attattgttatggcaaacgctaatgaacctgctactttacaatatttagctccttatactggcgctacgttagcagagtacttttatgtatacaggccgtgcaacgctggtaatctatgatgatttatcaaaacaag
cacaagcttaccgtgaaatgtcactattacttcgtcgtccaccaggacgtgaagcatacccgggtgatgtttttatcttcactcacgtcttttagaaagagctgcaaaattaagtacagcgctaggcgaagg
aagtatgacagcacttcctattgtagaaacacaagagggtgatgtatcggcctacattccaacaaacgttatttcaattacagacggtcagatattcttatctaatgatttatttaacgctgggatccgtcctgc
tattaacgtaggtatttcagtatcgcgcgtagggtcagctgcccaaccaaaagcaatgaaacaagttgctggacaattgaagcttgctcttgcacaattcgctgaattagaagcgttcagccaatttgcttct
gatttagatcaagcaacgcaaaatcaattagcgcgtggatcacgtttacgtgaaattcttaaacaagctcaatctgctcctttatcactagaagatcaagttgcaacaatttatgctggaaacaatggttatcta
gataaactggaagttagccaagtacgcgcatttgtaacaggtttacgtagcttgtagcttcaaaatatcctaaatacagtgaaattattaaaactactttaacatttaattcagaagctgaaggtttgttaaaac
aaggtattaatgagtatttaaatgaatatttagcaactaaaaaatcttaa

>Chlamydomonas_mutabilis-atpB
Atgagcgattcagtagaaacaaaaaatattggacgtattgtacaaatcatcggtccagttattgatattgtttttttcaaaaggtcaagttcctaacatttacaatgctttaataattcgttctaaaaacgccgctg
gttcagaggtaggtgtaacagtagaagtacaacagttacttggtgataactgcgttcgagcagtttcaatgaatccatcagatggtttaacacgtggaatggaagttttagacacaggtaaacctctatctgt
gcctgttggaaaagcgacgctaggccgtatttttaacgttcttggtgaacctgtagacaatctggggcctgtaaaggcagaaacaacattaccaattcaccgtacagctcctgctttcgtagatttagatact
cgtctttcaatttttgagacaggtattaaagttgtggatcttttagctccatatcgccgtggcggtaaaatcggattatttggaggggcgggcgtaggtaaaactgtattaattatggaacttattaataatattg
caaaagctcacgggggcgtttcagtatttgccggagtaggtgaaagaactcgtgaagggaatgaccttatactgaaatgaaagaatcaggtgttattgttgaaaagagtctttcagactctaaagtagctc
ttgtatatggacaaatgaatgaaccgccaggggctcgtatgcgagtagctctaacagctctaacaatggcagaatattttagagacttcaataaacaagacgttctattctttatcgataatatttttagattcgt
tcaagccggtgcagaagtttctgcg

ttattaggccgtatgccgtcagcggtaggctatcagccaactttagcaacagagatgggtaacttacaagagcgtatcacatctacgaaggacggttctattacgtcaatccaagctgtgtatgttccggcc
gacgaccttactgatccagcaccgtctgtaacgtttacgcacttagacgctactactgtactgtctcgcggcttggctagtaaaggtatttacccagccgtagatccgttagattcgacgtctactatgcttca
accttggattgttggcgaaaaacattatagtgttgcacaagatgtaaaaaaaaaccctcaaagatataaagaattgcaggacatcattgccatcttaggtttagacgaattgtctgaagaagatagattaatc
gttgctcgcgcaagaaaaattgaaagatatttatctcaaccttttcttttgttgctgaagttttcactggatcaccaggaaaatatgttagtttaacagaatcaatggaaggtttttctaaaattttctcaggagaactt
gatagtttacctgaacaagcttttttatttagtaggaaatataaatgaagtaattgcaaaagcagctacgttaaaa

>Chlamydomonas_mutabilis-atpE
atgagtttacaaatttcaattttaacaccagaacgtccttttggaatggtcaagcagaagaaattattcttcctacggaaactggagaaatgggcgttttaaaaaatcacgctccaattatcactggtttagat
gttggcgcaatgttgattcgtacaaaagacgaatggaattcttatgctattatgggaggatttgctttagtaaaacaaaatatagtgacaatttagcaaacgaagctgaatcagcagataacattgatcctga
agaagctaaaactagttttgatacagctaaaactaatttagaaaaagctgaaggtgtgaaagaaaaagttgaagcaaattttgcttacaaacgttcgaaagcccgttttcaagtagttaaagttttaaagcgt
agctctctt

>Chlamydomonas_mutabilis-atpF

atggaatctttaattaatggagttactataggacacggtggttttggatttaatggcaatatttttgaaacaaatatcataaacttagctgctgttgtaggcatagttgtgacatttgttggaggaaatcttacagc
attactagaagaccgtaaaaaaactatttatataataatttgcaagaagcaaatcaaagagctattgaagctcaagaaaaattaagccaagcacgagcacaattagaatctgctaaaaacaaagcaagagaa
attcgtgaagaaggaatttcaagagcaactatagaaattaataactgtgtaacacaacatgaaattagattagcaagattacaagagtttaaacaagaaactcttcaacaagctgaacaaaaagcttttaaa
caagcgtatatgtatgtaatttctcgtattttaacgagagttcgtgaaagattaaataatggattagattctacttatcatgtagttgttaataatttttacgtatctcgttttactgaatatacgttaaaa

>Chlamydomonas_mutabilis-atpH
atgaaccctatcgttgctgctgcatctgttgtttctgctggtttagctgttggtctggctgctattggtcctggaatgggtcaagggacagctgctggttacgctgttgagggaattgcgcgtcagccggaag
ctgaaggtaaaattcgtggggcgctattattaagttttgctttcatggaatcgctaacaatctatggtttagtagtagcacttgctttactatttgctaaccctttgcaggataa

>Chlamydomonas_mutabilis-atpI_partial
Ttagcaaaaacgcaaatcggcgaagaagaatatttaaaatgggttccctttttaggaactattttcttatttatttttgtttccaactggtctggtgctttattaccttggcgagttttagaattgccaaacgggga
attagcagctccaacaaacgacattaatacaacggtagctctagcgctattaacatcagtttcatatttctatgcaggcatcagaaaaaaggattaggatattttaaccggtatgtacagcccgctgctttct
tgctgccaatcaacatacttgaagatttcactaaaccgctatctttaagcttccgactgtttggtaacatcctggcagacgaacttgttgtaggagtgcttgtggcccttgtgccttttaattataccaatccctatt
atgttgttaggcgtgtttacaagtgctattcaagctttagtttttgcaacattagctggtgcatatattggtgaagcgcttgaagaccaccattaa

>Chlamydomonas_mutabilis-ccsA
atggacaccctattatttttaaattatgcttctagcaatggaagtaccgtagaaatttttttaagaaatttctcttttgctactttgtttatttcaatgcttttttattggatccaaaccgcttttagtactaagattaattat
ggttatggcgattctaaaaataaaaaaaaatatatatttttataaaaaaaaaagaaatctggagaagcagtcatcccgctattagaacaaaaaaagacttctgaattaaatttgtaccccccgtagttgccaacg
ttgttggcacgggggtaagggctaaatttcaaatagtcttaaatttttttttcttatcttaaaatggtcttcgggctcctaaaaaaaaaaaacttgttttctcttttttaacgctgctaaagagaattttggaaatagtact
atgattatttctaatttttttattattgttttttattaatttttcgttggactgaatccggacatttccattaagtaatttgtatgaatctttaatgttttatcatggagttgtacctctattcatcttattttagaatttaaaacag
aaaaagttagacgggaagctaatctttgacaaaattaataggttctataacgtctccgagtgctttttttactaatgctttgctacttttagtttgtcacccgaaatgcaaaaagcgtctccttttagttcccgcac
ttcaatctaattggttaatgatgcatgtgactgttatgataatcagttattctgctttaattcttggttctttattatctattgcttttctgattttgacatcttctaacccctctctttctcccctacgccttattgccgtcc
cttgctacgcttcccctacgggggatcccgtgtccccctccttttataataaagataataaaggggtgcgaagcacgcgaagcacccccttcttgcatgccgaaagcatactacaccctctatgatac
aagcttatacaatatggactacgaaatattagagcaaaaagctactgaaggtatgccacgtcgtaaaggaaatgaaacattaagtaatttagcttggaatttgataatttaagttatagagttatagggattg
ggtttccttttttaacaataggcatattgtcgggagcagtttgggctaatgaggcgtggggctcttattggagctgggatccaaaagaaacatgggctttattgacatggttaatttttgcaatctatttgcatgc
aagaataacaaaagggtggcagggcaaaaaacctgctctaatagcttcatttggttttttagccatctgggtttgtttttgggcgttaatttaataggagaaggtctacatagttatggttggttt

>Chlamydomonas_mutabilis-cemA
atgggttctcccctatcggggatgggtgaggggctgcgtagcaaaaaaaaaaacgctttttcttgtcttgaaagctaagaaaaaaaactataaaggggatcaacatccaatccccctttatgctacaagac
cacctgcttggagtgataagttacccgtgccaacaactacctcctacggaggtaaggcaactacgggggtggaggtggtactagaagaaggggataaagacggtttttggtttgcgaagctttttaaaaa
gctacgcagacaaagcaataataaaaccccccaaaagcaaaccttttatgtatccaaaaaagaaaaagaaactgcaaaagataaacaagttgtttctatgtatttttacatacgaagaaataggtttatttccc
agatcttttagccgagtatttgatagatttgttaaacagttgttctcagatgttgaaaatttggttcttcaagaatatcgtttttatcgatatttatttttaaccacagttaaatgtgtttttattcttttttttgtccccttttg
gttaatttattagctaagaattatttagttcggcccttaacagaatatttttggaatactaaacaaactgaaatctttttaaattcatatcaacaaaagcatgcttttgcctctttacaagattttgaagaaaaactttat
ttcgaatctttagtcttccctaatactcttttgtcagattctaaaaacttctcctctttaaacaaaacctcttcacctccgcaaatagttttttgttcccccctccaataaaaaaaaaaatacccacctccacctttccacc
cccaccttcccacctaccccccaaaggggggcgctacgcgcaaaggtgggggtgggaaaggtggaaaaggtggaggtggaaaaggtggaggtggaaaaggtggaggtggaaaaggtggaggtggaaagg
tggaaaaggtggtaagcaataaaagctttcttttttccaagataaaactgttgaattggctattagttataataaccaaagtattgaagcattacaaatttctttgcagatttaattagttttattacttttatggtattt
attagttgttatggaaattcaaataaatattacaaaatctttttgttagaagttttttttggattagatgatagcaaaaaatcacttttaatttttattggttactgatttattagtgggttatcattcgcctaatatttggga
acttttttccaatttttatttaatcattatggattaccagaaagtcaaaccggtattttcttattagtagccactctgccagttttactggatgttttatttaaatatttaatttttcgccatttaaacagagcatctccagc
aacagttgcaacttatcatgcaatgattgaatag

>Chlamydomonas_mutabilis-chlB
atgaaattagcatattggatgtatgccggaccagcgcatataggaactttacgtgttgctagttcatttaaaaatgtgcatgctattatgcacgctcccttaggtgatgactactttaatgttatgcgttcaatgtt
agaaagagaaagagattttactcctgtaacagcaagcatagttgatcgtcatgttttagcccgggggtcacaagaaaaagtggtagaaaatataacaagaaaagataaagaagaacgccctgatttaatc
gttctaactccaacatgtacttccagtattttacaagaagatttacaaaatttgttgatcgggctgcaattgctacgcaatcccaaagtgatgtgcttttagcggatgtaaatcattatagagttaatgaattaca
agcggctgatagaacattagagcaaattgttcgcttctatattgaaaaagcgaaaaagcaaaataatttgcatacaagaaaaactgaaaaaccatcagctaatattatagggtatttttaccttagggtttcata
atcaacatgattgtcgagaattaagacgtttattgaatgacttagggattgaagttaatgaagttttgcccgagggtggttcggtaaataatttaaaaaatctgccaaaagcgtggtttaattttattccttatcgc
gaagtgggcttaatgtctgctgtttatttagaaaaagaatttgatatgccctatgtagcaacaactccaatgggagttgtagacactgcggcttgtattcgcgaaattagcacgattttaaacaaaatttatata
agccaacctacggttgtacaagcttatacccaacattatgcatataatgataatgaaagagcttgttccccctccaatctataccccaaaaaataagggtataaatgataagcttgtagatgtctcaacggaaa
aaataaattttgagaactatattgataaacaaactcgttttgtatcccaagctgcttggttttcacgctcaattgactgtcaaaatttaacaggaaaaaaggccgttgtttttggggatgccactcacgcagcgt
ctataacgaaaattttagcacgagaaatgggaattcgtgttgcctgcgctggtacttattgtaaacatgacgccgattggtttagagagcaagttgtaggattttgtgatcaggtgttgataactgatgatcat
actatagtaggggatataattgcccgactggaaccagccgctatttttggaactcagatggagcgccatgttggaaaacgattagacattccttgcggtgttatttcagctccagttcatattcaaaacttcc
actaggttatcgtccctttttaggttatgaaggaacaaatcaaatagctgacttagtatataattcgtttacactgggtatggaggatcatctattagaaatatttgggggcatgataataaagaagttataaca
aaatcattgtccactgattctgatttaagctggtcgtctgacggttagcagaattaaacaaaatacaggttttgtacgtggcaaagtaaaacgtaatactgaaaaatttgcccgtcaaaagaatattggagt
tattacagtcgaagtaatgtttgctgctaaagaagcagccggtgcataa

>Chlamydomonas_mutabilis-chlL
atgaaattagcagtttatggtaaaggtggcattggaaaatcaacaacaagttgtaacatttcaattgctttagccagacgtggaaaaaaagttttacaaataggttgtgatcccaaacatgatagcacttttact
cttacaggttttttaattcccactataattgatacgctgcagcagcgcgattatcactacgaggatgtttggccagaagatgtcatttatcaaggttatggaggcgttgatagtgttgaagcgggcggccctc
cggcgggagctgggtgtgggggttatgtggtaggtgaaacagtaaaattgttaaaagaattaaatgcttttatgaatatgatattatttttatttgatgttcttggcgatgttgtttgtggtggatttgcagctccc
ttaaattatgctg

attattgtattattgttactgataatggttttgatgctctatttgcagcaaatcgtattgctgcttcggttcgtgaaaaagctcgtacccacccctttacgattagcaggtttaatcggtaaccgaacagcaaaaaga
gatttaattgacaaatatgtggaagcatgccctatgcctgttcttgaagttttacccctattgaagaaattagaatatcccgtgttaaaggtaaaacacttttgaaatggtagagtctgaaccagctcttcaata
tatttgcgatttttatttaaatattgcggatcaattactaacagaacctgaaggagttgttccacgagaactttctgaccgagaattatttagtttattatcagatttctatttaaacccagtggatcagactaaaaa
aacagaatccgcagaagcattagatttcttattagtt

>Chlamydomonas_mutabilis-ChlN
atgaaaaaccctaccaatactttgcaggttttaaatttaaacacttcgaataatacctcttctttaatttttgaatgcgaaactggaaattaccatactttttgtcctataagctgcgtagcctggctttaccaaaaa
atcgaagatagttttttttcttgttattgggacaaaaaacttgtggatatttttgcaaaacgctttaggggttatgatatttgctgaacctcgatatgctatggcagaattagaagaaagtgatatttcagctcaatta
aatgattataaagaattaaaaagactgtgtttgcaaataaaacaagatcgaaacccaagcgttattgtttggataggaacttgtacaacagaaattattaaaatggatttagagggaatggcccctcgtttag
aaacagaaattggaataccaatagttgttgctagagctaatggactagattatgcatttacacaaggcgaagatactgttcttgcagcgatggctcaaagatgcccagaaaaagtagtttctaactcttctttt
tctgctgaccactctatgcgtagcataaagcatatagatgaaataaaaaaagaaaatgtaagctatgcgaattctttaccccttcttggagggcaactttgtgaccctactacaccctttcttggaggggggaac
ttggaggggggaacccctacttcccctacgggggacgggacgggacgggacgggacgggacgtgtctttcccctacgggggaacgacggtacgagtggtaagcatacagaagtttatatagataaaa
ttagaggggggacattcttcaaataataaacaattggttttatttgggtccttaccaagtactgtaacttcgcaattaacaatggaattaaaaaaccaaggtattttagtctcaggttggttgccgtctcaacgttat
aacgatttgcctgttttaggtgaagaagtttatgtatgtggtgttaatccttttttaagtagaactgcaaccacttttaatgcgccgccgcgaaaatgtaaactcataggagcccccgtttccaataggccctgatgga

acacgtgcttgggtagaaaaaatttgttctatttttgggattgttccaaaaggattagaagagcgtgaagcaaaaatatggcaaggtttagaagattatttacaattaattcggggaaaatctgttttttttatgg
gtgataatcttttagaaatttcattagctagatttcttacgcgctgtgggatgattgtttatgaaattggaattccgtatatggataaaagattccaggccgccgaattggcattattagaaaaaacatgtaaaga
tatgaaagttcctatgccacgaattgtagaaaagcctgataattattatcaaattatacgaattcgagaactccaacctgatcttgcaataactggaatggcacatgcaaaccctttagaagcccgtgggatt
agcacaaaatggtctgtagaatttacatttgcccaaattcacggctttgccaatgcgcgtgatattttagagcttgttactcgacctttaagaagaaaccaaagtcttgaagctttaggttggataaaattggtt
gtttaa

>Chlamydomonas_mutabilis-ClpP
atgccaataggggtacctagaattatttattgttggggagaagaacttcctcctcaatggactgatatttataatttttattttcgtcgccgtatggtttttttaatgcaatatttagatgatgaactttgcaatcaaat
ttgtggattattaataaatattcatatggaagatcgatcaaaagaacttgaaaaaaaagaaattgaaaaaagtggcatttttaaaagtgctggaaaaggtggtaaagaatctgttacaactctaccatctcaaa
atatagctttaagtgaagttccccctggaggggttggaggggttaccccctctaacaaaaaaactaaaaaaagataaatcaatggaagatcttttgagttcttatggtaactctatttatgaagaagacttaact
atagatgaaaactatacaatagagcagtatactttacaaaaaataacattagaatggctaaattggaatgctcaatttttgactattcaaatgaaccctatttatttttatttagcagaaatttttatggaggggac
atctaaaaaagacgatacacaattattttatcaaactagcgcagctaaaactagcgcagctgttgacatgcgaaacatgctatcctcttctaactacgcatcccctcccccttctgtttgctgcgcaacaccta
cggtggcacagtacgcgaagcaaggaggtggaaaggtcttgaagcaggcgttatttgaagctttttctggattacaaaataaaaatttgggacaaatttttgtaccaaaatttaacttggccatcaaataaag
actatggttattttcaagcttggaatgataagcttgctcctttccgctcaatctcctctttggcaaatttaatttctaaagataaaaatttttagtatcaataatttgttaccagagtggagtgttaaagcaaaaatага
taaaaatggcacttctatttttaagcttaaaaggagaatgcgtagcattacaacaatataatatattagaacaaaaaccaggtttggccaagttaattggaggtttggcatggaataaaaaaaaagatttttaa
ataaaaatgtacctcttcaagcttgtacaaccgtaggttggctcttaaaccaactatctatttcagaaaaagttcttggacaaagaaatttaagaagagaatatgccttggataacacttataaaagtgcagct
acacttagctacgctgagccccaaagccaaaacagtcttgtaggatatagctcaaaaaacctttttagctccgctaagctaaaaaatcgtacaaaaaaaacaaagttttttaaaaatataaactatttagatttt
gatacaacaaatgacaatcaatcttatggttttactgactctaaaaacagaaatgactattcaagaaaacaaactaaaaaagcgcttcaggaagaagaatcgaaaaaagttttgttattattaattcatttggg
gggtctgtaggaaatggaattacagttcatgatgcccttcaatttattaaggctggatcgttgacgttgggcttaggtgtggcggcatccgcagcctcattagtacttgccggcggaaccatttcggagcgt
tacgttactgagggctgccacgtcatgatccaccagcccgaaggtggattgaatgggcaggcctcagacatatggattgatagccaggaaattatgaaaattcggctagacgtagcagaaatttattcttt
atcagcgcatcgtccgcgccataaaatattacgtgacttagacagagatttttatttaactgcaacagaaaccatccattatggtttagccgacgaaatcgcgactaatgaagtaatgcatgaaattattgaa
atgactagcaaagtttgggattatcatgatagcaaacaacaacgtttactcgaaacgcgtgacagcgctagttcaggtctggatacacaaacacaaaactaa

>Chlamydomonas_mutabilis-petA
Atgtctaataattttgcactctctttaaatacccgagatacatataaaaagactttgtcttctctttgttcggttttttttggaatttttatttttctaaattgtgatattttgtcagctaaagcgtatccagtttttgctcaa
caaaattatgaaaatcctcgtgaagcaaatgggcgaattgtttgtgcaaattgccatttagctcaaaaaccagcagaaattgaagttcctcaagcagttttacctgatactgtttttgaagcagttattcaaatc
ccttatgatcaacaattaaaacaagtacaagctaatggaaaaaaaaggcgatttaaatgttgggatggttttaattcttccagaaggatttgaacttgctcctccagatcgaattccagaagaactaaaaaaaa
aagttggaaatgtttattatcaaccttatagccctgaaaaaaaaaatattttagtagctgggcctcttccaggtaaaaaatatagcgaaatgattgttccactgctttctccggatcctgctaaaaataaaaatat
tccttatttaaaatacccaatttatttaggtgctaatagaggtcggggacaagtatatccagatggttcaaaatcaaataatactgtttataacgcatctgtaagtggtaaaattatttctataaaaccaggtgaa
aaaaaaggcgcacttgacataacaatagagaaatcgaatggagaaaatttttgttgaaaaaattcctccaggacctgaagttttagttcgtgaaggacaattaatagaagctgatcaaccattaacaaacaa
tccaaacgttggcggttttggacaaaaagatgtagaaatcgttttacaaaaccctgcacgtgttcaaggattattagttttctttacatttgtgcttttagctcttttagttcttaaaaagaaacaatttgaa
aaagtacaattagctgaaatgaatttc

>Chlamydomonas_mutabilis-petB
Atgagtaaagtttacgactggtttgaagaacgtttagagattcaatcgattgctgatgatatcagcagtaaatatgttccacctcatgttaacatttttttattgtttaggtggaattacatttacttgttttttagtaca
agttgcaacaggatttgctatgacttttattatagaccaacagtagccgaagcttttgcatctgttcaatatattatgactgatgtaaattttggttggttaattagatctattcaccgttggtcggctagtatgatg
gttctaatgatggtcctgcatgttttccgtgtttacttaacaggtggatttaaaaaaccacgtgaattaacatgggttagtggagttattatggctgtgtgtacggtttcatttggcgtaactggatactcattacct
tgggaccaggtgggttattgggctgttaaaattgttacaggagtaccagaggctttaccggtaattggtagtcctttagtagagcttctaagaggggtgtcggtgttggtcagtctacattaacacgttttta
cagtttgcatacttttgtattaccttagctacagctgcttttatgcttatgcatttcttaatgataagaaaacaaggtatttcaggacctctataa

>Chlamydomonas_mutabilis-petD
atgtcggtaactaaaaaaccagatttaacagatccagttttaaaagctaaattagctaaagggatgggtcacaacagctacggtgagccagcatggcctaatgatttacttttatattttccctgttgtaatttttg
gaacatttgcttgtgtagttggtttagctgttttagatcctgccgcgataggcgagcctgctaacccatttgcaactcctttagagatcttaccagaatggtatttttacccggtgtttcaacttctgcgaacagtt
ccaaataaacttttaggagttttattgatggctgctgttccagcaggcttaattacagttccttttattgaaaatattaataaattccaaaacccataccgcagaccaattgcaactatttttatttcttttaggaactgt
tgcagctatttggttaggtattggtgctacattcccaatcgatatttctctaacatttgggttattt

>Chlamydomonas_mutabilis-petG
atggttgaacctctattatctggaattgtgttaggattagtacctgtaacaatagcaggcttatttgtaactgcttatttacaatatcgccgtggcgatttagctacttttttaa

>Chlamydomonas_mutabilis-petL
atgttaactattacaagttatattgtattactagttggtgcattaggattcacgttaggtatttatcttggtcttttgaaaattgttaaattaatttaa

>Chlamydomonas_mutabilis-psaA
Atgacaattagttcaccgatccgtgaagcaaaaaaagtaaagattgctgttgatcgcaatccagttgaaacgagttttgaaagatgggctaaaccaggacatttttctcgtactctttctaaaggaccaagc
acaactacttggatttggaatcttcatgccgatgcgcatgactttgacagtcatactagtgatcttgaagaaatttcaagaaaagtatttagtgcgcactttggtcagcttggtattatttttatttggttgagtgg
gatgtatttccatggagcacgtttttctaattatgaagcttggttaagtgatccaacacacattaaaccaagtgctcaagtagtatggccgattgttggtcaagaaatcctaaacggggatgttggtggtggct
ttcaagggattcaaatcacgtctgggttcttccaacttggcgaggtgctggaatcacaagtgaattgcaacttttatagtacagctattggtggcttggtaatggctgcagcaatgtttttcgcaggctggttc
cactatcataaagctgctccaaaattagagtggtttcaaaacgtagaatctatgttaaaccaccacttagctgggttactaggattaggtagtttagcttgggctggccaccaaattcatatttcgcttccaatt
aataaattattagatgccggag

ttgatccaaaagaaattccactacctcatgatttaatgttaaacagaacttttatggcagatctttacccaagttttggaaaaggattagccccttttcttcactttaaattggagtgaatatagcgatttcttaacttt
taaaggtgggttaaacccagttactggtggattatggttaagtgacactgcacatcaccatttagctattgctgtattatttttagttgctggacatatgtatcgtacaaattggggtattggtcatagcatgaaa
gaaatttagaagctcatcgtggtccatttactggagcaggccatgtagggttatatgaaatttaactacatcttggcatgcacaattagctattaacttagctttatttggttcgctatctattatagtagcacat
catatgtattctatgcctccttatccttatcttgctactgactatgcaacacaacttttctttatttacacatcatatgtggattggtgggttctgtatcgtaggtggtgcagctcacgcagctatctttatggtgcgag
attacgatcctactaataattacaataacttacttgaccgtgttattcgtcaccgtgatgctattatttcgcatttaaactgggtttcaatcttccttggtttccattcttttggtttgtatattcataacgatactatgag
cgctttagggcgtccgcaagatatgttctctgatacagcgattcagcttcagccaatttttgcacaatggattcaaaatacacatttcttagcgccacaattaacagcaccaaacgctttagctgctacaagtg
ttacttggggtggcgatattgttgccgtgggtggaaaagtagcgatgatgccgatttcttaggaacatcagatttcttggttcatcatattcatgcgtttactattcacgtaactgtactaatccttctaaaaggt
gttttatttgctcgtagctcgagacttatcccggataaagcaaatttaggattccgtttcccgtgtgatggcccgggtcgcggagggacttgtcaagtttctgcttgggaccatgtgttttttaggccttttctgga
tgtacaatgccttatcagttactattttccatttcagctggaaaatgcaatcagatgtatggggtacagtaaatgcttcaggtgtgtcgcatattactggtggtaattttgcgcaaagtgcaaatacaattaatgg
ctggttacgtgatttcctttgggcacaatcaagtcaggttattcaatcatatggttcttcgttatcagcatatggcttaatcttcctaggtgcacacttcgtatgggcattctcattaatgttcttattcagtggacgt
ggttactggcaagaacttattgagtcaatagtatgggcgcataataaattaaaagtagcaccagcgattcaaccacgagctttaagtattactcaaggtcgtgcagtaggtgttgcccactatcttttaggtg
gcattgctacaacttggtcattctttgaagcacgtattattggtgtaggt

>Chlamydomonas_mutabilis-psaB

atggcaacaaagctgtttccgaaatttagccaaggtctggcccaggatccaactacgcgaagaatatggtacggactggctatggctcacgattttgaaagtcatgacggaatgactgaagaaaatcttt
atcaaaagatttttgcttctcactttggacaactggctattattttcctatggacgtcaggaaatcttttccacgtagcgtggcaaggtaattttgaacaatgggttacagatcctattcatgtacgcccaatagct
catgctatttgggatccacatttcggtcaaccggctgtagaagcatttacacgcggcggtgcctctgggcctgtgaatattgcaacttcaggtgtataccaatggtggtatacaattgggttacgtactaata
tagatttatatacaggatctgtgttccttgctttagtttctgccgttttttatttgctggttggcttcatttacaaccaaatttccagccatctttgtcttggtttaaagatgcagaatcaagattgaaccatcaccttc
aggtttatttggcgtaagttcgttagcctggacaggacatttagtacacgttgctattccagaatcacgtgggcagcacgttggctgggacaacttttattactgtattgccgcatccactagggttaactccttt
ttggactggaaactgggcggcttacgcgcaaaacccagattctgcagctcacgtattcgggacttctgaagggtcgggttctgctatttttaacttttttaggaggttttcatccacaaactcaaagtttatggtt
aactgatatggcacaccaccatttagcaattgcggttatttttattgtagccgggcacatgtatcgcactaactttgggattgggcaccgtatgaaagcaattctagattctcatgttccgccgggcggcggc
ctaggtgctggacataaaggcttatttgagactgtaaataactccctacatttccaattagggttagccttagcttctgtaggtacaattacttcattagtagctcaacatatctatgccttacctccttatgccttt
tagctattgattttactactcaagcttctctttatactcatcaccaatatattgcaggttttattatgtgtggagcctttgcacatggagcaatttttctttattcgagattatgatccagatcttaataaaggaaatgtttt
agcacgaatgctagagcataaagaagcaattatttcacatttaagctgggtttcgttattcctgggctttcacacgttaggcttgtatgttcataatgacgttatgcaagcttttggaactccagaaaaacaaatt
ttaattgaacctgtttttgcacaatggattcaatcagcacacgggaaatctctttatgggtttgatttattattatcttcttcatctagtccagctgcatctgcaagtcaaagcctatggttacctggctggttagat
gctattaataataaccaaaactctctttttcttaactattggcccaggtgacttcttagtacaccatgctattgctttagggttacacacaactacttttaattcttgtgaagggagctttagatgcccgtggctctaa
gttaatgccagataaaaaagattttggttacagcttcccatgtgatggtccaggccgtggtgggacatgtgatatttcggcttatgacgctttttatttagctgtttttggatgcttaatacaattggctgggtaa
cattctattggcactggaaacatttaacactgtggcaaggtaacgttgcacaatttgacgagtcatcgacttatctaatgggatggttacgagattatttatggttaaactcatcgcagctaattaacggatata
acccattcggtatgaacagtctatcagtttgggcttggatgttcctatttgggcatttaatctatgctactggcttcatgttcttaatctcttggcgcggctactggcaggagcttatcgagactcttgtttgggct
catgaaaaaactccattagcctctttagtttactggaaagataaacctgttgcactttctattgtacaggctcgtttagttggtcttgcgcatttctctgttggatatatctttacctacgcagctttccttatcgcttct
acatcagggagatttggataa

>Chlamydomonas_mutabilis-psaC
atggctcatttagttaaaatatatgacacatgcattggctgtacacaatgtgttcgtgcatgtcctttagacgtattagaaatggtaccttgggatggttgtaaagccaatcaaatggcctcagcgcctcgtac
agaagattgtgtcggctgtaaacgctgtgagacagcgtgccctactgatttttaagtgttagagtttatttaggctcagaaagtacgcgaagtatgggattagcttattaa

>Chlamydomonas_mutabilis-psaJ
atgaaagattttacaacttatttatcaacagctccagttgttagtttagcttggctagtattaacagcggttttattaattggttttaacaaaatatttccggatcctcttgtttttactttctaa

>Chlamydomonas_mutabilis-psbA
Atgactgctattcttgaaagacgtgaaaattctagcctatgggctcgtttctgcgaatggattacttctactgaaaaccgtatctacatcggttggtttggtgtaatcatgattcctactctattaactgctacttct
gtatttatcatcgctttcgttgctgctcctccagtagatatcgatggtatccgtgaaccagtttctggatctttacttttacggtaacaacatcatcactggtgctgtaatccctactagtaacgctatcggtcttcac
ttctacccaatttgggaagctgcttctttagacgaatggttatacaatggaggcccttaccaaatgatcgtatgccacttcttcttaggtgtatgctgctacatgggaagagaatgggaattatcttaccgttta
ggtatgagaccatggatcgctgttgcttactcagctccagtagctgctgcaactgctgtattcatcatctaccctatcggacaaggttctttctctgacggtatgccttttaggtatttctggtactttcaacttcat
gatcgtattccaagctgaacacaacatccttatgcacccttccacatgttaggtgttgctggtgtattcggtggttctttattctcagctatgcacggttcattagttacttcatctctaatccgtgaaacaactga
aaacgaatctgctaacgctggttacaaatttggtcaagaagaagaaacatacaacatcgttgctgctcacggttactttggtcgtttaatcttccaatacgcttcgttcaacaactctcgttcattacacttcttct
tagctgcttggccagttatcggtatttggttcactgctttaggtatttcaactatggctttcaaccttaacggtttcaacttcaaccaatcggttgttgactctcaaggacgtgttttaaacacttgggcagacgtg
cgaggtatcatcaaccgtgctaacttaggtatggaagttatgcacgaacgtaacgctcacaacttccctctagacttagcttctgttgaagctccagctgttaatgct

>Chlamydomonas_mutabilis-psbB
atgggattaccttggtatcgtgtacatactgtagttattaatgacccggggcgcttaatttcagtgcatttaatgcacacagctcttgtagctggttgggccggctctatgacactatttgaaatcgctgtttttga
cccttctgatcctgtacttaacccgatgtggcgtcaaggaatgtttgtacttccttttatgactcgtctaggtattacacaatcatggggcggctggactatcagtggtgaaacagctacaaatccaggcattt
ggagttatgagggcgttgctgcgtctcacattatcctatcaggccttctatttttagcttctgtctggcattggacatactgggacttagaacttttccgtgatccaagaactggtaaaactgcgttagatttacc
aaaaatttttggaattcatttatttttatctggccttctttgttttggttttggagctttccacgtaactggcgttttcggaccaggtatttgggtatctgatccttatggattaacaggaagcgtgcagccagtagcg
ccgtcatggggatcagatgggtttgacccttataaccctggtggtatcgcggcgcaccatattgcggcaggtattttaggcgttcttgcggggcttttccacctttgcgttcgtccgtcaatccgtctgtatttt
ggactatcaatgggtagtattgaaacagttctttcaagcagtattgctgcagtatttttgctgctttcgttgtagcaggaactatgtggtatggctctgctgcgactccaattgaacttttatggtccaacacgtta
tcaatgggatttaggcttttccaacaggaaattcaaaaacgtgttcaaactagcttaactgaaggttcttcgttaactgacgcttgggcaaaaattcctgaaaaactggcttttatgactatattggaaacaa
cccagcgaaaggtggacttttccgtacaggtgctatgaacagcggtgacggaattgcagttggctggttagggcacgctgtgtttaaagatcaagacggccgcgaattatttgtgcgtcgtatgcctacg
ttctttgaaacttcccggttattttaattgataaggacggagtagtaagagctgacgtaccttccgtaaagctgaatcaaaatatagtattgaacaggtaggtgtatctgtaacattctatggtggcgaattag
atggtttaacttttactgatccagcaacggtaaaaaaatatgctcgaaaagctcaattaggagaaatttttgaatttgatcgttcaactttacaatctgatggagttttccgcagtagtccacgaggatggtttac
gtttggtcatgtatgttttgcattattatttttctttggtcatatttggcatggtgctagaacaattttcagagatgtatttgccggtattgatgatgatgtgaatgagcaagtcgaatttgggaagtacaaaaaact
ggcgatacttcatcacttcgtgaagctttt

>Chlamydomonas_mutabilis-psbC
atgggagacgtaagaattattacaaatccaacaacaaatgctgcagttatttttggatatttacttaaatcacctttggaggtgacggttggattgtaagtgttgataatatggaagatattatcggcggacac
atctggatcggcacactagaaatttaggggggtatctggcatatttatacaacgccttggccttgggcaagacgtgcctttatttggtctggcgaggcatatttatcgtatagtttagcagctatttctttaatgg
gctttactgcttgctgtatgtcatggttcaacaatactgcttacccaagcgaattttacgggccgacggggcctgaagctagccaatcacaggcatttactttcttagtaagagaccaacgtttaggggcaa
acgttgcttccgctcaaggtcctacagggttaggtaaatatttaatgagatcccctactggagaaattatctttggcggtgaaactatgcgtttctgggatttccgtggcccttggttagaacctttaagaggtc
ctaacggtttagatttaaataaactaaaaaatgatattcaaccgtggcaagaacgccgtgctgcagaatatatgacacatgcacctcttgggtctttaaactctgttggtggagtagcaactgaaattaacgc
agttaatttcgttagtccacgttcatggttagctacttctcatttctgtttaggtttcttcttctttgtcggtcacttatggcatgctggccgtgctcgtgcagcggcagctggttttgaaaaaggaattgacagattt
gatgaaccagctctttcaatgagaccattagactaa

>Chlamydomonas_mutabilis-psbD
atgactatagcgattggaacatatcaagagaaacgtacttggtttgatgatgctgacgactggcttcgtcaagaccgtttcgtattcgtgggctggtcaggtcttttattacttccttgtgcttacttagctttag
gtggatggttaacaggaacaacatttgtaacttcttggtacactcatggattagctacatcatacctagagggttgtaacttcttaacagctgctgtttctacaccagcaaacagtatgggtcactcattactatt
tgtatggggtccagaagctcaaggtgattttacacgttggtgtcaacttggcggtttatggacttttgtggccttacacggcgcttttggtctaattggcttcatgcttcgtcagtttgaaattgcacgttctgtta
acctacgtccatataacgctatcgcgttttctgcaccaattgctgtatttgtttctgttttcctaatttacccgttaggtcaatcaggttggttctttgccccaagctttggtgtggcggctatttttccgtttcatcttatt
cttccaaggattccataactggactctaaacccattccacatgatgggtgtagctggtgtacttggagccgctcttctttgtgctattcatggtgcaactgtagaaaatactctattcgaagatggtgatggtgc
aaacactttccgtgcatttaacccaacgcaagcagaagaaacatattcaatggttacagcaaaccgtttctggtcgcaaatttttggagttgcattctcaaacaagcgctggttacacttctttatgttattcgtt
ccagtgacaggactttggatgagtgctatcggtgttgttggcttagctttaaatttacgtgcttacgatttcgtttcacaagaaatcagagcggcagaggatcctgagttcgagacgttctacactaaaaacat
tctacttaatgaaggtattcgtgcttggatggcagctcaagaccaaccacatgaaagactagttttcccagaagaagttttaccacgtggtaacgctctt

>Chlamydomonas_mutabilis-psbE
atggctggaaaaccagtagaacgtccttttttctgatatttttaacaagtattcgctattgggtaattcacagtatcacaattcctgcattatttattgcgggatggctttttgtcggaacagggctagcttatgacgt
atttggtagtccacgaccaaatgaatatttttacagaagatcgtcaagacgctccactaattactgatcgttttaatgctttaaatcaagttaaaaaattatctcaacag

>Chlamydomonas_mutabilis-psbF
atgtctacaaaagctgaaactattacatatcctatttttactgtacgttggctgtctattcatgctttagcagttccaactattttctttttaggtgcaattactgcaatgcaattcattcaacgttaa

>Chlamydomonas_mutabilis-psbH
atggcagtttttatgctccttttttgctgtgtttttacttattatccttgaaatttataacagttctctgattttagacgatgtgactatgagctgggaaactttaggcaaataa

>Chlamydomonas_mutabilis-psbI
atgctaacacttaaaatttttgtttatactgttgtaacatttttttgtatgcctgtttatctttggatttttgagtaatgaccctgcgcgttcacctggtaaaggtaac

>Chlamydomonas_mutabilis-psbK
atgtcttattacgctttttctattctacttgcaaaacttccagaagcttatgcacctttttgcacccattgttgatgtactgccaattattcctgttttatttatttttattagcgtttgtttggcaagcttcagtaagttttagataa

>Chlamydomonas_mutabilis-psbM
atggaagtaaacatttttggattaacagcaactgctttatttatcttaattccaacttctttttttattgattctatacgtaaaaacagcttcaacttctccagaagca

>Chlamydomonas_mutabilis-psbN
atgggaaattctgccttttttttttaccttttttttatggtttcttctgttaagcgtcactggttattccgtttatgtaagttttggcccctccatcaaaaaaattaagagatccgtttgaagaacacgaagattaa

>Chlamydomonas_mutabilis-psbZ
atgacatctattcttcaacttgctttatttgcattaattttagtttcttttggcttagtagttggcgttcctgtagttttcgcttcaccaaatggttggacagaaaacaaacgttttgttttttcgggtttaagcttatggcttctacttgtatttacagtgggtgtatttaactcatttgtagtgtaa

>Chlamydomonas_mutabilis-rbcL
atggttcctcaaacacaaactagggttggtgctggttttaaagctggtgttaaagattatcgtttaacatattacactcctgattacgttgtaaaagatacagacattcttgcagctttccgtatgactcctcaagcaggtgttcctgctgaagaagctggtgcagctgttgctgctgaatcttctacaggtacttggacaacagtatggactgatggtttaacgagccttgaccgttacaaagggcgttgttatgatatcgaaccggttgctggggaagaaaatcagtacattgcttacgttgcatacccaatcgacttatttgaagaaggttctgtaacaaacttatttacatcaattgtaggtaacgttttcggtttcaaagctcttcgtgcacttcgtcttgaagatcttcgtatttctactgcttactctaaaacattccaagggcctccacacggtattcaagtagaacgtgacaaattaaacaaatatggtcgtggtctttttaggttgtactattaaaccaaaattaggtctttcagctaaaaactatggacgtgcagtttatgaatgtttacgtggtggattagactttactaaagatgacgaaaacgttacttctcaaccttcatgcgttggagagaccgtttcgttttcgttgctgaagctatttacaaatctcaagcagaaacaggtgaggttaaaggtcactatttaaacgttactgcaggaacttctgaagaaatgttaaaacgtgctgaagtagctaaatctttaggtgtacctatcatcatgcacgattactaacaggtggtttaacatctaacacttcattagcacactactgtcgtgataatggtttattattacacattcacagagctatgcacgcggttattgaccgtcaaagaaaccacggtatccacttccgtgttttagctaaagctcttcgtatgtctggtggtgaccaccttcactctggaactgttgtaggtaaacttgaaggtgaacgtgaagtaactttaggtttcgttgacttaatgcgtgatgactacattgaaaaagatcgtagccgtggtatttactttacgcaagactggtgcggattacctggggttatgcctgtagcttctggtggtattcacgtatggcacatgcctgctttagtagaaattttcggtgatgacgcttgtcttcaattcggtggtggtactttaggtcacccttggggtaacgctccaggtgctgctgctaaccgtgtagcacttgaagcttgtactcaagctcgtaacgaaggtcgcgatttagctcgcgaaggtggcgatgttattcgttctgcatgcaaatggagtcctgaattagcggctgcttgtgaagtatggaaagaaattaaatttgaattcgatacaattgataaactg

>Chlamydomonas_mutabilis-rpl2
Atgggaattcgttttcttcaagcttatacaccagggactcgaaatcgttcagtttctgattttagtgaattaacaaatacaagacctaatagctcgttaacatctggtttacaacgagcaaaaggacgtaataatagaggcattattacttgccgtcatcggggaggtggccataaacgactatatcgtttaatagattttcgacgtgataaaattggtatgggtgcaaaagttgtaacgattgaatatgatccgaatcgaaatgcgcgaattgctcttcttcggtacgaagatggcgaaaaaagatatatattacatccgcgtgggctaagtgttggtgaaaatattatttcagatataaatgctgctataattgtaggaaattcacttccattacgcaatattccattaggggcagaaattcataacgtagagtttcaacccggttctggagggcaaatagcccgatcagcgggatctgtagtagagattttagcaaaagagggcaattttgtaacgctacgtttaccttccaaagaaatacgtttagtctcaaaaaattgttgggcaactataggtcaagtagggaatattgaagcttacaatttaacaataggaaaggcaggtcggacacgatggctaggaattcggccaacggtaaggggttctgttatgaaccctgtggatcaccctcatgggggtggtgaaggccgtgcaccaatcggccatagtcgaccattaactccttggggtaaacctgcgttaggtgttttaactagaacaccgaaaaaatatagtaatatttttattattcgtaaaagaaaa

>Chlamydomonas_mutabilis-rpl5
atgacacaaagattaaaaacatattatattgagacaattattccaaaattaaaaaataaatttaaatacgaaaatattcaacaagttcccaaaattgaaaaaatcgtaataaatagaggaattggtgcagcttcgcaaaatcaaaaaattgttgattcatctttaaaagaacttgcgataattgccggccaaaaaggaattatcacaagatcgaaaaaagcgatagcaggatttaaattacgagaacaaatgcctgtaggcgttgttgtaaatttaagaggcaatcgtatgtacggtttcttgatcgattaataaacttagcgcttcctcgtgttcgcgattttcaaggaattaatcccaaaagttttgataagcatggcaattatagtttaggtttagaagaacaattaatgtttcctgaaatagaatatgacaaaattgatcaagtacgaggaatggacatttcaattgtaacaactgctacaaaacaggcagaagggttagctcttctaaaagaatttggtttacccttaaagcttaa

>Chlamydomonas_mutabilis-rpl14
atgattaaacctcaatcttatcttaatgtagctgacaatagtggagcacgtaaacttatgtgtattcgtgttttaggtggaagtcaacgtgaatcaggaaatattggagatattattattgcagtagtaaaagattctattccaaatatgccattaaaaaaatcagatgttgttcgagcagtaattgtccgcactagcaaaggtttaaaacgtgataatggtatttcaatccgatttgacgacaacgctgctgtgataattaacaaagaaggaaatcctagaggcactagagtatttggtcctatcgctcgagaattacgagatcgaaactttactaaaatcgtatcattagctccagaagttctataa

>Chlamydomonas_mutabilis-rpl16
atgcttagtcctaaaagaacaaaatttcgtaaacagcatcgtggtagattaaatggaaaggcaactcgtgggaatgtaatcgcctttggtgattttgctttacaagcattagaaccttgttggctaacatctagacaaattgaagctggaagacgtgtcttgacgcgttatgttcgtagaggtggtaaattatggatcagaatttatccagacaaaccagttacaattcacccagctggaactcgtatgggttccggtaaaggatatccagaatattgggttgctgtagtacatcctggaaaaattatttatgaaatgaaaggcgtttcggaaataattgcaaaacaagctcttcgtattgctgcgtataaaatgccagttaaaacgaaatttttaacaaatcaaactatttaaag

>Chlamydomonas_mutabilis-rpl20
atgactcgtgtaaaacgtggcaatgtatctcgaaaacgtcataaaaaagtattaaaaatgtcaaaaggatttcgaggagcgggctctattttgtttagaacagcaaatcaacaaaatatgaaagctttaagatattcttatcgaaatcgacgtcaaaaaaagcgcgactttagacgtctttggatagcacgtgttaatgctgctgttcgttgttatggtctaaattatagtgaatttatgaattatttacaaaaacgcaatattaaattaaatagaaaaataatagctcagttagcttcatctgatactggtgccttattcaactactattattttaa

>Chlamydomonas_mutabilis-rpl23
atgattgatttaataaaatatcctattataacagaaaaatcttatttagctttatttaagaatcaacaatacacttttgatgtcgatcttagactaagtaaacctcaaataaaaaaattattctttaatttatttaatgtaaatgttatcgctgttaatacgcatatacctccacgtaaaaatatgcgtgttgggacaaagagaggatatcgagttcgatataaacgagctattttgactttaaaaaaaggccaatccttaaaattttctataccattaacaggtttcggttctactgctagtacaaccgtaggt

>Chlamydomonas_mutabilis-rpl36
atgaaagtacgtgcatcagtaaaagctatttgtgataaatgtcgtgttattcgtcgaaaaggaacagtaatggttatttgctcaaacccaaaacacaaacaacgtcaaggaacg

>Chlamydomonas_mutabilis-rpoA
atgataaaaattgttaaaaatgatttttttatttattgcaaagaagcacgtattgaaaataatcgaagtttttatggttcttttagtttaggtccttttgatgctggtcaaagtattacaatagcgaatgcattaagacg
aacattattatcagaattaaaaggtttagctattacttctgcagaaattgaaggagcttttcatgaatactcaacacttcctggagttcgagactctatattagatgttttattaaatttaaaagatattgttttgaaac
gtaactctcttttatgggaataaaaaaaaggggcacctctgagtctgcaacttttaaaactgaacaaatcggatatttgaaaaatagaggacctggggtagttacggcttcagaattaaaactacctccaggaa
ttcaatgtgtcgatccggatcaatatattgcaactctttctgaagaaggagttcttaatattaaatttttatataaatgaaggtaaaaattatgttattcaaagtccggaaaaaaaaaacttctggtcttaaagggttat
ggttaaaaaaaaataaagggaattcaatctagattaaattggtgcgtagctaaaaaaaaaaacaagctacgcaccccttttagagagaataactatataaaaaaatattacaacaacaaaaccttctatatgcttg
tttcccgtccctcttgcagcaagaaggggttccccctccaaggaggggtaacaaagaaaaaagatccgtgccaacaacgttggcaactactagccaaattaaaaaagaaaaaaaagcaaaggtaagg
caacaaaactttttagatttcaataacaaaataatagcaagcggagcctggctatccaatcctttagctatagatgctgttttttatgccagttaataaagtaaattatattattgaaattaaccagcttcgctgcttc
ggagccaaagataataactcctctctaaaaaaagacgctttttttagcacccctaaccccctctttgccacctacggtggcacgggggttaggggtcttaaatcgcacccctttcggggtacaaggggtta
aaaagaaaaaaatagaattgccaaaagcgccactgttttttttttccacggctaccacctttccaccccccacctttacgtgccacaaagtggcaacttcccctttattataaagcgcgtagcaagacaccct
acgggtgtcgggacgccacaaagtggcaaccaaggggtaggtgggaaaggtggaggtggtagagcctatttttttttggaaaaaagaaaaagatatattatataagttatctttttttttgtcgcaaagaag
cacttttctaccttcttacccgtccacctccccctactacaccctacttggaggggggaacttggaggggggaacttggaggggggaacttggaggggggaacttggaggggggaacttgcttgtttcccgtccc
tcttgcagcaagaaggggggaagtatgcgtagcatataacaaaaaaaaatatgctacgcatacacgggttgcgcagcaagggtcttgtagcaagggacgggacggggcgggtcttgtagcaagggac
gtgtctttggaggcaagcccctacttgctgcaagaggaacgaaaaggtggaagtaggggttagcccagcgtaagagatataagctgggctggtttaaaaatttttttagctcctttttttattaaaagcaaa
gttctaaaacctttttctacccttccgtttgcttatcactccaagcaacaccgtaggtggcaatacaagccaacctacggttgtacaagcttgtaaaaggggacagggtataaaaaagagttttaaggcaaca
ttgtttttttaagcgtattggaaaaaaaaaactaccacttacggtagtagggcaaaataagatagaacaaaaatttggtacaaatttttgtaccaaatttttatttttatcatcttatgcttttaaagccgtccctcttgtag
caagaaggggggaagcgtagcagaaagcttataaaaaaaaaaactttatttttacaagcttgctttacccttttaaaaagctacgcttgcatgcctccgaaggaccccgtgcccccgaaggaccccctactag
ccaaactccttttaaaaaagaaaaattatttttttagggtagcagggggaaagcggaaaggtgggggtggtcttgaagcaagaaggggggtgcttgcttccgaaggaaaacaaggttatattagtatgtcgtgtt
cactacaacccccctgcttgttttgctcccaccgtaagctacgcttatagaggtggtaaagtcttttttttttaaatctaagtaccctttcttgccactcttaccacctagagtgataagatacccccccaaaatataaag
gggggtaggtggtcttgctacaagacacgtctaaatttaaacacgcaaaccgtaaaaagaaaaggtttgtttagcagtagtactgctaaaaaaaagcagtactacaaataaagaatataatttttaaaagcttg
tatctaccccaaaggggtaaagctttttatttatacttaagaacaatatttttgttaattccttagatataagctccgcttatacttcttctttttagaaaacatatctctataaacaatttgtatacctccccttttattata
aaggagggcaacttttgttaacctactacatcctacttcccgtcccttgcgtagcaagtatgccccccccttacagggaaggcatacaagcgcagcaagtaggggggacgtgtctttggaggcaagccccta
cttggaggggggaacttggaggggggaacttggaggggggaacttggaggggggaacccctacttcccctacttggaggggggaacttgcttggagtgataagcaagggacgggacgggacgggacgg
ggtcttgtagcaagggacggaacgacgccacaaagtggcaaccaaaataaaacaaaaagtatattattgcttcaagatcaaaccaactccaatcttgctttttttaaaaaagcaacccttaccccctactctt
tttcaaattttaagcgtagctaataaaaaaaaaacaaaggttagaagcttagataaaggagttgcaccacctacggtttgctctacccgtgcctccgaaggaccccctactagcttagattttaaagcaagg
gcaaacgttggcaactacgggggtagtagtgagcacgtagggatacaagcaaaaacatccctaaccccccttggttttgctacaagacacgtccacccccacctttggaggggggaacctacaagc
ttggaatgataagcttgtagaaggtgggggtgggaaaggtggcgtccccctactacacctacgggtgtcgggaagggtaagggtttaaaaaaataatattattttagagatttggacaaat
ggcagtattcatccacgagaagcgctttatgaagcttgtaaaaatttatctagtctttttttcaaaattagaaaaaacaaaatttattatatctatgtataaagaaaaaacatcttataaaaagcaaaactttttagag
cccatatccagcgtttgtgctcacaacgtgagcaaaaccgtaggtggacatgctttaaaagtagtaaccttcttacccccttctttgccacctatggtggcacgggttagagacaagcagggtaaaaaaaa
gaccaaaataaatctctgcaaagtttcttgtgccaccgtaggtggcaaaactgtaactgccccccttttagagttgcg

>Chlamydomonas_mutabilis-rpoC2
atgcaatttaaaaaagaaaaaaaaaaatttatatttataaaccaagggcttatcactccaagcatattaaaaaaaacttttgttttttgtaccaatttaaaaaaaattaggttaaataattatctttttgtaaaataaa
aaaaaacataaataatcaaattctacccatacacccctacacccttcttgtgcgaagcaacaccttctacaagcttcgcaggtcaggcacgtgtatgggtgtagggtctaaaataaaaaaataataattttttg
ggctttataaaagttgttttaccgttcagcgattagagtacgctgagcgtataaaaacaaataaagctttaacgttgtataatcaagggcttttagcaattagggacgtgtcgggacgccacaaagtggcaa
ccaaggaggggggaacaaaaaaaaccctttctttgcctataagcttgttcccccctccaatggtaagcttgtactacggtggcacggataagaatatggatttttaaggtaagaaagaaaaacttagataaaata
actacatataatttttgtaaaaaaaataattaaagaacctttttggaaccaatgttttgataaaaatcgtttaaaaaatttttgtattatggtttttatctacttatggcgaatatagaacagtagaattaatagaacaattt
aaaaatattggttttgaatatgctaccaaagcaggtatttcgttaggaattgatgatttaaaaattcctccaaaaaaaatccttttttaatttttgacgcagaaaaaaaatactaattcggcactatatcaatataatcgc
ggcgaaattacaggtgtggaacgttttcaaaaaatgatagatacctggcatagaacaagtgaaatttttaaagcaagaagttattgattattttgaattaaccgatacgttaaatcccgtttatatgatggcctttt
cgggggctcgcggaaatatttcacaagttcgccaattagttggtatgagagggttaatggccaatccacaaggtcaaattttagactttccgattcgaagtaattttcgtgaaggattaagtttgactgaatat
attatttcaacatatggagcaagaaaaggaattgttgatacagctttaagaacagctaatgcaggctatttaacgagacgattagttgatgttgctcagcatgttattgtttctaattttgactgtggaactcgca
gaggtatttttttaaatgatatgaaagaaggaataaggttttatattcacttcaacatcgattaattggtcgagttttaggaaaagatgttttagtagtatacgcagccaacctacggttgtacaagctgaatc
gtttacaaaattagcatccccccctacttcccgtccccaaaggggaagtaggggacgtgtctttggaggcaagccccttacttggaggggggaaccccaaatataaaggggacggaacgaaaag
gtgttcccctccaagggacggaaggccactatgggggtagtcttgaaaggtcctatatataatagaaatttagaaatttcaaaagatttagcaatgaaacttttgtaaaatttcaaataaaatatttgttcgttc
aacgttaacttgtgaaactcataaatcagtttgtcagttatgttatggttggagtttagctcaaggtaatcttgtttctataggagaagcggtgggcgtagtagccgcccaatctattggcgagccgggtaccc
agttaaccatgagaacatttcatactggaggagtatttttctggtgatattaccgatcaaattaaagctctttataatggtattgtacaatatcttgaacctatagcgggcactttaattcgaacaccggaaggca
aaattgccttttttaaccaaaactgaaggctttatagaagttatttctaattctatttcttttagtgacgtcacccactacacaaaataatcagactcagttgtccttcaaaacctatgccttcggcatgcaaagggg
cgcccgtgtcgggacgccacaaagtggcaaccaaggggttaaaaaatgtctgcgtagcaaaaaaaatttataaaattccgtcttatacttttattatttgttcgtcatggagaaaaagttttgaaaaacaggta
attgctcaaatttctaatatttctcgacaacgaaacaaaagagatgacgcagaatttataataaattctgaaatagaaggagaatttttattccattaaattaaatttattagaaaaacgtgttggtccaaaacccc
catctacttctttaagcagcgtcagtgctctaccctacaagcttcgcttgtcgccaacaacgttggcaactacgggggtaaacgtgagcaaaaccgtaggtggacatgctaaactaaaaactttgcaaaat
aatgttaatagcacaatgcaagttaaacatgaagcaagcgactggggatacgcatggatttggcaggcaaaatttatcaacttcctgtcttatcgcgatttttcctcaaccgggcgattttattaatcgaaa
aagtttaatgaatcaaatccaattcaaagcaacaaaccatggaggtaatgttgggttttagagacaaaatttttattgctacgcacccctaccgcctacccccaaatataataaaggggtaacatcgga
ggtgttgaaaggcggtaaggggttggtgccactatgtggcaaatcttcttttagcagccacgcagctaaaactaagct

accaagacaaagtatttataagctatgctttaaaaagcaaagcgttttaacagcatataaaagcttagataaaaatatagattatgggagcctaaacaatatttttactaaacaaaccagttttatttttagaaatt
aataaaatatttttttaaaaaaaataggttatttttttaacattaccctcttctccttttttaaactatatttttttatcttctttttttttataatataaattctctttttgtttattctaaactatctaaactctctaaaaaaaaaaaagc
agattttatttcagaactatacagagctaaaaaaatataaagcttgcatccccttcttgccaacttctacccttgcttttttttctttttttaaattggcaagaaggggggtccttcggaggcgacaagcgaagcttgta
gggaaaggggggccttaccactctaaggctcttgtccatgcaacaggggggcaggaataagcatagatgcacaattattagcacgggctactaatagaatttttaaactaagaagaaaaccagtgttttt
tcttctgccttacctacacttccccttcttgccaaaagaagggggtgcttcgcaaccccctccaagaagggggtgccaaatataaaaaggagaggaaagcaaccccccccaatatgctacgcatacaa
gaaggggttcccccttgcatgccgaaggcatactacacccaatcccacctctggtgggaaagggtgttgggacgccacaaagtggcaaccaagaagcatatagaggtggcctaaatttttaaaggagct
agtaaggaatggacagggaatacgtatgacttgtctcagcgtagccacttggcagcaatttttttaaccttaaagttttaaggcaacaaaaattatttataacttccaatgatttattatttacacccacgtctttc
gataagagcacttctttttcccctttgggagagaaaaaaagctacgcaggcctaaacaatctcatttttagcggcgctaaaaatttttccgcctatataataataagcctctgccccccttttttgctaaaaaaaaaaag
tacgttccccttgcatgccgaaggcatactacacccctacgggtgtcgggacgccacaaagtggcaaccaaaggagcttagcgatagcgcctctttctaatttagattatataatagtacgaaatttaaaacaa
gttctgcaaattggagccaaaatcttagtttatttatggaatggtttccgtcccgtgccactatgtgtcttgtagcaaaagattttccaagacaaagaaatgaaatatatacgggcggtgcctttatatgtgattc
taaaacttaaaagcaatatcgtatatacctctatgtggcctactaaaattacgttatttaaaagcagcgcagcgacttttgtcttctatatataatttatttttcttataaacgatcttttcgtataaattcctttataaatt
taataagctctttttataacaaaaaacaaacaaggagaggttcccctctaacaaagaaaaaagaccctttaaaaagctacgcttgcgtgccaacgttgttggcaagaaggtaacgccaacaactactag
cttcaagaccactaccccccttctctttgggggtaacaccggaggtgttgaaaggtggtagggtaaggcaacaatttttaaaatttttaaatacaaaatactaaaacaaaccgtaagaacaaatttttacta
atgtacaaaaacgtttttattcatttaatatttctctagtctttttatctcttaatagtttagttgttttttttagatctacccgaagggtatctacggttgtacaagcttgtaggaagagacaagctaaaatatataattat
agttttttatttatttttgttaaatttaaaacatagcatggcttgcctcgaaaaaaatactttaactattagcgaaactaaagggttaaaaaataatttttataagcaaagcttatatcaagataattgtccgtcccgtc
aagctaaaaacctagggttttttatacaagcttgaagaagggggggtgcttgttccccctcagagtatacccgtagggtataagaggaaagcgctacgcaccccctccaagaaggtttagcttatcactctta
gctaaaaaaagacaagaaaatataaaaagtaaaaagaaacatggccgtaatacgccttttataaaaaaaaaatatatttagttttgcacaaagtcttaacttctattcttcaattatgaataggcatccccttcg
agattaaatttcgaaaaagctaacaattttagaaaatataatttttgtcttaaatttaaagttctatttataagtaaatcttatataaaaatagtttactaaaaaaaaaactatgcttataattctaagcatactttggagccg
tttataaaagcaaaaagttacttaccttttaccacctttttataccccctaaaaatttaaaggtattaaaaatttttaaggcaagctcgtttttaccagttaaagggtatggtaattctgtaaaaagtaaaaatatacagt
tgtttatccacatgaaaaaaaaaccttattatgatccatcgttgatgtttgcacaaagagaaaagctcttaaaaaattcggtgcatgataaactttttggatttctagacaattttatgaaatacccaaaagaacgt

ttttgtggaggggaaaatatagttttgatatggagttctggaaagatcataaagttctaccatgtctaaacaccagtatgcttttatatgcttgttccccctcctaccccctacttgctgcgcaagggttaggg
aagcaaaattttggtacaaaaagcgtagcttttaataagacagggattcacgcggagctagataaaaaaaatcgccaaggaaaaaaacaaccaatttatctctctttagcgcagctaaaaaaagataataa
ttttctgcttaccactctatgcgtagcataaagcatatagatggattgattcatattgctccaataattagacctacgcccttaaaaaaaaaccaaataacaaaatctattacacagccaataatggtatcgcct
accaccggaagtggtaaggggatattagctataccacccttttgcaaaacaggtggtactagaagatatctacccccttcgggggtatccgaaggataaaaaaagagcaaaaaaaaggttttcgtcttggct
taaaaaaaagaaacagtttgtttaaagcttactctattaaaaaacttaaaactcctatttcgaatactcactttagtatgggtcttacaaagcaaggattagccaaaaatctgctcaaaaaaaatacttttagtatt
gtatggaatatccctgttttgttccccctccaaaaaaagaaaccttttcataacctacaccctctcttgcctaccacctatccccaaaggggggcgctacgcgcaaaggtggtagaagtggcacgtgtaag
ggtattaaagagttttaaggcaacatttaaaaataatagaggcaagcgttttagttacaccaaaaactgttttgttggatataaaaagaatttattatttaataacgccggcattatcagcaatagcggttatagtt
tcgtaattaagggctcttgttcgtcccgttccccttcttgcatgccccccctacagggaagaagaacctgaaaggggggcatactttgctacaagacacatctaccacgtccaccttggtgaaaaag
gcggaatcggtggtcttgctacgcagcacgggacgggaagaaggggtcccctctccaagaagggcttgcttccaaagaaaaaggcaccttactccacaaagagagataacttcgtagggcttgctt
ccagggaaaataaaaataatagtttaaacaataatttattaattaaaaataattggctgcgccagcagcgccagctgcgctcgagatggacttttaaatttttagaaaaagattttagtacccatagtattaaaa
caactactaaacaactaactggccgaatatatactagctggtttagtagtagttttattgtttaaaaataaaaatataatatacaagcttgaaaaaggcttaagctacgcaggccttcgaagaaccccctctaagg
ttatttttagcatttttgttggttttttctttgttcttattcaaaaaagtccaagctcataccctccgtttgcttatcactccaagcaacacgtccacctctggtggaaacggtggcaatacaagccaacctacgg
ttgtacaagcttgtaaaaggggacagggtataatacaagcttaaatttaccatttagctttacaaaatttaatggtttatttagtcttaaaaagttttatataaaaaataaaaaacgttgcctccccccttacag
ggcttgcttccaaggccccttttaaaagttgtaaaagctaagcttgccccgaaggaccccctactagcttggaaagaaaaaagttgttaagaaatttaaataaaaaacaatttgttaaaaaaactatttcttg
taatccagtaaatactcttgcttggatgcaagccgtggtttttcttgaagctataaaagctcaaaaaaacagtgtacgtaacaattctatagttttagataaaaacaatttttggccaagcctctttccaccaacag
tggaacaggtacgaagtaagaaacaagttttataaaatatcttttaacgttcgaatagctctttatcctcatcatttaaaaaaaattaaaacaaatttgcctttaagttatacccttacacgtgcctccgaaggac
ctccgctgggggtagtgggtggcaacaggggtgtagggtattagaacaaaaagaaaatttggtataaaatttagaccaaatttttttttatgcttaccactctttcccttcttggaggggggaaccccttcttg
gaggggaaccccttcttcccgtgctgcgtagcaagaccaccgattctacctttttacgtgccacaaagtggcaacttccccttttattataaaggagggcaacttttgtgaacctactacaccctacttggagg
gggaacttggaggggaacttggagagggaacccaaatataaaggggacggaacgacgccacaaagtggcaaccaagggtaggcggacgtggtctaagacaagagcttggcaaaaaggg
ggttgcttcgcttcccctcttgtagcaaggaggggatgggacgggacgaacgtataagcatataagtgtaaaaaaaaacaaaatataaaaaaaatataaaatgcttttttaatgtttattgtatacataatttac
agtccaaattaaattatttaaaaatataaaaattctgtatataactactataaaataagacgtagctacgctgctttagctattacaataaaaaaaggttggttgtatttaccttttcagggttctttcggaggcctgcgta
gcagaaattttggtacaaaatttgtaccaaaagctaaaaagaataataatagttatactaaatataataaaagttttataagctcaggacatttcttactacatgatttaatttttgataataatatttacttttgaatgt
attccgtttaactatgccaatttttatttttaaagaaaaaattattcgtaccgtcccgtcaagccaaccgtaggttgtacaagcttgaagaacattctaaaaataaggttaaaaataaaaagcttgtataaaaacccta
ggtttttagcttttaaaaatatcagatttaaaaatttttttaaatctgctcttgccaacgcttcaagacaacctacggttgtagaatgtaataataatttacaaagtatggtaggtaataaaatttctaagcacaaaa
actttgcttataatcctaatcaagccaataaaataattacaaattttagttgggtatccttaaactacactgctaaaaaacttttagcatctctttcaaatacaagacaacaagttggaggggggaaaaaacaacttt
ctataaaaaacttaaacataaaataaaaaccaaattagttggtttagagaatatagttaataatttttaattatggtctctctattaaccgtagttatattaaaaatagaaaaaaaagccaagctattgccttaccctt
taaaaagcgtagcttgctaacccgtgccttacaagcttgttccccttcaatctatacccaaaaaataagggtataaatgataagctggtagaaggtggcaaggaaggggggtaagaaggtaacgccaac
aacgttggcaactactagcttgttcccctgtcccttgcgcagcaagtaggggggcttgcctccaaggaggggggaacaaaaaaaggcaagggtaaacgttggcacaggtaattatcgttttagttttaaaaact
aatagcagctccgcagctaacaatttttgggtaaccaataaaaaatgtcgttataaaagtaattataataatactaataatcatagagagtttagagagttgaaagctaataacatgattcttatacaaaaagttt
tcgaattaacattggatgatcctattgattataaaaatgaaattacgcaactgtataaaaaaaagtagtgcttgcctcgattctggtattacagaatctagagatttttttaatttcaagtatacatggtaaaagagct
aaaattgctttacctcccaaaggggggtgctacgcaggggaagataaaagctccgcttttaggttttaaacaagcttgaaccagttttatttaataaacacctatcctctgcttcctatcaacctccttttcgtg
cctccttagaggggttcttcgaaggccccgtgcataaaactctaagcaagaagaaggtagaagttggcaagaaggggttccccgtgccacaaagtggcaaccaagaaggggttcccctagaagaag
ggggacggcacaagagcgaagctcttaatttattaacgagttttaaggcaacaaataaaattagttttttttcaaaaaaaatatatatttggcaatcaattaataaaccaaaacttaataaatataatatatactta
aatcctatttgtattactattatccagttacgaaaaaataaagcatataattcttaatattaaaacaaaaatttcattcttctcttttaaataaacaaaaaaaaagtatacagtctagctggtctattccatactttgat
tttaaagttaaaaatgctttttaaaattaaaatcacttttatgcgaagctttggaagggaatccccttatacattaaaaaaaacctcttttataaaaaactgggctagttttccagatgcccttaccatgcctatcc
tttctaaaaagcaaagcttttaaagggtgggatgaaggtgttagaggctgctaaaaaatctttgttatgtttcgcttcccgtaccatcgatttatacccgaagggtatagattaatacgattgatacgtaggtg
tcttcccttacacgtgcctaacccccggggggggtaagaaaggtgttgcttcgcacaagaagggtgtagggttcccctttattataaaggaggggggaacaagcaaaaaaaaggggggatgtccaata
aaaatagccagttaaaacatcttctttttacccgtgcttaaaactctaagcaagaaggggtattatttctttgttccccctcctacacccgtagggtgtcggggaacaagcataaagcttatatataaataattata
ataaactcttgtgtcttgtagcaagaccacttacccccaaagggaggcgctacgcgcaaaggtggtagaagtggcaaaaaaaagattcttagaggggggcttgtttccaaggcaacaaagcactaaaa
attaacactgttatttttcaaaaaatctataaaatacaaatcctttattaattggttataatatgccctatggtgtggatctggacttttaatataccaaatgttcatttaaattctaaaaataatatgtataaaaacttaaaa
aagatagctacgttaaaaggttccatacacgtgtcttgtagcaagaccacctaccccccttttatatttggagggcgctacgcgcaaaggtggtaaaagtggcaagaaggtaacgcgtagcaaaaaaatc
tcttgttcccccctctagagataagcttttaaatttggcccatacacaaagacaatttgcttatatatttgcccaaccttgttttttcatttttttcttagtaaaaaattgaccaatttaaatccacttttttctgctacaagctt
gtataaaaacctaaaaagcgtctttttttttaaaaggttttagcttgtaaggaatataaaccaataaaaccgattataaaggttgccgaagtattaaataaaaaaaaatttggcacaaatttttgtaccaaaatttttt
tactaatacgatacatattaaaaacaaagtttcttttaggcctcgtctttaaagcgaaaaacgaataaactattaaactctaacaataagctttttttcacacttttattccaatggggattctttctataataacaaat
ttatttctgcgaatccatttatcaaaagttcttctttaagttcatttgaaggcgaatttctcaaaaccttgtcgttatcggtacagcctttagctgcaaaaccttaactacatcccctatactgtaactactgtaggt
ggtagggaagtaaatatagtaaggggagctaaaaaaaatgcagctttaccccttcttgtagggcaacacctccggagttaccctacttggaggggggaacttggaggggggaacttggaggggggaactg
gcatgccttcgaagaacctgaaaggggggcatacttgtcacatcgtggcgtccacctccccctactacaccctacgggtgtcgggaaaaggtggaagggacgggacgggaagaagtggcaagaa
gggggtaagaaggggatggggagtaccgaccaaacagttggaccatcttgtctaatttttaacaaaacaaaatttaataagttttttcttttaaaagaaaaatggttagaatttcctcggatagagctaatataa
actctctaataagattacaaaatgataattctaaaaaagcttttttataaaattgctcgagaaaatctttctaaagaaatcaccattgcttcgcaatcgtatcaatttaataatgttttaattaagttggacaaatatgg
tagtgattttaaaggaaatacactcaaatatagtccggttaacgttagccaagacaatattaaagatagtggagggaaaaccacctttgcgcgtagcgcccccctttgggggggtaggtggtcatctagaaa
attcgtcgtacgcgaattctaaaaatgaacctttttataccccaaggggtataaccagtttcaaggcaacaaatttttttaaatggtcgcaaagttttttaagatcaatggacttgtggccgggttacctaaacaa
aaaggtatatccttatttaaaaataaattatcgttgggtgagtttttttgtttatggcgataaaatattgccagctatgcttaaccattattcttatttttcatttgttttttaaaatcctctttttgtttcttctttttttatatataag
gatttaacg

gttcatactttattttctgtcccaagcgaagcaacccccttcttgccactaggtggcacgggaagaaggggaaaagaggggggaaatagagaccgaaaaaaagttaaaaataaatacacaaaagtgtggc
aattccaatttcaggtcaaattatccatataaatataaataaaattaccttaagacgagctcagccaatttttatatctccgaaaagtattttacatatttatgatggaagtttcataaaaaaataatgagcctgtaat
gactttaacttatcaaaggttaaaaacaggtgatattgttcagggtattccaaaagttgaacaatattttgaagctcgtataactaagcaaggacgattgtatcgagattgtttaccaaatttattaaaaggtttat
ttaaaagatatagatatcaattgccattagaaagagccgttagacaaagttttttataaaattcaacaaatcatagttgatggtgttttacgagtttatagatcccaaggggttacaattgctgataagcatttaga
ggtaattgtcaaacaaatgacttctaaagtaaaaatataaatggcgcccagacaggatttttccaggagaaatataagatttacaatttgtggaaagagtaaatataaacattaatgaaaaaaatccaatat
gaaccactaatattaggcataacaagagcagcattagaagtggatagttttttatcagcagcaagttttcaacaaacaacaagagtattaagtcttgctgctattgaaaaaaaaaaagactttttaaaaggatt
aaaagaaaatattcttttaggaaatttaatgccagcaggtactggatatttagtttatttagaagatatctctagcatg

>Chlamydomonas_mutabilis-rps2
atgcaggcctataatttaaataataaaataaatgctgaattagaacttattaagacaaaatatacatctgaacaacaaataatacgtattttaagaagaaagttaaaacgtcatgaagctgaaaaaaaactctt
acaatttttaccaaaattacgttatttaccaactccaaaaagtcaaatttctttagcaattcaaatttttaatgaaacgttttgttgatccgaaaatgaaatatccaatggatcttatttatgaagaaaaattaagaaaa
aaatctaaaaaagtggctgccgcaagaaaaaaaaaatggcagaggttagaaacatattttggtggaattgccaatatgacaaaattaaaaaaaacacaaatttctaaaaatgttgcaattattataggtcaa
caagaagaaatgaatgctgttcgtgaatgtaaaaaacttggtatcaaaatgttccatatagtggatacgaattgtaacccaggtttagctgaccattttattccatcaaatgatgattcaagaaattctattaaat
atcttttaacaaaaattgtcactcgtattcgattagctcaaaaaatgcgtttacggtttaagcgtaattttatgcaaagcaaagaaaaaaaacagttatctaaatcgttgcaatcaactttaaattatactaaagaa
aaacagttaaaaaacaaacaaagcagtcagaccaaaaacacactcgttaatagtgatgctatataaacaaaatttagtttcgcctaaaaatagaggcgttaacatgcttcgcgtaggtgatttttgtgaattaa
aagtaacagctttaggaccaaaaggtgttggcattgatgaattttcgtatggatacccgaagggtatattaattccaaatgcaaaattaggagaaacagttaaagctcaaataattaaaattaatttaaaaaga
gccaaatacgctgttgctaaaattacaagccgtattacgccaactaacctagaaaatactaatactttacaaaatttaaatgttggagatatattaacagtcaaattaaatcggatagctttcaaaactgacagt
acatgtattggcattgttaatttaaatttaccaataaaatacccaattaataattacattaaataatttagctactattggccaagaaataacaattaaagttactcgaataaaattaaaatatggctttgcaaaaatta
tatcaccagaaacacttagttccccctttattataaaggagttgcttaaaggggagcaaggaacataaaaataaagtaggacaacttacaaaaggaagtaaaataacttaacgataccttaaaactgca
aaaaaatatgctaaccatatagtactaaaaataaaaaataactacaataataacggagctcttttatgcacaattctcctctttatgcttataactctatgcgtagcataaaggatactcttttatcctttaccccttc
ttcccctacgggggacggtacgagtgataaggataaacaagataaaaaagttttatttattaaaaatgtattaggtgcaaagtacggagataaagttcgagtaaaatttacgcaaattggctcaacatttgc

tattggaaaagtagttagaatcgctccaatttctttgttaaaaaaaaaaataaaaattcgcaaaaatttaaatgaaatgcttctgaatggaatgcattatggagaaaaagttgtaaaatgccaagctagaatga
gaaatttttatctggaatcgaaaaaagaattcttgtaatatgaatcgtccactcctacctatgcgaaacttggaaaaggtagtaggatctaaaaaagatttttttaccaaataatatttctttcacacctcctcaagct
ttcccatccccttcttacatgcctgcggagcagaagaacctgaaaggggggcatactacaaccttcttggaggggcaaccggaggttggcttgacgtgtctttcccctacggggaacgacggaaggg
acttgaaattggacaggtaggaaaggggggcttagagaacaaagcttgacaagaggtagtaaaaggggacaactaacattcttatggaggggggacaaaaaaaccagagacaacctaaggttgt
cttgcagcaaataaaaaaaggtagacatttaattaatttgttaaagacacgtcgttgcttaaatgtagggctaaatatattaaccaaatatgcagcaaagggaagaacttttttatttgttggaactaaaaaacc
tgctgcaggtttaattgctcgagcagcattatttagtaaaacatcg

>Chlamydomonas_mutabilis-rps3
atgggacaaaaagtccatccaataggttttcgagttggtattactaaaagacatcagtcacaatggtttgccaggtttacaaaatatgcctatgcacaaagcatattagaagatcgtatgttaagaaatacgtt
aattagttttgctactgaagcattaaatacgaatacaggaaaaaaaaaggactcttcttttcaacaaaattccataacgccaaaaataactcatataaaaattgaacgcggcttgataccttatgaaattggaa
ttcaaattcatgctgaaaattgtcaaaaattaaaatcatctttaaataatttaaaaagtaatttagcatatgcgtcgcttggaggggggaacacaaagaggcagagatttattatgtaaattacaaaaaacccgtc
gttatctaattagtttagggatgggtatgactacaaaatctaattatttaaataaaaaagcagtaacaaatacaaataaagaaataactatataaaaaataaccgtgccgtccccaatataaaaaaggaagtgg
caagtatgccttagaggggttcttcgaaggcatgcaagaaggggttccctataaggggggggaaaaaaaaaggggttgcttcgcaaggaaaggggcaaatttttttatttagtattgcaaaaaagcaacct
gctataagaccctccgggtctagtaaacataagcgtagcttatattctaaaaagtttttaaaacaaccatacggagcgtctacttttattaaacgtggcaatattttaagtcaattatacaaaaaaaaaaaaatg
attattaaacgtttaagaaaacgccaagaaattcgtcgtcgttatagaagattaatgttaaaaggtcttttgtccaaaaaaaaggaaaaaatatttttttaaaaccttcttttttttttacgtttgcctacaagcgaaa
ctagtaaggcaagaataaaaaaaaaatacagttttatttaaaaataaatttaaaaataaaagtaccactgcttttaacaatagtgcaaaaactagaatttttaaatctaagtcagaatactactaaaactcctcttaa
agcgagaacagataataaaaatatgggtgtactaagctctgctaaaagaaaaaggatatacgaccaaccatttaaaaaagggagtttagataaaagcaataaatcatactcgctgcctctttggagtgata
agccaaacttatataagagtggaacaggggcttctggtaaagaaagacagattctcgtaggggataaatctaaaattggttacagaattaaaaaaaaaatttgtttctttatttttaaataaaataaacaaaaga
tttttaaaacaattaaaagatcatatgactacatggaataacaaaattatgcagcatcgccaagaacaactcaatctctatggatcacttcgttttgctccacttggttataatcaaaaatggagtttaaaaagat
taaatttttaaaaacaaaaccgttatttaaactgttaaaattagtatctattattgaaagaaaatctttaattaagttagaaactttacggaaagattttattgcatttgggacaatttctagaagcgaagcactaaa
ttactaccaaattatcaatttttttaaaaaatttaaaagaatatgtttcaaaattaaaaaaagaacaaaaacgcattttagttaaaactttaacaaaagtaaaatctaaaaattcttttttcgtccctgcgaagcaagg
agggggaagcgtagcacccccttctttgccaccttctacaagccaacctacggttgtacaagcttggaaggcacgggtaggagcaaattattccgagaaacaaaaaactgtaaagcagaacttttttaaa
ccttagagtttttaaggccagctctattaaacaaacttaatttaataagcggaactgagagcgttttaaacgctgaatgccgtaaaatatatttttattgaatatttaaaagaaattgtaaaaaaacatagaacaaat
aatttgtattattatttatcaacaatttcggatgcacgtaaaaatttaaaaaaaattaaacaatttactaaactccattcgaatttttttatttggacttgatttaaaagatcctgcttctgtgaatgcttcttctccgggat
taaaagagaaattacaagatcgaattaaaaatattattgatctttcttctttttaagcaaaaatctccttttgaaaaaggattaccagaaatttttcttagaacaaatagaaaaacaacgaacaatgtataaacagct
tgctcagcttggtgctaaaatttctataaaatttttattcagttaaatctgaaaatatgaaagtaaaagcatccattattactgattctgttatagatgctttagaaaaaagaaaagctttcagaaaagttataaaaga
tgctaaagaaaatttaatgcgaagcactaatgttaaaggagttaaaatacaagttgccgggcggttaaatggagctgaaattgcccgaactgaatgggtgcgagctggccgtgttcctttacaaacattac
gagctaatttagattattcttataaaacagctaaaacaatttatggtattattggagttaaagtgtggatatttaaaggttatactaagcttgtataa

>Chlamydomonas_mutabilis-rps4
atgtcacgctatttaggtccacgtttaagaattattcgaagaattgggaaattaagaggatttacacgtaaaaaaccttccgtcgagtattcaaaggccgaggtgctttacgtggtaaagttattccaccagg
tcagcatggaatggtaaaattatttaaaacacggccctatgattcctcagagtcagattatctaattcgtttaaaagtgaaacaaagattaagattcaactatggtttaacagaacgtcaattagtaaactatgtt
agaaaagcgaaaaaaaatataagaatcaactggtggtattttattacaattactggaaatgcgtttagataatattgtttttcgtttaaatatggctcctactattgtagcggctcgacaattaattagccatggtca
tattcaagtaaatataataaaaagtaaatatagtgcagttatttgtgtaaaccgaaagatgttgtatcagtagctatgaaacaaaaatcattaaaacttgttactaatttttttacaagattattataaacggatgcgctt
ttataaaaaacgtttagaaaaaacactatcttatatttttatttaaaggaaaagtagtttcaaatatggctgctgccattagtcttattcaaagcggtaattgtagaataaataatcgaaaagttgtaaaacctaatta
tatttgcggggaacgcgatataattacaatagcaactaaggcgggcattcgtgaagtaaaactg

>Chlamydomonas_mutabilis-rps7
atgcctcgtcgacctataacaaaaaaacgttctcttttaccagaccctatatataatagtatttcggttcatatgttagtaaaccgtgttttaaaaagcggaaaaaaatcggttgcttatcgtattgtctataatgct
ttaaaagaagtaggtgatattacaaaaaaaaatccagttgaagttttcgaaaaagctttagataacgtaacgccccgtgttgaagtaaaaccgcgtcgtcgtgccggaacagtacaacttgtaccacgagt
tcttcgcactggcgaccgtgctcgagcgtcggctcttagatggattttggaagcttgccaaaaaagatcaggtcaaccaatgattgcaaagttaaaaaatgaaatagtagaagcttataaaaaaacaggtt
ttgctgttcgaaaaaaagatgagcttcacaaaattgccatcaataatgccatgtatgctcgcaaacctcaaacagttattaatgcaattaatcaaattacaactaat

>Chlamydomonas_mutabilis-rps8
atggtaaatgatagtattagcgatatgttaactcgcatccgaaatgcttgtttagcaaaaaaatctactgtttctattcccttacacaattaaatcaccaaattgctcaaattttagaaaaagaaggatacatttta
acagtgcaaatgtcatttgattctagagatttaattatacgtcttaagtatagatcgaaaaaagtttataaaggaaaaataaaagaatcttgtttaacaaatttacgaagaattagtcgcccttctttacgaatttat
accaattcgagagaaattccacgagttttaggggaacaggaattatgatccttcaacaccaagtggggttttaaccgatcgtgaagcacgttcacgtggtattggtggtgaaatcttatgttcaatttggt
aa

>Chlamydomonas_mutabilis-rps9
atggaaattttagcaagagcagttggtcgtcgaaaagaagctgtggctcaagtgcaactagttcgcggaaatggtcaatttattattaatgacaaacccgcacaagcttatcttcaaaacaattcatgttccc
ttatatgcatcaaatcgccattagaagctgctttaagtgtttttctaatcaacctgattttgaaaaaactaataactctacttctattaagcaaagtttttagatgcaaataatgaaaaggaatacctacgggtg
ctagtttgcaaactgaacaaagtgaaaaaaaagtttctgacccccctccataatttgtacaagccaacctacggttggacaagcttgtagccgaagatcaaatttctgcgaagcatgacattgatgtaattat
aaaagtaaaaggtggaggacttataggtcaagctgaagccataaaactaggagttgcaagagctttttgtataattgaaaaagaggaagaaacttcacttcgaaaagttttgaaaactaaaggttatttaac
acaagactctcgtgttaaagaacgtagaaaatatggtttaaaaaaagcacgcaaagcttctcaatatcataaacgttaa

>Chlamydomonas_mutabilis-rps11
atggcaagacaaacgcgaaaggttgcacctagtaaagcaaaaaaaagaatttatcgtggggttgtgcatattcaagcaggttatcataatactattattacaataacaaacgttcgaggagacgttctttgtt
ggagttcagcaggggcttgcggttttaaagggaaacgtaaatctactagttttgcagcgaaaaaagcagcagaaacagctgcccgcaaatcaaaagattttgctatgagagaagccaaagtcttagttac
tggccctggccaaggccgagaaagtgccatcagagaaatttttaaagcaggtattaaagttaatgtaattcgtgaaaaaacaggtatccctcacaatggctgtcgtcctcctaaaaaacgaagtgtttaa

>Chlamydomonas_mutabilis-rps12
atgccaactattcaacaattaattaggtcagcacgaaaaaaactaacaaataaattaaaagcgcctgctcttaaatcttgtcctcaacgacgaggtatttgtcttagagtatatacaataaccccaaaaaaac
cgaattctgctcttcgaaaagtagctcgagttcggttaacgtcaggttacgaagttacagcatatattcctggaatagggcataatttgcaagaacatgcagttgttttagttcgtggcggacgggtaaaaga
tttaccaggagtacgttatcatattgtgcgtggtacattagatactgctggggttaaaaatcgtgtccaaagtcgctcaaaatatggagtaaaaatagcttctaaaactgcagcaaaaacagcggcaaaaaa
ataa

>Chlamydomonas_mutabilis-rps14
atggcaaaaaaaagtatgattcaacgagaattaaaacgtcaaaatttagttatgaaatatgcagaaaaacgtgcagatttaaaagaacagattaaacaaacctcttttttaaaagaaaaattagctttacatag
aaagttacaacaacttccgcgaaatagtgctgcggttagactgcacaaccgatgtatgataactggtcgccctaaagggtattatagagattttggattatcacgtcacgttctacgtgaaatggcccatga
aggtcttttaccaggagttcaaaaatcaagttggtaa

>Chlamydomonas_mutabilis-rps18
atgaataatttaaattctcaaaacaaaaacgaaactccattttcttcacctactggagctaaaaaaaatttttcttctacacctcctatgcgaagcataataaaagataattatattaatttggctaaaggtaaatat
aatagtgtatctactaaaaatggcaaaataaataagccaacaaaatttaataaaacatacacgtttacttccaatgcaaaatcttctaacattggtttaaaaataaaacgcgttttatcattatctcaaattctgggg
tcgtataagacaacaacgccaaaaaaaattagaacaacaaaaacgtcaaaaaccaataaaacctataattccaccaaaatccttacttattatttttaaaagataaacctgaaaaagctgtatataatcgtaga
attattgattataaacattgcggtttattgcaaagatatataggtttaggaggtaaaattttaccaagacgacaaactcgtttaactgcaaaacaacaaagatatgtagcaaaaacaattaaaagtgctcgaatt
atgggattattaccatttgtaagtaaagaacgtgggttttttagataa

>Chlamydomonas_mutabilis-rps19
atgccacgttctataaaaaagggtcctttcgtagcagatcatttattaaaaaaaattgaaaaattaaactctcaaggacaaaaaaaagttttaactacctgggctcgttcatctatgatactgcctcctatgattg
gtcatacaataggtgtttataatggtcgtgaacatattcccgtttttgtaacagatcaaatggtaggtcataaactaggagagttttctccaacacgaacatatcgaggtcatggtaaaacagataaaaaatct
aaacgataa

>Chlamydomonas_mutabilis-tufA
atggcacgtgctaaatttgaacgtaaaaaaccccatgttaatattggtacaataggtcatgtagaccacggtaaaacaactttaacagcggctattacaatgactctagcggctcgaggcggtggcgccg
gtaaacgatatgatgaaattgactctgcgcctgaagaaagagcacgcggtattactattaacaccgctcatgtggagtatgaaacagaaaatcgccactatgcgcacgttgattgtccaggtcatgccga
ctatgttaaaaatatgattacaggcgctgcccaaatggacggtgctattttagttgtgtcgggggctgatggtccgatgccgcaaactaaagagcatattttacttgcaaaacaagtaggtgttccaaatatt
gttgtttttttaaataaagaagatcaggttgatgacgcagaacttcttgaattagttgaattagaagttcgtgaaactctggataaatatgaatatccaggagacgaaattccaattgtacgagggtctgctttat
tagctttagaagctctggttgaaaatccaaaaattatacgagggcaacataaatgggtagacaagatttacgaattaatggataaagttgatagttatatcccaacaccagagcgccaaatcgataagccat
ttttattagccgttgaagatgttttatcaattactggtagaggaacagttgctactggccgtgttgaaagaggaactgttaaattaggtgattctgtagaaatagtaggcttaaaggatactaaaaatactactgt
aacgggccttgaaatgtttaaaaaaacattagaagaaagtattgctggagacaacgtaggtgtactactgcgtgggatccaaaaaaaggatatcgagcgtggtatggtaattgctaaaaccgggagcat
ccagcctcataaaaagtttgaagctcaagtgtacgtgcttactaaagaagaagggggacgccattccgctttcctagctgggtaccagccacagttttttgttcgtacaacagacgttactggtaaagtagt
tagctttagccatattcaaatgcgtaatccatcatcagttgccgaagaacattcgaataaaatggctatgcctggcgacagaattagcatgttagtagaacttatgtcgccgattgcaatcgaaaaaggtgta
agattcgctattcgtgagggcggtcgtacagtaggtgctggtgttgtaactgcaattcttgattctaaa

>Chlamydomonas_mutabilis-ycf3
atgccaagaactcaaaaaaatgataatttattgataaaactttactgttatagctgatatattacttaaagttttgccaacttcacaacgcgaaaaacaagctttttcatattatcgaaacggtatgtctgcacaa
gctgaaggcgaatacgctgaggctctacaaaattattacgaagcaatgcgtttagaaattgatgcctatgatagaagttatatttttatatatattggcttaattcatacaagtaatggcgaacatgggagagc
gttagaatactattatcaagctttagaaagaaatccatctttacctagcgcttaaataatattgctgttatttatcattatcgaggtgaacaagcaattatagatatcagcctgaaatttgtcaacttttatttgaaa
aagctgctgattattggaaagaagcaatacgtttagctcctactaattatattgaggctcaaaattggttaaaaatgactggccgagaataa

>Chlamydomonas_mutabilis-ycf4
Atgaaaaaagagagaacaaattttacgtcttcttcaaatgatcctctagaatcgcgctctaagcaattagaattaaatcgtcgctattatatagtcggggaacgaagactgagcaattattggtgggcatcg
gtcatttttttaggtggatctggctttttattaacaggaatttcgtcttatattaaatataatatttttagctttatccaatctaacgatatctttttttttccccaaggattacttttatgttttatgggagtttaggtgttttat
taagtttatactggtggctttttaatttattggaatgttggaggaggttttaacgagtttaacaaaaaagaaggatttatgcgtatttttagatggggatacccgggaaaaaatcggcgaattgatttagtatattc
attaaaagatattgaagctatacgagtagaatttaaacaatctcaaggtttattgtcagcagaacaaacaatttatgttcgattgcttcaactacttctgtagagcaaaaaaacgagaaaattaatcccgtaca
cccttctagcctaccacctaccccccaaagaggggcgctacgcgcaaaggtggtagaagtggcacgtgtaggggaagctaaaattaaattaaaagaaaagcgagaaattcctttaggtggaatcgga
caaccattaacattaaaagaaattgaaaaacaagcttctgaattagctaacttttgcaagtagctttagaaggtttataa

>Chloromonas_serbinowii-atpA
atggcaatgcgtactccagaagaattaagtaatttaattaaagacctaattgagcaatatactccagaagtgaaaatggttgattttggtattgtgtttcaagtaggggatggtattgctcgtgtttatggattag
aaaaagcaatgtctggagaacttttagaatttgaagatgggactcttggtattgctttaaatttagaagcaaataacgttggagctgttttacttggtgatggtgttaaaataacagaaggaagtcgagttcgtt
gtacaggaaaaatcgcagaaatccctgttggtgacggttatttaggtcgtgttgtggacgcattagctcgtcccgttgatggtaaaggagctatttcaacaaaagacacaagagccattgagtcaattgctc
caggtattattacacgtcgttctgtatatgagccttagctactggattagtttctattgatgcaatgattccagtaggtcggggtcagcgcgaattgattattggtgaccgtcaaactgggaaaacgtcaattg
cgattgatacaattttaaaccaaaaaggaaaaggtgttgtttgtgtttatgttgcaattggccaaaaagcatcatctgttgcacaagttttaaatacttttaaaagaacgtggcgcattagattatactattattgtta
tggcaaacgctaatgaacc

atcaacgttacaatatttagctccttatacaggtgcaacattagcagaattctttatgtatacaggacgtgctacattagtaatttatgatgatttatcaaaacaagcccaagcttaccgtgaaatgtcattattatt
acgccgtccaccaggacgtgaagcatatccaggtgacgtttttttatcttcattcacgtctttagaaagagctgctaaattaagtacagctttaggagaaggaagtatgactgcacttccaattattgaaacac
aagaaggtgatgtttcagcgtatatcccaactaatgttatttcaattacggatggtcagatatttttggccggggaccttttcaacgccggtattcgtccagcaattaacgttgggatttctgtatcccgtgtag
gatcagctgcgcaacctaaagctatgaaacaagttgcaggatctcttaaattgtccctagcacaatttgctgaattagaagcattttctcaatttgcgtcagatcttgaccaagcaactcaaaatcaattagca
cgcggtgtacgcttacgtgaaattcttaaacaagctcaatcagcgccttatctttagaagatcaagttttaacaatttatgcagggactcaaggctttcttgataaacttgaagtaaaccaagtacgagctttt
gttattggtttacgtagttttgtagcttctaaatatcctaaatttggtgaaatcatacaaaaaactttagctttaagttcagaagctgaaggtttattaaaacaaggaattaatgaatatttaaaagaatttttagcaa
caagcgctcctgctgtagcggctcgcccttа

>Chloromonas_serbinowii-atpB
atgaacgattctatagaaacaaaaaatattggacgtgttgtacaaattatcggtccagttttagacattgtttttttcaaaaggtcaagtacctaatatttacaatgctttagtgattcgttctaaaaacgctgctgg
attagaagttagtgttactgttgaagttcaacaacttcttggtgataattgtgtacgcgcggtttcaatgaatccaactgatggtttaactcgtggtgttgaagttatcgataccggtaaacctttaactgttcccg
ttggaaaagctactttaggtcgtatttttaacgttcttggtgaaccagtagacaatttaggtcccgtaaaagcggacacagcattaccaattcaccgtactgcaccagcttttgttgatttagatacacgtctatc
tatttttgaaacaggaattaaagtagttgaccttttagctccatatcgtcgtggtggaaaaattggcttatttggtggtgctggtgtaggaaaaactgtattgattatggagttaatcaataatattgcaaaagctc
atggaggtgtttccgttttgctggagtaggtgaaagaacacgtgaaggtaacgatctttatactgaaatgaaagaatctggtgttattgtagaaaaaagtctttctgattcaaaagtagcactggtttacggt
caaatgaatgaaccaccaggagctcgtatgcgtgttgcattaacagcattaacaatggctgaatattttagagatttcaataaacaagatgttcttttcttattgataatattttccgatttgttcaagctggtgct
gaagtttcggctttattaggtcgtatgccttctgctgtagggtatcaaccaacattagcaacagaaatgggtaatttacaagaacgtattacatctactaaagaaggttctattacatcaattcaagcagtatat
gtacctgctgatgatcttactgaccctgcgcctgcagtaacttttacgcatttagatgctacaactgtattatcaagaggtctagcaagtaaaggtatttatcctgctgttgatccttttagattcaacatcaacaat
gttacaaccttggattgttggtgaaaaacattatggtgtagcacaagacgttaaaaaaaacgcttcaaagatacaaagagcttactgatattattgcaattttaggcttagacgaattatctgaagaagatagat
tagtagttgctcgtgcacgcaaaattgaaagattccttagccaaccgttttttcgttgctgaagtttttactggatcgcctggaaaatatgttagtttaacagagagcatggagggatttggtaaaatttttactgg
ggaattagatagtttaccagaacaagcgttttatttagttggaaatattaatgaagtaattgcaaaagcagctacattaaaa

>Chloromonas_serbinowii-atpE
atgagtttacaaatttctattttaacaccagaacgccctttttggaatggtcaagcagaagaaatcatccttcctactgaaacaggagaaatgggtgttttaaaaaaccacgctccacttattacaggtttagat
gttggagcaatgttaatacgttctaaagatgagtggaattcatatgcgattatgggaggatttgctttagttaaacaaaatcaagtaactattttagcaaacgaagctgaatcagctgaaaatattgatccaga
agaagccaaaacaagttttgaaactgctaaagctaatttagaaaaagctgaaggtgtaaaagaaaaagtagaagccaattttgcttataaacgttcaaaagctcgatttcaaacttcaagaaaaaaa

>Chloromonas_serbinowii-atpF
atggaatgcgtaacatttattaacggagttactatagacatggtggttttggatttaatggcaatatttttgaaacaaatattattaacttagctgctgtgattggtattgttgtaacttttgttggaggaaatctta
aagcattattagaagaccgtaaaaaaacaattttaaataatttacaagaagcaaatcaaagagctattgaagctcaagaaaaattaagccaagcacgtacacaattagaatcagcaaaaaaaaaagctca

agaaattcgtgaagaaggaattttaagagcgactcaagaaataaataatttaaaatttaaccatgatattagattagcaagattacaagagtttaaacaagaaactcttcaacaagcggaacaaaaagctttt
aaacaagcttatatgtatttaatagcaaaaatttctaaaagagttcgtgaaagattaaatactggattagattctacttatcacgttgtagttaataatttttatgtatctcgttttacagattttaaa

>Chloromonas_serbinowii-atpH
atggcaattgatagtttaattggtgcagctagtgttttagctgctgggatcgctgtaggtgttggaagtattggccctggtactgggcaaggaactgctgctggatatgcagtagaaggtattgctcgtcaac
cagaagctgaaggtaaaatccgtggtgctcttttacttttcttttgcctttatggaatccctaacaatttatggtttagttgttgctttagctttactatttgctaaccctttttaggctaa

>Chloromonas_serbinowii-atpI
atgattaatcctttattagagattggtgaagtttcagtgggacaacatttttattggaacataggagaatatcaagttcatgggcaggttttgataacttcatggattgtttttaacaataataggaacgttaagtttt
ttaggaaatagcaatttaaaatcaacacccgatggatttcaaaactttacagaattgatcactgaatttattcgtgatctagcaaaaacgcagatcggagaatctcatggcgactatctacgttgggttcctttt
cttgggactatatttttatttatttttcgtatcaaattggtctggagctttaattccttggaaaattattgaattaccaaacggtgaattagcagcaccgacaaatgacattaatactacagtggcattagctctttttaa
cttcaatctcttatttttacgctgggattaaaaaaaaaggcttaggttactttaatagatatgttcaaccagctgccttcttattgcctatcaacgtacttgaagattttacgaaaccattatcactatctttttcgtctttt
tggaaatatcctagcagatgaactagttgttggagttcttgtagctcttgtccctcttattgtccctattccgcttatggttcttggtctctttacaagtgcaattcaagctttagtatttgcaacacttgcaggcgct
tatataggagaagctctagaagatcatcattaa

>Chloromonas_serbinowii-ccsA
atggcttttaacttttatttctatattactaagcaactctgttacaaatttctttaatttttttacaccatttttaggctctgtagttggaactgctggagtagaaaatgaaagtataatttataacaacacttacccattctt
attactagaaactaatttgactcaaagctctgctctaattattgaaagcaacgctttaagaaattgttctttttcttgtctctttctttcaatgatttttattggattcaaacagcttttagttctccacccttaaagtgct
aggggtctctacctcaaaaggcttggttggggccttttatggacatactattttcgaaaaattaaaacaagttccgaattttgcaggcacaacttaatgatcatttctaatttttattaatgttttttattgatttttcg
atggaaagaatctggccattttccattaagtaacttgtatgaatcctaatgttttatcttggagttgtacgtttattcatcttgttttagagtttaaaacagaactaattagatctgatgcaatacttttaaaacgatg
cttaggttctataacttctccaatagctttatttacaaatgcttttgcaacttttaatttgccaacagaaatgcaaaaagcatctccttttagttcccgcacttcaatccaattggttaatgatgcatgttactgttatga
ttatcagttactcagccttaattcttggctcattattatcaatagcgtttctaatattgataagttcttttttcttttctgttttgattttagttttttaaaacaaagtttacaaaaacaagtaaatcaagcttcttcggaacg
aagcaaggttcccctaatccttatccatcagataaaccccttaatttgtttccattctcttctgagaagcccagcaagacggcttctttgttggttttgggaaccaagaagcaagaagagccccctaccttcctc
tgctttggttggttgggacctttttttagatatagatggtatggttcccaaacaacccaagtccgaaggagccagtagctttaaagaagcgccttttgagcaaggaaaggagaacaagcagcaaggaaaa
ggacagcaagaagaaccctttggttttcgacccgcttctctcaaaatgcgtagcaagaaaggcgacgcaagaaagatgctaggaagtgtaggaaatcaaggggaaaggacaaaagcaaaaatgact
ttagcttggaattttgataatttaagttatcgtgttttaggaataggatttccttttttaactataggtattttatcgggcgctgtatgggctaatgaagcatgggggtcctattggagttgggatccaaaagaaac
atgggctttattaacatggttaatttttgctatctatttacatgccagaatcacaaaaggttggcaaggcaaaaaaccggctattatagcttctttggttttttagcaatttggatttgttttttaggagtcaatttaat
tggtgaaggtttacatagttatggttggttttcaaattca

>Chloromonas_serbinowii-cemA
Gtgtcatttacttataattttaatgaatccaataaagtcaggttaaaaaatattaatttttagctaatgcaaacaaagaaaaagtttctttaaaagataaacaagtagtttctattacatatgaagaaatagggtta
tttccaagatcatttagccgtgtatttgatagatttattaaacaattgtttttgatgtcgaaaatttagttattcaagaatatcgtttttatagatatttatttttaacaacagttaaatgttttttattcttctttttgttccttt
tttagtaaacgtagctagtaaaaattatttaatcagacctctaacagaatattgttggaatcaaaaacaaagtgagatttttttaaatgcttatcaacaaaaacatgcttttgctgaattacaagattttgaagaaa
aagtttatttttgaatctctagttatccctaaaactgaaattgataaaacggataatggaaaacaaatttttacaagaaaaaactattcaattagctattgattatataataatgccagcattgaggctattagttgtatgt
ttgctgacttaataagtctcagtgtgtttggatggttacttattcttatggaggtacaaattagcgtaagtcgtttatttatacttgaagttttttttgggcttgatgatacaaaaaaatctcttattattttattagtcacc
gatctcttagtaggctatcattcagttggcccatggtcaattttatttgaatctttatttaatcgctatggtttgccaaatagccaagctgcaatttatttactcacaggaagcctaccagtaattatggatgtgttttt
taaatacttgatttttagacatttaaaccgagcttcaccagcttcagttgcgacctatcatgcaatgattgaataa

>Chloromonas_serbinowii-ChlB
Atgaaattagcgtattggatgtatgctgggccagctcatattggaactttacgtgtggcaagttcatttaaaaacgtacatgcaattatgcatgctccattaggagatgattatttttaacgttatgcgttcgatgt
tagaaagagaaagagattttaccccagttactgcaagtattgtcgatcgtcatgtattagctcgtggctctcaagaaaaagttgttgaaaatatttactcgtaaagacaaagaggaacaacccgatttaatagt
tcttacccccacatgcacttcttcaattttgcaagaagacctccaaaacttttgttgatcgtgcttcaactgaatcaaaatgtgatgttttattagcagatgttaaccactaccgtgtaaatgaactgcaaagtgct
gatagaacatta

gagcaaattgttcgtttttatattgaaaaagcaagaaagcaaaataatttacaaactaacaaaacagaaaaaccttctgccaatattataggtatttttactttaggttttcacaatcagcatgattgtcgcgaatt
gagacgtttattaaatgatttaggcattgaggttaatgaggttttaccagaaggtggttcagttaataacttgaaaaatctaccaaaagcctggtttaattttattccttatcgtgaagtgggcttaatgtccgcta
tttatcttgaaaaagaatttaatatgccttatgttgccattactccaatgggtgttgtagacactgcagcttgtattcgagaaattggcactattttaacaaaaatagatccatctttaacaatcaacgggggttt
attaaaaacgaatatgacaattatattgataaacaaacacgttttgtatcacaagctgcgtggtttcacgttctatagattgtcagaatttaactggaaaaaaagctgtagtttttggagatgcaactcatgctg
cttctatgactaaaatttagcacgtgaaatgggaattcgtgttgcttgtgctggaacttattgcaaacatgatgcggattggtttagagagcaagttgttggttttgtgaccaagttttaattactgatgaccac
actttaataggagacgttattgctcgaatggaaccagctgctatttttgggacacagatggaaagaatgattggtaaaaggctagatatcccttgcggtgttatatcagcacctatacacattcaaaatttcc
acttggttatcgcccttttctaggttatgaaggaactaatcaaatcgctgatttagtgtacaattcatttaccttaggaatggaagaccatttgttagaaatttttggtggacatgacaatagaggaagtaattacaa
aatcattatcaactgattctgatttaagttggtcatctgatggtttagccgaattaaataaaatacctggttttgttcgtggtaaagtgaaacgtaatacagaaaaatttgcacgtcaaaaaaatattggagttata
acaattgaagtgatgtttgctgctaaagaagcagctggtgcataa

>Chloromonas_serbinowii-chlL
atgaaattagcagtttatggtaaaggtggtattggtaaatcaactacaagttgtaacatttcaatagctttagctagacgtggcaaaaaagttttacagattggttgtgatccaaaacatgatagtacttttaccc
ttacaggttttttaattcctacaataattgatactcttcaacaaaaggattaccattatgaagatgtttggccagaagatgtaatttatcaaggttatggaggtgttgacagtgtcgaagctggaggtccacctgc
tggtgcaggctgtggtggatatgttgtaggtgaaacggtaaagttattaaaagaattaaatgcttttatgaatatgatattatttttatttgacgttttaggagatgttgtatgtggtggatttgcagcaccgttaaa
ctatgcagattattgtattattgtaacagataatggttttgatgcattatttgcagccaaccgtattgctgcttcagtgcgcgaaaaagctcgtacacatccattaagattagcgggcttaatcggaaatcgaac
agctaaaagagattt
aattgataaatatgttgaagcttgtccaatgcctgttctggaagttttacctctgattgaagaaatcagagtttcccgtgttaaaggtaaaacattatttgaaatggttgaatctgaaccagctcttcaatatatttg
tgatttttatttaaatattgcagatcaactattaaccgaaccagaaggggttgttccaagagaattgtctgaccgagaattatttagtttactatcagatttctatttaaatcctgttgatccgacaaaaaaaacag
aatcggttgaaactttagacttttattagtt

>Chloromonas_serbinowii-ChlN
atgtcaagcaatctgcttatgtctaataaatcggctactttaaacaattcagactccgttagttttattgaagaagtttctgttccaaacacagatactttagcagtctctacgcagaccaatgatgattcgttaac
ttttgaatgtgaaacaggtaattaccatactttttgcccttattagttgtgtcgcatggctttatcaaaaaattgaagatagctttttcttggtaataggtactaaaacatgtggttacttcttgcaaaacgctttagga
gttatgatttttgctgaaccgcgctatgctatggcggagttagaagaaagcgatatttcagcgcaattaaatgactataaagaattaaaaagattgtgtttacaaataaaacaagatagaaatccaagtgtgg
tcgtttggatagggacttgtactacagaaattataaaaatggatttagagggtatggctcctcgtttagaaactgaaattggaataccaattgtcgttgctagagctaatggtctagattatgcgtttacacaag
gcgaagacaccgttttagctgcaatggctcaaagatgtccccaagctttaagttctgaaaataaaaaaaaataatacagtaggagttgcttcatctcttccagttaacggagctttccaagagcaaccaaaaa
ctttaaaaaatacaaaacaattagttttatttggttctttacctagcacagtaacttctcaattatcaatggaattaaaaaaccaaggcattctagtttcaggttggttaccttcgcaacggtataatgatttacctg
ctttaggtgaagatgtttatgtttgtggtgttaatccatttctgagtagaactgctacaacttcaagcgaatatcaatttaaagcctaa

>Chloromonas_serbinowii-ClpP

atgccaattggagtaccaagaattatttattgttggggtgaagaacttcctccacaatggactgatatttataatttttattttcgtcgacgaatggttttttaatgcaatatttagatgatgaactttgtaatcaaatt
tgtggattattaattaatattcatatggaagatcggtcaaaagaactagaaaaaaaagaaatagaaaaaagtggcatttttaaaagtgctaataaaagtggtaaagaatctagtccaaatttctctgggcagc
aaggaggaggaccgtcgcaagctttatttaatactaaaaaagacaaacaatttttagatctaaaaaatagcgaagttataaatgatagtaaaaagttaaaaaaagataaatctatggaagatcttttgagttct
tatggaaatgctctttatgaagaagatttagctattgatgaaaactatactttagaacagtatactttacaaaaaataacattagaatggttaacttggaatgctcaatttttgactattcagatgaaccatatttat
tttatttagctgaaatattatctaaagattttacaaaagatcaagcttcgcttaaattatttttataatttaaattaccatccattcccaaattatgtagatacattccctttgcttgggatttaggaaccaagaagcaa
ggggccccaactcctagcgtagcgaatgcaagattatgcaaagcattaaataaaattaataaatttgggcctacattagaaagtctaggaagtattataaaaccaacagttgactctaaaaaatttagcggc
actaaaaagaatgcttttcaggaatattggccgtttcgatcaattttaaatttagctaatttcatttctaaagatcagaattttaatataaagaagttattacctcaactgagtaattcccctatccattcccttagcttt
tatcatccatatccattaaaaatgaaccaagcaaccccctcttctgccaacttttgataaggaccccaaccatgaaatagagcaaagtagaggtgtttctgcaacgcagggaatggatgtaggagacgcag
ccttcttgttacaaggtttgcccttagctttagtacccaaacaaggaacaggaggggagttattaaatcaatttaatcaattaaatcaaaaatctagtataacaaaattactgggagttttagcttcaaaagaatt
tgtaaataaaaatacaaattttcaaaagcaacaatctgtggttacaaaagttctaggacaaagaaatttaagaaaagagtatgctttagataataatttttaataattatggttccttgggagggcaaagcccatc
ggagtctggtcccttacaagtgccccccttcccttagttcagactagcgcaactaaagctcgaagaactcgtgttgtcaattcttttccacacccctctgctcttcccatcaaaaaaggaagtaggctcggt
gaggaacaagctatgctggcatataatttacaaaaaccacaaaaaaagaaaatagcgttaaattatttagattttgattctgttggtaataaaaaaggaaataaaaaaagcttctcaatgagtgccgcaaca
agaaaacaaacaaaaagagctcttcaagaagaagaatctaaaaaagttttgtgattattaattcttttggtggatctgttgggaatggtattactgttcatgatgctctacagtttataaaagctggatctttaac
tttaggtcttggagttgctgctagcgcagcctctttagtattagctggaggaactatttctgagcgttatgttactgaaggttgtcactgtatgatccatcagcccgaaggtggcctaacggtcaagcatcag
atatctggattgatagtcaagaaattatgaaaattcgcttagatgtagcagagatttattcattatctgctcatcgacctcgtcataaaatttttacgtgatttagaccgagatttttatttaactgcaacagaaacta
ttcattatggtttagctgacgaaatagctactaacgaagttatgtcgcaaatcattgaaatgacaagcaaagtatgggattatcacgacagtaaacaacaacgtttacttgagacacgaaatagcgcaaata
atgttgacactcaaactcaaagttaa

>Chloromonas_serbinowii-petL_partial
atgttaacaattacaagttacgtggtattattagttggtgcgttaggttttacgttagggatttatcttggtcttttaaaagtggttaaattgatttaa

>Chloromonas_serbinowii_-_petA
Atgtctaaaaattcaattggctactcgcttcttgctactcacttcttacgtagcaagaaagatagtaatttaagagttactttattagatgggagttttgcaaaaacatttttttgttcaatctttttggaatttttctttt
tctaaatggtgcagtttctgactccgctattgcttacccagtcttcgcacaacaaaattatgaaaatccgcgtgaagctaatgggcgtattgtttgtgcaaattgtcacttagcacaaaaacctgctgaaattga
agttcctcaagctgtattacctgatactgtttttgaagcagtagttcaaattccttatgataaacaattacaacaagtacaagctaatggtaaaaaaggggatttaaatgtaggtatggttttaattttgcctgaag
gatttgaacttgcaccaccagatcgcattccagaagaattaaaaaaaaaagttggaaatatttattatcaaccgtatagtcctgaaaaaaaaaatattttagtagcaggtccacttccaggtaaaaaatatagt
gacatgattattccaattctttctcctgatcctgctaaaaattcaaaagttgcttatttaaaatatcctatttttttaggtgctaaccgtgggagagggcaagtatacgctgacggttctaaatcaaacaacaccgt
atataacgcctctgtaagtggtaaaattgtatcaattaatccgggtgagaaaaaaggatctattaacattacaattgaaaaagcaaatggagaaaatttttattgaaaaaattcctgcaggacctgaagttttag
tacgagaaggacaattaattcaagcggatcaaccttaacaaataatccaaatattggtggttttggacaaaaagatgtagagttggtgttgcaaaacccggcacgtgttcaggggttattagttttctttgca
tttattcttatagcacaagttttattagttcttaaaaagaaacaatttgaaaaagttcaattagctgaaatgaatttt

>Chloromonas_serbinowii-petD
atgtcagtaacaaaaaaacctgatctaactgatccagttttaaaagctaaattagcaaaaggtatgggtcacaatagttatggtgaaccagcatggcctaatgatttacttttatattttcccagttgttatttttgg
tacgtttgcttgtgttgttggtttagctgtgttagatcctgctgctataggtgaacccgctaatccatttgcaacaccgttagaaatttttaccagaatggtatttttatcctgtattccaacttctacgaacagttcca
aataaactgttaggtgttttgttgatggccgcagtgccttgctgattgggatcgtaccttttgttgaaaatattaataaattccaaaaccctttccgtcgaccaattgctaccattgtcttccttgtaggtactata
gtagctatttggttaggaattggtgctacttttcctattgatatttcattaacttttggtttattt

>Chloromonas_serbinowii-petG
atggttgaacctttattatctggaattgttttaggattagtacccgtaacaatagcaggcttgtttgttactgcttatttacaatatcgtcgtggtgatttagccacttttaa

>Chloromonas_serbinowii-psaA
atgacgattagttcaccagaacgtgaagcaaaaaaagtaaagattgcagtagatcgcaatcctgttgaaacaagttttgaaagatgggccaaaccagggcatttttcgcgtactctttctaaaggacctaa
cactactacttggatttggaatcttcatgccgatgctcatgactttgatagtcatacaagtgatcttgaagaaatttcaagaaaagtatttagtgcacattttggtcaacttggaattatctttatttggttaagtgg
g

>Chloromonas_serbinowii-psaB
atggcaacaaagctgtttccaaaatttagccaaggtttggctcaagatccaactacacgaagaatttggtttggtcttgctacagcgcacgattttgaaagtcatgacggaatgactgaagaaaatctttatc
aaaagattttcgcgtctcactttggacaactttcagtaattttcctttggacttcaggaaatttgttccatgtagcatggcaaggaaatttttgaacaatgggttactgatcctattcatgttcgtccaatcgcacatg
ctatttgggatccgcactttggccaaccagcagttgaagcatttacacgtggtggagcttcagggcctgtaaacattgcaacatcaggtgtttatcagtggtggtatactgtaggtttacgcacaaattctga
tttatatactggatcggttttccttgcattagtttctgcaatttttttatttgcaggttggcttcatttacaaccaaatttttcaaccttcattatcttggtttaaagatgcagaatcaagattaaatcaccatctttcaggtt
tatttggtgtaagttcattagcgtggacaggccatttagtacatgttgctattccagaatcacgaggacaacacgtaggttgggataatttttattacagtattaccacatccacttggcttaactccttttggaca
ggtaattgggcagcttatgcacaaaacccggattctgcggctcatgttttggtacttcagaaggttcaggggatgcaatttttaacttttttaggggggtttccatcctcaaacgcaatcgttatggctaacggat
atggctcaccaccatttagctattgctgtaattttcattgttgctggtcatatgtatcgtacaaatttggtattggtcatcgtatgaaagcaattcttgaagctcatgttgctcccagcaatcgaatgggtgcagg
gcacaaagggttatttgatactgtaaataattcattacattttcaattaggtttagctttagcttctgttggaacaattacatcattagtagcacaacatatgtatgctttaccgccttatgcttttttagctgttgatttt
acaacacaagcgtcactttatactcatcatcagtatatcgcgggatttattatgtgtggtaaacaactgtggagttatttaaataaatatggtatttatggtatttttttgggacttttagttttaattaatttttttagttttt
tttataaaaaataaccataaaggttttagttttttgtaagttccaatgaaccttttaaacaactcagataatgaggcgttccccaatcaagagactccaacgttttcaacaccctcttctcaaccaacaaaaaaag
acaacaaaaaaaaagcgttgtatatcccatccgcctcgtcaaattttacctgatacaagtagtcgccctttaaacaaaactcacggtggacttgatttagctaagtttagtaaacatcctgctatttatcgattgtt
ccataaactaagtcaacgtaattactacggagaaaccgagagtcctttagggcgtgcttcacaacattttaaagatcttgcaaaaggttctcatagccaaaaagcacttcaagaagcgtggaatttgacac
caggtgtagaaaattttgagttctatattatagaacaaggccccccatttaacagatccattatacgtaaatcacgccaaaatctttatattgccaaaagtcccctgcccacctttaacacttaa

>Chloromonas_serbinowii-psaC
Atggctcatttagttaaaatatacgatacttgtattggttgtacacaatgtgttcgtgcatgtcctttagacgtcttagaaatggtaccatgggatggttgtaaagcaaatcaaatggcttcagctcctcgcact
gaggattgtgttggttgtaaacgttgtgaaacagcatgtcctactgacttttaagtgttagagtttatttaggttctgaaagtacacgcagtatgggattagcttattaa

>Chloromonas_serbinowii-psaJ
atgaaagattttacaacttatttatcgactgctcctgtagtaagtttagcatggttagtattaactgcggtattattaattggttttaacaaagtattccctgatcctcttgtttttacttttttaa

>Chloromonas_serbinowii-psbA
atgacagcgattatcaacaaacaagacgtttcaactagcttatgggctcgtttctgcgaatgggtgacttcaactgaaaaccgcatctacgtgggatggtttggtacaattatgttcccaactttattaactgc
aacatcagtatatattattgctttcgttgcggctcctccagtagatatcgatggtatccgtgaaccagtttctggttctttattatacggaaacaacatcatttctggtgctgtaatccctacaagtaacgctattgg
tcttcacttctacccgatctgggaagcagcttctcttgacgaatggttatacaatggtggtccttaccaaatgatcgtatgccacttcttcatcggtatctgtgcgtatatgggtagagaatgggaattatcttat
cgtttaggtatgcgtccatggatctgtgtagcttactcagcaccagttgctgcagctactgctgtattcatcatctaccctatcggacaaggtagtttttcagatggtatgccattaggtatttcagggactttca
acttcatgatcgtattccaagcagagcataatatccttatgcatccgttccacatgttaggcgttgcgggtgtatttggtggttctttattctcagctatgcatggttcattagttacttcatctctaatccgtgaaa
caactgaaaacgaatctgctaatgctggttacaaatttggtcaagaagaagaaacgtacaatattgttgctgctcacggctactttggtagactgatcttccaatatgcttcgttcaacaactctcgttcattac

acttcttcttagctgcgtggccggtagtaggtatttggttcacagctctaggaatttcaactatggctttcaacttaaacgggtttaatttcaaccaatctgtagtagactctcaaggacgtgtattaaacacttg
ggctgatatcatcaacagagcgaatttaggtatggaagttatgcacgaacgcaatgctcacaacttcccactagacttagcgtcagtggaagctccttcagtaaatgct

>Chloromonas_serbinowii-psbB
atgggattaccttggtatcgtgtacatactgtagttattaatgacccgggccgcttaatttcagtgcatttaatgcacacagctcttgtagctggttgggcgggttcgatgacacttttttgaaattgctgtttttgat
ccatcagatccagttttaaatcctatgtggcgtcaaggaatgtttgtacttccttttatgactcgtttaggaattacacaatcttggggtggttggacaattagtggagaaactgcatcaaatccaggcatttgg
agctatgaaggtgtagcagcttctcatatcgttctttcagggcttctttttttagcttctgtttggcattgggtttattgggaccttgaattatttcgtgatccaagaacaggtaaaacagcattagatttaccaaaa
atttttggaattcatttattcttatcaggtcttctttgttttgctttggcgcttttcatgtaactggtgtttttggtcctggtatttgggtttctgatccttatggattaacaggaagcgtgcaaccagtttctccttcttgg
ggttcggatggcttcgacccttataaccctggaggaatcgcagctcatcacatcgctgcgggaatttttaggtgtactagctggtctttccacctttgtgttcgtccatcaattcgtctttatttttggactttcaat
gggtagcattgaaacagtattatctagtagtatagcagctgttttttgggcagctttcgttgttgcaggcacaatgtggtatggttcagcagctactccaattgaacttacggcccaacacgttatcaatgggg
atttaggctttttccaacaagaaatacaaaaacgtgttcaaactagtttaagtgaaggatcttctttaccaactgcttgggcaaaaattccggagaaattagcttttatgattacattggtaataaccctgcaaa
aggtgggcttttccgtacaggagctatgaacagtggcgacggcattgctgttggatggttaggtcatgcagttttaaagatcaagatggacgtgacttatttgttcgtcgtatgccaacattctttgaaacgt
tccctgttatttttaattgacaaagatggtgtagttcgtgctgacgttcctttccgtaaagcagaatctaaatatagtatcgaacaagtaggagtttctgttacttttctatggtggtgaattagatggtttaacttttaa
tgatccagcaactgttaaaaaatatgcgcgtaaagctcaattaggtgaaatttttgaatttgatcgttcaactttacaatcggatggggttttccgtagtagcccacgcggatggtttacttttgggcacgtttgt
tttgccttattattcttctttggtcatatttggcatggtgcaagaactattttcagagatgtttttgccggaattgatgacgatttaaatgaacaattagaatttggtaaatacaaaaaacttggtgatacttcatcact
tcgtgaagctttt

>Chloromonas_serbinowii-psbD
atgactatagcgattggaacatatcaagaaaaacgtacttggtttgatgatgctgatgactggcttcgccaagatcgttttgttttattggatggtcaggtcttttattattaccttgcgcttatttagcacttggtg
gttggtttacaggaactacttttgtaacttcatggtatacacatggattagcaacgtcatacctagaaggttgtaatttttaacagctgctgtgtctactcctgctaacagtatgggtcactctttacttttttgtttg
gggtccagaagctcaaggagattttactcgttggtgccaacttggtggtttatggacttttgttgctctacatggcgcatttggcttaattggatttatgcttcgtcagtttgaaatcgctcgttctgtaaatctac
gcccatataacgcaatcgctttctctgcgccaattgcagtgtttacttctgttttcctaatttatccattaggtcaatcaggttggttttttgcaccaagcttcggtgtggctgcaattttccgattcatccttttcttcc
aagga

>Chloromonas_serbinowii-psbE
atggctggaaaaccagtagaacgtccgttttctgatatcttaactagtattcgttattgggtaattcatagtattactattcccgcattatttattgctggctggttattcgttggtacaggattagcttatgacgtgtt
tggtagtccaagaccaaacgaatattttacagaagatcgtcaagatgctccactaattacagatcgttttgatgctttaaatcaagtgaaaaaattatcacaacaa

>Chloromonas_serbinowii_-_psbF
atgtcaacaaaagctgaaactattacatatcctatttttactgtacgttggctgtctattcatgctttagcagtgccaacagttttctttttaggtgctattactgcaatgcaattcattcaacgttaa

>Chloromonas_serbinowii_-_psbH
atggcaacaggaacaacttctaaagttaaattaaatcaaactgataattcaaattttcaagaaccagggatttctactcctttaggtactttattacgcccattaaattcagaagctgggaaagttttacctggat
ggggtactactgtcttaatggccgtttttatgctacttttgcagtatttttactaattattttagaaatctataacagttctcttatcctagacgatgtagtgaataattgggattattcagctaaataa

>Chloromonas_serbinowii_-_psbI
atgttaacactaaaaatttttgtttatactgttgtaacatttttgtatgtttatttctttcgggtttctttctaatgaccctgcacgtaacccaggaaaaggtaat

>Chloromonas_serbinowii_-_psbK
atgtcagctttttctattttacttgcaaaacttccagaagcttatgcaccttttgctccaatcgttgatgttatgccagttattcctgttttatttatttttattagcctttgtttggcaagcttcagtaagttttagataa

>Chloromonas_serbinowii_-_psbM
atggaagtaaacatttttggattaacagcaactgctttatttatcttaattccaacttctttttctattaattttatatggagcgactcgtttaccc

>Chloromonas_serbinowii_-_psbN
atgggaagctcagctttttttctttacctttttttttatggtttctgctgttaagcgtaacaggttattcagtatatataagttttggtcctccttcaaaaaaattaagagatccttttgaagaacatgaagattaa

>Chloromonas_serbinowii_-_psbZ_partial
cctgttgttttcgcgtcacctaatggttggacagaaaataaaggtttcgtttttcaggtctgagcttatgggcagttctagtattcactgttggtgttttaattcatttgttgtttaa

>Chloromonas_serbinowii_-_rbcL
atgaacgaaccaacgcagatcactgacaccaagaagcgctatgccgccggtgtcctgaagtacgcccagatgggctactggaacggcgactacgtgccgaaggacaccgacatcctggcgctgttc
cgcatcacgccgcaggagggcgtcgaccccatcgaggctgccgccgcggtggccggcgagtcgagcaccgcgacctggaccgtggtgtggaccgatcgcctgaccgcctgcgacatgtaccgc
gccaaggcctacaaggtcgagccggtgccgaacaccccggggcagtacttctgctacgtcgcctacgacctgagcctgttcgaggaaggctcgatcgccaacgtcaccgcgtcgatcatcggcaac
gtcttcagcttcaagccgctgaaggcggcgcggctcgaggacatgaagttcccggtctcgtacgtgaagaccttcgccggcccgccgaccggcatcgtcgtcgagcgcgagcgcctggacaagttc
ggccgcccgctgctgggcgccaccaccaagcccaagctgggcctgtcagggcgcaactacggccgcgtggtctacgagggcctgaagggcggcctcgatttcatgaaggacgacgagaacatca
actcgcagcccttcatgcactggcgtgaccgcttcctgttcgtgatggacgcggtcaacaaggcgagcgcggcgaccggcgaagtcaagggcagctacctgaacgtcaccgccggcacgatggag
gagatgtaccgtcgtgcggagttcgcgaaggagctcggctcggtcatcatcatgatcgacctggtgatcggctacaccgccatccagagcatgagcaactggtgccggcagaacgacatggtgctgc
acctgcaccgcgccggccacggcacctacacccggcagaagaaccacggcgtgagcttccgcgtgatcgccaagtggatgcgcatggcgggtgtcgatcacatccatgcgggtacggccgtcgg
caagctcgaaggcgatccgatgacggtgcagggctactacaacgtctgccgcgataccacaccaaggtcgacctgccgcgcggcatcttcttcgaccaggactggggcgcactgaagaaggtgat
gcctgtggcctcgggcggcatccacgccggccagatgcaccagctgatcgacctgttcggcgacgacgtggtgctgcagttcggcggcggcacgatcggccacccgcaaggtatccaggccggc
gccaccgccaaccgcgtcgcgctcgaagccatggtgctggcgcgcaacgaaggccgcgacatcaagaacgaaggtccgcagatcctgcgcgatgccgccaagagctgcacgccgctggccgc
cgcactcgatacctggggcgacatcaccttcaactacacgcctaccgacacgtcggactacgtgcccacgccgtcggtcgcc

>Chloromonas_serbinowii_-_rpl2
atgggaattcgttttcttcaagcatttacaccaggaactagaaatcgttcagtttctgattttagtgaattaacaacaactaaaccagagagttcgttaacctataatttacaaagagcaaaaggacgaaatca
ccgaggcgtgattacatcgcgtcatcgtgggggtggtcataagcgtctttatagacttatcgattttcgtcgtgacaaaattggaatggaagcaaaagttattacaattgaatatgatcctaatcgtaatgcac
gaatagctctcttcgttatgatgatggtgaaaaaagatatatattacatccacgtggattaaatattggtgaaaaaatcatttcagaaataaatgctccaattattattggaaattcacttccattacgtaatattc
cgttaggtgctgaaattcataacgtagaatttcaacctggttctgggggccaaattgcccgttctgctggagctgttgttgaaatttttagcaaaagaaggcaattttgttactttacgtttaccctctaaagaaat
ccgtttagtttcaaaaaattgttgggcaactataggtcaagtaggaaatattgaagcgtataatttaacttttaggaaaagctggtcgaacacgttggttaggaattagaccaactgtaagaggttcagttatga
accctgttgatcacccgcatgggggtggagaaggccgtactccaattgggcatagtcgtccattaactccttggggcaaacctgctttaggtgttttaactcgcacgccaaaaaaatatagtaatcaatttat
tattcgtaaaagaaaacaa

>Chloromonas_serbinowii_-_rpl5

atgacacaaagactcaaaacatattatacagaaactattattccaaaattacaaaaacaatttcaatatgaaaatattcaccaagtgcctaaaatagaaaaaattgtaatcaatagaggtattggagcagcttc
tcaaaatcaaaaaattgtagattcgtctttaaaagaactggctatcattgcaggtcaaaaaggtatcataacacgctcaaaaaaagctatcgcagggtttaaagtaagagaaaaaatgccagttggaattgt
agtgagcttaagaggtgatcgtatgtacagttttttagatcgattaataaacttagctttgcctcgggtacgggattttcaaggaattaatcccaaaagttttgataaaaatggcaattatagtttaggtttagaa
gaacaattaatgtttcctgaaattgaatatgataaaattgatcaagttcgaggtatggacatttcaatcgttactactgcaacaaaacaagctgaaggtttagctcttttaaaagaatttggtttaccgtttaaagc
ttag

>Chloromonas_serbinowii_-_rpl14
atgattaaacctcaatcttatcttaatgttgctgacaatagcggagcacgaaaattaatgtgtattcgtgttttaggtggaagtcaacgtgaagctggaaatattggagatattattattggagttgttaaagatt
ctattccgaatatgccattaaaaaaagctgatgttgttcgagcagtaattgtacgaactagtaaagggttaaaacgtcaaaacggaatttccattcgttttgatgataatgctgctgtaattataaataaagaag
gaaatcctagaggcacacgagtttttggtccgatagctcgggaattaagagatcgtaatttcactaaaatcgtttcattagctccagaagttatttaa

>Chloromonas_serbinowii_-_rpl16
atgcttagtccgaaaagaacaaaatttcgtaaacaacatcgtggtagattaaatggaaaagcaactcgtggtaataaaatttcttttggtgattttgctttacaagcattagaaccgtgttggattacttcgaga
caaattgaagctggaagacgcgtcttgactcgttatgtacgtagaggtggaaaactatggataagaatttttcctgataaacctgttacacttcatccagctggaactcggatgggctctggaaaaggatat
cctgaatattgggttgctgtcgtgcgcccaggaactataatctatgaaatgaaaggggtttctgaaataattgcaaaacaagcttttcgaattgctggccataaaatgccagttaaaaccaaatttatattacg
cacacaaacactttgatcccaaaaagtactcctattcttcttccgacaccaattcctgca

>Chloromonas_serbinowii_-_rpl20
atgactcgtgttaaacgtggtaatgtatctcgaaaacgtcataaaaaagtattaaatatgtctaaaggttttcgcggagcgggctctgttttatttagaacagcaaatcaacagaatattaaagcattacgatat
tcttatcgtaatcgccgacaaaaaaaacgtgattttagacggctttggattgcacgtttaaatgctgctgttcgttgttatggtcttaattataatgagtttatgaattatttaaaaatacgcagtattatattaaatcg
aaaaattatagctcaattagcaacagctgatactgaagcatttatgcaattacttttattttaa

>Chloromonas_serbinowii_-_rpl23
atgattgatttaataaaatatccaattattacagaaaaatcttatttaacttttatttaaaaatcaacaatatacttttgatgtagatttacgattaagtaaacctcaaattaaaaaattatttgaaaatttatttaacgtaa
gtgtgattgcagtaaatactcatataccgccacgtaaaaatcttcgtgttggtacaactaaaggatatcgagctcgttataaacgagctatattgaccttaaaaaagggtcaatcattaaaatttgctatacctttt
aacatta

>Chloromonas_serbinowii_-_rpl36
atgaaagtccgtgcatctgtaaaagctatttgtgataaatgtcgtgttattcgtcgaaaaggtaaagtaatggttatttgttcaaatccaaaacataaacaacgtcaaggaacg

>Chloromonas_serbinowii_-_rpoA
atgattaaatttgtaaaaaatgatttttttatttattgtaaagaagcacgtattgaaagtcatcgaagtttttatgggtctttagtcttggcccatttgaggctggacaaagtattacaattgcaaatgctttaagac
ggactttattatcggaattaaaaggtttagcaattacctctgctgaaattgaaggagcgtttcatgaatacgctacacttccaggagttcgagattcgatattagatattttattaaatttaaaggatattgtcttaa
aatccattaattatgggaataaaaaaagaggaacttcttatacagcaacttttaaaaacgaacaacttggatatttaaaaagcaaaggaccgggaattgttacagcgtccgatttaaaactccctcccggaa
ttcagtgtgttgatccagatcaatatatagcaactctgtcagaagaaggaattcttaatattaaatttcatataaatgaaggagcaaacgctaatactgatcaaagtctttcccataataaaccaaaatctcctg
ctgcttcttttgcttatccttctgacagctccataaaattagaagtaaagaagttagaaggagcagaccggttaggggcctcaagcatagccgccccgctacgcgcacctctggtgttaacgtagcaaaa
aaaattgcctcaagcaaccccaatatattaggagggaatggaatggatgggcagagcccatattatggttttactataattccacctcaagtatggcaaagcccttccagcaagctacatacaaaggatag
ttttttaaaagaaataaaaaaccaagcttcacctaatgttgtttttccactgttccagtagatgtcaaaacatataaaaatttaaatctaaatgccaagccagttcctcctaacaacaacgtttatatggatgggc
tcacaaaaactataaataaattatttttttctcctcgaccatctaagctgcgcttatctaacttagatccgctagtctcgatacacttgctcttacaaacaataaagagcaaaaaaaaaaaaagagcctttggcc
ttggttatcgaaagaaaatgagcaaaaaaaaatagacattgggtcccttagctatggctataataattcgtctaatcaaattaatggaattaaaatgggggtgaggtcaatataaaacgtttatggttaaaaa
aaataaaagcaattcaagctagttatgtccataaaattaatagacaaaaaaaagcttttagtttaaataaaattaataaaacagttttaaaagagcttggtaacttaataaaaaaaaaaaattcagatttaaaga
agaaaattccttataatgaagccataatagcattgtcaactcctttagcaattcagcaaagctttatgcctgtgaataaagttaattatattattgaaattaatgaacaatcagttatagcaaaagatagtttaggt
aacgtaaataataatataggtaccgtaaataataattataatactaatataagcaataactctataaaattaccatttcaagttattggctttgccaataacttgacagttaaaaaaacgccttcgtctaaggctaa
ttttgaacaaatgacgcgctccttgctttttggtttgcaagcaaaagcataccaaaccaaccaaagcaaaggggtaggaaagcttggcttgcctagcactatagatgtaggggccccaaacaaggtatgt
acggcaaaaagccaaaggacaagtgcttctagctgcaatgtcacttctaatggccttactaatgtagaagcagcctattataaaaataaacaagctatttataaaaataacaccatttaataaacaagcttcgc
ctgctttttttcagttaaaaaaaaacaagcatggcaaatcgaaccaaaagcgcttttaaaaaaaagccatattgttttgatgaaagcaagttgtttaagaacttttttcttctagctctcaagttaactacattaaa
gctccgcttaattatcattatgtaaagggccctaaagaaaatactaaaaaatttactagaggaactttgctttctaactatcctcacataaaatctaccttaagtgtactacaaacatctttaaaaaataatattgtt
ttagaaatttggacaaatggaagtattcatccccgccaagcacttttatgatgcatgtaaaaatttattaagtatttttttcaaaattagaaaaagcccgttttactagagcaatttttcaatcagaaaaaacatatga
gaaattaggtttatcactcatctcgaaggggactacgtctatggctaaaggcccgggccctaagccatggtggaaatcaaatttaattaagcaaagcttaattaaaaatgaatggcaaacttctaatggctc
tttaccttttcagtagtcaacaaattaaatcagacccctttctttaatcttatgggggaaagcttatactactttaaggctccttctgctttgcttgggcagcgcagcagccaacccgttctctctgcttgggttg
gttcccaagcgcagctgccaacctcaagaaagctagcaaagagaaaggaagcagcagagcagataatcagcaaaacagataggtattgtttacctccattagaaatagcaggttaccctagtgttttat
taaaaaaaacattaatagacaaattattacaatttcattattggccgctctttcaatcaaattaaaaaatcaaaaagtattaaaaaatattagattttgttttactagtagtatgttccgcatgaactctgtatatgga
ccaaaaaaaggttctaaaattcaagcacagcaagaatttaacatagcgagtagcgggcaaaaaactagaagcacacaaaaaaaacaacagagctttttttatgaatcttgggccatattatctagttaccc
caaaatcaataattttattaagcaaaacttaaaccaaatagagggtagaaacaactctgttaacagagttagcaaaagtcacccttagctgcgctaaaagtgaattttaatgccttgcataacacctgcttcg
cgggagctgaaatgccaaccaggctaagcctaagtggattagataagggaatttttaatagtaagcaatttttcttttatttcgttattaaaaagttataaatcttataaagaaattgcttttaaatataactatttag
acccttttaaatcaaaaaaaagatttaaaaacactgcccaactattagactatttgggagtagcaaaaaaaaattactatgcatggcttgatttagaaaaaaaaaagaaagcgcccttaagaaataatagtca
attttcttttgctcctatttggggagttatgtttaacaaaaataaacaatttttttttagctatgccaaaaaatcatcaaatagcttttttcaaaatataaactatataaaaaaagatttttgactcatcaacgcagcagctt
tgctaggtctgctgcgcttgggaaccaaccaaagcaaggagtcattaagggcagttataaaaatggatttttatttatccaaattagcgaatacgagatccaagaataacccactctatccctcccctaataaa
aagggtttgggaaccaaaagaggagctatttaaatattggtttgctcaacccttcggagaagcaaggatggggctcgttgcttgggttggttcccaagcgcagcagccaaccccattattcaccaaagca
aaaaaaatggcccctccttgcttggaatgggaaccaaccaaagcaaagggcgctgcgagtctcgaa
agaaaagaaagaaaagaaagaaaagaagaaaacagtaaaggaatgagtaattataaaacattagcttctagtattttatctctaagtgtaggggctcctagttctctgccaaaagcagatgcgtccttagt
ggtaaatactcaactgtgcttagctaacaccctaaaagcaaatcaagttcctttatggcgagaagctaaaatgattaaagtactaggttcagaagcctatccttcctcaaaaagcgcagcttcctgggagaa
aaaggggtccataattattaaaaactaaaaaaaattgatgttttttaaaacaaaaacaaatcaattggaatataaaaatgtctctaactgattataatatcttaaaaacaaaacccgaatttgttgatattggaa
aacttaatatttcacttcgtccatatacttgtttaaaaagagctaatattaatactttaggagaacttttaaaaaaatctaaagaagaattattagctttaaaaaattttggaaaacaatcattaaaagaggttgaag
aaaatttaaaacagcttggactctctttaaaatct

>Chloromonas_serbinowii_-_rpoC2
ttgttaaaaaaaaatattactaaactctatcatgggaaaatcttgctacgctcaagtgtagcttacatgctagattgtatggcctctgttttaaaaaatgattcatccaaaattaagtccaataagttaaactttaaa
aaaagtattgcattatttaaacaaagcaatcaccttgtagcctttggctataatagcaaaattaaccctgttaattttaattccgttggcactttagaagctattaaaccgaagttaaaaatttataattacacaaat
aaaatgatttcttctttttaattcaaattttaaaacatcaaaagaccaatttaagtcatctaaaaaagaaaatattaatggaaaaaacaaaaaccaaatatctccactaaccttaaaaaatccatttttggaatcaatg
ttttgataaaaatcgtttaaaaaatttttgttttatggttttttagttacctatggagaatatcgcacagtagaattaattgaacaattaaaaaatgttggttttgaatatgctactctagcaggaatttctttaggaattga
cgatttaaaaattcctccaaaaaaatctttttttaatttatgatgcagaaaaaactacaaacttagctcttaatcaatataatcgcggtgaaattactggggtagaacggtttcaaaaaatgattgatacatggcat
agaacaagtgaatttttaaaacaagaagttattgattattttgaaattacagatacctttaaatcccgtttatatgatggcattttccggagctcgggggaatatttcccaagttcgtcaattagtagggatgcgag
gattaatggctaatccccaaggtcaaattttagattatccaattcgaagcaattttcgtgaagggttaagtttaactgaatatattatttccacatatggagcaagaaaaggaattgtggatacagctttacgaa
cagcaaatgcaggatatttaactagacgattagttgatgttgctcagcatgtaattgtttcaaattttgactgtgggactcaaagaggtatttttttaaatgatatgaaagagggtaataaagttttatattccttgc

agaatcgattaattggtcgagttttaggaaaagatatttttcataatcaagaggcttttaaaaagtatcatgtattaaaggggagccatgataaagaatctcctctcatgctttataaaggtctaggcaaagcct
ctaatgaaagacccataatgaaccaaagcatagataaaaatttgtctttgaggcttctgcgaaaaaaataacaaaaaaaagaactaattggacgactaaaagtggaggagaggttgctaatttaccactc
caaatagaagcttcgctaggtagcgctgtccctgcttctagtggaactaggttgattagagaaacgaagctcggggtaaatagcccctactctccttttatttttttagaaatagtgaaataacaagtgatttagc
aagtaaacttagtaaactctctaataaaatttttgttcgttcaactttaacttgtgaaactcaaaaatcagtttgtcaattatgttatggttggagtttagctcaaggaaatcttgtttcgataggcgaagctgttggt
gttgtagctgcccaatcaataggtgaaccggggacacagttaacaatgagaacatttcatacaggtggcgtcttttcaggtgatataacggatcaaataaaagctctttataatggcattgttaaatatttagg
acctatagctggtactttgatccgaacaccggaaggaaaaatttccttttttaactaaaacagaaggatttctagaagtcctttcaataactactgaaaaagaatttgcaaaaaaaaaaagtctataaaattccttc
ttttacattattatacgttcgacatggcgaaaatgtttttgaaaaacaagttattgctcaaatttcaaatatttcgcgacaacaaaataaacgagatgatgcagaatttaccataaattctgaaatggaaggccaa
tttcgtgtcatgaccttaaatttattagaaaaacaagttggtccaaaaccaacttttagctgtcaaaaaaaaaatcagtcccaagcaagcctctaagcaaagcttagtgaaaaaatctagttttgtcttaaatgta
ggagccccccctaacctttataatgggagtcaaagttcgttagatgtcaattccacaatggaaacattacttgaagctagcgattggggttatgcttggattttagctggtaaaatctatcaacttcctatttcat
ccaatatattacctaaatctggtgattatgttaataaaaaaagtcttgtaaattctgtagctttaaaaccatcactgcagggagtgaattctggttttttagaaactaaattaaataataatttgctgttaaataaacc
tattttatttttagagattaataaaatgttttatacaaaaattggatatttttaaaattgccagaatatcctttttttaattccaatatttaccatcttttgttaatgatgtatggctagcaagtgtagcttaccccttgctt
gggatgggaaccaaccaaagcaaggagcgcaacaagaccccaccttgcgttgcttggttcattatcattatcattatcattatcattatcattatcatgggaaacccataatgagaatgataataaactctag
ctttgccatcaccgaacgttctaaggttgcaaaaaaaaattacttttttaagcaaccgttttttttcttttggtgcaaattcgtacaaaatccaaactaagccgtctaattatatttctcgctacgctcaaactccctac
aatcgcttcgtggttcccaactataaagcgaagcttagcaaagcaaatttaaagttaatgtaattaaaagcactgtagaaaccataaaagaattctacccagaatgggtagaccccggaattatcagtttaa
aaaataatcttataaaaaaaaaaagaaaattctttgggttttaatcgtggaggtcaaaccgttttaactctaaatctttataaaaatttaaacttaaaacgcttaaacttttttaaaaaagtagaagttaaaaagagg
aatacttttttaaaagctaactatttaaataataagttttatttcaatatagacatataaacttattatatagaaataaaatttattataacctattagctccgctagaagcaaaatcttatttacatataaaatgatttgt
tatttacacctctagcttcgcaacatttaacttcttttctaacttttttgggtcgcctatttgtgtagccaaaaaaactgtaaacactattccatttttatactactagtcagagtactggtctagtttattcaaaagttgc
gtctatagggaaaagaagttccgactctgctcctatccttgctttggttgctggagcccttcccttatttcataaaagccatagggtttagaagccataggcttgatggtttgttcccaaacaaaagcttaaaaac
caagcaaagtaagctacgctgggtacctagtagcagccctaaactacaaattctaaataaagcctctaatttgagttataactatgctgatagtaatttattaactaatttaaaatctaataataaaaactctt
gcagaaaaactacaaacaggacgagccttaaatactaatagatttttagtggttctaaaactataggtaaaaatagttcagcaaattggagcaaaaatttttagtttatttttagaatggtttccacattctaaaa
aagttacaagtggtgcattacttagtgattcacaaactttacaaactagtacttatttaccttcaatgtggccaaaaaaaaattacttttttaatctttttaaaaaatctttttttcttttattatttacgtcttctctagcgtt
ttttcttaactctaacgttagttacactaaagctctaaagtgcaaagcacttttaaaaaaatttaataaagcttctaaagtgccaacgaagttaaaattaacagagttaattttgcttacatcttttttttcaaaagaa
gagcctttaacatttacttcctttacctagctataaattaggtaaactatctaatacttttttcaatttctaaagtgctttggactataaaatttttggttcccaagaagcaaagccagaggcttacagctaagcttagc
ttggttgggcttgggtacctagcaaaagcttacaaggataggattagttgcactagaaaattgggcttaataaaaaaaaaatttaaaaaacccagctagcgctagttatatttttggtgcaaagacaccaaac
aaaatatttgttaatagttgtgtaagattgagtgttagaacccttaactttctaactcacttttagttccactaagttaaatagtaaatatataaatagaagcttattgctaaaaaaaaagagtttaaaaaaatttcttg
ctttagctatgcaaagagtcgttacaaaaaaaaattattttggcaaagcctttaatgacttctttttttaacaaatgataagtctcatagcaataacctagattcctcaaaagctgtagtaagcaaaattaaccctgtt
aattttaaagtgctgcgcactttagaagctttcttttgtaacttctaatgaaataacattgttagagagacctaaagacgtacaagttccttatttgagaaatatagcaaaaagaataaaaggtatattaaattact
aaatttttttaaatcgtagtagttttgaacttgtggtcgcttttaaacagcaaaaaaattcttattataattatgcaaaattaattgcaaaaaaagaacaaatttttaaatgaaacgttatatgaaaaattttttttatatttca
agacaatttttttgaaattccacaagaacgtttaactatatctaacatttttttttaaaaataagatgtcttatatattttttatgacactttcaaaaaaaaatcgacaaggaattaaaaaagcacttttttcttaattctagc
aacgctagcgggacttttggtcaaactagaagctttgccgatgaaaccacacttaacatagtgaaaagagttagccctagctaccctaataaagctaatggtggaggcccaagtagcagaaggggtttg
cttcctcaccaactcagcgtggttggcagagcatcaacaacgttgtggaggggttcttctaacaaaattagcaatacgaatttttaaagttaatgttattaaaagcactttagaagttacacagaaggatagca
agcagcaagggaaagg
agctgagcattcccaaagaatgctactaggttcccaaccattgtgggaaagaagcaactccaagcctaggttgatacaatttaaagctcacgtattaaaagattcggctttgataagccatgctttggtacc
caagcatgttcattacacggactatggccatactttcctaaaaaaaaaacactttaaagaatggacaataaaacaatcaatttttgtcaaaacgtcttcagctacgctaaaacaagtcaccccgtatccaatta
aactctataaagggagattgaaaaaaaagtgaggctttacaagatagagttaaaaagattacgagtggattaatttttcccttagattatttaaaattgtttaaacttattaaaatttttttaactaacgattataa
aaaatttaatactgcgttaactgcgttaaaaataaacatggtatcaagttattaactttattgatggctaaaggtccttatcctaagccatggttactaatccactatggctttattggggctgctacgcttaccc
caaaggtgtatggggaagcctattatggggtccaatatagaagccaaaaagatagttaaaaaaaacttgagtcttactttctcaatagggagcaaaacttagtccaaaaagcttccttcgtgactttctc
attatcaaggggttggctgctgcgctgcccaagcgacgctgcccattatttaccaaagcaaggggtactctagctagaaataaattaaataatttcaaatttctaaaaaactaaataatattgctctttttata
acaagcaaggaccttaatgaaaaaacaagcatttttttgcttaaataaaccagttttaaatttgccgcaaaactttttaaaaagaaaatagtaatttatcctcatcatttaaaaaaattaacaaataatttttctaa
ttttaaagaaaataagcgtagctttaatggttttcccataattagtaggggcacgtattgtaaattatttttgatagcccattatgtaaaacgtgggccaagcgttggtaatcaaaaactaatttcattatctactctt
tactgtcatgcaaaaaaacaagtaacaactataatccaaaaaaaattcatttattaagcttcctgtttggcatgcaatgcttacaagacgcaaggggctgctacgcggcgtagcttaactaagccaaactta
agttacactagtgcagctcttccttcactttgcttccctgctagcgtagctttcaagagcagagaaagaaagcttctaggttggttttgggaaccaagaagcaagaggggcctctcccaaaccagcgtag
cagaggaataacggaaccaactaacaacgttgtagtaaggaaaggaaaatctcgaacctctacgataactaacccttatccaattacaagttctgagtctattaatcatattcctggattaataactattttgc
aaaaaaaaggttgggttttttttccttctatgtctataaaaatgaaaagtagtatttctatctcacaatcatataattatcaagaggaaccctaataataggtgctttccctaagcaaacccgtagtctagtgaag
ctaactggactacaaaacttacagtttataaggccaggacaatttttgttgccttatttaatgtttgataataataatatttatgttgaatggtttaataatttttagttctgcaaactatattcataagcaccggttatttc
ccttttatggttctttattaaaaacgtggaccataaaaaaagttattaattctaaaactgcaaaaaaaactgttagagttataaacacaggaattaaaaaatctcaagctctagttaaaaataaattgtttaaagat
tcttatagtttagacccattattttggtgctcacgtttatgggcatatgtctatttaaaccctttggtttgctctattccctctgcttggaatccaactatcttgcctatagcttggaaggtaaaaggaaagtcaaaca
aaggagaggaacaggccctttttgaaaaaaaaaagagaaaaaaaaaataaaatttatttaaccatttttaattctgttctccccccccctcctttttggtgtaaagccaaaagaagatataattatacggagtaat
acattgctgcgtaaccatgtacatagcaaaaatcaaggggccctaacaagtacggagtgggcaggaagaaaaaagtagagacttccagcgtagcaggcctttatgcaaagcttaagtcttctaatttttt
gcctaatctgaaacaatcttaccccaataatggaagccttttagattcgaaagcaatttcttcttctcttaatgaattaaatttttaatataaaaacaagtgattcctgttgggtacaaaatcaaaattgtcgttatcta
ccaaatcattccgctagttctataacttttaaagatttaactagttatggtcatgcgaagcataaacctatcttaattattcaaaaagctttagaattgactattgcgaagcctattaattataaaaatgaaattgca
gaattatatgcctctaacagggatttagtgatgactggaacaaagattttttaatttcaagtttacttggaaaaacgcttcaatcaccatataaacctgttttttctgagaaaattaaaggttctaatcggttaagt
catttttaataacaatcttagttcttgtttatctatgcttatgttaaaaaaaaaatactataacaaacttaaaaaaaactcaaactataggcacacatctcctaaaggcagaaggggaggagtctttaacttatactg
ctcatgcttgggttggttcccatccaaagcctagccttccttttgctttatcatcgtataggcttcacttataattattattacggcaaagccttgatcatgaagatgataatgaaccaaagcaaggaaaggatac
ttgaaaaatcagttaagaagattaaaataaaaccatttatgtccttccttttcttttgccttccctctccttctgcgcatccattccatagaggcgctagggtaccaaagctactttcccaaaggataggaaaaag
atcttttatgttatctgttaaaagtatctggaaaaagagcttagtcagcttcctgtctttgcttttatcattatcattatcattatcattatcatggggaacccataatgataataataatcaggctaagcctaatgaac
ttagcaaagcaaggcagggtaccctagcaaacgtagctttaacaggagcttcttttctgtgactcaaaacaaaataagcaaattaaaggttccttttagaggcaaagcttatactttgaagagactaaaaa
aacataatattagttttttttctaaaaaaatttatctttggcaatcaattagtaaaagaaagttaaatttagtacaagcagttaagggcttgctctgggtaccaagcagaaggactcaaaagtctactaaaagtaa
attaaaaaataatccttgactaaaataactattattcaattgcgtaaaaataaatcttataattcagcattagttagaaagcaatttttgggtctttacttaataaacaaaataaacaatagaaggtgtagctag
tacaatgccttattttgattttaaagtgaaacataatgaatttttacacaaaaatggtctcactttaaaattaaaagcaagtagtggagttaaagttttagcgaaactagcaaaaagccctcgttttaaaggcatc
caagaggttaccccagcttctaaattaggggtttcgcacagaaagatgaagccaagctacgcatgaagctccctacggttgggctgctacgctgggggccccctagctttgcctctaaagcggagcgtaac
tttgttagcgcagctgaagttaactcttgggaaccaaaagctcttcagcgaagtttaaccactaaaagcactttagaagagcaagcaaagcttagaaaaggttcatttctgtcttctaatgggcaaagtccat
taaaaaagtcattattgaaaaaaaatagttatgcacataaactagggtcttctatctcaacacctcaagctacacttgcaaaaagtctcagctcaacaaaagctagtaatttaatttttaaaaaaccaattaatct
aactcctttattaataggttataatatgccttatggttttgatttagatttaatgtgccaaatattcattttaactctaatgttgcaagttgtaaacttcttggtgcgaagcaaggaactttacccacttttcaatttttaa
caaaaaaacaaaaattgaatttttaaaagagagtaacggagctaaaatgcttttacagagtaaaaattcttggttttttattttgctcaattacaagcacttaaaattccgtcgttatttgcacaaccatgtttttct
ttttttatgagcaaaaaattaaccaattttataaaacgttttaaacaaagcttaaaatgggttcaagaatctgatccttttgcttctttaactaatcataatttttcttataaaataaaaaggagtataggaataagaaa
ggcaaattcttcatttattaatttaattctactgcgttaattaaagctcattcatttttaaatttaaaaaaaaaaaaagcaaaaaaatcttctgtttatgaaaagttacattggaagttttttttgacttttattcttatgg
agcatctttatacaaaaataaattgctttttctagtggaacaaagagctcttcttccttttctacccttcctctggagaaaaaaggaaaagaaaaaagagagcaaagtatcaaacaaagcttaataaaatttc
attcttaagttctttgaaggtgaatttgttaagctcaattcacatctaaactcacacctaaattcaccagcttccctaagtgaattgaaaaaagcttctaattcaactgcattaattcatttaagaagcttaaaaag
ttatgaatataaaaacctcactagtgttttaaatttaaatagttataataagttagatctagtaaagcaacaatggacaaaacctattatggacaaagcattctttccttcgctctgtaagagctttccgcgtcatg
ccagaaatggaaaaactaaagctaatttttttaaatttttttgggtttttcaggtcgcaaaaagttttctaggagcagagcaaaggaatgggaaggaaatttttatggcaaagccatgaaccaagcagagcggt
tcaggggcagcgtcgcagcttcttaatgcgtagcataaacccagcaacagcttacaagagaaaacttacaacgttaggatggtctttggtagggcaaacaagagttcgaacaggccagccacttaata
aacgccttcaaataaaactctgacttcttttaataagtttcttggaccatcaagtcttatttttaacaaaagctaatttaataagttattcgtttaatttttaactcttagctccgctaagaaaaaagcttaattctaaag
tcaaacataacgtatattcttataaaattgctcgagaaaatctttctaaagaattaattaactttactagatattataaatataatatattttaattaaagttagttcgaatcttttgaggttggttatggcgaaattct
aagtgaaagtaataaaaattctcaagctagaggcgaagccttaaagcgaagctagcaacgcaaaatttaaagttaatgtaattaaaagcactttagaagctatacaagcatcactgaaaaaggtggag

caagcaaaccttaataatgttaaatttcacaatgtgccacaagaaaatcctttagccaatcattacaacgaatccacattaaacggacaaactaagcaaagcttaaagctttattcgcagaaatccaatggat
ttttaacaaaaaatttttatgcaggtgaagataaaattggttttttataaagcaaagctcttcttcggcaaaacattccaaagaatctctaacaacaacgttattaattaaaaatgaatatatgataatacaaaattt
aattagtgtagctaaagcaaaaaacattttattagctacggtaagcaacgcttgtccacatccttctagtttaagcacacaagttccaaaaatagatattggggtaggaacttataattttagtgttaattccgag
attaatgttggtagttacggcgaagccattatcagagcaaaaatttttaaaattaatggattaattgctggtttaccaaaacaacaaaatgtatcattatttaaaaataaattatctttaggggaatttttttgtttatg
gcgacaaaataatgtctattccctttattaaccacaatacaaatttcccattattaggggaaagatcaaaacataaaatcgctacattaaaaaataaaaattttggaattccttttgcgggtcaaataattcatatt
aatcaaaataaaattacattaagacgcgcacaaccaatatttatttcacccaaaagtattttacattttaatgatgggagttttgtgcgtaataatgatccagttatgactttaacttatcaacaattgaaaacagg
tgacattgttcaaggaatacctaaagtggaacaatattttgaagcgcggataaccaaacaaggaagattatttcgtgattgtttgccaaatttattaaaaggcttatttaaaagatatagatctcaattaccatta
gaaaaagctgttagacaaagtttttataaaattcaacaaatcattgtagatggcgtattacgtgtttatagatctcaaggagtaacaattgctgataaacatttagaagttattgtaagacaaatgacttctaaag
tgaaaattataaatggagcgcaaactggtttttttcctggtgaaataattgatttacaatttgtcgaaagagttaatttatcattaatgaaaaaaatacaatatgaaccattaattttaggtataacgcgagcagctt
tagaagtggatagtttttatcggcagcaagttttcaacaaactaccagagttttaagtagtgcggctattgaaaaaagaaaagatttttttaaaaggtttaaaagaaaatattcttttaggaaatttaatgccagc
tgggactggttatttagtttcttttgaagatcttaaaaataaaaca

>Chloromonas_serbinowii_-_rps2
atggagttaaaaatcaatttaaaatatttaaaaaaaacattaacagattatcttttattaaataataaaataaatagtgaattagaacttattaaaacaaaatatttatctgaacaacaaattattcgtgtcttaagaa
gaaaattaaaacgtcatgaagctgaaaaaaaaattgttacaatttttacctaaattacgttatttacctacacctaaaagtcaaatatctgtcgctattcaattattaatgaaacgatttgttgatcctaaaatgaaat
atccgttggatctaatttatgaccaagcttcgcttagaaaaaaatcaaaaaaagtagctgcagctagaaaaaaaaaaatggcaacgattagaaacttattttggtggtattgcaaacatgactaaattaaagaa
aacacaaattgctaaaaatgttgctataatcataggtcaacaagaagaaataaatgctattcgtgaatgtaaaaaacttggtattaaaatgtttcatattgttgatactaattgtaatcctggtttagccgatcattt
tattccagcaaatgatgattctagaaattcaataaaatatttattatcaaaaatggttacacgtattagattagctcaaaaaattcgtttacgattaaaaaaaaaaaatgcaaagtaaagaaaaaaaaggtaatat
gtttctaaataaagttaaaaaccaaacaatatcaaataaacagaaaacaattaaaagtaaggcaatttcatttagcaaaacaaagcctaatagcccgaaagtaggtgatatttgtaatttaattgtaacagcttt
aggtcccaaaaatattggtattgatgaattttcatacggatattctattttaattccaaatgccaaattaggagaaaaggttaaagcacaaattatcaaaaaaaatttaaacaaagcaaaatacgctgttgctaa
aattataatgcctaactctcaagtattacataaagctaaacttaaatctttagataaagacaaaactttccgtgatggttctgcaatccttgaagatggaaaagaaaaaataaatgttggtgatattttaacagta
aaattaaataagattgctattcaaactgataattcttgtattggtatagtgaattcagatttcaaattaccgataaaataccaattcataattcctaacatagttgagactaaggttggtcaagaaatccaaattca
agtaactaaaatgaaattaaaatatgggtttgctttaatgaaatccagcgtagttaactcagatcaagtaaaacctaatattaagcaaaacttgggtacttcaaataagcttcctgcaactcgactaaaccttct
ttcatcccctcaaaaagcgcaacaaacgaaccctttggtccgtcgcaaaggcgtttatccattgcctcctacggttgcttatcctactgcattggctccaattaagcctttggaagagcgaggaaagcagc
aaggaattaatggtcaaatatggctaacaaagttacaggaaggggccggaaaaaccagtgtaatagctccaagtcataaagaatcgactaatttattaaaaccttttatttctttaagctcacgtccaagcg
aagcttatgtttttggaacctctctttccaaaattggctttccaagtcttaacattccaaaagtgcaaaattatggtcaaaaaatcactcctaataatagttctttttcatcttttatttctgggagccttgcagatttaa
ataagactaaaaaaacaaattacaaaccttttaacgttgcgaagatgacaaaaggaagtaaaattactttaacattacctaaaactgcaaaaaaatatgctaaatatttagtaattaacattttaaatcaaaatat
tgaaacgggggggggtttagtgatctctagtcatccaactactcttaatgatggaaaaactcgctataattcttatattttatttataaaaaatgtattaagagcaaaattaggggataaagtccgagtaaaattc
acaaaaattggttcttcatttgctattggtgaagttattcgtatttcacccaattcatcattcacaaaattagctcgtgttcgaaaaaatttaaatgaaatgcttgtaaatggaatgcattatggagaaaaagttgtc
aaatgccaagctaaaatgcgtaattttatttggaatcgaaaaaaaatctccttctctcccctttcggcaacttatcctttaaattccaatagttataaaggaagtccagttaaagcccgcaattttctggcaaagc
catgggttcaaggctctgccttaaagccggccaaaggctctgtgacttttgctgggaagtccctagaagctttaaagcaaggagggctagtaaatacaaactcggcagggagcaaaaaaagctattcttt
tggtactaggggccccaaacaaaaagttcttctaactatgcgcagcaaggccagtgaaacagagcaccctaaggagttacacagaagccctttagaagcctctaaaaatcaatctttaaacaaaggcgt
tgctttaataaaaaaaggtagacatttaattaatctattaaaaacccgtcgttgtttaaatataggacttaatcaattaacaaaatatgctgctaaaggaagaactttttttatttgttggaacaaaaaaaccagcag
caggcttaatagctcgcgctgctttatttagtaaaacctcgtttttcgtcaatacaagatggttaggtggtatgcttactaattggaaaactattattaaatcgatttcaaaaatccgtcctattcttaaagaaaaac
aaaaagtcatttcaaatattcttcaaaaacgttcaaaaattcaaaatcgtttactaaaaaaagtgttttctttaagaaaaaaaagtaaactcttaatcaaaaaaggtaaacaacttatagctaaagtgaaaattgc
tagtgtaactgattctgttaataataaaaatagcttcgctaagctccgcttaaagctaaatcagattaatttaaagcgaaaagaactttttgaaaaaggtcaaagtttattagaaaaacgccaaaatcttttaaat
aaacaaaaagaattaaaagaaaaaactttaaatattaaagaaaatggtttttaatcgccaataaa

>Chloromonas_serbinowii_-_rps3
atgggacaaaaagtccatccaataggatttcgcgttggtattacgaaaagacatcaatcacagtggtttgcaaggtttacaaaatatgcttatgctcaaagtttattagaagatcgtatgctacgaaatacgtt
aatcagtgtagctaatgaagctttaaattctaatttaataaagaaacgtgactctgcatttcaacaaaacgtaataacacctaaagtaacacatataaaaattgaacgaggattaattccgtatgaaattggaat
tcaaattcatgctcaagctccgctaaaaaaactgaaatcatcattaaataatttaaaaattaacagagaattattatgtaaattacaaaaaactcgtcggtacttaatgagtttaggaaagcagcaaagttcaaa
actaaaaagtgaagtgcctttaaatcaagataatctaattaatagctctccgaaagaaggagctgttaataaagatgaaacggtgcttgttggaaaaaaaaataattttaaaaaagttaatcccctaaaaaaa
aaactttttttgttttagttctgctaacactacaggtaagcaaaacttaacaggcaacactcctttttttttctccattaccaacgggaaaggagcgtggcaaagcaggtacacagggaatcattttacagaagg
gtaagctcccaaccaaacaaagtttagaagttaaaggtaagtttagctttaaccctaaatttaaaaaacgtagacgtcctaatttttttaagtctagtctacaaaaaaaaaaaaatgataataaaacgtttacaaa
aacgtcaaaaaatacgtcaacgttatagaagattaatgttaaaaggcttatttgttaaaaaaaaaggaaaacgaattagttgtgttgcttctttattttttagtaaaaagcaaataaaaaaagctagtgctgttaa
aaaaaaaaaaagtaataaaggggctaaaaaaatgaacattaaatcagtaaaatctcctttaacaaaaaatactagtttttttaaaaaagggttctttgccagtgaaacagaataccaatagattagttgcttcgt
catttgaagggagcggaattgagcagcagaaggacgggctcaaacaatcaaaaactgctaatttttttagctgcacctaaaaaattaaataaacaacaaactagcctgggaacaagtggaacgacgtgg
ccttcattattaacaaattctaaatctaataaaaatataaaagaaaaacttctacttttagcaaaattacgcagtagacagagctctcctcaaacaacaacatcaagtgctagcaaagcttggttgttagcttcg
cctaaaaagttaggatattccgcattaggaagcttaaataagataaatgctagttctttaacaaaattaaacagttctccctacgcaccaagttggaatgttggttatagaatcaaaaaaaaaatttgtctcattat
ttttaaataaaattaacaaacgctttttaaaacaactaaaagagcaaatgactcaatggaataagtctcttaaacaacatcgacaagaccaaatcaaactttatggcgcacttcggtttgctcctcttggttaca
atcaaaaatggagtttaaaaagattaaatttttttgaaaaaaaaaccgttatttaagcttttaaaattaattaaaattatcgagaaaaaatctttaataaagttagaaacgttaagaaaagattttatagcttttggca
caatttctaaaagtgaagctttgagttattatcaaataataaatttttttaaaaaatttaaaaattgcttctttaaaattaaaaaaagaccaaaaacgtctgttagtaaaaactttaacaaaagattattctaaaatatc
agtagtaacatcatcaaatcaagaaatacgtagtccaggtgaaaataaagatgtagatacgttaaaacgtggcaataagcaaagtataatgaaaaccctacctcgcttaggttgggtaccaaaccaagca
agggcaatgaatccagaaataagagtttctggaaaatctaactttgcaactttttttgacatcttctgttgatgtagataaatctcaaggcaagccatttttagaagctattaataatgaatgccgtaaaattaatttt
attgaatatttaaaagaaatggttaaaaaacatagaacaaataatttgttttattatttgtcaaccatttctgatgctcgtaaaaatttaaaaaaaattaaacaatttactaaagttcattcaaaattcttatttgggttt
gaatttaaccaaccaacagccattttagagacttctactctctcaaaaaatactgatagcaaagatctattagtttttaataaaattaaaactgtaattgatttagcttcaaaaaaatcggattttgaaaaaggatt
accagaaattttcttagaacaactggaaaagcaacgcaatatgtataaacaactagctaagttacgcgctaaaatttcaattaaattttattcagtaaaatcagaaaacgtaaaaataaaagcggcaataata
tctgattctgttattgatgctttagaaaaacggaaagctttcagaaaagttataaaagatgctaaagaaaatttaatgcgaaattctaaagtaaaaggtgttaaaatacaagtatctggacgtttaaatggagct
gaaattgctcgaactgaatgggtaagaagtggtcgagtgccattacaaactttaagagcaaatttagattattcatataaaacagctaaaacaatttatgggatcattggtgttaaagtatggatatttaaaggt
tatactaagctagtttaa

>Chloromonas_serbinowii_-_rps4
atgtcacgttatcttggtccacgattaagaattattcgaagaattggcaaattaagaggattcacacgaaaaaaaccttttcgcagagtttttaaaggtcgcggtgctttacgtggtaaagttattccaccaggt
caacatggaatggtaaaattatttaaaacacgaccttatgactcttctgagtctgattatctaattcgtttaaaagtaaaacagagattacgttttaattatggtttaactgaacgtcaattagtaaacgtagttaga
aaagcaaaaaaaaaataaagaatccactggtggtattttattacaattattagaaatgcgtttagataatatagtttttcgtttaaatatggccccaacaattgttgcagctcgacaactaatttgccatggtcatat
aaaagtaaacaacaagaaagttaatataggcagctacttgtgtaaacaaaaagatgttatttcggtttcaatgaaagaaaaatcattaaaacttattacaaacaatttacaaaattactatcaacgcgtacgctt
ctataaagaacgtttgaagaaaactttgtcttttattcttctgaaatcgcaaatagttaccaatatggcagctgccattagtcttattcaaagcggaaattgtaaaataaacaatctagctacacggaaacctaat
tatatttgtggcgaacgagatgtcattacaatcgttactaaagctggaattcgtcaaattaaactaaac

>Chloromonas_serbinowii_-_rps7
atgcctcgtcgccctataaaaaaaaaacgttctcttttaccagatccaatctataatagtatttctgttcatatgttagttaaccgtgttttaaaaaatggaaaaaaatctattgcttatcggattgtctataatgcttt
aaaagaagtaggggacataactaaaaaaaatccagttgaaatttttgaaaaagctttagataatgtcacaccacgcgttgaagttaaacctcgtcgacgggctggaacggttcaacttgtaccacgcgttc
ttcgtactggtgacaaagctcgcgcatcagccctaagatggattttagaagcttgtcaaaaacgatcaggtcaatcaatgattgcaaagttaaaaaatgaaattgttgaagcatataaaaaaaccgggtttg
ctgttcgaaaaaaagatgagcttcataaaattgctatgaataatgcaatgtatgctcgaaaaccgcaaactgtaattaatgctataaatcaaagtgcgactaat

>Chloromonas_serbinowii_-_rps8
atgattaatgatactattagtgatatgttaactcgtattcgaaatgcaagtttagcaaaaaaatcaacagtttgtattccatatacgcacttaaatcagcaaattgctcaaatttttagaaaaagaaggatatatatt
aactgttcaaatgtcattagattcaaaacatttaattgtccgtcttaagtatagatcaaaaaaaatttataacggaaaaactaaagagtcatgtttaacaaatttaagacgaataagtcgtccttctcttcgaatttta
tactaattctaaagaaattccaagagttttaggaggaacaggaattataattctttcaacaccaagtggacttttaactgatcgtgaagctcgttctcgtggtattgggggtgaaatattatgctctatatggtaa

>Chloromonas_serbinowii_-_rps9
atggaaattttagcaagagcggttggacgtcgaaaagaagctgtagctcaaattcagcttgttcgtggaaatgggaaatttataattaatgataaacctgcccaagtttatttgcaaaataattcatgttctctt
atttgtattaaatcaccattagaagcagcattaagtctttattctaataatttttcaaatatccaaagtacttctcaagactctgccttaataaaagatcaagctatagttggccccttcggagcatagcatactaga
agccaatcaagcggttaatgaaaaattagaagaacaaagtacaacaagtgtctcaactgaaaaaagtgaaatttttaaataacaaagagagtggagtagaagccaacaaaggggggaagctttgagttga
atctttctggtgacgaaactcaaaatttaaatgattctacacataaactcagcgtgcttaagttaaataatttagtttcgttagaggaaattgatgtacttatataaagtaaaaggtggaggacttattggtcaagc
agaggctatcaaattgggaattgcccgagctttttgtctaatgcaagttgcgcttaccgaagggatttttaaaaaaatttttaaaaactaaaggttatttaactcaagattctcgtgttaaagaacgtagaaaatat
ggtttaaaaaaagcacgtaaagcttcgcaatatcataaacgttaa

>Chloromonas_serbinowii_-_rps11
atggcaagacaaacacgaaaagttgcaccaaataaagcaaaaaaaaaaatttatcggggagttgtacatatccaagcaggttaccataatactatcattacaataactaatgtaagaggagacgttctttgt
tggagttccgcaggtgcttgtggatttaaaggcaaacgaaaatctacaagttttgcagcaaaaaaagcagcggaaacagctgcgcgcaagtcaaaagatgctgctatgagagaagctaaagttttagta
acaggtcctggtcaaggtcgggaaagtgctattagagaaattttcaaagcaggtataaaagttaatgttattcgtgaaaaaactggtatccctcataatggatgtcgtccgccaaaaaaacgaagtgtttaa

>Chloromonas_serbinowii_-_rps12
atgccaactattcaacaattaattcgttcagcacgaaaaaaactaacaaataaacttaaagctccggctctaaaatcttgtcctcaacgaagaggtatttgtcttagagtttatacaattaccccaaaaaagcc
caactcagctcttcgaaaagttgctcgagttcgtttaaccacaggttacgaagtaacagcatatattcctggaattggtcacaatttacaagaacacgctgttgttttagttcgaggtggccgagtaaaagat
ctacctggagttcgttatcatattgttcgtggcaccttagataccgctggagtaaaaaaccgtgtacaaagtcgctcaaaatatggagtaaaaatagcttcgaaaacagcagcaaaaacatcctctaaaaa
atag

>Chloromonas_serbinowii_-_rps14
atggcaaaaaaaagtatgattcaacgtgagttaaaacgacaaaatttagtaatgaaatatgctgaaaaacgagcttctttaaaagaacagattaaacaaacatctttttttaaaagaaaaattagcgttacatc
gtaagttacaacaacttccacgtaatagtgctgctgttagattacataaccgttgtatgattactggtcgtcctaaaggatattatagagattttggattatcgcgacatgttttacgggaaatggctcatgaag
gtcttttaccaggagtacaaaaatcaagttggtaa

>Chloromonas_serbinowii_-_rps18
atgaatcaacccattcctcgttcacagaataatattaatggaacttttattagcccgagtactcctaaaaaaaatttttttttcaaattctccaaataagagtttagtaaaagaaaatacttttaaaggtacaaaatcc
aattataatagtttcacaactaaaacaagcaatattaataaatcaaataaattaaacaaagttaatggaattaaaactcagtcaaaaggtaaattaaaaaaaaatctaaaaattaaacgtgttttatcattatctca
aattcttggtcggattagaatactacgtcaaaaaaaagtagagcaacaaaaacgccaaaaaccaataaaaccaataattccacccaaatcattaattattatcttaaaagataaaccagaaaaagcagttta
taatcgtagaataattgattataagcattgtggtttattacaaagatatatcggtttaggtggtaaaatttgccaagaagacaaacgcgattaactgcaaaacaacaacgatatgtagcaaaaacaattaaaa
gtgctagaataatgggattattacctttgtaagtaaagaacgaggatttttagataa

>Chloromonas_serbinowii_-_rps19
atgccacgttcgattaaaaaaggtccgtttgttgctgatcacttattaaaaaaaattgaaaaattaaatgctcaaggtcaaaaaaaagttttaacaacatggtcgcgttcttcaatgattttacctccaatgatcg
gccatacgattggagtttataatggtcgtgaacatattcctgtttttattacagatcaaatggttggccataaattaggggaattttctcctactcgaacttatcgaggccatggtaaaacagataaaaaatctaa
acgttaa

>Chloromonas_serbinowii_-_tufA
atggcacgtgctaaatttgaacgtaaaaaacctcatgttaatattgggactattggtcatgttgaccatgggaaaacaacattaacagctgcaattacaatgactctagctgctcggggaggtggtgctggt
aaacgttatgatgaaattgactcttctccagaagaacgtgcacgcggtattactattaatacagcgcatgttgaatatgaaacagaaaatcgtcattacgcgcatgttgattgccctggccacgctgattatgt
aaaaaatatgattacaggagcagctcaaatggatggagcaatttttagtagtatctggcgcagatggtccaatgccacaaacaaaagaacatatcttattagcaaaacaagtaggtgttccaaatattgttgt
atttttaaataaagaagatcaagtagatgatgcagaacttctggaattagtagaattagaagttcgtgaaactttagataaatatgaatttccaggtgacgaaattccaattgttcgtggatcagcactattagct
cttgaagcacttgtagaaaatccgaaaattcaacgaggtgagcataaatgggttgataaaatctacgaattgatggcagcagtagatagttatattccaacacctatccgccaaactgataagccgtttttat
tggcagttgaatcaactgtatctattacaggccgtggaactgtagcaactggccgtgttgaaagagggacagttaaattaggtgaagttgttgaaattgtaggcttaagagaaacaaaaaatactattgtta
ctggtttagaaatgttcaaaaaatccctagacgagagcgttgcgggagataatgtcggagttttgttaagaggaattcaaaaaaaagatatcgaacgtggtatggttctttgtaaaccaaaaagtattcttcct
tacaagaaatttgaagcacaagtttatattcttacaaaagaagaaggtggtcgacattctgctttttagcaggttatcaacctcaatttttgttcgaacaacagacgttactggaaaagttgttagttttagtcat
attcaaatgcgtaatccttcttctgttgcagaagaacattctaataaaatggcaatgccaggtgaccgaattagtatgttagttgaacttatgtcaccaattgctattgaaaaaggtgttagatttgcaattcgtg
aaggtggacgtactgtaggggctggtgttattactgcaattattgaagaaaaa

>Chloromonas_serbinowii_-_ycf3
atgccaagaactcaaaaaaatgataatttattgacaaaactttacggtaatcgcagacattttactcaaagttttaccaacctcacaacgagaaaaacaagctttttcatattatcgaaatggtatgtctgccc
aagcagaaggtgaatatgctgaagctcttcaaaattattatgaagcaatgcgtttagaaattgatgcctatgatcgcagttatattctatataatataggtttaattcatacaagtaatggcgaacatggtagag
cattagaatactattatcaagctttagaaagaaatccttctttacctagtgcattaaataatattgctgtgatttatcattatcgaggtgaacaagcgattcaagataatcagccagaaatttgtcaacttttatttga
aaaagcagctgattattggaaagaagctattcgtttagctcctacaaactatattgaagctcaaaattggttaaaaatgactggtcgagaataa

>Chloromonas_serbinowii_-_ycf4
atgaacaattctcttttatcgcaagagagctcttcaaaggtagccccttagagacaaactcaaagcaaacagaactaaatcggcgttattttattgtcggagaacgtagactgagcaattattggtgggctt
ctgtaatttttttaggtggctttggttttttattaacaggaatttcgtcgtatctgaattataatattttagcaaatgcatttaacatatttaatgttactacgttaatgcaaactttacctaatggggaacctctaaatcc
gaatttgactagcaatttaacttctagctttagtaacgctacgttaaacgaaaattcagttattgctttttttccccaaggattactaatgtgttttacggcagcttaggtgttctattaagtgtatattggtggcttct
aatttattgggatgttggtggtggtttaatgaatttaataaaaaagaaggctttatgagaatttttcgctggggatacccaggaaaaaatcgtcgaattgatctaaaagtaagtttagctgatcttgaagctatt
cgagttgaaacacaaggtttatctgaacaaactatttatgttcgttggcgttccatcactcctaccgaagaccctaaagggttgccttctttttgccccaatctaaaactcaacaggggccagcgcgctgc
ttggtttggggaccaaccaaccaaaggtaaggggagcttcactcctacaaaaacaagaacaagaaaagcgtgaaattcctttaggtgggattggccaaccattaactttaaaagaaattgaaaaacaag
cagctgacttagctaattttctacaaattgaactaataggcttatag

>Chloromonas_rosae__-_atpA
atggcaatgcgtactccagaagaattaagtaatttaattaaagacctaattgagcaatatactccagaagtgaaaatggttgattttggtattgtgtttcaagtaggggatggtattgctcgtgtttatggattag
aaaaagcaatgtctggagaactttagaatttgaagatgggactcttggtattgctttaaatttagaagcaaataacgttggagctgttttacttggtgatggtgttaaaataacagaaggaagtcgagttcgtt
gtacaggaaaaatcgcagaaatccctgttggtgacggttatttaggtcgtgttgtggacgcattagctcgtcccgttgatggtaaaggagctatttcaacaaaagacacaagagccattgagtcaattgctc

caggtattattacacgtcgttctgtatatgagcctttagctactggattagtttctattgatgcaatgattccagtaggtcggggtcagcgcgaattgattattggtgaccgtcaaactgggaaaacgtcaattg
cgattgatacaattttaaaccaaaaaggaaaaggtgttgtttgtgtttatgttgcaattggccaaaaagcatcatctgttgcacaagttttaaatactttaaaagaacgtggcgcattagattatactattattgtta
tggcaaacgctaatgaaccatcaacgttacaatatttagctccttatacaggtgcaacattagcagaattctttatgtatacaggacgtgctacattagtaatttatgatgatttatcaaaacaagcccaagctta
ccgtgaaatgtcattattattacgccgtccaccaggacgtgaagcatatccaggtgacgttttttatcttcattcacgtctttagaaagagctgctaaattaagtacagctttaggagaaggaagtatgactgc
acttccaattattgaaacacaagaaggtgatgtttcagcgtatatcccaactaatgttatttcaattacggatggtcagatatttttggccggggacctttcaacgccggtattcgtccagcaattaacgttgg
gatttctgtatcccgtgtaggatcagctgcgcaacctaaagctatgaaacaagttgcaggatctcttaaattgtccctagcacaatttgctgaattagaagcattttctcaatttgcgtcagatcttgaccaagc
aactcaaaatcaattagcacgcggtgtacgcttacgtgaaattcttaaacaagctcaatcagcgccttttatctttagaagatcaagttttaacaatttatgcagggactcaaggctttcttgataaacttgaagt
aaaccaagtacgagcttttgttattggtttacgtagttttgtagcttctaaatatcctaaatttggtgaaatcatacaaaaaactttagctttaagttcagaagctgaaggtttattaaaacaaggaattaatgaata
tttaaaagaatttttagcaacaagcgctcctgctgtagcggctcgcccttaa

>Chloromonas_rosae__-_atpB
atgaacgattctatagaaacaaaaaatattggacgtgttgtacaaattatcggtccagttttagacattgtttttttcaaaaggtcaagtacctaatatttacaatgctttagtgattcgttctaaaaacgctgctgg
attagaagttagtgttactgttgaagttcaacaacttcttggtgataattgtgtacgcgcggtttcaatgaatccaactgatggtttaactcgtggtgttgaagttatcgataccggtaaaccttttaactgttcccg
ttggaaaagctactttaggtcgtatttttaacgttcttggtgaaccagtagacaatttaggtcccgtaaaagcggacacagcattaccaattcaccgtactgcaccagcttttgttgatttagatacacgtctatc
tatttttgaaacaggaattaaagtagttgaccttttagctccatatcgtcgtggtggaaaaattggcttatttggtggtgctggtgtaggaaaaactgtattgattatggagttaatcaataatattgcaaaagctc
atggaggtgtttccgtttttgctggagtaggtgaaagaacacgtgaaggtaacgatctttatactgaaatgaaagaatctggtgttattgtagaaaaaagtctttctgattcaaaagtagcactggtttacggt
caaatgaatgaaccaccaggagctcgtatgcgtgttgcattaacagcattaacaatggctgaatattttagagatttcaataaacaagatgttcttttcttattgataatattttccgatttgttcaagctggtgct
gaagtttcggctttattaggtcgtatgccttctgctgtagggtatcaaccaacattagcaacagaaatgggtaatttacaagaacgtattacatctactaaagaaggttctattacatcaattcaagcagtatat
gtacctgctgatgatcttactgaccctgcgcctgcagtaacttttacgcatttagatgctacaactgtattatcaagaggtctagcaagtaaaggtatttatcctgctgttgatccttagattcaacatcaacaat
gttacaaccttggattgttggtgaaaaacattatggtgtagcacaagacgttaaaaaaacgcttcaaagatacaaagagcttactgatattattgcaattttaggcttagacgaattatctgaagaagatagat
tagtagttgctcgtgcacgcaaaattgaaagattccttagccaaccgtttttcgttgctgaagtttttactggatcgcctggaaaatatgttagtttaacagagagcatggagggatttggtaaaatttttactgg
ggaattagatagtttaccagaacaagcgttttatttagttggaaatattaatgaagtaattgcaaaagcagctacattaaaa

>Chloromonas_rosae__-_atpE
atgagtttacaaatttctatttttaacaccagaacgccctttttggaatggtcaagcagaagaaatcatccttcctactgaaacaggagaaatgggtgttttaaaaaaccacgctccacttattacaggtttagat
gttggagcaatgttaatacgttctaaagatgagtggaattcatatgcgattatgggaggatttgctttagttaaacaaaatcaagtaactatttttagcaaacgaagctgaatcagctgaaaatattgatccaga
agaagccaaaacaagttttgaaactgctaaagctaatttagaaaaagctgaaggtgtaaaagaaaaagtagaagccaattttgcttataaacgttcaaaagctcgatttcaaacttcaagaaaaaaa

>Chloromonas_rosae__-_atpF
atggaatgcgtaacatttattaacggagttactatataggacatggtggttttggatttaatggcaatatttttgaaacaaatattattaacttagctgctgtgattggtattgttgtaacttttgttggaggaaatctta
aagcattattagaagaccgtaaaaaaacaatttttaaataatttacaagaagcaaatcaaagagctattgaagctcaagaaaaattaagccaagcacgtacacaattagaatcagcaaaaaaaaaaagctca
agaaattcgtgaagaaggaattttaagagcgactcaagaaataaataatttaaaatttaaccatgatattagattagcaagattacaagagtttaaacaagaaactcttcaacaagcggaacaaaaagctttt
aaacaagcttatatgtatttaatagcaaaaatttctaaaagagttcgtgaaagattaaatactggattagattctact

>Chloromonas_rosae__-_atpH
atggcaattgatagtttaattggtgcagctagtgttttagctgctgggatcgctgtaggtgttggaagtattggccctggtactgggcaaggaactgctgctggatatgcagtagaaggtattgctcgtcaac
cagaagctgaaggtaaaatccgtggtgctcttttacttctttgccttatggaatccctaacaatttatggtttagttgttgctttagctttactatttgctaaccctttttaggctaa

>Chloromonas_rosae__-_atpI
atgattaatcctttattagagattggtgaagtttcagtgggacaacatttttattggaacataggagaatatcaagttcatgggcaggttttgataacttcatggattgttttaacaataataggaacgttaagtttt
ttaggaaatagcaatttaaaatcaacacccgatggatttcaaaactttacagaattgatcactgaatttattcgtgatctagcaaaaacgcagatcggagaatctcatggcgactatctacgttgggttcctttt
cttgggactatatttttatttatttttcgtatcaaattggtctggagctttaattccttggaaaattattgaattaccaaacggtgaattagcagcaccgacaaatgacattaatactacagtggcattagctcttttaa
cttcaatctcttatttttacgctgggattaaaaaaaaaggcttaggttacttttaatagatatgttcaaccagctgccttcttattgcctatcaacgtacttgaagattttacgaaaccattatcactatcttttcgtctttt
tggaaatatcctagcagatgaactagttgttggagttcttgtagctcttgtccctcttattgtccctattccgcttatggttcttggtctctttacaagtgcaattcaagctttagtatttgcaacacttgcaggcgct
tatataggagaagctctagaagatcatcattaa

>Chloromonas_rosae__-_ccsA
atggcttttaacttttatttctatattactaagcaactctgttacaaatttctttaatttttttacaccatttttaggctctgtagttggaactgctggagtagaaaatgaaagtataatttataacaacacttacccattctt
attactagaaactaatttgactcaaagctctgctctaattattgaaagcaacgctttaagaaattgttctttttctttgctctttctttcaatgatttttttattggattcaaacagcttttagttctccacccttttaaagtgct
aggggtctctacctcaaaaggcttggttggggccttttatggacatactatttttcgaaaaattaaaacaagttccgaattttgcaggcacaacttttaatgatcatttctaatttttttattaatgttttttattgatttttcg
atggaaagaatctggccattttccattaagtaacttgtatgaatccttaatgttttttatcttggagttgtacgtttattcatcttgttttagagtttaaaacagaactaattagatctgatgcaatacttttaaaacgatg
cttaggttctataacttctccaatagctttatttacaaatgcttttgcaacttttaatttgccaacagaaatgcaaaaagcatctccttttagttcccgcacttcaatccaattggttaatgatgcatgttactgttatga
ttatcagttactcagccttaattcttggctcattattatcaatagcgtttctaatattgataagttcttttttcttttttctgttttttgatttagttttttaaaacaaagtttacaaaaacaagtaaatcaagcttcttcggaacg
aagcaaggttccccttaatccttatccatcagataaacccctttaatttgtttccattctcttctgagaagcccagcaagacggcttcttttgttggttttgggaaccaagaagcaagaagagccccctaccttcctc
tgctttggttggttgggacctttttttagatatagatggtatggttcccaaacaacccaagtccgaaggagccagtagcttttaaagaagcgccttttgagcaaggaaaggagaacaagcagcaaggaaaa
ggacagcaagaagaaccctttggttttcgacccgcttctctcaaaatgcgtagcaagaaaggcgacgcaagaaagatgctaggaagtgtaggaaatcaaggggaaaggacaaaagcaaaaatgact
ttagcttggaatttgataatttaagttatcgtgttttaggaataggatttccttttttaactatagtattttatcgggcgctgtatgggctaatgaagcatggggtcctattggagttgggatccaaaagaaac
atgggctttattaacatggttaatttttgctatctatttacatgccagaatcacaaaaggttggcaaggcaaaaaaccggctattatagcttctttggttttttagcaatttggatttgttttttaggagtcaatttaat
tggtgaaggtttacatagttatggttggtttcaaattca

>Chloromonas_rosae__-_cemA
gtgtcatttacttataatttttaatgaatccaataaagtcaggttaaaaaatattaatttttttagctaatgcaaacaaagaaaaagtttctttaaaagataaacaagtagtttctattacatatgaagaaatagggttat
ttccaagatcatttagccgtgtatttgatagatttattaaacaattgttttttgatgtcgaaaatttagttattcaagaatatcgtttttatagatatttatttttaacaacagttaaatgttttttattcttcttttgttccttt
ttagtaaacgtagctagtaaaaattatttaatcagacctctaacagaatattgttggaatcaaaaacaaagtgagatttttttaaatgcttatcaacaaaaacatgcttttgctgaattacaagattttgaagaaaa
agtttattttgaatctctagttatcccctaaaactgaaattgataaaacggataatggaaaacaaatttacaagaaaaaactattcaattagctattgattataataatgccagcattgaggctattagttgtatgttt
gctgacttaataagtctcagtgtgtttggatggttacttattctttatggaggtacaaattagcgtaagtcgtttatttatacttgaagttttttttgggcttgatgatacaaaaaaatctcttattattttattagtcaccg
atctcttagtaggctatcattcagttggcccatggtcaattttatttgaatctttatttaatcgctatggtttgccaaatagccaagctgcaatttatttactcacaggaagcctaccagtaattatggatgtgttttt
aaatacttgatttttagacatttaaaccgagcttcaccagcttcagttgcgacctatcatgcaatgattgaataa

>Chloromonas_rosae__-_ChlB
atgaaattagcgtattggatgtatgctgggccagctcatattggaactttacgtgtggcaagttcatttaaaaacgtacatgcaattatgcatgctccattaggagatgattatttttaacgttatgcgttcgatgtt
agaaagagaaagagattttaccccagttactgcaagtattgtcgatcgtcatgtattagctcgtggctctcaagaaaaagttgttgaaaatattactcgtaaagacaaagaggaacaacccgatttaatagtt
cttacccccacatgcacttcttcaattttgcaagaagacctccaaaacttttgttgatcgtgcttcaactgaatcaaaatgtgatgttttattagcagatgttaaccactaccgtgtaaatgaactgcaaagtgctg

atagaacattagagcaaattgttcgtttttatattgaaaaagcaagaaagcaaaataatttacaaactaacaaaacagaaaaaccttctgccaatattatagtattttttactttaggttttcacaatcagcatgat
tgtcgcgaattgagacgtttattaaatgatttaggcattgaggttaatgaggttttaccagaaggtggttcagttaataacttgaaaaatctaccaaaagcctggtttaattttattccttatcgtgaagtgggctt
aatgtccgctatttatcttgaaaaagaatttaatatgccttatgttgccattactccaatgggtgttgtagacactgcagcttgtattcgagaaattggcactattttaacaaaaatagatccatctttaacaatcaa
cgggggggtttattaaaaacgaatatgacaattatattgataaacaaacacgttttgtatcacaagctgcgtggttttcacgttctatagattgtcagaatttaactggaaaaaaagctgtagttttggagatgca
actcatgctgcttctatgactaaaatttagcacgtgaaatgggaattcgtgttgcttgtgctggaacttattgcaaacatgatgcggattggtttagagagcaagttgttggttttgtgaccaagttttaattact
gatgaccacactttaataggagacgttattgctcgaatggaaccagctgctatttttgggacacagatggaaagaatgattggtaaaaggctagatatcccttgcggtgttatatcagcacctatacacattc
aaaattttccacttggttatcgcccttttctaggttatgaaggaactaatcaaatcgctgatttagtgtacaattcatttaccttaggaatggaagaccatttgttagaaatttttggtggacatgacaataaggaa
gtaattacaaaatcattatcaactgattctgatttaagttggtcatctgatggtttagccgaattaaataaaatacctggttttgttcgtggtaaagtgaaacgtaatacagaaaaatttgcacgtcaaaaaaatat
tggagttataacaattgaagtgatgtttgctgctaaagaagcagctggtgcataa

>Chloromonas_rosae__-_ChlL
atgaaattagcagtttatggtaaaggtggtattggtaaatcaactacaagttgtaacatttcaatagctttagctagacgtggcaaaaaagttttacagattggttgtgatccaaaacatgatagtactttttaccc
ttacaggttttttaattcctacaataattgatactcttcaacaaaaggattaccattatgaagatgtttggccagaagatgtaatttatcaaggttatggaggtgttgacagtgtcgaagctggaggtccacctgc
tggtgcaggctgtggtggatatgttgtaggtgaaacggtaaagttattaaaagaattaaatgctttttatgaatatgatattattttatttgacgttttaggagatgttgtatgtggtggatttgcagcaccgttaaa
ctatgcagattattgtattattgtaacagataatggttttgatgcattatttgcagccaaccgtattgctgcttcagtgcgcgaaaaagctcgtacacatccattaagattagcgggcttaatcggaaatcgaac
agctaaaagagatttaattgataaatatgttgaagcttgtccaatgcctgttctggaagttttacctctgattgaagaaatcagagtttcccgtgttaaaggtaaaacattatttgaaatggttgaatctgaacca
gctcttcaatatatttgtgatttttatttaaatattgcagatcaactattaaccgaaccagaaggggttgttccaagagaattgtctgaccgagaattatttagtttactatcagatttctatttaaatcctgttgatcc
gacaaaaaaaacagaatcggttgaaacttagacttttattagtt

>Chloromonas_rosae__-_ChlN
Atgtctaataaatcggctactttaaacaattcagactccgttagttttattgaagaagtttctgttccaaacacagatactttagcagtctctacgcagaccaatgatgattcgttaacttttgaatgtgaaacag
gtaattaccatacttttgccctattagttgtgtcgcatggctttatcaaaaaattgaagatagctttttcttggtaataggtactaaaacatgtggttacttcttgcaaaacgctttaggagttatgatttttgctgaa
ccgcgctatgctatggcggagttagaagaaagcgatatttcagcgcaattaaatgactataaagaattaaaaagattgtgtttacaaataaaacaagatagaaatccaagtgtggtcgtttggatagggact
tgtactacagaaattataaaaatggatttagagggtatggctcctcgtttagaaactgaaattggaataccaattgtcgttgctagagctaatggtctagattatgcgtttacacaaggcgaagacaccgtttt
agctgcaatggctcaaagatgtccccaagctttaagttctgaaaataaaaaaaataatacagtaggagttgcttcatctcttccagttaacggagctttccaagagcaaccaaaaactttaaaaaatacaaa
acaattagttttatttggttctttacctagcacagtaacttctcaattatcaatggaattaaaaaaccaaggcattctagtttcaggttggttaccttcgcaacggtataatgatttacctgctttaggtgaagatgtt
tatgtttgtggtgttaatccatttctgagtagaactgctacaacttcaagcgaatatcaatttaaagcctaa

>Chloromonas_rosae__-_clpP
atgccaattggagtaccaagaattatttattgttggggtgaagaacttcctccacaatggactgatatttataatttttattttcgtcgacgaatggttttttaatgcaatatttagatgatgaactttgtaatcaaatt
tgtggattattaattaatattcatatggaagatcggtcaaaagaactagaaaaaaaagaaatagaaaaaagtggcatttttaaaagtgctaataaaagtggtaaagaatctagtccaaatttctctgggcagc
aaggaggaggaccgtcgcaagctttatttaatactaaaaaagacaaacaatttttagatctaaaaaatagcgaagttataaatgatagtaaaaagttaaaaaaagataaatctatggaagatcttttgagttct
tatggaaatgctctttatgaagaagatttagctattgatgaaaactatactttagaacagtatactttacaaaaaataacattagaatggttaacttggaatgctcaatttttgactattcagatgaaccatatttat
tttatttagctgaaatattatctaaagattttacaaaagatcaagcttcgcttaaattattttataatttaaattaccatccattcccaaattatgtagatacattccctttgcttgggattttaggaaccaagaagcaa
ggggcccaactcctagcgtagcgaatgcaagattatgcaaagcattaaataaaattaataaatttgggcctacattagaaagtctaggaagtattataaaaccaacagttgactctaaaaaatttagcggc
actaaaaagaatgcttttcaggaatattggccgtttcgatcaattttaaatttagctaatttcatttctaaagatcagaattttaatataaagaagttattacctcaactgagtaattcccctatccattccctagctt
tatcatccatatccattaaaaatgaaccaagcaacccctcttctgccaacttttgataaggaccccaaccatgaaatagagcaaagtagaggtgtttctgcaacgcagggaatggatgtaggagacgcag
ccttcttgttacaaggtttgcccttagctttagtacccaaacaaggaacaggaggggagttattaaatcaatttaatcaattaaatcaaaaatctagtataacaaaattactgggagttttagcttcaaaagaatt
tgtaaataaaaatacaaattttcaaaagcaacaatctgtggttacaaaagttctaggacaaagaaatttaagaaaagagtatgctttagataataatttttaataattatggttccttgggagggcaaagcccatc
ggagtctggtccctttacaagtgccccccttcccttagttcagactagcgcaactaaagctcgaagaactcgtgttgtcaattctttttccacaccctctgctcttcccatcaaaaaaggaagtaggctcggt
gaggaacaagctatgctggcatataatttacaaaaaccacaaaaaaagaaaatagcgttaaattatttagattttgattctgttggtaataaaaaaggaaatataaaaaagcttctcaatgagtgccgcaaca
agaaaacaaacaaaaagagctcttcaagaagaagaatctaaaaaagttttgtgattattaattcttttggtggatctgttgggaatggtattactgttcatgatgctctacagtttataaaagctggatctttaac
tttaggtcttggagttgctgctagcgcagcctctttagtattagctggaggaactatttctgagcgttatgttactgaaggttgtcactgtatgatccatcagcccgaaggtggccttaacggtcaagcatcag
atatctggattgatagtcaagaaattatgaaaattcgcttagatgtagcagagatttattcattatctgctcatcgacctcgtcataaaattttacgtgatttagaccgagatttttatttaactgcaacagaaacta
ttcattatggtttagctgacgaaatagctactaacgaagttatgtcgcaaatcattgaaatgacaagcaaagtatgggattatcacgacagtaaacaacaacgtttacttgagacacgaaatagcgcaaata
atgttgacactcaaactcaaagttaa

>Chloromonas_rosae__-_petA
atgtctaaaaattcaattggctactcgcttcttgctactcacttcttacgtagcaagaaagatagtaatttaagagttactttattagatgggagttttgcaaaaacatttttttgttcaatctttttggaatttttctttt
ctaaatggtgcagtttctgactccgctattgcttacccagtcttcgcacaacaaaattatgaaaatccgcgtgaagctaatgggcgtattgtttgtgcaaattgtcacttagcacaaaaacctgctgaaattga
agttcctcaagctgtattacctgatactgttttgaagcagtagttcaaattccttatgataaacaattacaacaagtacaagctaatggtaaaaaaggggatttaaatgtaggtatggttttaattttgcctgaag
gatttgaacttgcaccaccagatcgcattccagaagaattaaaaaaaaaaagttggaaatatttattatcaaccgtatagtcctgaaaaaaaaaatattttagtagcaggtccacttccaggtaaaaatatagt
gacatgattattccaattctttctcctgatcctgctaaaaattcaaaagttgcttatttaaaatatcctatttttttaggtgctaaccgtgggagagggcaagtataccctgacggttctaaatcaaacaacaccgt
atataacgcctctgtaagtggtaaaattgtatcaattaatccgggtgagaaaaaaggatctattaacattacaattgaaaaagcaaatggagaaaattttattgaaaaaattcctgcaggacctgaagttttag
tacgagaaggacaattaattcaagcggatcaacctttaacaaataatccaaatattggtggttttggacaaaaagatgtagagttggtgttgcaaaacccggcacgtgttcaggggttattagttttctttgca
tttattcttatagcacaagttttattagttcttaaaaagaaacaatttgaaaaagttcaattagctgaaatgaatttt

>Chloromonas_rosae__-_petB
atgagtaaagtttacgactggtttgaagaacgtttagaaattcaatcaattgctgatgatattagcagtaaatatgttccaccgcatgttaatatttttattgtttaggtggcattacctttacatgtttttagttcaa
gtggcaacaggctttgcaatgacttttattatagaccaacagtagcagaagctttgcttctgttcaatatataatgacagatgtgaattttggttggttaattcgctctattcatcgttggtcggcaagcatgat
ggttttaatgatgattttacatgtttccgtgtttacttaactggtggttttaaaaaaccacgtgagttaacatgggttagtggggttattatggctgtctgtactgtttctttcggtgtaacaggctattcattaccttg
ggaccaaatcggatattgggctgttaaaattgttacaggtgttcctgaagctattcctgttattggtggccctttagtagaattattaagaggtggtgttggcgttggccaaagtacattaacgcgttttatagt
cttcatacttttgtcttaccattagctacagctgcttttatgttaatgcatttcttaatgatcagaaaacaaggtatttcaggaccactataa

>Chloromonas_rosae__-_petD
atgtcagtaacaaaaaaacctgatctaactgatccagttttaaaagctaaattagcaaaaggtatgggtcacaatagttatggtgaaccagcatggcctaatgatttacttttatattttcccagttgttatttttgg
tacgtttgcttgtgttgttggtttagctgtgttagatcctgctgctataggtgaacccgctaatccatttgcaacaccgttagaaatttttaccagaatggtatttttatcctgtattccaacttctacgaacagttcca
aataaactgttaggtgttttgttgatggccgcagtgccttgctgattgggatcgtaccttttgttgaaaatattaataaattccaaaaccctttccgtcgaccaattgctaccattgtcttccttgtaggtactata
gtagctatttggttaggaattggtgctacttttcctattgatatttcattaacttttggtttattt

>Chloromonas_rosae__-_petG
atggttgaacctttattatctggaattgttttaggattagtacccgtaacaatagcaggcttgtttgttactgcttatttacaatatcgtcgtggtgatttagccacttttttaa

>Chloromonas_rosae__-_petL
atgttaacaattacaagttacgtggtattattagttggtgcgttaggttttacgttagggatttatcttggtcttttaaaagtggttaaattgatttaa

>Chloromonas_rosae__-_psaA
atgacgattagttcaccagaacgtgaagcaaaaaaagtaaagattgcagtagatcgcaatcctgttgaaacaagttttgaaagatgggccaaaccagggcatttttcgcgtactctttctaaaggacctaa
cactactacttggatttggaatcttcatgccgatgctcatgactttgatagtcatacaagtgatcttgaagaaatttcaagaaaagtatttagtgcacattttggtcaacttggaattatctttatttggttaagtgg
gatgtattttcatggcgcacgattctctaactatgaagcttggttaagtgatccaactcatattaaaccaagtgctcaagtggtttggcctattgttggtcaagaaatttttaaatggtgatgttggtggaggtttcc
aagggattcaaataacttctggtttttttccaacttttggcgcggcgctggaattacaagtgaactacaattatacagtactgctattggtggattagttatggcagcagctatgttctttgcaggctggtttcactt
ccataaggctgctcctaagctcgaatggtttcaaaacgttgagtcaatgttaaatcaccatttaggaggtttacttggattaggaagcttagcttgggcggggcatcaaattcatatttcattgcctattaataa
attactagacgcaggagtagatccaaaagaaattccattacctcatgatttaatgtttaacagaacatttatggcagatctttatccaagttttggaaaaggattagcaccattctttacttttaaattggagtgaat
atagtgatttttttaactttcaaaggaggattaaatccggttactggtggtttatggctaagtgatacagcgcatcatcacgttgcaatcgcagttttatttttagttgctggtcatatgtatcgtacaaattgggga
attggccatagcatcaaagaaatcctggattataaaaaagtagtccctttagagagaaatctctatcgaaaaatctgtttaataaagagaatctctgagtttatatttttcaaagttgaaaatattaactcagatg
actcttcgccaagttatgaaataacttctgttccttctttatcagaaaagaacaacttaaatccccccctctggaggagaccaaaattcacctcaatctttacaacttccgtttgtttttaatccttatttttttaccaggt
ctttaa

>Chloromonas_rosae__-_psaB
atggcaacaaagctgtttccaaaatttagccaaggtttggctcaagatccaactacacgaagaatttggtttggtcttgctacagcgcacgattttgaaagtcatgacggaatgactgaagaaaatctttatc
aaaagattttcgcgtctcactttggacaactttcagtaattttccttttggacttcaggaaatttgttccatgtagcatggcaaggaaatttttgaacaatgggttactgatcctattcatgttcgtccaatcgcacatg
ctatttgggatccgcactttggccaaccagcagttgaagcatttacacgtggtggagcttcagggcctgtaaacattgcaacatcaggtgtttatcagtggtggtatactgtaggtttacgcacaaattctga
tttatatactggatcggttttccttgcattagtttctgcaatttttttatttgcaggttggcttcatttacaaccaaatttttcaaccttcattatcttggtttaaagatgcagaatcaagattaaatcaccatctttcaggtt
tatttggtgtaagttcattagcgtggacaggccatttagtacatgttgctattccagaatcacgaggacaacacgtaggttgggataatttttattacagtattaccacatccacttggcttaactccttttggaca
ggtaattgggcagcttatgcacaaaacccggattctgcggctcatgttttggtacttcagaaggttcaggggatgcaatttttaacttttttaggggggtttccatcctcaaacgcaatcgttatggctaacggat
atggctcaccaccatttagctattgctgtaattttcattgttgctggtcatatgtatcgtacaaatttggtattggtcatcgtatgaaagcaattcttgaagctcatgttgctcccagcaatcgaatgggtgcagg
gcacaaagggttatttgatactgtaaatattcattacattttcaattaggtttagctttagcttctgttggaacaattacatcattagtagcacaacatatgtatgctttaccgccttatgcttttttagctgttgatttt
acaacacaagcgtcactttatactcatcatca

gtatatcgcgggatttattatgtgtggtaaacaactgtggagttatttaaataaatatggtatttatggtatttttttgggacttagttttaattaatttttagttttttttataaaaaataaccataaaggttttagttttt
tgtaagttccaatgaaccttaaacaactcagataatgaggcgttccccaatcaagagactccaacgttttcaacaccctcttctcaaccaacaaaaaaagacaacaaaaaaaagcgttgtatatcccatcc
gcctcgtcaaattttacctgatacaagtagtcgcccttaaacaaaactcacggtggacttgatttagctaagtttagtaaacatcctgctatttatcgattgttccataaactaagtcaacgtaattactacgga
gaaaccgagagtcctttagggcgtgcttcacaacatttttaaagatcttgcaaaaggttctcatagccaaaaagcacttcaagaagcgtggaatttgacaccaggtgtagaaaatttgagttctatattatag
aacaaggcccccatttaacagatccattatacgtaaatcacgccaaaatctttatattgccaaaagtcccctgcccacctttaacacttaa

>Chloromonas_rosae__-_psaC
atggctcatttagttaaaatatacgatacttgtattggttgtacacaatgtgttcgtgcatgtcctttagacgtcttagaaatggtaccatgggatggttgtaaagcaaatcaaatggcttcagctcctcgcactg
aggattgtgttggttgtaaacgttgtgaaacagcatgtcctactgacttttttaagtgttagagtttatttaggttctgaaagtacacgcagtatgggattagcttattaa

>Chloromonas_rosae__-_psaJ
atgaaagattttacaacttatttatcgactgctcctgtagtaagtttagcatggttagtattaactgcggtattattaattggttttaacaaagtattccctgatcctcttgttttttacttttttaa

>Chloromonas_rosae__-_psbA
atgacagcgattatcaacaaacaagacgtttcaactagcttatgggctcgtttctgcgaatgggtgacttcaactgaaaaccgcatctacgtgggatggtttggtacaattatgttcccaactttattaactgc
aacatcagtatatattattgctttcgttgcggctcctccagtagatatcgatggtatccgtgaaccagtttctggttctttattatacggaaacaacatcatttctggtgctgtaatccctacaagtaacgctattgg
tcttcacttctacccgatctgggaagcagcttctcttgacgaatggttatacaatggtggtccttaccaaatgatcgtatgccacttcttcatcggtatctgtgcgtatatgggtagagaatgggaattatcttat
cgtttaggtatgcgtccatggatctgtgtagcttactcagcaccagttgctgcagctactgctgtattcatcatctaccctatcggacaaggtagtttttcagatggtatgccattaggtatttcagggactttca
acttcatgatcgtattccaagcagagcataatatccttatgcatccgttccacatgttaggcgttgcgggtgtatttggtggttctttattctcagctatgcatggttcattagttacttcatctctaatccgtgaaa
caactgaaaacgaatctgctaatgctggttacaaatttggtcaagaagaagaaacgtacaatattgttgctgctcacggctactttggtagactgatcttccaatatgcttcgttcaacaactctcgttcattac
acttcttcttagctgcgtggccggtagtaggtatttggttcacagctctaggaatttcaactatggctttcaacttaaacgggtttaatttcaaccaatctgtagtagactctcaaggacgtgtattaaacacttg
ggctgatatcatcaacagagcgaatttaggtatggaagttatgcacgaacgcaatgctcacaacttcccactagacttagcgtcagtggaagctccttcagtaaatgct

>Chloromonas_rosae__-_psbB
atgggattaccttggtatcgtgtacatactgtagttattaatgacccgggccgcttaatttcagtgcatttaatgcacacagctcttgtagctggttgggcgggttcgatgacacttttgaaattgctgtttttgat
ccatcagatccagttttaaatcctatgtggcgtcaaggaatgtttgtacttccttttatgactcgtttaggaattacacaatcttggggtggttggacaattagtggagaaactgcatcaaatccaggcatttgg
agctatgaaggtgtagcagcttctcatatcgttctttcagggcttctttttttagcttctgtttggcattgggtttattgggaccttgaattatttcgtgatccaagaacaggtaaaacagcattagatttaccaaaa
atttttggaattcatttattcttatcaggtcttctttgttttgctttggcgctttcatgtaactggtgtttttggtcctggtatttgggtttctgatccttatggattaacaggaagcgtgcaaccagtttctccttctttgg
ggttcggatggcttcgacccttataaccctggaggaatcgcagctcatcacatcgctgcgggaattttaggtgtactagctggtcttttccacctttgtgttcgtccatcaattcgtctttattttggactttcaat
gggtagcattgaaacagtattatctagtagtatagcagctgttttttgggcagctttcgttgttgcaggcacaatgtggtatggttcagcagctactccaattgaacttacggcccaacacgttatcaatgggg
atttaggctttttccaacaagaaatacaaaaacgtgttcaaactagtttaagtgaaggatcttctttaccaactgcttgggcaaaaattccggagaaattagctttttatgattacattggtaataaccctgcaaa
aggtgggcttttccgtacaggagctatgaacagtggcgacggcattgctgttggatggttaggtcatgcagttttaaagatcaagatggacgtgacttatttgttcgtcgtatgccaacattctttgaaacgt
tccctgttatttttaattgacaaagatggtgtagttcgtgctgacgttcctttccgtaaagcagaatctaaatatagtatcgaacaagtaggagtttctgttacttttctatggtggtgaattagatggtttaacttttaa
tgatccagcaactgttaaaaaatatgcgcgtaaagctcaattaggtgaaatttttgaatttgatcgttcaactttacaatcggatggggttttccgtagtagcccacgcggatggtttacttttgggcacgtttgt
tttgccttattattcttctttggtcatatttggcatggtgcaagaactattttcagagatgtttttgccggaattgatgacgatttaaatgaacaattagaatttggtaaatacaaaaaacttggtgatacttcatcact
tcgtgaagctttt

>Chloromonas_rosae__-_psbC
atggaaacacttttaacggggtcctatcaataggtgggcgaactcaagaagagactggctttgcatggtgggctggtaatgctcgattaattaatctttcgggcaaacttttaggagctcatgttgctcatg
ccggattaattgtttttgggctggggctatgaatttatttgaagtttcgcacttttgtacctgaaaaacctatgtatgaacaaggtcttattttattaccacatatcgcaacattaggttacggtgttggcccgggt
ggtgaagttattgatacattcccatattttgtatctggcgttttacatttaatatcatctgctgttttaggttttggtggagtttttcactcactaattgggccagaaacattagaagaatctttcccattctttggttac
gtttggaaagacaagaacaagatgactaatattttaggttatcaccttattatgttaggtcttggtgcttggcttcttgtgattaaagcaatgtttttaggaggggtttatgatacttgggcgccgggtggtaaaa
aacgtctattgccaccctccagcgaaagctggatttaa

>Chloromonas_rosae__-_psbD
atgactatagcgattggaacatatcaagaaaaacgtacttggtttgatgatgctgatgactggcttcgccaagatcgttttgttttttattggatggtcaggtcttttattattaccttgcgcttatttagcacttggtg
gttggtttacaggaactacttttgtaacttcatggtatacacatggattagcaacgtcatacctagaaggttgtaatttttttaacagctgctgtgtctactcctgctaacagtatgggtcactctttacttttttgtttg
gggtccagaagctcaaggagattttactcgttggtgccaacttggtggttttatggacttttgttgctctacatggcgcatttggcttaattggatttatgcttcgtcagtttgaaatcgctcgttctgtaaatctac
gcccatataacgcaatcgctttctctgcgccaattgcagtgtttacttctgttttcctaatttatccattaggtcaatcaggttggttttttgcaccaagcttcggtgtggctgcaattttccgattcatcctttttcttcc
aagga

>Chloromonas_rosae__-_psbE
atggctggaaaaccagtagaacgtccgttttctgatatcttaactagtattcgttattgggtaattcatagtattactattcccgcattatttattgctggctggttattcgttggtacaggattagcttatgacgtgtt
tggtagtccaagaccaaacgaatattttacagaagatcgtcaagatgctccactaattacagatcgttttgatgctttaaatcaagtgaaaaaattatcacaacaa

>Chloromonas_rosae__-_psbF
atgtcaacaaaagctgaaactattacatatcctatttttactgtacgttggctgtctattcatgctttagcagtgccaacagttttctttttaggtgctattactgcaatgcaattcattcaacgttaa

>Chloromonas_rosae__-_psbH
Atggcaacaggaacaacttctaaagttaaattaaatcaaactgataattcaaatttttcaagaaccagggatttctactcctttaggtactttattacgcccattaaattcagaagctgggaaagttttacctgga
tggggtactactgtcttaatggccgtttttatgctactttttgcagtatttttactaattatttttagaaatctataacagttctcttatcctagacgatgtagtgaataattgggattattcagctaaataa

>Chloromonas_rosae__-_psbI
atgttaacactaaaaatttttgtttatactgttgtaacatttttttgtatgtttatttctttttcgggtttctttctaatgaccctgcacgtaacccaggaaaaggtaat

>Chloromonas_rosae__-_psbK
atgtcagctttttctattttacttgcaaaacttccagaagcttatgcaccttttgctccaatcgttgatgttatgccagttattcctgttttatttatttttattagcctttgtttggcaagcttcagtaagttttagataa

>Chloromonas_rosae__-_psbM
atggaagtaaacatttttggattaacagcaactgctttatttatcttaattccaacttctttttctattaattttatatggagcgactcgtttaccc

>Chloromonas_rosae__-_psbN
atgggaagctcagctttttttctttaccttttttttatggtttctgctgttaagcgtaacaggttattcagtatatataagttttggtcctccttcaaaaaaattaagagatccttttgaagaacatgaagattaa

>Chloromonas_rosae__-_psbZ
cctgttgttttcgcgtcacctaatggttggacagaaaataaaggtttcgtttttcaggtctgagcttatgggcagttctagtattcactgttggtgttttaattcatttgttgtttaa

>Chloromonas_rosae__-_rbcL
atggttcctcaaacacaaactaaggttggtgcaggttttaaagctggtgttaaagattatcgtttaacatattatactcctgattacgttgtaaaagaaacagacattcttgctgcattccgtatgactccacaag
ctggtgttcctattgaagaagctggtgccgcggtagctgctgaatcttcaacaggtacatggacaactgtatggactgatggtttaactagtcttgaccgttataaaggtcgttgttatgacatcgaaccagt
agcaggtgaagacaatcaatatatcgcttatgttgcataccctatcgaattatttgaagaaggttctgttactagtttattaacatctattgtaggtaacgtttttggtttcaaagctcttcgtgctctacgtcttgaa
gatttacgtatttctccagcatatgctaaaacattccaaggacctccacacggtattcaagtagaacgtgacaaattaaacaaatatggtcgtggtcttttaggttgtactatcaaaccaaaattaggtctttctg
ctaaaaactacggacgtgctgtttatgaatgtttacgtggtggtcttgacttcacgaaagatgatgaaaacgtaacttctcaatcgtttatgcgttggagagaccgttttattttcgtttctgaagctctttataaat
ctcaagctgaaactggtaaacttgaaattaaaggtcactatttaaatgctacttcaggaacatctgaagaaatgttaaaacgtgctgaagttgctaaaaatttaggtgtacctattattatgcatgactatttaac
aggtggtttaactgctaacacttcattagcacactactgtcgtgataatggtttattattacacattcacagagctatgcacgcggttattgaccgtcaaagaaatcatggtattcacttccgtgttttagctaaa
gttctacgtctatcaggtggtgaccaccttcactctggtactgttgtaggtaaacttgaaggtgaacgtgaagtaacttaggtttcgttgatttaatgcgtgatgaatacatcgaaaaagaccgtagccgtgg
tatttatttcactcaagattggtgtggtttaggtggtgtaatgccagtagcttctggtggtatccacgtatggcatatgcctgctttagtagaaatctttggtgatgacgcttgtcttcaattcggtggtaaatgg

>Chloromonas_rosae__-_rpl2
atgggaattcgttttcttcaagcatttacaccaggaactagaaatcgttcagtttctgattttagtgaattaacaacaactaaaccagagagttcgttaacctataatttacaaagagcaaaaggacgaaatca
ccgaggcgtgattacatcgcgtcatcgtgggggtggtcataagcgtctttatagacttatcgattttcgtcgtgacaaaattggaatggaagcaaaagttattacaattgaatatgatcctaatcgtaatgcac
gaatagctctccttcgttatgatgatggtgaaaaaagatatatattacatccacgtggattaaatattggtgaaaaaatcatttcagaaataaatgctccaattattattggaaattcacttccattacgtaatattc
cgttaggtgctgaaattcataacgtagaatttcaacctggttctgggggccaaattgcccgttctgctggagctgttgttgaaatttagcaaaagaaggcaattttgttacttttacgtttaccctctaaagaaat
ccgtttagtttcaaaaaattgttgggcaactatagtcaagtaggaaatattgaagcgtataatttaacttttaggaaaagctggtcgaacacgttggttaggaattagaccaactgtaagaggttcagttatga
accctgttgatcacccgcatgggggtggagaaggccgtactccaattgggcatagtcgtccattaactccttggggcaaacctgctttaggtgttttaactcgcacgccaaaaaaatatagtaatcaatttat
tattcgtaaaagaaaacaa

>Chloromonas_rosae__-_rpl5
atgacacaaagactcaaaacatattatacagaaactattattccaaaattacaaaaacaatttcaatatgaaaatattcaccaagtgcctaaaatagaaaaaattgtaatcaatagaggtattggagcagcttc
tcaaaatcaaaaaattgtagattcgtctttaaaagaactggctatcattgcaggtcaaaaaggtatcataacacgctcaaaaaaagctatcgcagggtttaaagtaagagaaaaaatgccagttggaattgt
agtgagcttaagaggtgatcgtatgtacagttttttagatcgattaataaacttagctttgcctcgggtacgggattttcaaggaattaatcccaaaagttttgataaaaatggcaattatagtttaggtttagaa
gaacaattaatgtttcctgaaattgaatatgataaaattgatcaagttcgaggtatggacatttcaatcgttactactgcaacaaaacaagctgaaggtttagctcttttaaaagaatttggtttaccgtttaaagc
ttag

>Chloromonas_rosae__-_rpl14
atgattaaacctcaatcttatcttaatgttgctgacaatagcggagcacgaaaattaatgtgtattcgtgttttaggtggaagtcaacgtgaagctggaaatattggagatattattattggagttgttaaagatt
ctattccgaatatgccattaaaaaaagctgatgttgttcgagcagtaattgtacgaactagtaaagggttaaaacgtcaaaacggaatttccattcgttttgatgataatgctgctgtaattataaataaagaag
gaaatcctagaggcacacgagtttttggtccgatagctcgggaattaagagatcgtaatttcactaaaatcgtttcattagctccagaagttatttaa

>Chloromonas_rosae__-_rpl16
atgcttagtccgaaaagaacaaaatttcgtaaacaacatcgtggtagattaaatggaaaagcaactcgtggtaataaaatttcttttggtgattttgctttacaagcattagaaccgtgttggattacttcgaga
caaattgaagctggaagacgcgtcttgactcgttatgtacgtagaggtggaaaactatggataagaatttttcctgataaacctgttacacttcatccagctggaactcggatgggctctggaaaaggatat
cctgaatattgggttgctgtcgtgcgcccaggaactataatctatgaaatgaaaggggtttctgaaataattgcaaaacaagcttttcgaattgctggccataaaatgccagttaaaaccaaatttatattacg
cacacaaacacttttgatcccaaaaagtactcctattcttcttccgacaccaattcctgca

>Chloromonas_rosae__-_rpl20
atgactcgtgttaaacgtggtaatgtatctcgaaaacgtcataaaaaagtattaaatatgtctaaaggttttcgcggagcgggctctgttttatttagaacagcaaatcaacagaatattaaagcattacgatat
tcttatcgtaatcgccgacaaaaaaaacgtgattttagacggctttggattgcacgtttaaatgctgctgttcgttgttatggtcttaattataatgagtttatgaattatttaaaaatacgcagtattatattaaatcg
aaaaattatagctcaattagcaacagctgatactgaagcatttatgcaattacttttattttaa

>Chloromonas_rosae__-_rpl23
atgattgatttaataaaatatccaattattacagaaaaatcttatttaacttttatttaaaaatcaacaatatacttttgatgtagatttacgattaagtaaacctcaaattaaaaaattatttgaaaatttatttaacgtaa
gtgtgattgcagtaaatactcatataccgccacgtaaaaatcttcgtgttggtacaactaaaggatatcgagctcgttataaacgagctatattgaccttaaaaaagggtcaatcattaaaatttgctataccttt
aacatta

>Chloromonas_rosae__-_rpl36

atgaaagtccgtgcatctgtaaaagctatttgtgataaatgtcgtgttattcgtcgaaaaggtaaagtaatggttatttgttcaaatccaaaacataaacaacgtcaaggaacg

>Chloromonas_rosae__-_rpoA
atgattaaatttgtaaaaaatgatttttttatttattgtaaagaagcacgtattgaaagtcatcgaagtttttatgggtcttttagtcttggcccatttgaggctggacaaagtattacaattgcaaatgctttaagac
ggacttttattatcggaattaaaaggtttagcaattacctctgctgaaattgaaggagcgtttcatgaatacgctacacttccaggagttcgagattcgatattagatatttattaaatttaaaggatattgtcttaa
aatccattaattatgggaataaaaaaagaggaacttcttatacagcaacttttaaaaacgaacaacttggatatttaaaaagcaaaggaccgggaattgttacagcgtccgatttaaaactccctcccggaa
ttcagtgtgttgatccagatcaatatatagcaactctgtcagaagaaggaattcttaatattaaatttcatataaatgaaggagcaaacgctaatactgatcaaagtctttcccataataaaccaaaatctcctg
ctgcttcttttgcttatccttctgacagctccataaaattagaagtaaagaagttagaaggagcagaccggttaggggcctcaagcatagccgccccgctacgcgcacctctggtgttaacgtagcaaaa
aaaattgcctcaagcaaccccaatatattaggagggaatggaatggatgggcagagcccatattatggttttactataattccacctcaagtatggcaaagcccttccagcaagctacatacaaaggatag
ttttttaaaagaaataaaaaaccaagcttcacctaatgttgttttttccactgttccagtagatgtcaaaacatataaaaatttaaatctaaatgccaagccagttcctcctaacaacaacgtttatatggatgggc
tcacaaaaactataaataaattatttttttctcctcgaccatctaagctgcgcttatctaacttttagatccgctagtctcgatacacttgctctaacaaacaataaagagcaaaaaaaaaaaaagagccctttggcc
ttggttatcgaaagaaaatgagcaaaaaaaaatagacattgggtcccttagctatggctataataattcgtctaatcaaattaatggaatttaaaatgggggtgaggtcaatataaaacgtttatggttaaaaa
aaataaaagcaattcaagctagttatgtccataaaattaatagacaaaaaaaagcttttagtttaaataaaattaataaaacagttttaaaagagcttggtaacttaataaaaaaaaaaaattcagatttaaaga
agaaaattccttatatgaagccataatagcattgtcaactcctttagcaattcagcaaagctttatgcctgtgaataaagttaattatattattgaaattaatgaacaatcagttatagcaaaagatagtttaggt
aacgtaaataataatatatggtaccgtaaataataattataatactaatataagcaataactctataaaattaccatttcaagttattggctttgccaataacttgacagttaaaaaaacgccttcgtctaaggctaa
ttttgaacaaatgacgcgctccttgctttttggtttgcaagcaaaagcataccaaaccaaccaaagcaaaggggtaggaaagcttggcttgcctagcactatagatgtaggggccccaaacaaggtatgt
acggcaaaaagccaaaggacaagtgcttctagctgcaatgtcacttctaatggccttactaatgtagaagcagcctattataaaaataaacaagctatttataaaataacaccatttaataaacaagcttcgc
ctgctttttttcagttaaaaaaaaacaagcatggcaaatcgaaccaaaagcgcttttaaaaaaaagccatattgttttgatgaaagcaagttgtttaagaacttttttcttctagctctcaagttaactacattaaa
gctccgcttaattatcattatgtaaaggccctaaagaaaatactaaaaaatttactagaggaacttgctttctaactatcctcacataaaatctaccttaagtgtactacaaacatctttaaaaaataatattgtt
ttagaaatttggacaaatggaagtattcatccccgccaagcacttttatgatgcatgtaaaaatttattaagtatttttttcaaaattagaaaaagcccgttttactagagcaatttttcaatcagaaaaaacatatga
gaaattaggtttatcactcatctcgaaggggactacgtctatggctaaaggcccgggccctaagccatggtggaaatcaaatttaattaagcaaagcttaattaaaaatgaatggcaaacttctaatggctc
tttacctttttcagtagtcaacaaattaaatcagacccctttctttaatcttatgggggaaagcttatactacttttaaggctccttctgctttgcttgggcagcgcagcagccaacccgttctctctgcttgggttg
gttcccaagcgcagctgccaacctcaagaaagctagcaaagagaaaggaagcagcagagcagataatcagcaaaacagataggtattgtttacctccattagaaatagcaggttaccctagtgttttat
taaaaaaaacatttaatagacaaattattacaatttcattattggccgctctttcaatcaaattaaaaaatcaaaaagtattaaaaaatattagattttgttttactagtagtatgttccgcatgaactctgtatatgga
ccaaaaaaaggttctaaaattcaagcacagcaagaatttaacatagcgagtagcgggcaaaaaactagaagcacacaaaaaaaacaacagagctttttttatgaatcttgggccatattatctagttaccc
caaaatcaataatttttattaagcaaaacttaaaccaaatagagggtagaaacaactctgttaacagagttagcaaaagtcaccctttagctgcgctaaaagtgaattttaatgccttgcataacacctgcttcg
cgggagctgaaatgccaaccaggctaagcctaagtggattagataagggaatttttaatagtaagcaattttcttttatttcgttattaaaaagttataaatcttataaagaaattgcttttaaaatataactatttag
acccttttaaatcaaaaaaaagatttaaaaacactgcccaactattagactatttgggagtagcaaaaaaaaattactatgcatggcttgatttagaaaaaaaaaagaaagcgccttaagaaataatagtca
atttctttttgctcctatttggggagttatgtttaacaaaaataaacaatttttttttagctatgccaaaaaatcatcaaatagctttttttcaaaataaaactatataaaaaaagatttttgactcatcaacgcagcagctt
tgctaggtctgctgcgcttgggaaccaaccaaagcaaggagtcattaagggcagttataaaaatggatttttatttatccaaattagcgaatacgagatccaagaataacccactctatcccctccctaataaa
aagggtttgggaaccaaaagaggagctatttaaatattggtttgctcaacccttcggagaagcaaggatggggctcgttgcttgggttggttcccaagcgcagcagccaaccccattattcaccaaagca
aaaaaaatggcccctccttgcttggaatgggaaccaaccaaagcaaagggcgctgcgagtctcgaaagaaaagaaagaaaagaaagaaaagaagaaaacagtaaaggaatgagtaattataaaac
attagcttctagtattttatctctaagtgtaggggctcctagttctctgccaaaagcagatgcgtccttagtggtaaatactcaactgtgcttagctaacaccctaaaagcaaatcaagttcctttatggcgaga
agctaaaatgattaaagtactaggttcagaagcctatccttcctcaaaaagcgcagcttcctgggagaaaaaggggtccataattattaaaaactaaaaaaaattgatgttttttttaaaacaaaaacaaatc
aattggaatataaaaatgtctctaactgattataatatcttaaaaacaaaaccgaatttgttgatattggaaaacttaatatttcacttcgtccatatactttgtttaaaaagagctaatattaatactttaggagaac
ttttaaaaaaatctaaagaagaattattagctttaaaaaattttggaaaacaatcattaaaagaggttgaagaaaatttaaaacagcttggactctctttaaaatctt

>Chloromonas_rosae__-_rpoC2
atgctagattgtatggcctctgttttaaaaaatgattcatccaaaattaagtccaataagttaaacttttaaaaaaagtattgcattatttaaacaaagcaatcaccttgtagcctttggctataatagcaaaattaac
cctgttaattttaattccgttggcacttagaagctattaaaccgaagttaaaaatttataattacacaaataaaatgatttcttctttttaattcaaatttttaaaacatcaaaagaccaatttaagtcatctaaaaaaga
aaatattaatggaaaaaacaaaaaccaaatatctccactaaccttaaaaaatccattttggaatcaatgttttgataaaaatcgtttaaaaaatttttgttttatggttttagttacctatggagaatatcgcacagt
agaattaattgaacaattaaaaaatgttggttttgaatatgctactctagcaggaatttctttaggaattgacgatttaaaaattcctccaaaaaaatctttttttaatttatgatgcagaaaaaactacaaacttagct
cttaatcaatataatcgcggtgaaattactggggtagaacggtttcaaaaaatgattgatacatggcatggaacaagtgaatttttaaaacaagaagttattgattatttttgaaattacagatacctttaaatcccg
tttatatgatggcattttccggagctcgggggaatatttcccaagttcgtcaattagtagggatgcgaggattaatggctaatccccaaggtcaaattttagattatccaattcgaagcaattttcgtgaagggt
taagtttaactgaatatattatttccacatatggagcaagaaaaggaattgtggatacagctttacgaacagcaaatgcaggatatttaactagacgattagttgatgttgctcagcatgtaattgtttcaaatttt
gactgtgggactcaaagaggtatttttttaaatgatatgaaagagggtaataaagttttatattccttgcagaatcgattaattggtcgagttttaggaaaagatatttttcataatcaagaggcttttaaaaagta
tcatgtattaaaggggagccatgataaagaatctcctctcatgctttataaaggtctaggcaaagcctctaatgaaagacccataatgaaccaaagcatagataaaaatttgtctttggaggcttctgcgaaa
aaaataacaaaaaaaagaactaattggacgactaaaagtggaggagagggttgctaatttaccactccaaatagaagcttcgctaggtagcgctgtccctgcttctagtggaactaggttgattagagaaa
cgaagctcggggtaaatagcccctactctccttttatttttagaaatagtgaaataacaagtgatttagcaagtaaacttagtaaactctctaataaaatttttgttcgttcaacttaacttgtgaaactcaaaaat
cagtttgtcaattatgttatggttggagtttagctcaaggaaatcttgtttcgataggcgaagctgttggtgttgtagctgcccaatcaataggtgaaccggggacacagttaacaatgagaacatttcataca
ggtggcgtcttttcaggtgatataacggatcaaataaaagctctttataatggcattgttaaatatttaggacctatagctggtacttttgatccgaacaccggaaggaaaaatttccttttttaactaaaacagaa
ggatttctagaagtcctttcaataactactgaaaaagaatttgcaaaaaaaaaagtctatataaaattccttctttttacattattatacgttcgacatggcgaaaatgtttttgaaaaacaagttattgctcaaatttca
aatatttcgcgacaacaaaataaacgagatgatgcagaatttaccataaattctgaaatggaaggccaatttcgtgtcatgaccttaaatttattagaaaaacaagttggtccaaaaccaacttttagctgtcaa
aaaaaaaatcagtcccaagcaagcctctaagcaaagcttagtgaaaaaatctagttttgtcttaaatgtaggagccccccctaacctttataatgggagtcaaagttcgttagatgtcaattccacaatggaa
acattacttgaagctagcgattggggttatgcttggattttagctggtaaaatctatcaacttcctatttcatccaatatattacctaaatctggtgattatgttaataaaaaaagtcttgtaaattctgtagctttaaa
accatcactgcagggagtgaattctggttttttagaaactaaattaaataataatttgctgttaaataaacctattttatttttagagattaataaaatgttttatacaaaaattggatatttttaaaattgccagaatat
ccttttttaattccaatattttaccatcttttgttaatgatgtatggctagcaagtgtagcttacccctttgcttgggatgggaaccaaccaaagcaaggagcgcaacaagaccccaccttgcgttgcttggttc
attatcattatcattatcattatcattatcatgggaaacccataatgagaatgataataaactctagctttgccatcaccgaacgttctaaggttgcaaaaaaaattactttttttaagcaaccgttttttcttt
ttggtgcaaattcgtacaaaatccaaactaagccgtctaattatatttctcgctacgctcaaactccctacaatcgcttcgtggttcccaactataaagcgaagcttagcaaagcaaatttaaagttaatgtaa
ttaaaagcactgtagaaaccataaaagaattctaccgagaatgggtagaccccggaattatcagtttaaaaaataatcttataaaaaaaaaagaaaattctttgggttttaatcgtggaggtcaaaccgtttta
actctaaatctttataaaaatttaaactaaaacgcttaaacttttttaaaaaagtagaagttaaaaagaggaatacttttttaaaagctaactatttaaataataagttttatttcaatataggacatataaacttatta
tatagaaataaaatttattataacctattagctccgctagaagcaaaatcttatttacataaaaatgatttgttatttacacctctagcttcgcaacatttaacttcttttctaacttttttgggtcgcctatttgtgtagc
caaaaaaactgtaaacactattccatttatactactagtcagagtactggtctagtttattcaaaagttgcgtctatagggaaaagaagttccgactctgctcctatccttgctttggttgctggagcccttccct
tatttcataaaagccataggtttagaagccataggcttgatggtttgttcccaaacaaaagcttaaaaaccaagcaaagtaagctacgctgggtacctagtagcagccctaaactacaaattctaaataaag
cctctaatttgagttataactatgctgatagtaatttattaactaatttaaaatctaataataaaaaactctttgcagaaaaactacaaacaggacgagccttaaatactaatagatttttagtggttctaaaactat
aggtaaaaatagttcagcaaattggagcaaaaatttagtttatttttagaatggtttccacattctaaaaaagttacaagtggtgcattacttagtgattcacaaactttacaaactagtacttatttaccttcaatg
tggccaaaaaaaattacttttttaatcttttaaaaatcttttttcttttattatttacgtcttctctagcgttttttcttaactctaacgttagttacactaaagctctaaagtgcaaagcactttaaaaaaatttaata
aagcttctaaagtgccaacgaagttaaaattaacagagttaattttgcttacatctttttttcaaaagaagagcctttaacatttacttcctacctagctataaattaggtaaactatctaatacttttcaatttcta
aagtgctttggactataaaatttttggttcccaagaagcaaagccagaggcttacagctaagcttagcttggttgggcttgggtacctagcaaaagcttacaaggataggattagttgcactagaaaattgg
gcttaataaaaaaaaaatttaaaaaacccagctagcgctagttatatttttggtgcaaagacaccaaacaaaatatttgttaatagttgtgtaagattgagtgttagaaccctttaacttttctaactcacttttagttcc
actaagttaaatagtaaatatataaatagaagcttattgctaaaaaaaaagagtttaaaaaaatttcttgctttagctatgcaaagagtcgttacaaaaaaaaattattttggcaaagcctttaatgacttctttttaa
caaatgataagtctcatagcaataacctagattcctcaaaagctgtagtaagcaaaattaaccctgttaattttaaagtgctgcgcactttagaagctttctttgtaacttctaatgaaataacattgttagagag
acctaaagacgtacaagttccttatttgagaaatatagcaaaaagaataaaagggtatattaaattactaaatttttaaatcgtagtagtttgaactttgtggtcgctttaaacagcaaaaaaattcttattataa
ttatgcaaaattaattgcaaaaaaagaacaaattttaaatgaaacgttatatgaaaaattttttatatttcaagacaatttttgaaattccacaagaacgtttaactatatctaacatttttttaaaaataagatgtct
tataatatttttatgacactttcaaaaaaaaatcgacaaggaattaaaaaagcacttttctaattctagcaacgctagcgggactttggtcaaactagaagctttgccgatgaaaccacacttaacatagtg

aaaagagttagccctagctaccctaataaagctaatggtggaggcccaagtagcagaaggggtttgcttcctcaccaactcagcgtggttggcagagcatcaacaacgttgtggaggggttcttctaac
aaaattagcaatacgaattttaaagttaatgttattaaaagcactttagaagttacacagaaggatagcaagcagcaagggaaaggagctgagcattcccaaagaatgctactaggttcccaaccattgtg
ggaaagaagcaactccaagcctaggttgatacaatttaaagctcacgtattaaagattcggctttgataagccatgctttggtacccaagcatgttcattacacggactatggccatactttcctaaaaaaa
aaacactttaaagaatggacaataaaacaatcaatttttgtcaaaacgtcttcagctacgctaaaacaagtcaccccgtatccaattaaactctataaagggagattgaaaaaaaaagtgaggcttttacaa
gatagagttaaaaagattacgagtggattaatttttccctagattatttaaaattgtttaaacttattaaaatttttttaactaacgattataaaaaatttaatactgcgttaactgcgttaaaaaataaacatggtatc
aagttattaactttattgatggctaaaggtccttatcctaagccatggttactaatccactatggcttttattggggctgctacgcttaccccaaaggtgtatggggaagcctattatggggtccaataatagaa
gccaaaaagatagttaaaaaaaacttgagtcttacttttctcaatagggagcaaaacttttagtccaaaaagcttccttcgtgactttctcattatcaaggggttggctgctgcgctgcccaagcgacgctgcc
cattatttaccaaagcaaggggtactctagctagaaataaattaaataatttcaaattttctaaaaaactaaataatattgctctttttataacaagcaaggacctttaatgaaaaaacaagcatttttttgcttaaa
taaaccagttttaaatttgccgcaaaactttttaaaaagaaaatagtaatttatcctcatcatttaaaaaaattaacaaataatttttctaatttttaaagaaaataagcgtagcttttaatggttttcccataattagtag
gggcacgtattgtaaattatttttgatagcccattatgtaaaacgtgggccaagcgttggtaatcaaaaactaatttcattatctactctttactgtcatgcaaaaaaacaagtaacaactataatccaaaaaaat
tcatttattaagcttcctgtttggcatgcaatgcttacaagacgcaaggggggctgctacgcggcgtagcttaactaagccaaacttaagttacactagtgcagctcttcctcacttttgcttccctgctagcgta
gcttttcaagagcagagaaagaaagcttctaggttggttttgggaaccaagaagcaagaggggcctctcccaaaccagcgtagcagaggaataacggaaccaactaacaacgttgtagtaaggaaag
gaaaatctcgaacctctacgataactaacccttatccaattacaagttctgagtctattaatcatatttcctggattaataactattttgcaaaaaaaaggttgggttttttttccttctatgtctataaaaatgaaaag
tagtatttctatctcacaatcatataattatcaagaggaaccctaataataggtgctttccctaagcaaacccgtagtctagtgaagctaactggactacaaaacttacagtttataaggccaggacaatttttg
ttgccttatttaatgtttgataataataatatttatgttgaatggtttaataatttttagttctgcaaactatattcataagcaccggttatttccctttttatggttctttattaaaaacgtggaccataaaaaaagttattaat
tctaaaactgcaaaaaaaactgttagagttataaacacaggaattaaaaaatctcaagctctagttaaaaataaattgtttaaagattcttatagtttagacccattatttggtgctcacgtttatgggcatatgtc
tatttaaaccctttggtttgctctattccctctgcttggaatccaactatcttgcctatagcttggaaggtaaaaggaaagtcaaacaaaggagaggaacaggccctttttgaaaaaaaaaagagaaaaaaa
aataaaatttatttaaccatttttaattctgttctccccccccctcctttggtgtaaagccaaaagaagatataattatacggagtaatacattgctgcgtaaccatgtacatagcaaaaatcaaggggccccta
acaagtacggagtgggcaggaagaaaaaagtagagacttccagcgtagcaggccttatgcaaagcttaagtcttctaatttttgcctaatctgaaacaatcttacccaataatggaagccttttagattc
gaaagcaatttcttcttctcttaatgaattaaatttaaataaaacaagtgattcctgttgggtacaaaatcaaaattgtcgttatctaccaaatcattccgctagttctataacttttaaagatttaactagttatgg
tcatgcgaagcataaacctatcttaattattcaaaaagctttagaattgactattgcgaagcctattaattataaaaatgaaattgcagaattatatgcctctaacagggatttagtgatgactggaacaaaaga
ttttttaatttcaagtttacttggaaaaacgcttcaatcaccatataaacctgttttttctgagaaaattaaaggttctaatcggttaagtcattttaataacaatcttagttcttgtttatctatgcttatgttaaaaaaaa
atactatatacaaacttaaaaaaaactcaaactataggcacacatctcctaaaggcagaaggggaggagtctttaacttatactgctcatgcttgggttggttcccatccaaagcctagccttccttgctttat
catcgtataggcttcacttataattattattacggcaaagccttgatcatgaagatgataatgaaccaaagcaaggaaaggatacttgaaaaatcagttaagaagattaaaataaaaccatttatgtccttcctt
tcttttgccttccctctccttctgcgcatccattccatagaggcgctagggtaccaaagctactttcccaaaggataggaaaaagatcttttatgttatctgttaaaagtatctggaaaaagagcttagtcagct
tcctgtctttgcttttatcattatcattatcattatcattatcatgggggaacccataatgataataataatcaggctaagcctaatgaacttagcaaagcaaggcagggtaccctagcaaacgtagctttaacag
gagcttctttttctgtgactcaaaacaaaataagcaaattaaaggttccttttagaggcaaagcttatacttgaagagactaaaaaaacatatattagtttttttctaaaaaaatttatctttggcaatcaattag
taaaagaaagttaaatttagtacaagcagttaagggcttgctctgggtaccaagcagaaggactcaaaagtctactaaaagtaaattaaaaaaataatcctttgactaaaataactattattcaattgcgtaaaa
ataaatcttataattcagcattagttagaaagcaatttttttgggtctttacttaataaacaaaataaaacaatagaaggtgtagctagtacaatgccttattttgattttaaagtgaaacataatgaatttttacacaaa
aatggtctcacttttaaaattaaaagcaagtagtggagttaaagttttagcgaaactagcaaaaagccctcgttttaaaggcatccaagaggttaccccagcttctaaattaggggtttcgcacagaaagata
agccaagctacgcatgaagctccctacggttgggctgctacgctgggggcccctagctttgcctctaaagcggagcgtaactttgttagcgcagctgaagttaactcttgggaaccaaaagctcttcagc
gaagtttaaccactaaaagcactttagaagagcaagcaaagcttagaaaaggttcatttctgtcttctaatgggcaaagtccattaaaaaagtcattattgaaaaaaaatagttatgcacataaactagggtc
ttctatctcaacacctcaagctacacttgcaaaaagtctcagctcaacaaaagctagtaatttaatttttaaaaaaccaattaatctaactcctttattaataggttataatatgccttatggttttgatttagattttaa
tgtgccaaatattcattttaactctaatgttgcaagttgtaaacttcttggtgcgaagcaaggaactttacccacttttcaatttttaacaaaaaaacaaaaattgaatttttttaaaagagagtaacggagctaaaa
tgcttttacagagtaaaaattcttggttttttattttgctcaattacaagcacttaaaattccgtcgttatttgcacaaccatgttttttcttttttatgagcaaaaaattaccaatttttataaaacgttttaaacaaagc
ttaaaatgggttcaagaatctgatccttttgcttctttaactaatcataatttttcttataaaataaaaaggagtatagggaataagaaaggcaaattcttcatttattaattttaattctactgcgttaattaaagctcatt
catttttaaatttaaaaaaaaaaaaagcaaaaaaatcttctgtttatgaaaagttacattggaagttttttttgacttttattcttatggagcatctttatacaaaaataaattgcttttttctagtggaacaaagagct
cttcttccttttctacccttcctctggagaaaaaaggaaaagaaaaaagagagcaaagtatcaaacaaagcttaataaaattttcattcttaagttcttttgaaggtgaatttgttaagctcaattcacatctaaac
tcacacctaaattcaccagcttccctaagtgaattgaaaaaagcttctaattcaactgcattaattcatttaagaagcttaaaaagttatgaatataaaaacctcactagtgttttaaatttaaatagttataataag
ttagatctagtaaagcaacaatggacaaaacctattatggacaaagcattctttccttcgctctgtaagagctttccgcgtcatgccagaaatggaaaaactaaagctaatttttttaaatttttttgggtttttcagg
tcgcaaaaagttttctaggagcagagcaaaggaatgggaaggaaattttttatggcaaagccatgaaccaagcagagcggttcaggggcagcgtcgcagcttcttaatgcgtagcataaacccagca
acagcttacaagagaaaacttacaacgttaggatggtctttggtagggcaaacaagagttcgaacaggccagccacttaataaacgcccttcaaataaaactctgacttcttttaataagtttcttggaccat
caagtcttattttaacaaaagctaatttaataagttattcgtttaatttttaactctttagctccgctaagaaaaagcttaattctaaagtcaaacataacgtatattcttataaaattgctcgagaaaatctttctaaa
gaattaattaacttactagatattataaatataataatatttttaattaaagttagttcgaatctttttgaggttggttatggcgaaattctaagtgaaagtaataaaaattctcaagctagaggcgaagcctttaaag
cgaagcttagcaacgcaaaatttaaagttaatgtaattaaaagcactttagaagctatacaagcatcactgaaaaaggtggagcaagcaaacctaataatgttaaatttcacaatgtgccacaagaaaatc
ctttagccaatcattacaacgaatccacattaaacggacaaactaagcaaagcttaaagctttattcgcagaaatccaatggatttttaacaaaaaatttttatgcaggtgaagataaaattggttttttataaa
gcaaagctcttcttcggcaaaacattccaaagaatctctaacaacaacgttattaattaaaaatgaatatatgataatacaaaatttaattagtgtagctaaagcaaaaaacattttattagctacggtaagcaa
cgcttgtccacatccttctagtttaagcacacaagttccaaaaatagatattggggtaggaacttataatttagtgttaattccgagattaatgttggtagttacggcgaagccattatcagagcaaaaatttttt
aaaattaatggattaattgctggtttaccaaaacaacaaaatgtatcattatttaaaaataaattatctttaggggaatttttttgtttatggcgacaaaataatgtctattccctttattaaccacaatacaaatttccc
attattaggggaaagatcaaaacataaaatcgctacattaaaaaataaaaatttggaattccttttgcgggtcaaataattcatattaatcaaaataaaattacattaagacgcgcacaaccaatatttatttca
cccaaaagtattttacattttaatgatgggagttttgtgcgtaataatgatccagttatgacttttaacttatcaacaattgaaaacaggtgacattgttcaaggaatacctaaagtggaacaatattttgaagcgc
ggataaccaaacaaggaagattatttcgtgattgtttgccaaatttattaaaaggcttatttaaaagatatagatctcaattaccattagaaaaagctgttagacaaagtttttataaaattcaacaaatcattgta
gatggcgtattacgtgtttatagatctcaaggagtaacaattgctgataaacatttagaagttattgtaagacaaatgacttctaaagtgaaaattataaatggagcgcaaactggttttttttcctggtgaaataa
ttgatttacaatttgtcgaaagagttaatttatcattaatgaaaaaaatacaatatgaaccattaattttaggtataacgcgagcagctttagaagtggatagttttttatcggcagcaagttttcaacaaactacc
agagttttaagtagtgcggctattgaaaaaagaaaagatttttttaaaaggtttaaaagaaaatattcttttaggaaatttaatgccagctgggactggttatttagtttctttttgaagatcttaaaaataaaaca

>Chloromonas_rosae__-_rps2
atgcttactaattggaaaactattattaaatcgatttcaaaaatccgtcctattcttaaagaaaaacaaaaagtcatttcaaatattcttcaaaaacgttcaaaaattcaaaatcgtttactaaaaaaagtgttttctt
taagaaaaaaaagtaaactcttaatcaaaaaaggtaaacaacttatagctaaagtgaaaattgctagtgtaactgattctgttaataataaaaatagcttcgctaagctccgcttaaagctaaatcagattaatt
taaagcgaaaagaacttttgaaaaaggtcaaagtttattagaaaaacgccaaaatcttttaaataaacaaaaagaattaaaagaaaaaacttttaaatattaaagaaaatggttttttaatcgccaataaatata
aagttttattaacgcaattaagaacaacacatactaaacttcatgatttaagacttcttttttcattaaataaacaacttcgacaaattaaaaaattagctaaagaccaaggaaaaaataaatttatacttttcttatg
ggaaatttaaacaattaactacacttgctacagaaaaaaacaatcaaggtactcttttcacttcaagcaaccaagcttctcatacactagatggctcaccattatttaatggaaataaaagtcaagacatcaaa
aactttaatactgactttattaattcgaaaaaattttggcttattccaaaccctcctcaagaaatattaaataaacttctaagcaaaatcaattatgtacagaaaacttcttcttcttacatagcaagccaaaagatcta
tcctctcctacttggcaacaaaacagagccccccttggaggcgaagcctctaaagctaagtttacaaaacaagctatagtgataaaacaaaactaaaaaaacaagcagagcttactgttaaaaaaaactt
ctttttaaattctatggcctgctacgcttacaccaaaggtgtatggggttatgctacgcatgatactttagctgcagaccacaaccagtttagtattaaagaaaaaaatacaacaagtcttttaaaggaaagta
aattaattattttaactaaacttttaagtaaatttagtcgtttttaccttatattaaaaattgtattaatttaatcaaaaacttaatgcagaatatagaaaaaaaaattatggagttaaaaatcaatttaaaatatttaaaa
aaaacattaacagattatctttattaaatataaaataaatagtgaattagaacttattaaaacaaaatatttatctgaacaacaaattattcgtgtcttaagaagaaaattaaaacgtcatgaagctgaaaaaaa
attgttacaatttttacctaaattacgttatttacctacacctaaaagtcaaatatctgtcgctattcaattattaatgaaacgatttgttgatcctaaaatgaaatatccgttggatctaatttatgaccaagcttcgct
tagaaaaaaatcaaaaaaagtagctgcagctagaaaaaaaaaatggcaacgattagaaacttattttggtggtattgcaaacatgactaaattaaagaaaacacaaattgctaaaaatgttgctataatcat
aggtcaacaagaagaaataaatgctattcgtgaatgtaaaaaacttggtattaaaatgtttcatattgttgatactaattgtaatcctggtttagccgatcattttattccagcaaatgatgattctagaaattcaat
aaaatatttattatcaaaaatggttacacgtattagattagctcaaaaaattcgtttacgattaaaaaaaaaaatgcaaagtaaagaaaaaaaaggtaatatgtttctaaataaagttaaaaaccaaacaatatc
aaataaacagaaaacaattaaaagtaaggcaatttcatttagcaaaacaaagcctaatagcccgaaagtaggtgatatttgtaatttaattgtaacagctttaggtcccaaaaatattggtattgatgaattttc
atacggatattctattttaattccaaatgccaaattaggagaaaaggttaaagcacaaattatcaaaaaaaatttaaacaaagcaaaatacgctgttgctaaaattataatgcctaactctcaagtattacataa
agctaaacttaaatctttagataaagacaaaactttccgtgatggttctgcaatccttgaagatggaaaagaaaaaataaatgttggtgatatttttaacagtaaaattaaataagattgctattcaaactgataat
tcttgtattggtatagtgaattcagatttcaaattaccgataaaataccaattcataattcctaacatagttgagactaaggttggtcaagaaatccaaattcaagtaactaaaatgaaattaaaatatgggtttgc

tttaatgaaatccagcgtagttaactcagatcaagtaaaacctaatattaagcaaaacttgggtacttcaaataagcttcctgcaactcgactaaaccttctttcatcccctcaaaaagcgcaacaaacgaac
cctttggtccgtcgcaaaggcgtttatccattgcctcctacggttgcttatcctactgcattggctccaattaagcctttggaagagcgaggaaagcagcaaggaattaatggtcaaatatggctaacaaag
ttacaggaagggccggaaaaaccagtgtaatagctccaagtcataaagaatcgactaatttattaaaacctttattttctttaagctcacgtccaagcgaagcttatgtttttggaacctctctttccaaaattg
gctttccaagtcttaacattccaaaagtgcaaaattatggtcaaaaaatcactcctaataatagttctttttcatctttttatttctgggagccttgcagatttaaataagactaaaaaaacaaattacaaacctttaa
cgttgcgaagatgacaaaaggaagtaaaattactttaacattacctaaaactgcaaaaaaatatgctaaatatttagtaattaacatttttaaatcaaaatattgaaacgggggggggttttagtgatctctagtc
atccaactactcttaatgatggaaaaactcgctataattcttatatttttatttataaaaaatgtattaagagcaaaattaggggataaagtccgagtaaaattcacaaaaattggttcttcatttgctattggtgaag
ttattcgtatttcacccaattcatcattcacaaaattagctcgtgttcgaaaaaatttaaatgaaatgcttgtaaatggaatgcattatggagaaaaagttgtcaaatgccaagctaaaatgcgtaattttatttgg
aatcgaaaaaaaatctccttctctccccttttcggcaacttatcctttaaattccaatagttataaaggaagtccagttaaagcccgcaatttttctggcaaagccatgggttcaaggctctgccttaaagccggc
caaaggctctgtgacttttgctgggaagtccctagaagctttaaagcaaggagggctagtaaatacaaactcggcagggagcaaaaaaagctattcttttggtactaggggccccaaacaaaaagttctt
ctaactatgcgcagcaaggccagtgaaacagagcaccctaaggagttacacagaagccctttagaagcctctaaaaatcaatctttaaacaaaggcgttgcttaataaaaaaaggtagacatttaattaa
tctattaaaaacccgtcgttgtttaaaatataggacttaatcaattaacaaaatatgctgctaaaggaagaacttttttatttgttggaacaaaaaaaccagcagcaggcttaatagctcgcgctgctttatttagta
aaacctcgtttttcgtcaatacaagatggttaggtggtatgcttactaattggaaaactattattaaatcgatttcaaaaatccgtcctattcttaaagaaaaacaaaaagtcatttcaaatattcttcaaaaacgtt
caaaaattcaaaatcgtttactaaaaaaagtgttttcttaagaaaaaaaagtaaactcttaatcaaaaaggtaaacaacttatagctaaagtgaaaattgctagtgtaactgattctgttaataataaaaatag
cttcgctaagctccgcttaaagctaaatcagattaatttaaagcgaaaagaacttttgaaaaaggtcaaagtttattagaaaaacgccaaaatcttttaaataaacaaaaagaattaaaagaaaaaacttaa
atattaaagaaaatggttttttaatcgccaataaatataaagttttattaacgcaattaagaacaacacatactaaacttcatgatttaagacttcttttttcattaaataaacaacttcgacaaattaaaaaattagc
taaagaccaaggaaaaaataaatttatactttcttatgggaaatttaaacaattaactacacttgctacagaaaaaaacaatcaaggtactcttttcacttcaagcaaccaagcttctcatacactagatggctc
accattatttaatggaaataaaagtcaagacatcaaaaactttaatactgactttattaattcgaaaaaatttggcttattccaaaccctcctcaagaaatattaaataaacttctaagcaaaatcaattatgtaca
gaaaacttctttcttacatagcaagccaaaagatctatcctctcctacttggcaacaaaacagagccccccttggaggcgaagcctctaaagctaagtttacaaaacaagctatagtgataaaacaaaac
taaaaaaacaagcagagcttactgttaaaaaaacttctttttaaattctatggcctgctacgcttacaccaaaggtgtatgggggttatgctacgcatgatacttttagctgcagaccacaaccagtttagtatta
aagaaaaaaatacaacaagtctttttaaaggaaagtaaattaattatttttaactaaactttaagtaaatttagtcgtttttttaccttatattaaaaattgtattaatttaatcaaaaacttaatgcagaatatagaaaaa
aaaattatggagttaaaaatcaatttaaaatatttaaaaaaaacattaacagattatctttttattaaataataaaataaatagtgaattagaacttattaaaacaaaatatttatctgaacaacaaattattcgtgtctt
aagaagaaaattaaaacgtcatgaagctgaaaaaaaattgttacaatttttacctaaattacgttatttacctacacctaaaagtcaaatatctgtcgctattcaattattaatgaaacgatttgttgatcctaaaat
gaaatatccgttggatctaatttatgaccaagcttcgcttagaaaaaaatcaaaaaaagtagctgcagctagaaaaaaaaaatggcaacgattagaaacttatttttggtggtattgcaaacatgactaaatta
aagaaaacacaaattgctaaaaatgttgctataatcataggtcaacaagaagaaataaatgctattcgtgaatgtaaaaaacttggtattaaaatgtttcatattgttgatactaattgtaatcctggttttagccg
atcattttattccagcaaatgatgattctagaaattcaataaaatatttattatcaaaaatggttacacgtattagattagctcaaaaaattcgtttacgattaaaaaaaaaa

>Chloromonas_rosae__-_rps3
atgggacaaaaagtccatccaataggatttcgcgttggtattacgaaaagacatcaatcacagtggtttgcaaggtttacaaaatatgcttatgctcaaagtttattagaagatcgtatgctacgaaatacgtt
aatcagtgtagctaatgaagctttaaattctaatttaataaagaaacgtgactctgcatttcaacaaaacgtaataacacctaaagtaacacatataaaaattgaacgaggattaattccgtatgaaattggaat
tcaaattcatgctcaagctccgctaaaaaaactgaaatcatcattaaataatttaaaaattaacagagaattattatgtaaattacaaaaaactcgtcggtacttaatgagtttaggaaagcagcaaagttcaaa
actaaaaagtgaagtgcctttaaatcaagataatctaattaatagctctccgaaagaaggagctgttaataaagatgaaacggtgcttgttggaaaaaaaaataatttttaaaaaagttaatcccctaaaaaaa
aaacttttttgttttagttctgctaacactacaggtaagcaaaacttaacaggcaacactccttttttttctccattaccaacgggaaaggagcgtggcaaagcaggtacacagggaatcattttacagaagg
gtaagctcccaaccaaacaaagtttagaagttaaaggtaagtttagctttaaccctaaatttaaaaaacgtagacgtcctaatttttttaagtctagtctacaaaaaaaaaaaaaatgataataaaacgtttacaaa
aacgtcaaaaaatacgtcaacgttatagaagattaatgttaaaaggcttatttgttaaaaaaaaaggaaaacgaattagttgtgttgcttctttatttttttagtaaaaagcaaataaaaaaagctagtgctgttaa
aaaaaaaaaaagtaataaaggggctaaaaaaatgaacattaaatcagtaaaatctcctttaacaaaaaatactagttttttaaaaaagggttctttgccagtgaaacagaatacaatagattagttgcttcgt
catttgaagggagcggaattgagcagcagaaggacgggctcaaacaatcaaaaactgctaatttttttagctgcacctaaaaaattaaataaacaacaaactagcctgggaacaagtggaacgacgtgg
ccttcattattaacaaattctaaatctaataaaaatataaaagaaaaacttctacttttagcaaaattacgcagtagacagagctctcctcaaacaacaacatcaagtgctagcaaagcttggttgttagcttcg
cctaaaaagttaggatattccgcattaggaagcttaaataagataaatgctagttctttaacaaaattaaacagttctccctacgcaccaagttggaatgttggttatagaatcaaaaaaaaatttgtctcattat
ttttaaataaaattaacaaacgctttttaaaacaactaaaagagcaaatgactcaatggaataagtctcttaaacaacatcgacaagaccaaatcaaacttttatggcgcacttcggtttgctcctcttggttaca
atcaaaaatggagtttaaaaagattaaatttttttgaaaaaaaaaccgttatttaagctttaaaattaattaaaattatcgagaaaaaatctttaataaagttagaaacgttaagaaaagatttttatagcttttggca
caatttctaaaagtgaagctttgagttattatcaaataataaatttttttaaaaaatttaaaaattgcttctttaaaattaaaaaaagaccaaaaacgtctgttagtaaaaacttttaacaaaagattattctaaaatatc
agtagtaacatcatcaaatcaagaaatacgtagtccaggtgaaaataaagatgtagatacgttaaaacgtggcaataagcaaagtataatgaaaaccctacctcgcttaggttgggtaccaaaccaagca
agggcaatgaatccagaaataagagtttctggaaaatctaacttttgcaacttttttgacatcttctgttgatgtagataaatctcaaggcaagccatttttagaagctattaatgaatgccgtaaaattaatttt
attgaatatttaaaagaaatggttaaaaaacatagaacaaataatttgttttattatttgtcaaccatttctgatgctcgtaaaaatttaaaaaaaaattaaacaatttactaaagttcattcaaaattcttatttgggttt
gaatttaaccaaccaacagccatttttagagacttctactctctcaaaaaatactgatagcaaagatctattagttttttaataaaattaaaactgtaattgatttagcttcaaaaaaatcggattttgaaaaaggatt
accagaaattttcttagaacaactggaaaagcaacgcaatatgtataaacaactagctaagttacgcgctaaaatttcaattaaattttattcagtaaaatcagaaaacgtaaaaataaaagcggcaataata
tctgattctgttattgatgctttagaaaaacggaaagctttcagaaaagttataaaagatgctaaagaaaatttaatgcgaaattctaaagtaaaaggtgttaaaatacaagtatctggacgtttaaatggagct
gaaattgctcgaactgaatgggtaagaagtggtcgagtgccattacaaacttaagagcaaatttagattattcatataaaacagctaaaacaatttatgggatcattggtgttaaagtatggatatttaaaggt
tatactaagctagtttaa

>Chloromonas_rosae__-_rps4
atgtcacgttatcttggtccacgattaagaattattcgaagaattggcaaattaagaggattcacacgaaaaaaacctttcgcagagtttttaaaggtcgcggtgctttacgtggtaaagttattccaccaggt
caacatggaatggtaaaattatttaaaacacgaccttatgactcttctgagtctgattatctaattcgtttaaaagtaaaacagagattacgttttaattatggtttaactgaacgtcaattagtaaacgtagttaga
aaagcaaaaaaaaataaagaatccactggtggtattttattacaattattagaaatgcgtttagataatatagttttcgtttaaatatggccccaacaattgttgcagctcgacaactaatttgccatggtcatat
aaaagtaaacaacaagaaagttaatatatggcagctacttgtgtaaacaaaaagatgttatttcggtttcaatgaaagaaaaatcattaaaacttattacaaacaatttacaaaattactatcaacgcgtacgctt
ctataaagaacgtttgaagaaaacttgtcttttattcttctgaaatcgcaaatagttaccaatatggcagctgccattagtcttattcaaagcggaaattgtaaaataaacaatctagctacacggaaacctaat
tatatttgtggcgaacgagatgtcattacaatcgttactaaagctggaattcgtcaaattaaactaaac

>Chloromonas_rosae__-_rps7
atgatttgcgtatttggtaataaattagatcagttaaaaaaaatgcctcgtcgcccctataaaaaaaaaacgttctcttttaccagatccaatctataatagtatttctgttcatatgttagttaaccgtgttttaaaaa
atggaaaaaaaatctattgcttatcggattgtctataatgctttaaaagaagtaggggacataactaaaaaaaatccagttgaaatttttgaaaaagctttagataatgtcacaccacgcgttgaagttaaacctc
gtcgacgggctggaacggttcaacttgtaccacgcgttcttcgtactggtgacaaagctcgcgcatcagcccttaagatggatttttagaagcttgtcaaaaacgatcaggtcaatcaatgattgcaaagtta
aaaaatgaaattgttgaagcatataaaaaaaaccgggtttgctgttcgaaaaaaagatgagcttcataaaattgctatgaataatgcaatgtatgctcgaaaaccgcaaactgtaattaatgctataaatcaaa
gtgcgactaat

>Chloromonas_rosae__-_rps8
atgattaatgatactattagtgatatgttaactcgtattcgaaatgcaagtttagcaaaaaaatcaacagtttgtattccatatacgcacttaaatcagcaaattgctcaaatttagaaaaagaaggatatatatt
aactgttcaaatgtcattagattcaaaacatttaattgtccgtcttaagtatagatcaaaaaaaatttataacggaaaaactaaagagtcatgtttaacaaatttaagacgaataagtcgtccttctcttcgaattta
tactaattctaaagaaattccaagagttttaggaggaacaggaattataattctttcaacaccaagtggacttttaactgatcgtgaagctcgttctcgtggtattgggggtgaaatattatgctctatatggtaa

>Chloromonas_rosae__-_rps9
atggaaattttagcaagagcggttggacgtcgaaaagaagctgtagctcaaattcagcttgttcgtggaaatgggaaatttataattaatgataaacctgcccaagtttatttgcaaaataattcatgttctctt
atttgtattaaatcaccattagaagcagcattaagtctttattctaataatttttcaaatatccaaagtacttctcaagactctgccttaataaaagatcaagctatagttggcccttcggagcatagcatactaga
agccaatcaagcggttaatgaaaaattagaagaacaaagtacaacaagtgtctcaactgaaaaaagtgaaatttttaaataacaaagagagtggagtagaagccaacaaaggggggaagctttgagttga
atctttctggtgacgaaactcaaaatttaaatgattctacacataaactcagcgtgcttaagttaaataatttagtttcgttagaggaaattgatgtacttatataaagtaaaaggtggaggacttattggtcaagc

agaggctatcaaattgggaattgcccgagctttttgtctaatgcaagttgcgcttaccgaagggatttttaaaaaaatttttaaaaactaaaggttatttaactcaagattctcgtgttaaagaacgtagaaaatat
ggtttaaaaaaagcacgtaaagcttcgcaatatcataaacgttaa

>Chloromonas_rosae__-_rps11
atggcaagacaaacacgaaaagttgcaccaaataaagcaaaaaaaaaaatttatcggggagttgtacatatccaagcaggttaccataatactatcattacaataactaatgtaagaggagacgttctttgt
tggagttccgcaggtgcttgtggatttaaaggcaaacgaaaatctacaagttttgcagcaaaaaaagcagcggaaacagctgcgcgcaagtcaaaagatgctgctatgagagaagctaaagttttagta
acaggtcctggtcaaggtcgggaaagtgctattagagaaattttcaaagcaggtataaaagttaatgttattcgtgaaaaaactggtatccctcataatggatgtcgtccgccaaaaaaacgaagtgtttaa

>Chloromonas_rosae__-_rps12
atgccaactattcaacaattaattcgttcagcacgaaaaaaactaacaaataaacttaaagctccggctctaaaatcttgtcctcaacgaagaggtatttgtcttagagtttatacaattaccccaaaaaagcc
caactcagctcttcgaaaagttgctcgagttcgtttaaccacaggttacgaagtaacagcatatattcctggaattggtcacaatttacaagaacacgctgttgttttagttcgaggtggccgagtaaaagat
ctacctggagttcgttatcatattgttcgtggcaccttagataccgctggagtaaaaaaccgtgtacaaagtcgctcaaaatatggagtaaaaatagcttcgaaaacagcagcaaaaacatcctctaaaaa
atag

>Chloromonas_rosae__-_rps14
atggcaaaaaaaagtatgattcaacgtgagttaaaacgacaaaatttagtaatgaaatatgctgaaaaacgagcttctttaaaagaacagattaaacaaacatctttttttaaaagaaaaattagcgttacatc
gtaagttacaacaacttccacgtaatagtgctgctgttagattacataaccgttgtatgattactggtcgtcctaaaggatattatagagattttggattatcgcgacatgttttacgggaaatggctcatgaag
gtcttttaccaggagtacaaaaatcaagttggtaa

>Chloromonas_rosae-rps18
atgaatcaacccattcctcgttcacagaataatattaatggaacttttattagcccgagtactcctaaaaaaaattttttttcaaattctccaaataagagtttagtaaaagaaaatacttttaaaggtacaaaatcc
aattataatagtttcacaactaaaacaagcaatattaataaatcaaataaattaaacaaagttaatggaattaaaactcagtcaaaaggtaaattaaaaaaaaatctaaaaattaaacgtgttttatcattatctca
aattcttggtcggattagaatactacgtcaaaaaaaagtagagcaacaaaaacgccaaaaaccaataaaaccaataattccacccaaatcattaattattatcttaaaagataaaccagaaaaagcagttta
taatcgtagaataattgattataagcattgtggtttattacaaagatatatcggtttaggtggtaaaattttgccaagaagacaaacgcgattaactgcaaaacaacaacgatatgtagcaaaaacaatttaaaa
gtgctagaataatgggattattacctttgtaagtaaagaacgaggatttttagataa

>Chloromonas_rosae-rps19
atgccacgttcgattaaaaaaggtccgtttgttgctgatcacttattaaaaaaaattgaaaaattaaatgctcaaggtcaaaaaaaagttttaacaacatggtcgcgttcttcaatgattttacctccaatgatcg
gccatacgattggagtttataatggtcgtgaacatattcctgtttttattacagatcaaatggttggccataaattaggggaattttctcctactcgaacttatcgaggccatggtaaaacagataaaaaatctaa
acgttaa

>Chloromonas_rosae__-_ycf3
atgccaagaactcaaaaaaatgataatttattgacaaaacttttacggtaatcgcagacattttactcaaagtttaccaacctcacaacgagaaaaacaagctttttcatattatcgaaatggtatgtctgccc
aagcagaaggtgaatatgctgaagctcttcaaaattattatgaagcaatgcgtttagaaattgatgcctatgatcgcagttatattctatataatataggtttaattcatacaagtaatggcgaacatggtagag
cattagaatactattatcaagctttagaaagaaatccttctttacctagtgcattaaataatattgctgtgatttatcattatcgaggtgaacaagcgattcaagataatcagccagaaatttgtcaacttttatttga
aaaagcagctgattattggaaagaagctattcgtttagctcctacaaactatattgaagctcaaaattggttaaaaatgactggtcgagaataa

>Chloromonas_rosae__-_ycf4
atgaacaattctcttttatcgcaagagagctcttcaaaggtagccccttttagagacaaactcaaagcaaacagaactaaatcggcgttattttattgtcggagaacgtagactgagcaattattggtgggctt
ctgtaatttttttaggtggctttggttttttattaacaggaatttcgtcgtatctgaattataatattttagcaaatgcatttaacatatttaatgttactacgttaatgcaaactttacctaatggggaacctctaaatcc
gaatttgactagcaatttaacttctagctttagtaacgctacgttaaacgaaaattcagttattgctttttttccccaaggattactaatgtgttttacggcagcttaggtgttctattaagtgtatattggtggcttct
aatttattgggatgttggtggtggttttaatgaatttaataaaaaagaaggctttatgagaatttttcgctggggatacccaggaaaaaatcgtcgaattgatctaaaagtaagtttagctgatcttgaagctatt
cgagttgaaacacaaggtttatctgaacaaactatttatgttcgttggcgttccatcactcctaccgaagaccctaaagggttgccttctttttgccccaatctaaaactcaacaggggccagcgcgctgc
ttggtttggggaccaaccaaccaaaggtaaggggagcttcactcctacaaaaacaagaacaagaaaagcgtgaaattcctttaggtgggattggccaaccattaactttaaaagaaattgaaaaacaag
cagctgacttagctaattttctacaaattgaactaataggcttatag

>Chloromonas_clathrata_-_atpA
atggcaatgcgtactccagaagaattaagtaatttaattaaagacttaattgagcaatatactccggaagttaaaatggttgattttggtattgtgtttcaagtaggggatggtattgctcgtgtttatggattaga
aaaagcaatgtctggagaacttttagaatttgaagatgggactcttggtattgctttaaacttagaaactaataacgttggagctgttttacttggtgatggcgttaaaattacagaaggcagtcgagttcgttg
tacaggaaaaattgctgaaatacctgttggtgacggttatttaggtcgtgttgtggatgctttagctcgcccagttgatggaaaaggcgcaatttctacaaaagacacaagagcaattgaatctgttgcacca
ggtattatcacgcgtcgttctgtatatgaacctttagcaactggattagtgtctattgatgcaatgattccagtaggtcgtggtcaacgtgagttaattattggtgaccgccaaacagggaaaacgtcaattgc
aattgatacaatttaaaccaaaaaggaaaaggtgttgtttgtgtttatgttgctattggccaaaaagcatcatctgttgcgcaagttttaaatactttaaaagaacgtggtgcattagattatactattattgttatg
gcaaacgcaaatgaaccatcaactttacaatatttagctccttatacaggtgcaacattagctgaatactttatgtatacaggacgtgctacattagttatttatgatgatttatcaaaacaagctcaagcttacc
gtgaaatgtcattattacttcgccgtccaccaggacgtgaagcatacccaggtgacgttttttatcttcactcacgtcttttagaaagagctgcaaaattaagtacagctttaggagaaggaagtatgactgc
acttccaatcattgaaacacaagaaggtgatgtatcagcttatattccaacaaatgttatttcgattaccgatggtcaaatcttttggccggggacttgtttaacgccggtattcgtcctgctattaacgtaggg
atttctgtatctcgtgtaggatcagctgcgcaacctaaagcaatgaagcaagttgcaggatctcttaaattggccttagcacaatttgccgaattagaagctttttctcaattttcttcagatttagaccaagcaa
ctcaaaaccaattagcgcgcggtgtacgcttacgtgaaattcttaaacaagctcaatcagcgccttatctttagaagatcaagttttaacgatttatgcaggtactcaaggctttcttgataaacttgaagtaa
gccaagttcgagcttttgtagctggcttacgtagttttgttgcttcaaaatacccttaaatacagtgaaattatataaaagcgacttttagcttttaatgcagaagctgaaggaatagtaaaacaagcaattaaagaat
atttaaatgaatttgttgcaagcagcgcaacatctgctaaaaaataa

>Chloromonas_clathrata_-_atpB
atgagcgattctgtagaaacaaaaaatattggacgtgttgtacaaattattggtcccgttttagacgttgttttttcgaaaggtcaagtacccaatatttacaatgctttagttattcgttctaaaaacgctgctgg
attagaagtaagcgttactgttgaagttcaacaacttcttggtgataattgtgtacgtgcagtttcaatgaatccaacagatggtttaacgcgcggtgtggaagttattgatacagggaaaccttttaactgttcc
tgttggaaaagcaacttttaggtcgtatttttaacgttcttggtgaaccagtagataatttaggtcccgtaaaagctgaaacagcattaccaattcaccgtacagcaccagcttttgttgatttagatacacgtctt
tcaatttttgaaacaggaattaaagtagttgatctttagcaccatatcgtcgtggtgggaaaattggcttatttggtggtgctggtgtaggaaaaactgtattaattatggagttaattaataatattgcaaaagc
tcatggaggtgtttctgtttttgctggagtaggggaaagaacacgtgaagggaatgaccttatacagaaatgaaagaatctggtgttattgtagaaaaaagtctttctgattcaaaagtagcacttgtttatg
gtcaaatgaatgaaccaccaggagctcgtatgcgtgttgctttaacagctttaacaatggctgaatattttagagatttcaataaacaagatgttctttttctttattgataatattttccgatttgttcaagctggag
ctgaagtttctgccctattaggtcgtatgccttctgctgtagggtatcaaccaacgttagcaacagaaatgggtaatttacaagaacgtattacatctactaaagaaggttctattacatcaattcaagcagttt
atgtacctgctgatgatcttactgatccagcgccagctgtaacttttactcatttagatgctacaactgtattatcaagaggtttagctagtaaaggtatttatcctgcagttgacccttagattcaacatcaaca
atgttacaaccttggattgttggtgaaaaacattatggcgtagcacaagatgttaaaaaaacgcttcaaagatacaaagagcttactgatattattgcaattttaggtttagacgaattatctgaagaagatag
attagtagttgctcgtgcacgcaaaattgaaagattccttagtcaaccgtttttcgttgctgaagttttcacaggatcgcctggaaaatatgttagtttaacagagtcaatggaaggttttggtaaaatttttactg
gggaattagacagtttaccagaacaagcgttttatttagttggaaatattaatgaagtaattgccaaagcagctacattaaaa

>Chloromonas_clathrata_-_atpE
atgagtttacaaatttcaatttttaacaccagaacgtccttttttggaatggtcaagcagaagaaataattcttcctactgaaacaggagaaatgggagttttaaaaaaccacgcaccacttattacaggtttagat
gttggagcaatgttgattcgttctaaagatgaatggaattcatatgctattatgggaggatttgctttagttaaacaaaatcaagtaacaattttagcaaacgaagctgaatcagctgaaaatattgatccagaa
gaagcaaaaactagttttgaaactgctaaagctaatttagaaaaagctgaaggtgtaaaagaaaaagtagaggcaaattttgcctataaacgtgcaaaagctcgctttatgactgtaagaaaaaaa

>Chloromonas_clathrata_-_atpF
atggaatgcgtaacatttattaacggagttactataggacatggtggttttggatttaatggcaatatttttgaaacaaatattattaacttagctgctgtgattggcgttgttgtaacatttgttggaggaaatctta
ctgcattattagaagaccgtaaaaaaactattttaaataatttacaagaagcaaatcaaagagcgattgaagctcaagaaaaattaagtcaagcacgtgcacaattagaatcagcaaaaaaaaaagctcaa
gaaattcgtgaagaaggcattttaagagcgactcaagaaattaataattgtgtaacacaacatgatattagattatcaagattacaagagtttaaacaagaaactcttcaacaagctgaacaaaaagctttta
aacaagcttacatgtatttaatagcaaaaatttcaaaaagagttcgtgaaagattaaataatggattagattctacttatcatgttgtagttaataacttttatgtatctcgttttactgattttaaa

>Chloromonas_clathrata_-_atpH
atggcaattgatagtttaattggtgcagctagtgttttagctgctggtatcgctgtaggtgttggtagtattggccctggaactggacaaggaactgctgcaggatatgcagttgaaggtattgctcgtcaac
cagaagctgaaggtaaaatccgtggtgctcttttacttttgcttttatggaatcactaacaatttatggtttagttgttgctttagctttattatttgctaaccctttttttaggttaa

>Chloromonas_clathrata_-_atpI
atgattaatcctttattagaaattgctgaagtatctgttggacaacatttttattggcaaataggagactacagtgttcacggacaagttcttttaacgtcatggttcgtattagcactaattggaattttaagtctttt
aggaaatagtaatttaaaacaaacaccagacggtatacaaaacttcacagaattgattactgagtttattcgtgatttagcaaaaactcaagttggagctcatgactacttatcatgggtacctttttttaggaac
tattttcttatttatctttgtctcaaattggtctggagctttgattccttggaaacttattgaattaccaaatggcgaattagctgcaccaacaaatgatattaatacaacagttgcattagcactattaacttcaatct
cgtatttttacgcgggtatcaaaaaaaaaggtttaggctactttaatagatatgttcaaccagctgctttcttactacctattaacgtgctggaagactttactaaaccccttcacttttcttttcgattatttggaaat
attttagcagacgaattagttgttggagttcttgttgcccttgttccgcttattgttcccattcccctttatgcttcttggtctatttaccagcggaattcaagcgttagtttttgcaacacttgcaggtgcttatattggt
gaagctcttgaagatcatcattaa

>Chloromonas_clathrata_-_ccsA
atggcttcgccattagttatggtttcgccattagttatggcttcgcctttagttatggcatcgccattagttatggcttcgccattagttattgaaatttttttaagaaattgttcttttttttgtctctttctttcaatgatttt
ttattggattcaaacagcttttagttggggtaagggcccctatccaacttttggaagtgctagggcctcctacctctttaaaatctgttgaggttgtaaagcaaggggcccctacctctaaaaaaaagaagtta
ggcatgcatagcatacctaacccttattaattaactacactaaaagcactttaacgcttttcctactgccaatccctctgctagctctgtttcacccaattcttcgtaccccctcttttcctgggataggtaacgg
gatagataaaggggcttcttcccttgaatagcaaaaccccttattgtccctccacccgttcttttatcgtacccctagagggtagaagccctttctttttaaaagaggaaaagcaaagcattttaatttatttttttga
aatactaaaacaagttcccaattttgcaggcacagccttaatgattatttcaaatttttattaatgttttattaatttttcgatggaaagaatctggacattttccattaagtaacttatatgaatctttaatgttttatc
ttggagttgtacgtttattcatcttgttttagagttaaaaacagaactaattagatctgatgcaatactttaaagcgatgcttaggttctataacttctccaattgcttttttacaagtgctttgcaacttttaatttgc
caacggaaatgcaaaaagcatctccttttagttccggccctacaatcaaattggttaatgatgcatgttactgttatgattatcagttattctgccttaattcttggctcattattatcaattgcttttctaatttgataa
gttctttttcaatttttgattttagttttttaaaaaaattaagcgcttctagtgaaactaaagttagcttaaaaggcgaagcgcttagtttcaacttaccttcaaaagcaaaaatgactttagcttggaattttgataatt
taagttatcgtgttttaggaataggatttccttttttaactatagggaattttatcgggtgctgtatgggctaatgaagcatgggggtcatattggagttgggatccaaaagaaacatgggcattattaacttggtt
gatttttgctatctatttacatgctagaataacaaaaggttggcaaggcaaaaaaccagctattatagcttctttttggttttttagctatttggatttgttttttaggagtaaatttaattggtgaaggtttacatagtta
tggttggttttcaaatttg

>Chloromonas_clathrata_-_cemA
gtgtcatttacttataatttaatgaatcaaattatgggcaaagcccattaaataaagtcaggctaaaaaaaaataattgtttaactttgtcacctgttaaggacttttggtctaaaagtgcaaagcacttttaaaa
aatgcttcgcattttggttgttaggggaatctctaacacttatcctcatacatggcaaagccataatagatcgttattaaaagttgatgccaacaaagaaaaagtatctttaaaagatаaacaagtagtttcta
ttacatatgaagaaatаggtttatttccaagatcatttagtcgtgtatttgatagatttattaaacaattgttttctgatgtagaaaatttagttattcaagaatаccgtttttatcggtatttatttttaacaacagttaaa
tgttttttattcttctttttgttccctttttagtcaatgtagctagtaaaaattatttaatcagacctctaacagaatattgctggaatcaaaaacaaggtgaaatttttttaaatgcatatcaacaaaaacatgctttg
ctgaattacaagatttcgaagaaaaagtttattttgaatccttagttatccctaaaactacaaatgggaaaattgataatggaaaagcgggggagctttacaagaaaaaacgattcaattggctattgattata
ataatgctagcattgaagctattagttgtatgtttgctgacttaattagtctcagcgtttttggatggttacttattcttatggaggtacaaataagcgtaagccgtttatttatacttgaagtcttctttgggcttgatg
acactaaaaaatctctgattatttttattagtaaccgatctcttagtaggatatcattcagttggcccatggtcgattttgtttgaatctttatttaatcgctatggcttgccaaatagccaagctgcaatttatttactc
acaggaagcttaccagtaattatggatgttttttttaaatatttgatctttagacatttaaatcgagcttcacctgcttcagttgctacttatcatgcaatgattgaataa

>Chloromonas_clathrata_-_ChlB
atgaaattagcttattggatgtatgctgggccagctcatattggaactttacgtgtggcaagttcatttaaaaacgtacatgcgattatgcatgctccattaggagatgattattttaacgttatgcgttcaatgtt
agaaagagaaagagattttaccccagttactgcaagtattgttgatcgtcacgtattagctcgtggatctcaagaaaaagtcgttgagaatattactcgtaaagacaaagaagaacaacccgatttaattgtt
cttacccccacatgcacttcttcaatttttacaagaagatttacagaattttgtagatcgagctgcaatggaatcaaaatgtgatgttttattagcagatgttaatcactaccgagtaaatgaactgcaaagtgctg
atagaacattagaacaaattgttcgtttttatatcgaaaaagcaagaaaacaaaataatttacagactaataaaacagaaaaacctctgcgaatattataggtatttttacttaggttttcacaatcaacatgat
tgtcgcgaattaagacgtttattaaatgatttaggcattgaagttaatgaggttttaccagagggcggctcagttaataatttaaaaaatctacccaaagcatggtttaattttattccttatcgtgaagtaggctt
aatgtctgctatttatctagaaaaagaatttaatatgccttatgttgcaattactccaatgggtgttgttgatactgcagcttgtattcgagaaatcggatctattttaacaaaaatagatccatctttaaaaatcaac
ggggtaggtattaaaaaagaatatgacaattatattgataaacaaacacgttttgtatcgcaagcagcgtggttctcacgttctattgattgtcaaaatttaactggcaaaaaagctgtagttttggagatgca
actcacgctgcttctatgactaaaattttagcacgtgaaatgggaattcgtgttgtttgtgctggaacttattgtaaacatgatgcggattggtttagagagcaagtggttggtttttgtgaccaagttttaattact
gatgaccacactctagtaggagacattattgctaaaatggaaccagcagccatttttgggacacagatggagagacatgttggaaaacgattggatataccttgtggtgttatatcagcacctatacacatt
caaaattttccacttggttatcgcccttttttaggttatgaaggaactaatcaaatcgctgatttagtttacaattcatttaccttaggaatggaagaccatttgttagaaatttttggtggacatgacaataaagaa
gtaattacaaaatcattatcaactgattcagatttaagctggtcatctgatggtttagccgaattaaataaaatacctggttttgttcgtggtaaagttaaacgtaatacagaaaaatttgcacgtcaaaaaaatct
tggaattataactgtggaagttatgtttgctgctaaagaagcagctggtgcataa

>Chloromonas_clathrata_-_ChlL
atgaaattagcagtttatgggaaaggtggtattggtaaatcaacaacaagttgtaatatttcaattgctttagcgagacgtggaaaaaaagttttacaaattggttgtgatccaaaacatgatagcacttttacc
cttacaggttttttaattcccacaattatagatactcttcaacaaaaggattaccattatgaagatgtttggccagaagatgtaatttaccaaggttatggaggtgttgatagtgttgaagctggaggtccccca
gctggtgccggctgtggtggatatgttgtaggtgaaacagtaaagttattaaaagaattaaatgcttttttatgaatatgatattatttttgtttgatgttttaggggatgttgtatgtggtggatttgcagcaccttaa
actatgcagattattgtattattgtaacagacaatggttttgatgcattatttgcagcaaatcgtattgcagcttcagttcgcgaaaaagctcgtacacacccattaagattagcaggcttaatcggaaatcgaa
cagctaaaagagatttaattgataaatatgttgaagcttgtccaatgcctgttcttgaagttttacctctgattgaagaaattagaatttctcgtgttaaaggtaaaacattatttgaaatggttgaatctgaaccaa
ctcttcaatatatttgtgatttttatttaaatattgcagatcaacttttaacagaaccagaaggggttgttccaagagaattgtctgaccgagaattatttagtttactatcagatttctatttaaatccagttgatcca
acaaaaaaaacagaatctgttgaaacattagacttttattagtt

>Chloromonas_clathrata_-_ChlN
Atgaggcttaatttaaatattagtaaagttttatatcaaagaaaaatgtcaagcaatattcttatggcgaataaatcagttactttaaacaattcaagctctgttagtttgattgaagaagtttatgctccaaacac
agatactattgcagtttctacacaaacaaatgatgactcgttaacatttgaatgtgaaacaggcaattatcatacttttgccctattagttgtgttgcatggctttatcaaaaaattgaagatagctttttcttggta
attggaactaaaacttgtggttacttcttgcaaaatgctttaggagtaatgattttttgctgaaccgcgctatgctatggcggagttagaagaaagcgatatttcagcgcaattaaatgattataaagaattaaaa

agactgtgtttacaaataaaacaagatcgaaatccaagtgtggtagtttggatagggacttgtactacagaaattatcaaaatggatttagagggcatggctcctcgtttagaaactgaaattggaataccaa
ttgtcgttgctagagctaatgggctagattatgcgtttacacaaggcgaagacactgttttagctgcaatggctcaaagatgtcctgaaagaagcgctttaagttctgaaaataaaaaaaaataatacagtag
gaaccagttcacagtctattgttaattttgctacttcaaaagaatcgaaaaaaccgaagggtaaaatagagcaactagggggttgcgaagcaaggggtagggttggggcccaaccctactacctctggca
ctttaaaaaatactaatactaaacaattagttttattcggttctttgcctagcacagttacatctcaattatcaatggaattaaaaaatcaaggcattatagtttcaggttggttaccttcacaacgttataatgattta
cccgttttaggtgaagatgtttatgtttgtggtgttaatccatttttaagtcgcactgctacgactctaatgcgtcgtcgaaaatgtaaactgataggagctccatttccaataggtccggatggaactcgtgctt
gggtagaaaaaatttgtagtatttttgatattgttccacaaggtttagaagaacgtgaagcaaaaatatggcaaagtttagaagattatttacaattagttcgtggcaaatctgtattctttatgggagattcgctt
ttagaagttccttggccagatttttaacacgttgtggaatgactgtatttgagataggaattccctatatggataaaagattccaagccgctgaacttgctttattagaaaaaacttgcagagatatgaaagttc
ccatgccgcgaattgttgaaaaaccagacaattattatcaacttcaacgaattcgtgaacttttaccagatctcgtaattactggaatggcccatgccaatcctttagaagctcgcggcattagtactaagtgg
tcagtagaatttacatttgctcaaattcatggatttgctaatgctcgtgatatttagaacttgttactcgaccattaagaagaaataaaagtttagaagctttaggttggatgaaattggtaaaataa

>Chloromonas_clathrata_-_ClpP
atgccaattggagtaccaagaattatttattgttggggtgaagaacttcctcctcaatggactgatatttataatttttattttcgtcgacgaatggttttttaatgcaatatttagatgatgagctttgtaatcaaatt
tgtggattattaattaacattcatatggaagatcgctcaaaagaacttgaaaaaaaagaaatagaaaaaagtggcatttttaaaagtactaataaaagtggtaaagattctagtccaaattcctctgtaggagg
gacgtcggaaagaagcactttatttgatacaaaaaataacttaaatgatggtaaaaaattaaaaaaagataaatcaatggaagatcttttaagttcttatggcaattctctttatgaagaagatttagctattgat
gaaaactatactttagaacaatatactttacaaaaaataacattagaatggttaaattggaatgctcaatttttgactattcggatgaaccatatttatttttatttagctgaaattttatctaaagattttacaaaagat
gatacaggtcaattatttataatttaattacgctaaaacgaatccaaataatcaatttgggtctacattagaaagtttcgtaagctctaacttgataaaaccaacatttgacactaaaaaaagtagtggcattaa
aaaaaatgcttttcaggaatattggccgtttcgatcaatattaaatttagcaaatttaatttctaaagatcaaaacttaagataaataatttattacctaagtttataggtaacgcctataagctaagtaataacaa
agctataagagggacactagcttcgaaccctactagactctttccatccgaacaccgttgttgcaaaagaaaacgttgtagaggtagccacaacaaatacaaaagaaacacaatcaactgtgcaagg
gatgcaacttttttctaaaaagtcctttacaaccccgttaccaaaaagcaatggacaaaatttattaaatcaatttaatcaattaaaccaaaaacctggtttaactaaattacttggagtgttagcatcaaaagaa
tttgtaaaaaaaaatacaaactctcaaaaacaacaatctgttgcaacaaaagttttaggacagagaaaatttaagaaaagaatatgctttagataataatttttaataattatggttcttttgaaggtcttttatggg
cgaagcccataaaagccctttaatatctgaatatggttcttttagaagtgctgccttaaaacaggacatcgcgaaagaaaggaagaacaaaacgttctagatagttcacaaaaattacaaaaaaataaaa
gagcattaaattatttagaatttgattctgctagtgcaaaacaaatcaataaaaaaggaaataaaaaaagtttcagcgctgctacaagaaaacaaacaaaaagagcacttcaagaagaagaatctaaaaaa
gtttttgttataataaattcttttggtggatcagttggaaatggtattactgttcatgatgcgttacaatttataaaagctggatctttaacttgggtctgggagttgctgctagcgctgcttctttagtattagccgg
aggaacaatttcagagcgttatgtaactgaaggctgtcacgtcatgatacatcagccggaaggtggactcaacggtcaagcatcagatatttggattgatagtcaagaaataatgaaaattcgcttagatgt
agcagaaatttattctttatctgcacatcgccccgccataaaatattacgtgatttagaccgagatttttatttaactgcaactgaaactattcattatggtttagctgatgaaattgcaacaaacgatgttatgc
accaaattattgaaatgacaagcaaagtttgggattatcacgatagcaaacaacaacgtttacttgaaactcgaaacagcgcaaatgctgttgacactcaaacgcaaagttaa

>Chloromonas_clathrata_-_petA
atgtctacaaattcggcacgaaatttaagagttactttattagaaggcaattttgcaaaaacatttttttgttcaatattatttggaatttttctttttataaatggtgcagttactgactacgctagcgcttatccagtc
tttgcacaacaaaactatgaaaatccgcgtgaagctaatggccgtattgtttgtgcaaattgtcacttagctcaaaaaccagctgaaatcgaagttcctcaagcagtattacctgatactgttttgaagcagt
agttcaaattccttatgataaacaattacaacaagtacaagctaatggtaaaaaaggcgatttaaatgtaggtatggtattaattctacctgaaggatttgaacttgcgccaccagatcgtattccagaagaat
taaaaaaaaaagttggaaatatttattatcaatcttatagtcctgaaaaaaaaaatattttagtagcaggtccacttccaggtaaaaaatatagtgacatgataattccaattctttctccagatcctgctaaaaat
tcaaaagtcgcttatttaaaatatcctatttttttaggtgctaaccgtgggagaggtcaagtataccctgacggttctaaatcaaacaacaccgtatataacgcatctgtaagtggtaaaattatttcaattattcc
gggtgagaaaaaaggatctattaacattacaattgaaaaagcaaatggagaaaatgttattgaaaaaattcctgccggacctgaagttttagtacgtgaaggacaattaatccaagctgatcaaccttaac
aaataatccaaatatcggtggttttggacaaaaagatgtagaaatcgtccttcaaaacagctcaagaattcaaggcttattaatttttcttgcatttattcttatagcacaagttcttttagttcttaaaaagaaaca
atttgaaaaagtacaattagctgaaatgaatttt

>Chloromonas_clathrata_-_petB
atgagtaaagtttacgactggtttgaagaacgttttagaaattcaatcaattgctgatgatattagcagtaaatatgttccaccgcatgttaatatttttattgtttaggtggtattacctttacgtgtttttttagttcaa
gttgcaacaggctttgcaatgactttctattatagaccaacagtagcagaagcttttgcatctgttcaatatattatgacagatgttaatttttggttggttaattcgatctattcatcgttggtcagctagcatgatg
gttttaatgatgattttacatgtcttccgtgtttacttaacaggtggttttaaaaaaccacgtgagctaacatgggttagtggagttattatggcagtatgtactgtttcttttggtgttacaggttattcattaccttg
ggaccaaattggatattgggctgttaaaattgttacaggtgttcctgaagctattcctgtaattggtggtcctttagtagaactattaagagggggtgttggcgttggccaaagtacattaacgcgttttatagt
cttcatacttttgtattaccattagcaacggctgcatttatgttaatgcatttcttaatgatcagaaaacaaggtatttcaggaccactataa

>Chloromonas_clathrata_-_petD
atgtcagtaacaaaaaaacctgatctaactgatccagttttaaaagctaaattagcaaaaggtatgggtcacaatagttatggtgaaccagcatggcctaatgatttacttttatatttttccagttgttatttttggt
acgtttgcttgtgttgttggtttagctgttttagatcctgcagctataggtgaaccagcaaatccatttgcaacaccgttagaaattttaccggaatggtatttttatccagtattccaacttttacgaacagttcca
aataaactgttaggtgttttgttaatggctgcagttccttttggactaggattaatacctttattgaaaatatcaataaatttcaaaatccatatcgtagaccaattgcaacaatttttatttctttttagggactgttgct
gctatttggttaggtattggcgcaacatttccaattgatatttcattaacttttggttttatt

>Chloromonas_clathrata_-_petG
atggttgaaccttattatcgggaattgttttaggattagtacctgtaacgatagcaggcttgtttgttactgcttatttacaatatcgtcgtggtgatttagctacttttttaa

>Chloromonas_clathrata_-_petL
atgttaacaattacaagttacgttgtattattagtgggtgcgttaggttttacgttagggatttatcttggtctttttaaaagtggttaaattaatttaa

>Chloromonas_clathrata_-_psaA
atgacaattagttcaccagaacgtgaagcaaaaaaagtaaagattgcagtagatcgcaatcctgttgaaacaagttttgaaagatgggccaaaccaggtcatttttcccgtactcttgctaaaggaccaaa
cacaactacttggatttggaatcttcatgcagatgctcatgacttgatagtcatacaagtgatcttgaagaaatttcacgaaaagtatttagtgcacatttggtcaacttgggattatttttattggttaagtgg
gatgtattttcatggcgcacgattctctaattatgaagcttggttaagtgatccaacacatattaaaccaagtgctcaagtagtttggcctattgttggtcaagaaatttaaatggtgatgttggtggaggtttcc
aaggaattcaaattacttctggtttcttccaacttggcgaggtgctggaattactagtgaattacaattatatagtactgctattggtggattagtgatggctgctgctatgttcttcgcgggatggtttcattttc
acaaggctgcaccgaaactagagtggtttcaaaacgttgaatcaatgttaaaccaccatttaggtggtttacttggattaggaagtttaggttgggctgggcatcaaattcatatttcattgcctattaataaatt
actagacgcaggagtagatccaaaagaaattccattacctcatgatttaatgtttaacagaacttttatggcagatctttatccaagttttggaaaaggattagcaccattctttacttttaaattggagtgaatata
gtgacttttttaactttcaaaggaggattaaatcctgtaactggtggtttatggttaagtgatacagcgcatcaccatgttgcaattgcagttttattcttagttgctggtcatatgtatcgtacaaattggggaattg
gccatagtatgaaagaaatattagaagctcatcgtggccctttttacaggagcaggtcacgttggtttatatgaaatttttaacaacttcatggcatgctcaattagctattaacttagctttatttggttcattgtcta
ttatcgttgcacatcacatgtatgcaatgccaccatacccatatcttgctactgactatgcaacacaacttttctttatttacacaccatatgtggattggtggattctgtattgttggaggtgcagctcacggggc
tatttttatggtacgtgattatgatcctacaaataactataataacttattagaccgtgttattcgtcatcgtgatgctattatttcacatttaaactgggtatcaattttcttagggttccattctttggtttatatatcca
taatgatactatgagcgcattaggtcgtcctcaagatatgttctctgatacagctattcagcttcaaccaatttttgctcaatggattcaaaatacacatttttagctccacaatttacagctccaataatgctttagc
agcaacaagtgtatcttggggtggcgatgttgtagctgttggtggtaaagtagctatgatgcctatttcttaggcacttctgatttcttggttcatcatattcatgcgtttacaattcatgttactgtattaatcctttt
aaaaggtgttttatttgcacgtagttcacgtttaattccagataaagcaaatttaggtttccgtttcccttgtgatggtccaggacgtggaggtacatgtcaagtttctgcttgggaccatgtgttccttggcctat
tctggatgtataattcattatctattgttctttccattttagttggaaaatgcaatctgatgtttggggaactgtaaatgcttcaggtatttctcatattacaggtggtaatttcgcacaaagtgctaatacaattaat
ggttggcttcgtgacttcttatgggcacaatcaagtcaagttattcaatcgtatggttcatcactttctgcgtatggtcttatcttttttaggcgcacacttcgtgtgggcatttagtcttatgttccttttctcaggac
gtggttactggcaagaattaattgagtctattgtttgggctcataataaaatgaaagttgcgccatctattcaaccacgtgctttaagtattactcaaggtagagctgttggagttgcacattatcttttaggagg
tattgcaactacatggtctttctttgaagctcgtattattggtgtagga

>Chloromonas_clathrata_-_psaB
atggcaacaaagctgtttccaaaatttagccaaggtttggctcaagatccaactacacgaagaatttggtttggtcttgctacagcgcacgattttgaaattcatgacggaatgactgaagaatatctttatca
aaagatttttgcgtctcactttggacaacttttcagtaattttccttttggacctcagggaatttattccatgtagcatggcaaggaaatttttgaacaatgggttactgatcctattcatgttcgtccaatcgcacatgc
tatttgggatcctcactttggccaaccagctgtagaagcttttacacgtggtggagcttcagggccagtaaacattgcaacatcaggtgtttatcagtggtggtatactgtaggtttacgcacaaattcagatt
tatatactggatctgtgtttcttgcattagttttctgcaatttttttatttgcaggttggcttcatttacaacccaattttcaaccatcattatcatggtttaaagatgctgaatcacgattaaatcaccatcttgcaggttt
atttggtgtaagttcattagcttggacaggccatttagtgcacgttgcaattccagaatcacgtgggcaacatgtaggttgggataatttttattacagtattaccacatccacttggcttaactccttttttggaca
ggcaattgggctgcttatgcacaaaacccggattctgcagctcatatttttggtacttcagaaggttcaggagatgcaattttaacttttttaggtggttttcatccacaaacgcaatctctttggttaactgatat
ggctcaccaccatttagctattgctgtaattttcattgttgcaggtcatatgtatcgtacaaatttcggtattggtcatcgtatgaaagcaattcttgatgctcatgttgctccttcaaatagactgggtgctgggc
acaaaggattatttgatactgtaaataattcattacactttcaattaggtttagctttagcttctgtaggtacaatcacatcattagtagcacaacatatgtatgctctacctccttatgctttttagctgttgatttca
caacacaagcttctctttatactcatcatcagtacattgcgggctttattatgtgtggtaaacaactgtggagttttttagataacaattgtatttttgttgttaccccttggggggttttaattaatttt

>Chloromonas_clathrata_-_psaC
atggctcatttagttaaaatatacgatacttgtattggttgtacacaatgtgttcgtgcatgtccgttagatgtattagaaatggtaccatgggatggttgtaaagcaaatcaaatggcatcagcgcctcgcact
gaggattgtgtaggttgtaaacgttgtgaaactgcttgtcctactgactttctaagtattagagtttatttaggttcagaaagtacacgtagtatgggattagcttattaa

>Chloromonas_clathrata_-_psaJ
atgaaaaattttacaatttatttatcaactgctcctgtggtaagcttagcatggttagtttttaactgcagttttattgattggtttttaataaagtattccctgatcctcttgtttttacttttttaa

>Chloromonas_clathrata_-_psbA
atgactgctattatcaacaaacaagatgtttcaactagcttatgggctcgtttctgcgaatgggtgacttcaactgaaaaccgtatctatgtaggatggtttggaacgattatgttcccaactcttttaactgcaa
cttcagtttatattattgctttcgttgctgctcctccagtagatatcgatggtatccgtgaaccagtttctggttctttattatacggtaacaacatcatttcaggtgctgtaatccctacaagtaacgcaattggtct
gcatttttatcctatctgggaagcagcttcgcttgacgaatggttatataacggcggtccttaccaaatgatcgtttgccacttcttcatcggtatctgtgcttatatgggtagagaatgggaattatcataccgt
ttaggtatgcgtccatggatctgtgtagcttactcagcaccagttgctgcagctacagctgtattcatcatctatccaatcggtcaaggttcgttctcagatggtatgcctcttggtatctctggaacgttcaact
ttatgatcgtgttccaagcagaacataacattcttatgcacccattccacatgttaggtgttgctggtgtattcggtggttcattattctcagctatgcatggttcattagttacttcatctttaatccgcgaaacaa
ctgaaaacgaatcagctaacgctggttacaaattcggtcaagaagaagaaacttataacatcgttgctgctcatatttggttcactgctttaggtatttcaactatggctttcaacttaaacggatttaacttcaa
ccaatcagttgttgactctcaaggtcgtgtattaaacacttgggctgacatcatcaaccgtgctaacttaggtatggaagttatgcatgaaagaaatgctcataatttcccgttagacttagcttctgttgaagc
tcctagcgtaaacgct

>Chloromonas_clathrata_-_psbB
atgggattaccttggtatcgtgtacatactgtagttattaatgatccgggccgcttaatttcagtgcatttaatgcatacagctcttgtagctggttgggcgggttcgatgacacttttgaaattgctgtttttgat
ccatcagatccagttttaaatcctatgtggcgtcaaggaatgtttgtacttccttttatgactcgtttaggaattacacaatcttggggtggttggactattagtggagaaacagcatcaaatccaggcatttgg
agctatgaaggtgtagcagcttctcatatcgttctttcagggcttctttttttagcttctgtttggcattgggtttattgggaccttgaattattccgtgatccaagaacaggtaaaacagcgttagatttaccaaaa
atttttggaattcatttattcttatctggtcttctttgttttggttttggtgcttttcatgtaactggtgtatttggtcctggtatttgggtttctgatccttatggattaacaggaagcgtgcaacctgtttcaccttcttgg
ggttcagatggttttgaccctataaccctggaggtattgcggctcatcacatcgctgcgggtattttaggtgtactggctggtctttccacctttgtgttcgtccatcaattcgtctttattttggacttttcaatgg
gtagcattgaaacagtattatcaagtagtattgcagcagttttttgggcagctttttgttgttgcaggcacaatgtggtatggttcagcagcgactccaattgaactttacggtccaacacgttatcaatgggatt
taggtttttttccaacaagaaatccaaaaacgtgtacaaactagtttaagtgaaggttcttctttacctgctgcttgggcaaaaattccagaaaaattagctttctatgattatattggtaataacccagcaaaagg
tggtcttttccgtactggagctatgaacagtggtgatggcattgctgttggatggttaggtcatgcagttttaaagatcaagatggacgtgatttatttgttcgtcgtatgccaacattctttgaaactttccctgt
tattttaattgacaaagacggtgttgttcgtgctgacgtaccttccgtaaagcagaatctaaatatagtattgaacaagtaggtgtttctgtttctttctatggtggtgaattagatggtttaactttcaatgatcca
gctactgttaaaaaatatgctcgtaaagctcaattaggtgaaatttttgaattcgatcgttcaactttacaatctgatgggggttttccgtagtagtccacgtgggtggtttacgtttggtcacgtttgttttgccttat
tattcttctttggtcatatttggcatggtgcaagaactattttcagagatgtttttgccggaattgatgacgatttaaatgaacaattagaatttggtaaatacaaaaaacttggtgatacttcatcacttcgtgaag
ctttt

>Chloromonas_clathrata_-_psbC_partial
atggaaacacttttaacggagttctatcaataggtgggcgaactcaagaagagactggctttgcatggtgggctggtaatgctagattaattaatctttctggcaaacttttaggagcgcatgttgcacatg
caggattaattgttttctgggcaggggctatgaacttatttgaagtttcgcactttgttcctgaaaaaccaatgtatgaacaaggtcttattcttttaccacatatcgctacattaggttacggtgtaggcccaggt
ggtgaagttattgatacattcccatattttgtgtcaggcgttttacacttaatttcatctgctgttttaggttttggcggagttttccactcactaattggtcctgaaacattagaagaatctttcccattctttggttac
gtttggaaagataaaaataagatgactaacattttagggtatcaccttatcatgttaggtcttggtgcttggcttctggttattaaagctatgttcttaggtggagtttatgatacttgggctccaggtggtaaaaa
aacgtctattgccaccattcagccaaagctgaat

>Chloromonas_clathrata_-_psbD
atgactatagcgattggaacatatcaagaaaaacgtacttggtttgatgatgctgatgactggcttcgccaagaccgttttgtttttattggatggtcaggtcttttattattaccttgtgcatatttagcacttggt
ggttggtttacaggaactacttttgtaacttcttggtatacacacggattagcaacgtcatacctagaaggttgtaatttttttaacagcagctgtttctacaccagctaacagtatgggtcactctttactatttgttt
ggggtccagaagctcaaggagattttactcgttggtgccaacttggtggtttatggccttttgtagctctacatggcgcatttggcttaattggatttatgcttcgtcagtttgaaattgctcgttctgtaaatcta
cgtccatataacgctatcgctttctctgctccaattgcagtgtttacttctgttttcctaatttatccattaggtcaatcaggttggttttttgcaccaagcttcggtgtggctgcaattttccgattcattctttttcttca
aggattaaatttgtcccactat

>Chloromonas_clathrata_-_psbE
atggctggaaaaccagtagaacgtccgttttctgatatcttaacaagtattcgttattgggtaattcatagtatcactattcctgcattatttattgctggatggttattcgttggaacaggattagcatatgatgttt
ttggtagtccaagaccaaacgaatattttacagaagatcgtcaagatgcaccactaattaccgatcgttttgatgctttaaatcaagttaaaaaattatcacaacaa

>Chloromonas_clathrata_-_psbF
atgtcaacaaaagctgaaactattacatatcctatttttacagtacgttggttgtctattcatgcgttagcagtgccaacagttttctttttaggtgcgattactgcaatgcaattcattcaacgttaa

>Chloromonas_clathrata_-_psbH
atggcaacaggaacaacttctaaagttaaattaaatcaaactgataattcaaatattcaagaaccagggatttctactcctttaggtactttattacgtccattaaattcagaagctggtaaagttttacctggat
ggggtacaactgttttaatggccgtttttatgctccttttgcagtatttttactaattattttagaaatttacaatagttcattaattttagatgatgttgtaaataattgggattattcagctaaataa

>Chloromonas_clathrata_-_psbI
atgttaacactaaaaatttttgtttatactgttgtaacatttttgtatgtttatttctttttgggtttctttctaatgaccctgcacgtaacccaggaaaaggtaat

>Chloromonas_clathrata_-_psbK
atgtcagcttttctattttacttgcaaaacttccagaagcttatgcaccttttgctccaatcgttgatgttatgccagttattcctgttttatttatttttattagcctttgtttggcaagcttcagtaagttttagataa

>Chloromonas_clathrata_-_psbM

atggaagtaaacatttttggattaacagcaactgctttatttatttttaattccaacttcttttctattaattttatatggagcgactcgtttaccc

>Chloromonas_clathrata_-_psbN
atgggaagctctgctttttttctttaccttttttttgtggtttctgctgttaagcgtaacaggttattcagtgtatataagttttggtcctccttcaaaaaaattacgagatccttttgaagaacatgaagattaa

>Chloromonas_clathrata_-_psbZ
atgacatctattcttcaactgacttttatttgcattaattttagtttcatttggattagtagtgggtgtacctgttgtttttgcatctccaaatggttggacagaaaataaaggttttgtttttttcaggtctaagcttatggg
cagttcttgtattcactgttggtgtttttaattctttgttgtttaa

>Chloromonas_clathrata_-_rbcL
atgggccacatgaccgaagccacgcagatcaccgacgccaagaagcgctattcggccggcgtcctcaagtacgcccagatgggctactgggacgccgactaccagcccaaggagaccgacatcc
tggcgctgttccgcatcacgccgcaggacggcgtggaccccatcgaggccgccgccgcggtggccggcgagtccagcaccgcgacctggacggtggtgtggaccgaccgcctgaccgcctgcg
acatgtaccgcgccaaggcctaccgcgtcgagcccgtgcccaacaacccgggccagtggttctgctacgtcgcctacgacctgagcctgttcgaggaaggctccatcaccaacgtcgccgcgtcgg
tcatcggcaacgtgttcagtttcaagccgctgaaggcggcgcggctggaggacatgaagttcccggtcgcctacgtgaagaccttcgccggcccgcccaccggcatcgtggtcgagcgcgagcgg
ctggacaagttcggccggccgctgctgggcgccacgaccaagcccaagctgggcctgtcgggccgcaactacggccgtgtcgtctacgaaggcctcaagggcggcctggacttcatgaaggacg
acgagaacatcaactcgcagcccttcatgcactggcgcgaccgcttcctgttcgtcatggacgccgtcaacaaggccagcgccgccaccggcgaggtcaagggcagctacctcaacgtgaccgcc
ggcacgatggaggagatgtaccgccgcgccgagttcgccaagtcgctgggctcggtgatcatcatggtggacctggtggtgggctacaccggcatccagagcctgagccactggtgccgccagaa
cgacatgatcctgcacctgcaccgcgcaggccacggcacctacacgcgccagaagaaccatggcgtgagcttccgcgtcatcgccaagtggatgcgcctggtgggcgtggaccacatccatgcgg
gcaccgcggtcggcaagctggagggtgacccgctgacggtgcagggctactacaacatctgccgcgacacgcacacggccatcgacctgccgcgcggcgtgtacttcgaccaggactggggtgc
gatcaagaaggtcatgcccgtcgcgtccggcggcatccatgccggccagatgcaccagctcatcgacctgttcggcgacgacgtggtgctgcagttcggcggcggcaccatcggccacccgcagg
gcatccaggccggcgccaccgccaaccgcgtcgcgctggaagccatggtgctggcccgcaacgagggccgcgacatcgccaacgaggggccgcagatcctgcgcgacgcggccaggtggtg
cagcccgctggccgcggccctggatacctggggcgacatcaccttcaactacacgtcgaccgacacgtcggactacgtcgccacgccgtccgtggcc

>Chloromonas_clathrata_-_rpl2
atgggcattcgttttcttcaagcatttacaccaggaacaagaaatcgttcagtttctgattttagtgaattaacaacaactaaacctgaaagttcgttaacgtataatatacaaagagcaaaaggccgaaatca
ccgaggtgttattacttcgcgccatcgtgggggtggtcataagcgtctttatagacttatagatttccgtcgagacaaaattggaatagaagcaaaagttattacaattgaatatgatcctaatcgaaatgcac
gtattgctctacttcgttatcaagatggagaaaaaagatatattacatccacgtggattaaatattggtgaaaaaatcatttcagaaataaatgctccaattattattggaaattcacttccattacgtaatattc
cattaggtgctgaaattcataacgtagaatttcaaccaggttctggtggccaaattgctcgttcagctggagctgtagtagaaatttagcaaaagaaggcaatttttgttacttacgtttaccttcaaaagaaa
tccgtttagtatcaaaaaattgttgggcaactataggtcaagtaggaaatattgaagcttataatttaactttaggaaaagcaggtcgaacacgttggttaggaattagaccaacagtaagaggttcagttatg
aaccctgtggatcacccgcatggtggaggagaaggccgtactccaattgggcatagtcgtccattaacgccttggggcaaacctgctttaggggttttaacccgaacacctaaaaaatatagtaatcaatt
tattattcgtaaacgaaaacaa

>Chloromonas_clathrata_-_rpl5
atgacacaaagacttaaaacatattatacagaaaacattattccaaaattacaaaaacaatttcaattcgaaaatattcaccaagtccctaaaatagaaaaaattgttattaatagaggtattggagctgcttct
caaaatcaaaagattgtagattcttcttttaaaagaattggctataatttgctggacaaaaaggaatcataacacgatcaaaaaaagcgattgcagggtttaaagtaagagaaaaaatgccagtaggtattgta
gttagcttaagaggcgatcgtatgtacagttttctagatagattaataaacttagctttgcctcgtgttcgagattttcaaggaatttaatccaaaaagttttgataaaaatggcaattacagtttaggtttagaaga
acaattaatgtttcctgaaattgaatatgataaaattgatcaagttcgaggtatggacatttcaattgttactacagcaacaaaacaagccgaaggtttagctcttttaaaagaatttggtttaccatttaaagctt
aa

>Chloromonas_clathrata_-_rpl14
atgattaaacctcaatcttatcttaatgttgctgacaatagtggagcacgtaaattaatgtgtattcgtgttttaggtggaagtcaacgtgaagcaggaaatattggagatattattattggagtagtgaaagatt
ctattccaaatatgccattaaaaaaagctgatgttgttcgagctgtaattgtaagaacaagtaaagggttaaaacgtcaaaacggaatttcaattcgttttgatgataacgctgctgtaattataaataaagaag
gaaatcctagaggcacacgagtttttgggcctatagctcgagaattaagagatcggaattttactaaaatagtttcattagctccagaagtaatttaa

>Chloromonas_clathrata_-_rpl16
atgctgagcccaaaaagaacaaaatttcgtaaacaccatcgtggtagattaaatggaaaagcaactcgtggtaataaaatttctttttggtgattttgctttacaagcattagaaccatgttggataacttcgag
acaaatcgaagctggaagacgtgtattaactcgttatgttcgtagaggtggaaaattgtggataagaatttttcctgacaaacccattacacttcatccagctggaactcgtatgggctctggaaagggatat
cctgaatattgggttgctgttgcgcgcccaggaactataatctatgaaatgaaaggggtttctgaaataattgcaaaacaagcttttcgaattgctggccataaaatgccagttaaaaccaaatttatattacg
cacacaaacacttttgatcccaaaaagtactcctgttcttccgacacaaattccttca

>Chloromonas_clathrata_-_rpl20
atgactcgtgttaaacgtggtaatgtatctcgaaaacgtcataaaaaagtattaaatatgtctaaaggttttcgcggagctggttctgttttatttagaacagcaaatcaacagaatatgaaagcattacgatatt
cttatcgaaatcgccgtcaaaaaaaacgtgattttagacgtctttggattgcacgtttaaatgctgctgttcgttgttatggtcttaattataatgagtttatgaattatttaaaaatccgtagtattaaattaaatcga
aaaattatatctcaattagcaacagctgatactgaagcttttatgcaattacttttatttaa

>Chloromonas_clathrata_-_rpl23
atgattgatttaataaaatatccaattattacagaaaaatcttatttaactttatttaaaaatcaacaatatacttttgatgtagatttacgattaagtaaacctcaaattaaaaaattatttgaaaatttatttaatgtaa
gtgtgattgcagtaaatactcatataccccacggaaaaatcttcgtgttggcactactaaaggatatcgagctcgttataaacgagctatattgaccttaaaaaaaggtcaagcattaaaatttgctatacctt
taacagtt

>Chloromonas_clathrata_-_rpl36
atgaaagttcgtgcctcagtaaaagctatttgtgataaatgtcgtgttattcgtcgtaaaggtacagtaatggttatttgttctaatccaaaacataaacaacgtcaaggg

>Chloromonas_clathrata_-_rpoA
Atgattaaatttgtcaaaaatgatttttttattattgtaaagaagcacgtattgaaagtcatcgaagtttttatgggtcttttgtctcggcccatttgaggctggacaaagtattacaatcgcaaatgcattaaga
cgtactttgttatcagaattaaaaggtttagcaatcacttctgctgaaattgaaggagccttttcatgaatacgctacacttccaggagttcgagattcaatattagatattttattaaatttaaaggatattgtcttaa
aatcaattaattatggaaataaaaaaagcggaacttcatatacagcaacttttaaaaacgaacaacgtggatatttaaaaagcagaggacctggagttgttacagcgtcagatttaaaattacctcccggaa
ttcagtgtgttgatccagatcaatatattgcaactttatcagaagaaggaattctgaatattaaatttaatattaatcaaggagttaacgttaatactgatacaagtctttccctacgtagcaataaaaaatcgact
aaaggagcaaataacgataagcaagaacatagcccatattatggctttacaataaatcaaccacccagtttttccacagttcccaaagttcccaaagttccagtagatataaagacttataatagttttgccat
gaatcaagctccagtagggcttggggtatctttagataaacgatttataacaaaaactctaaataaattatttttttctccgagatatttacttggggtaggagcccctagttcatctaattttagatcttctacctc
acagattttgatgctgggttatctcaacacaaaaaaaactttggccttggttctccagagcaagtgagcaaaaaacaattgcttttttaaaaaattcggtatataacgcttcacttagtttaccccctcccatcc
acctttggtaaaaaggggcgtagcaaaaaaaatgcttcgcatattttatcgtacccctagatgggtttggcgtagcagtagaaaagggacaggagaaggaacaagaaagtagtataaaacgtttatggt
taaaaaaaataaaagcaattcaatctagttatgtgcataaaattaataaaaaaaaaaaatgcttttagtataaataaaattaataaaactgttctcaaagaactcggtaacttgataaaaaaaaaaaattcagattt
aaacaccaaaattccttataacgaagcaataattgcattgtctactcctttaacaattgatcctgttttttatgcctgtaaataaagtcaattatattattgaaataaatgaacaatcagttatttcaaaagatagttta
agtaacataaataatagtagcaataatcctataaaaatacctatactaatttattgttacttacaaagcaaaaccaagaatcaattaaacaattaagtatatatggcaaagccatttctaaaaaatattataaaca

aatagctccattaaagaaagatcttagtccgtttttttcttaccccaaagtcgtgcgtagcaaaataaaaaatacctacattaagggtgaaacaaagcaagtccgtttccctaaacgaagtaaaggagcatg
gcaaatcgaaccaaaagctcttttaaaaaaaagcaatcttcttttttttaaagccagttttttaagaacttttctactagcgatgcattcaagctctgtaagttagcaaatcaaattccttttggcgttctcaattaa
aaataaaaactggttttttaaaaaaacaacccaagcaaaaagggcatatttcggtaacttcgctaaaaaaattacaaattcaattgtcttttttaaaccttttataacattacgttcggcattatggccattatca
attaatgttgctaaaattaattacatgatagctcttaaaaaagttagttatttttatggggcccctaagagttttaagcaaagctttaagttatttaatggaattgaaaaaacattatatagaactactatagctaaaa
ctaatttaaaagctccccccgcttaaacacccggctttttccttgctgctaacctttttagagtaagtttgctattaaaaaagtcttccttattacttgcttcgcaaacccacggggacaaaggaaaagggagag
tcaaaaagttttttttcaaatatagcaaagcttttaagaaaaagactttttttacccgtacgtaggaagcaaaaccaagaacttttgaccctacctaacaactccactaaaatcgtacccccttagaaaggggaac
aagagtaaatttttaggtccacaatttttttataaaaggccctaaaaaaaactaatagacactttagtcgaggaactttgttgtctaactatcctcacagctacgctcctttaagtgtaattcaaacatctttaaaaaat
aatattattttagaaatttggacaaatggaagtattcatccacgccaagctctttatgatgcatgtaaaaatttattaagtatttttttcaaaattagaaaaagctcgctttactagagcgatttttcaatcagaaaaa
acatatgagaaattagggttatcactcagttctaatttagaagatttgagggctaaaaaaagggatgctcgaggcccctaccctttaatatcaaatatggcggagccattaaagataaagcaaagaaagga
atatctaaggggtttgataaaccctactccaacctctacttttcttgctacggatccaataatttcgccttctaatgatttttaccattttcggtcatcaacaaactgaatcagacaacttctttagattatgggag
aagacttataaaacttttaaggccccccaatctttctactactccgtccagttacaaccttattgaatgttttagatttataaagcaagatccttcttctcttttttatctaaaaagaaaaaacaaaaagcaaattgtc
agttaagacaattgtctagaaaaacagagcaggtatataaatctactttaactccactagaaaaagcaaattaccctagtgttttattagcaaaaacactaataaaattattaataagtccttttctttagttcttt
attttcttacaaaagttctaaatctgataaaaaaaatagcttttaaagctaactatttagaaaaattgaaatctaccaaaatatttaaaaacaagaatagcaatttttctttgctcctatttggggagcaatgtttaaa
acaaataaacaatttttcttaactatgccaaaaaatcaccagatagctttttttactactagtgaaagtaaaatcaactccaaaaaaaataaaaaagaatttttaaccgtgaattgccctgcaattcaaggctgtt
ataaaaatggttttttttatccagattagctaatacaagagctcaagcttcgcttaaatttaacaacatctagcgtaatcaaaacttagagcctctttatatgaggggggcaaagtctttttctacaggcgaagc
ctataaaaaaacatttacaacaacggcatcttcttctcccgttaatttacaaagtaaacccatatttgtcaactctataatacctataaaaaaaaaaagagcgagatttaaaggccccggatgcttcccgtttgg
ggcaagtaagaggtccccttgctttgctaagcaaacaaagggagcaaccccaacagcaagggaaagaaaaaagtcacttgattatttatggctctgccataaaaaaaaattgatgttttttaaaacaaaaa
caaataaattggaatataaaaatgtctctaactgattataatctcttaaaaacaaaaccagaatttgttgatattggaaaacttaatatttcacttcgtccgtatacttgtttaaaaagagctaatattaatactttag
gagagcttttaaaaaaatcaaaagcagaattattagcattaaaaaatttggaacacaatcattaaaagaagttcaagaaaatttaaaacagcttggactttctttaaaatct

>Chloromonas_clathrata_-_rpoC2
atggcctacgttttaaattatgattcctacaaaatcaagtccaataagttaaaatttgaaaaaagtgttgcaactttagaacaaagtctttgcaataacttttaaacattggctataattgcaaagttataaaacct
aagttaaaaatttataattatagtcacaaaatgatgtcttctttttaattcaaatttaatggcaaagccattaaatgaacgacttcagttatctaaaaaagatatattaatggacaaaaaaaaaatttatatctccc
ctaatcgttaaaaatccatttttggaatcaatgttttgataaaaaccgtttgaaaaatttttgttttatggttttagtaacctacggagaatatcgcacagtagaattacttgaacaattaaaaaatgttggttttgaata
tgctactttagcaggaatttctttaggaatagatgatttaaaaattcctcctaagaaatctttttttaatttttgatgctgaaaaaaatacaaatttagctcttaatcaatataatcgtggtgaaattactggggttgaac
gttttcaaaaaatgattgatacatggcatagaaccagtgaaatttttaaaacaagaagttattgactattttgaagaaacggatacattaaatccagtctatatgatggcattttcaggagctcgagggaatattt
ctcaagttcgtcaattagtaggaatgcgaggattaatggcgaatcctcaaggtcaaatttttagattttccaattcgaagtaatttttcgtgagggggttaagtttaacagaatatattatttctacatatggagctag
aaaaggaattgtggatacagctttaagaaccgctaatgcaggatatttaacaagacgattagtagatgttgctcaacaggtaattgtttctaattttgactgtggtactccaagaggtatttttttaaatgatatg
aaagagggtaataaagttttatattccttgcagaatagattaattggtcgagttttaggaaaagatgttttttaataatcaagaagctttaaaaactctaaaacaaaaatttttaatgaatctccccttttcttacataa
aaagtcataaaatgtccttttggaggctcctgccattacctctgcttcttcctctgtttcaaaagatttgtcagaaagacttcccattccctcaactaatatagagactccttttttttactggcacagctttaaatttt
aaaaaggatggaatgcaaaaggcctactctccgttgttttttagaaatagtgaaatttctagtgatttagcaagtaaacttagtaaacttttctactaaaatttttgttcgttcaagtttaacttgtgaaactcaaaaa
tctgtttgtcaattatgttatggttggagtttagctcaaggaaatcttgtttcaattggcgaagctgtcggcgttgttgctgctcaatctataggtgaaccaggtacacaattaacaatgagaacatttcataccg
gtggggtatttacaggtgatacaacagatcaaataaaagcacttttataatggtattgttaaatatttagaacctattgcagggactttgatacgaacatcggaaggaaaaatttcttttttaacaaaaacagaag
gttttttagaagtcatttcaactgctaacccattcccttctatagctttgcaagaaaaaggaaaaaaagaatttgctactactaaaacaaaaaaagtctatataaattccttctttttacttattatatacgttcggcatg
gcgaaaatgttttgaaaaacaagttattgcacaaatttccaatatttcacgacaacagaataaacgagatgatgcagaattcacaataaattctgaaatggaaggtcaatttcgcataaagacctttgaatcta
ttaaaaaaacaagttggtccaaaaccgacttagcagcaaaaaaaaataggtcctactcacaatctagactctttagaagtcaattctacaatggaaacattactcgaagctagcaattgggttatgcttgga
ttttagctggaaaaatatatcaacttcctatttcatcgaatctattacctcaatctggtgattatgttaataaaaaaagtcttgtaaatgccatagcattaaaaccgtcacttcaaggagtcaatactggttttttaga
aacaaaattcaataataatttattattaaataaacctattttatttttagagattaataaaatgttttatacaaaaattggatattttctctctttaccggaatatcctttttttaaatcaaatattttaccttctttttgttctact
aaaaaacccggagacagcaattttcgccatagcttaccctatcacctctttacataaaaaagatttgttgtttacacctctagcttcccaatatttaacttctttttcaaacttctttagttcacctatttgcatcgc
caaaaaaagtgtaaggactcataaaacaaaactctttggtgaaaaactacaaacaggacgggcgttaaattccaatcgatttttagtggttctaaaactatagtaaaaatagttctgccaattggagcaaa
aattttagtttatttttagaatggtttccccattcgaaaaaagtttcaagtggtgctttagtgactgattttaaaacttacaaactagttcttatctaccttcaatgtggccaaataaaaaaaacttttctttaatccttta
aaaaatctttattctgtttattagttacgtcttctttagagttttttttatggcaaagccatacccccacatctgttttttcttccccttttactttagatagtgcagcaaacctatccgttatttcatccctactctattagg
gagtcggcttagcatgcccccctaactccaggggatgggggagagaagaggatcaagtatattatttttcaataatgttaattccttttcctttgcttttcgaagaaatcgtactgaaaaggtcaaagctttattc
ctgttccacctaaggacgtgctaaacccaaactttgttgcacaaaaacaaagaggtttaaaaggagtaagaacctctatttctcttctcattactcaatccaaaaaggtatttttttcacccctatagggggaat
gggtaacaaagcaacccctttactgacttatttcccttacgaagtaaaaagcttaggggccccccactttttcccatttccacttgctatggtttctataggctctgcctataaaaagtttagtaatttattacgtatag
caaggaaagaagcaactatttttttgccaagaaaacaaagcaaggtatctttagattgttgggcttctcaatttgcttttgctatgcaaagaatccttaataaaaaaaaaatttatctcgaagtccttaatcatgtgc
aacatgcgcttaaaaaataataaatttaatacaaactccatctattcttttaatactacttttagcaaaaagtaaaaaaggacaatttttcttatttaattaatcgtgagaaaaaaataaaagggtatattaaattagta
aatttttttaaattttataaattctgaaaagcttcgttgcttttaaatatcaaaaaagttcttattataactatgaaaaattcatttcaaaaaaagaaaaactttttaaatgaaacttttatatgaaaaatttttttatatttcaag
acaatttgttgagattccacaaaaacatttaactatatctaatatttttttgctactaatgaaacaaaagttaaccccaaagaaaataaactgtcttacaatatatatatgacacttttcaaaacgaaatcgacaag
gaatgaaaaaagcatttttttcgaaaatctagtgacgctagtggaatttttaggtcaaacaacacttaatatagtgaaaagagtcaatacttttttctaatgacgctattttcaaaagaagccccttttttatgcctaat
gtcatgcctatttttttcattaccaaatagaaataaaaaagcgggggagatttttaaaaaaaaatatcagaagcgtgactccactcaaaacaagctggaaggacactagcttattattctcatttacactattaag
agctcaagtattacaaaattcggccttttataaaaaactcttggccttgtaatgtttccaccaaagaacccctttttaaattaggtatctttagctctatcataccccctgcgtagcaaaaaaaaggggggtgttaa
aggtgtagaggctcctatgttatctaaatttaggctttctcaagagactccaacaatagaaaaattcactaataaattcataaagtttaatttaaaaaaaaaaatactccgaaaatactataaaacaatctatttttg
taaaaacatcttgtaattcattcatacaagttaccccgtaccctattcatctaaaaaggcatagccttctttttaaaaaagcatagctttttttagtaccccaaaacaaagtatttttctttggggtacgtagcaaaaaa
aaaaaagtaaaatttttacaaagtaaagtgaaaaagattaaggttatgcttggatttccttagattataaaaaattatttaaaatgattaaaacttttgtagctaacaattataaaatttataaaaaaatatctctctttt
aaaaatcaaattggaattaagttcttaacttttatttatggctaaccccaaaaaaagtcttatcctaagccatggttactcataaacaatggttattatggcgggttactattatggaatccaatttatatatatacaa
tgataaaagccataaatatacttaaaaaacgcttatttaaaaaacttggttaggtaagtattttaagataaaccctagtctgttttataaaaggacagcaaaggtaaaaatggaaaaaacaaaagcttcttaca
aatggctttgccataaaaagaaattaaatctttaatggctttgccaaatgactttgccataaaaaaattaaagagtaacaaaccaatttttaaattcacatcaaaacttttaaaaagaaaatcgttatttattctc
atcatttaaaaaaagtcactaataatttttcaaattgtaaagaaaataatttttaaatttaatgggtttgcggtatctaaagatattttaatagtaaagatgtatcttgtaaattattattgatggcgcattatataaaac
ggggtccatttaaagaacatcagcaagctatggcaaagccattattatctattatttaccattattcaaaaaaacacataacaaccctaatacccaaaaaattcatttattagggggcccaaacgcgggggag
ctttttcacaaagcagaaattaaatggaactgatgctgttttccctccaaccctaagctcagaagttgcaaatttggttctcctacttttttgcttaattcatcagataataagcaaccctgtccatagcagcta
tgctgccaacaaaaaaaaatataaaaggaccccttgaaaattaaagctttaccaaaagttcataaaaaaaaagtcccttcggttgttagaggccgctacccccaaaagctagaaaaggattcagcaaa
caaggacaaatacggtgtagccataaaagataaatcctctatcatgaaaaacccatatcaaattacaagtaatcaacatcctggattaattactatttttgcaaaaaaaaggttggatttttttccttggtgctcc
tcactgtctccaaaaaggaacagtattatctctcaatcttataagagacaaattcggtctttccttacgaagtaaaaaaaaccaccttgcaaggaactactagcccaaggggtacgaaaaaaggtgacaa
ggaatgggtaggggcccctaattccttgggttgtatagtaaaaaagggaagctaagggcacctacccctaagcatttttttagtccaggacaaatttttttgccttttttaatgtttgataataataatatttatgt
tgaatggtttaataatttagttctgtatgcttcgcatgggttgataagcaccggttatttccttttatgggtcattattaaaaacgtgggcgatcaaaaaaataattaattataaaaaagtaaaaaaaacagttaa
tgttttaaaaaataataatgggattagaaagtctgacgctcttattaaaaagaaagtcttaaagagtctttatagtttacacccattatttggtgctcacgtttatggacttatgtatattttaaacccttgatttgct
ccccttcaaagaacccttttttgtttaggagccccaacccaaaagtaagggacttgcaacttttcccttgctacttgcaaagcaaaaccaccaaagcaccccctaaaagaggtaaaagcaatcctaaaaa
gggtaaaagcaccctaaaagaggtaaaagcactcctaaaaagggtaaaagcacttgtaaaaaaaaataaaaggttgtatgttacgaaaattaaaggccccttttttaagtttaaaaaaaaaaaagtatttaa
ctatttttgattctgttgttagggttccatttctaaaagcacccataaacaaggtcaaagtaaaaggcaggttaggggggcctctacctccaattctcgcaatctactaggaatgggattaactttagttacacag
aaggcgcctctacatttatgtttatacatagagacttaccttaccaaaaactatctaaatcaaataataattcggtagataaatcttttgctactagtgaaactaaaattaacccattaaataatctaagttgtttag
actataaaactatttcttctttttcctatgagtgaattaaatttttaagataaaaacaaatgattattcttgggtacaaaatcaacattgtcgttatctaccaaatcagtttataggcaacgccaataaattaactagttct
ataaatttaacaaaagagactagtaagcacacaatacaaaaaccgatcttaattattcaaaaagctttagaatttactatagctaagccaataaattataaaaatgaaattgcagaattatatgcttctaacaga
gatttagtaactactggaacaaagaatttttttaatttctggtttacaggggaaaatgctagagtcttgcaaagcagtaagagggttatctaaagctccgccgctttcccttatttcatccattatttcatgccccaa
ccctaaatcaccggggtacgataaacctatttatgattttaggggttataacactaaaggttctaaaatagagagtaatttttggaggggtcaggcgaagcctataaaacaaaaaaaaatcgtttcggttaaa

ataaaaccgaaaagaaaaatttttagcttcgctaaagaaatttagtgttgtttttccttccctccagtttggggttaggttaggatattcgatgcaaccctttttgctttgcacccaactaaaaactggaagtatt
taaaaggaaagcctatagttaacgattgttttttaaagcaaataacttttttattaaggcttctctatttttagtctctaaaaacaagcttcgcttagtacccctggcaagggtaaaaattcatggcaaagccata
aatttttaaaatttctggcaaagccagaaatttttaccctagatgtctctaaagatacccaagccataaggtatggcaaagccatttttaaaaaactaaaaaaagtaatataagttttttttctaaaaaaatttatt
gttggcaatcaattactaaatatggcaaagccataaatttagtacaaacaattaacaatttacaaaggaatcctttaacacgttgtggaactattattcaattgcgtaaaaataaatcttataattcagtcttagtta
gaaagcaatttttgggttttactctcaaaacaaaataaaacaattgaaagtgttgctagtacaatgccttattttgatttaaagtaaaacataatgaatttttacactaaaatggtatcacttttaaaagtaaaagc
aaatagggggggttaaaattttaaggatgctagctaaaagccctctttttaaaatcaaaaccaattattaaagttgatccattattacgaaatgtgttttcttgtctatttccttcgctatgtaagagctattctagtaag
gataattcccgcaaaaaaaaaagcaacaaaaaagaatctttttgaaagattcctttatctcttctaatgtacaacatccatttaaaaagccccaactcccatctaaggggtacgataaaagcacccctaaaa
ggggtaaaaattctaaagataaaagcacccctaaaaggggtaaaaaattagattcgttcttgcgtagcaaaccaaaggtagttgcgcataaaaaaaagaggggttgcatcgaagatactaccctacccta
aaagacttgaatttttacaaaaaagtcttagctctacaaaagctaataatttagttttttaaaaaaccaattaatttaactcctttattaattggttataatatgccttatggttttgatttagattttaatctacctaatattc
actttagcgaagctaacttgaattctaaggtaggttataataaacttcttggttctggaagtttacctacattgaatttttaaaaaaaaaaaaaaaattaaagcaaagcttagaatcaattagccgagctaaaac
gagtccaagttttcaaacttgcaacatttgaaaaagctcaattcttggttttttattttgctcaattacaaacatttaaaaatccatcttatttgcacaaccgtgttttgcttttttatgagtaaaaaattaactaattt
tataaaacgttttaaacaaagtttaaaatggattcaaaattctgatccttttacttcttttaactaataataattttttcttttaaaattaaaaaaagtatagggataagaaaggctaatttttcatttctgagtttaaattcat
ttttaaattataaaaaaaaaaaaaatagcaacaaattcttttttttctaaaaagctactaaaagggcaaatccctgttaaaaaacaagttccctgtttgagttggaagctttttttgacatttttattcttctaatttatata
aaaataaattacttttttcgagtacaagttctccatccgaatgtttgccagttaaaaagcataacagaaggttactgattatgggggctccgaagaaaataaaaaaaaaacagtaataagaaggaagcaaatg
taaagggagtgccttggtgtatacaaagtaacaaacaatactaatgaaatttttcttttttaagttttttgaaggtgagtttattaaaacaaaaaatattatataatataagagtcttactagtattgttgattttaata
gttataataatttagatctaataaagcaacaaggtacaatacccataataggcacagctttccttccatctcttaagtcctttgcttactcagtaaaaaaaaaaaacgggggggtactatgctaggggcccc
tgcccctactaagtttcgctgttacgcttacacgtttggtgttaagaaaagtttgcaagagcgtcaccacaggcatataactaaaacttcttttttaaaatttatgggttttcaggtattacaaatccaaagccaa
ctaaagttgcaaaaatgaaagtcaattcttccatacaaacaagcggtaacataaaattaaaaaaatatgtaaaatggccttgttaagacgaattattgccaatattagtcaaccactgaatcaactaaaaag
gcaaagcacttttagcgttttaaaaaaaaaaaaaaataaaaactcagactaatttcaataatttttcttggaccatcaagtctttatttaacaaaagcaaatttaattagttattctttaaaatttggagtaggtttggg
gttaggggcaccttctctggctccgaaggaaggggttaaacaaaacccaaataatcaaaatggcacttataaacgtgctagagaaaatctttcttctgaattcataagctttgccaattattataaatataata
atattttaataaaagtcagttcaaagcttgtagaagaaccctaaaagaggtagaaggtaaaagtatccccctaatcaaaatataagcaacgcttatacaaggaatactatgatacatataaaaagggtatgat
aaatggggaacataaaagtataccattagataatactcgaccaaatcatttatcattaaatgaacagacttataaacaaaccatggcaaagccatgttttacaaatttttctattaataactttttgaaagtcaat
atattagctccttcactaagcaactaaaagtacttgaaacttttagctcagaaaaacatttaaaagtgcaagacactttactcgatacatctatggttgggggagctttacatggcaagccatatttattaataaa
aaatttaatttgtttagctaaagccacaaacttttttatttgctactggaaaaatcaattctttactcccttctcaatttagtacacaagttaagggtagtctaaaaattggggtaggaactggaagtataaattccga
gagaaatttgatgagtgatggcaataaaatagatttatctacaaatattcaagcaattactagagcaaaaatttttaaaataaatggattaattgcaggtttaccaaaacaacaaggtgtttcattatttaaaaat
aaattatctttaggggaatttttgtttatggcgataaaatcatgtctcctttcttaaacccatcgaattgttttcctatgcagtcctttaagggaaatgaaacccccggtgacgaagcacgtactttgggagcaat
taatacaaagaaaacgaatattaaacaattaattccatttacaccttaaggcgtacatttattagccatactaaaagttttccattattagggggaaagagcacaacatcaaaaaagaaatacaaaaaaataaga
gtgttgctgttacgttagcaggtcaaattattcatattaatcaacacaaaattaccttaagacgggcacaacctatttttatttcacctaaaagtattttacattttaatgatggtagttttgtgcgtaataatgatcca
gtaatgacattaacatcaccaacgattgaaaacgggtgatattgttcaaggaatacctaaagtggaacaatatttttgaagctcggatcactaaacaaggaagattatttcgtgattgtttgccaaatttattaaaa
ggcttatttaaaagatacagagctcaattaccattagacaaagctgttagacaaagttttataaaattcaacaaatcatagttgatggagttttacgtgtttatagatcccaaggagttactattgcagataaac
atttagaggttattgtcagacaaatgacttctaaagttaaaattataaatggagcacaaactggttttttccaggtgaaataattgatttacaatttgtggaaaaagtcaatttatcattaatgaaaaaaattcaat
atgaaccgttaattttaggtattacacgagcagctttagaagtggatagttttttatcggcagcaagtttccaacaaactacaagagttttaagtagtgctgctattgaaaaaagaaaagatttttaaaaggttt
aaaagaaaatatacttttaggaaatttaatgccagctggaactggttatttagtttctttgaagatcttaaaaat

>Chloromonas_clathrata_-_rps2
atgcagaatatagaaaaaaaaattatagaattaaaaaataatttaaaatatttaaaaaaaacattaacaaattatctgttcttaaataataaaattaatagtgaattagaacttattaaaacaaaatatttatctgaa
caacaaattattcgtatcttaagaagaaaattaaaacgtcatgaagctgaaaaaaaattgttacaatttttaccaaaattacgttatttacctactcctaaaagtcaaatctctgtagctattcaattattaatgaaa
cgatttgttgatcctaaaatgaaatatccattagatctaatttatgaccaagcttcgcttaaaaaaaaatcaaaaaaagtagctgctgctagaaaaaaaaaatggcaacgattagaaacttatttggtggtatt
gcaaatatgactaaattaaagaaaacacaaattgctaaaaatgttgctatcataataggtcaacaagaagaacttaatgctattcgtgaatgtaaaaaacttggtattaaaatgtttcatattgttgatactaatt
gtaatcctggtttagcggatcattttattcctgcaaatgatgattctagaaattcaataaaatatttattatcaaaaatggtaacacgtattagattagctcacaaaattcgtttacgattaaaaaaaaaataaaaaaa
tgcaaagtaaagaaacaaaaggtaatatatttctaaataaagctaaaaaccaaataatagcaaataaaaaaaaaccaattaaaaataatgcaattccatttagccctactttgggagttaaaccaaatagcc
ccaatgttggtgatatttgtaatttaattgtaacagctttaggtccaaaaaatattggtattgatgaattttcatatggttattcaatttttaattccaaatgctaaattaggagaaaaagttaaagtacaaattatcaaa
aaaaatttaaacaaagctaaatatgctgttgctaaaagaattataccaagctctaacaaaaaccatgacgttgccataaaaagtaatggcaaagccatacatagctctgccatacttgaagatggaaaaga
aaaaattaatgttggtgatattttaactgttaaattaaataagattgttattcaaacagataattcttctattggtatagttaattcagatttaaaattacctataaaatacccaattaataattcctaacgtagctgaaac
taaagttggtcaagaaatccaaattcaagttactaaaatgaaattaaaatatggatttgcttttaattaaatctagtgtagctaaattggatggggtaggatcccctaaatctctaagttcagtctataaaacaaaa
cttacaactcaacgtcctaccgttcaatttttagataaagtaggggcccctttccctttatttcatcccccacaagaagctacatcaatttcaaaagggtctaataatatagggggaaaagacggagtagtaggat
caagcctatctttcgatactaatttagcaaaaccttttatttctttaagctcacgtccaagtgaagcttatgtttttggaacctctaattcttttaccccttgcgcaacaaacaggggggtaactctgttgctaaacaag
ggacttttacaaaaaatggctttccaagtgttaacggtcctaaagttaataattatggtaaaaaaaacaatgtaaataatagttcttttttcatcatttatttcagtaggggccccaagccttgctgatttaaataata
ctaaaaacataaattacaaaccttttaacgtagctaatatgacaaaaggaagtaaaattacattaacattacctaaaacagcaaaaaaaatatgctaaatatttagtcattaacgttttaaatcaaaatcttgataa
aaagttagatttaggtataggccaagcctataatcaatcaacgcttaataatggaaaagctagccataattcttatattttattttttaaaaaatttattaggagcaaaattaggcgataaagtacgagtaaaattc
acaaaaattggttcttcatttgcaattggtgaaattatttgtatttcacctaattcatctttaacaaaattagttcgtgttcgaaaaaatttaaatgaaatgcttgtaaatggaatgcattatggagaaaaagttgttaa
atgccaagctaaaatgcgtaattttatttggaatagaaaaaaaaataatcaataattcgacaactcatccattagattctaatagttataatcgaaatgcaactaaaacttcgaaccttggcgcagtatcttaccg
agggcttgtgagtacatatcctacttttggcaaaaaaaaactactccttcaatcccattccaagctactcgttattagcttctaatggagccccaaaacaaggggcttcttctttatctaaagataaggcgaagc
ctgcaaagaatggcttgagtaaaaaacaagatagtaagaagcaaggaataaatacattagaaggttctaaaacttaaataaaggcgttgctttaataaaaaaaggtagacatttaattaatttattaaaaact
cgtcgttgtttaaatataggacttaatcaattaacaaaatatgcagctaaaggaagaactttttattgttggaacaaaaaaaccagcagcaggcttaattgctcgcgcagctttatttagtaaaacatca

>Chloromonas_clathrata_-_rps3
Atgggacaaaaagttcatccaataggatttcgtgttggtattactaaaagacatcaatcacagtggtttgcaagatttacaaaatatgcttatgctcaaagtttattagaagatcgtatgctacgaaatacatta
atcagtgtagcaaatgaagctttaaattctaatttattgaaaaacgtgattctacatttcaacaaaatgtaataatacctaaagtgacacatattaaaattgaacgaggattaattccatatgaaatcggcatac
aaatacatgctcaagctccacttaaaaaattgaaagcatcattaaataatttaaaaattaacagagaattattatgtaaattacaaaaaactcgtcgctatttaatgagtttaggaaagcaacaaagttctaaac
caaaaagtgaagggcttgcgcagcttttaaataaagaaaaagagagttccattcctggtcttgcaaaacaatcaaaggggggaaaataaaggagtaattaataaagatgattcttcgctttcttttaccaaag
acgaaagaggagcaaataaagtaggagttacttcgctcatgcgtagcaaaccaaaggtgggatcagaagatagagttggaaaaaaaaaataatttttaaaaatgtaaataatcttaaaaaaaaactgttttgt
tttagttcaactaatgcaactagtaaaagtaagcaaaacttaaaagaaaattctatttttttcgtctgcatcttcattctctcttgctgcgcaacccaaacaaagtttagaaggtaaaggtctttttatttaactctaaat
ttaaaaaacgtagacgtcctaattatttaagtatagtttacaaacaaaaaaaaattattataaaacgtttacgaaaacgtcaagaaatacgtcaacgttatagaagattaatgttaaaaggcttatttgttcaaaa
aaaaggaaaattaattagttgtgttgcttctttatttttttagtaaaaaacaaataaaaaaagctagtgtttgcttccaacgtacagctttaaaaaaaaataataaatttgcaaaaaaagttaacataaaatcaagac
catcttcttttacaaaaaaaactaatttgcttaaaaaaaaaagtctagtttcgctagaagaagcttcatctccatcgttcgatacctggttaactccaacacgaagggttataaaaaaccttccattctttttttg
gggttaagaagcaaaaagacgggttaaaacaatcaaaaactgataatctttcaaaattagctggaccgttaaaatttaataaaagtaataacttagctaattcaatcacgtggccttcattattaaccaattcta
aatctaataaaaatataaaagaaaaacttcttcttttagcaaaattacgtaatagacagagctctcctctaacaccaacgttgttagcttcgccgttttctaagcgtaacttggataataagattaaagcaaattc
tcaaactatagggcgaagcctataaaacgaagttacctacctactcgcttggggtaggagcccctattgctattaataaggagcaagaaggggggctacgcaccccttggaatgtaggttctagaataaag
aaaaaatttgtttcattgttttttaaataaaattaacaaacgttttttaaaacaactaaaagaacaaatgactcaatggaataagtctcttaaacaacatcgacaagaacaaataaaacttttatggtgcacttcgtttt
gctcctcttggttacaatcaaaaatggagtttaaaaagattaaatttttttgaaaaataaaccgttatttaagcttttaaaattaattaaaattattgagagaaaatctttaataaagttagaaacttaagaaaagatt
ttatagcatttggcacaatttcaaaaagtgaagctttgagttattaccaaatcattaatttcttaaaaaatttaaaaattgcttctttgaaattaaaaaaagagcaaaaacgtctgttagtaaaaacttaacaaaag
attattctaaaatatcaatattaacatcagcaaatcaagaaatacctaattcaggtgagaataaagaaggttcttctttccctttaaaaagtggtaataaggtatggcaaggtgctcctaatcttgttggggcctc
taaacaaaagagcaatgagctaaacttaataaaaacatgggaaatcaatcacactatcggcaaaactcctaaagcaataggagtctctggaaattctagcttcgcaacttattgacttcttctgttggtgtg
ggtacctctaatcaaggcaagcccttattagaagcgatcaataatgaatgccgtaagattaattttattgaatatttaaaagaaatggttaaaaaacatagaacaaataatttgttttattatttatcaacaatttcc

gatgcccgtaaaaatttaaaaaaaattaaacaatttactaaagttcattcaaaatttttatttggatttgagtttaaccaaccaactgctatgttagggacttctactactttaaaaaatacagaaaatacagaaaa
tacagatggtaaagatttattagttttaaataaaattaaaactgtaattgatttagcttctaaaaaatcagattttgaaaaaggattaccagatattttcttagaacaactggaaaagcaacgtaatatgtataaac
aactagctaagctgcgtgctaaaatttcaattaaatttttattcagttaaatctgaaaacgtcaaaataaaagcatcaataatctcagattctgttattgatgctttagaaaaacgaaaagctttcagaaaagttata
aaagatgctaaagaaaatttaatgcgcaattcaaaagtaaaaggtgttaaaatacaagtatctggacgattaaatggagctgaaattgctcgaactgaatgggtgcgaagtggtcgagtgccattacaaac
tttaagagcaaatttagattattcgtataaaacagctaaaacaatttatgggataattggtgttaaagtatggatatttaaaggttatactaagctagtttaa

>Chloromonas_clathrata_-_rps4
atgtcacgttatcttggtccacgattaagaattattcgaagaattgggaaattaagaggctttacacgaaaaaaaccttttcgcagagtctttaaagggcgaggtgctttacgtggtaaagtaattccaccag
gtcaacatggaatggtaaaattatttaaaacacgaccatatgattcttctgagtctgattatctaattcgtttaaaagtaaaacagagattacgttttaattatggtttaactgaacgtcaattagtaaacgttgttc
gaaaagctaaaaaaaataaagaatctacagggggtattttattacaattattagaaatgcgtttagataatatagtttttcgtttaaatatggccccaacaattgtggcagctcgacaacttatttgccatggtca
tattaaagtgaacaacaaaaaagttaatataggcagctacttgtgtaaacaaaaagatgttatttcagtttcaatgaaagaaaaatcattaaaacttattacaaataatatttacaaaattactatcaacgcgtgcg
cttctataaagaacgtttgaaaaaaacttttgtcttttattcttctaaaatcgcaaatagtcacaaatatggcagctgcaattagtcttattcaaagcggaaattgtaaaataaataatctcgctacacgaaaacct
aattatatttgtggcgaacgtgatgtcattacaatagttacaaaagcaggaattcgtcaagtaaaacaaaat

>Chloromonas_clathrata_-_rps7
atgcctcgtcgccctataacaaaaaaacgttctcttttgccagatcctatctataatagtatttctgttcatatgttagttaatcgtgttttaaaaaatggaaaaaaatctattgcttatcgtattgtatataatgcttta
aaagaagtaggggacataactaaaaaaaatccagttgaaatttttgaaaaagctttagataatgttacaccacgcgttgaagttaaacctcgtcgacgagctggaacggttcagcttgtaccacgagttctg
cgcactggtgacaaagctcgtgcatcagcccttagatggatttagaagcttgtcaaaaacgatcaggtcaatcaatgattgcaaaattaaaaaatgaaattgttgaagcgtataaaaaaaaccgggtttgct
gttcgaaaaaaagatgaacttcataaaattgctattaataatgcgatgtatgcgcgaaaaccgcaaacagtaattaatgctgtaaatcaagtgatagct

>Chloromonas_clathrata_-_rps8
atggttaatgatactattagtgatatgttaactcgtattagaaatgcaagtttagctaaaaaatcaacagtttgtattccatatacacacttaaatcagcaaattgctcaaattttagaaaaagaaggatatatttta
acagttcaaatgtcattagattcaaaaaatctaattgtccgtcttaagtatagatcaaaaaaaatttataacggaaaaactaaagaatcatgtttaacaaatttaagacgaattagtcgtccttctcttcgaatttat
acaaattctagagaaattccaagaattttaggggggaacaggaattataattctttcaacaccaagtggacttttaacggatcgtgaagctcgttctcgtggtattggtggtgaaatttttatgctctatctggtaa

>Chloromonas_clathrata_-_rps9
atggaaattttagcaagagcggtcggacgtcgaaaagaagctgtagctcaaattcagcttgttcgtggaaatgggaaatttttaattaatgataaacctgcccaagtttatcttcaaaataattcttgttctctta
tttgtattaaatcaccattagaagcagccttaagtctttattctaataatctttcaaatattcaaagtacttctcaagattctgccttaattaaaacagttaatgaccaattagaagaacaaagtactacaattggcc
aaactgaaaacagtaatggcaaagccataaataacaaagagggtggagtagaagccaatgtagcaaatagttctgtcaatactcacggttcgccttttattgtcagtgaaacagagcaaccttacaacttc
gacgttgttgtaaaggaaagctttgagttgaatatttctggtgacgaaacacaaaatttaaatgattctaaagataagctcagcttgctccctttaaataatttagtttcgttagatgaaattgacgtacttataaa
agtaaaaggtggcggacttattggtcaagcagaagctataaaattgggaattgcacgagctttttgtctaatgcaagcagcgcttactgaagggcttttaaaaaaagttttaaaaactaaaggttatttaactc
aagattctcgtgttaaagaacgtagaaaatatggtctaaaaaaagcacgtaaagcttcgcaatatcataaacgttaa

>Chloromonas_clathrata_-_rps11
atggcaagacaaacacgaaaagttgcaccgaataaagcaaaaaaaaaaatttatcggggagttgtacatatccaagcaggttaccataatacaataattaccataactaatgtacgaggagacgttctttg
ttggagttcagcaggtgcttgtggatttaaaggaaaacgaaaatctacaagttttgcagcaaaaaaagcagctgaaacagcagcgcgcaagtcaaaagatgctgctatgagagaagctaaagttttagta
actggtcctggtcaaggtcgggaaagtgctattagagaaattttcaaagcaggtataaaagttaatgttattcgtgaaaaaacaggtattccccataacggatgtcgtccgccaaaaaaaagacgtgtttaa

>Chloromonas_clathrata_-_rps12
atgccaactattcaacaattaattcgttcagcacgaaaaaaactaacaaataaacttaaagctcctgctctaaaatcttgtcctcaacgaagaggtatttgtcttagagtttatactattaccccaaaaaagccc
aactcagctcttcgaaaagtagctagagttcgtttaacttcaggttacgaagtaacggcatatattcctggaattggtcacaatttacaagaacatgctgttgttttagttcgaggtggtagagttaaagatttac
ctggagttcgttatcatattgttcggggcactttggataccgctggagttaaaaatcgtgtacaaagtcgctcaaaatatggggttaaaatagcttctaaaacagcagcaaaaacagcatctaaaaaataa

>Chloromonas_clathrata_-_rps14
atggcaaaaaaaagtatgattcaacgtgaattaaaacgacaaaatttagtaatgaaatatgctgaaaaacgagcttctttaaaagaacagattaaacaaacatctttttttaaaagaaaaattagcgttacatcg
taagttacaacaacttccgcgtaatagtgctgctgttagattgcataaccgttgtatggtgacaggtcgtcctaaaggatattatagagattttggattatcgcgacatgtgttacgtgaaatggctcatgaag
gtcttttaccaggagtacaaaaatctagttggtaa

>Chloromonas_clathrata_-_rps18
atgaatcaacccgttcctcgttctcaaaataatattaatggaacttttattagcccaagtactcctaaaaaaccttttttttcaaattctccaaataagagtttcaaagaaaattcttttaaaggtacgaaattaaata
ataatagttttacaacaaaaacaagcaattttaataaatcaaataaattaaacaaaggaaatacctcttctcagtcaaaaggtaaattaagaaaaaatttaaaaataaaacgtgttttatcattatctcaaattctt
ggtcggattagaatacaacgtcaaaaaaaggtagagctacaaaaacgccaaaaaccaataaaaccaataattccaccaaaatcattaattattattttaaaagataaaccagaaaaagcagtttataatcgt
agaataatagattataaacattgtggtttattacaaagatatataggtttaggtggtaaaattttgccaagacgacaaacgcgattaactgcaaaacaacaacgatatgtagcaaaaactattaaaagtgctc
gaataatgggattattacctttttgtaagtaaagaacgaggctttttagataa

>Chloromonas_clathrata_-_rps19
atgccacgttcgattaaaaaaggtccgtttgttgctgatcacttattaaaaaagattgaaaaattaaatgctcaaggtcaaaaaaaagttttaacaacatggtctcgttcttcaatgattttaccaccaatgatag
gtcatacgattggagtttataatggtcgtgaacatattcctgtttttattactgatcaaatggttggccataaattaggggaattttctcctactcgaacttatcgaggtcatggtaaaacagataaaaaatcaaa
acgttaa

>Chloromonas_clathrata_-_tufA
atggcacgtgctaaatttgaacgtaaaaaacctcatgttaatattgggacaattggtcatgttgaccacggaaaaacaacgttaacagctgcaattactatgactcttgctgctcgcggaggtggtgctgga
aaacgttatgatgaaattgattctgctccagaagaaagagcacgtggtattacaattaatacggctcatgttgaatatgaaacagaaaatcgccattatgcgcatgttgattgtccaggccatgctgattatgt
aaaaaatatgattacaggggcagctcaaatggatggagctatttttagttgtatctggtgcagatggtccaatgccccaaacaaaagaacatattttgttagctaaacaagtaggtgttccaaacattgttgtttt
tttaaataaagaagatcaagtagatgatgcagaacttctagaattagtagaattagaagttcgtgaaacgttagataaatatgaatatccaggtgacgaaattccaattgttcgtggatcagctttattagctctt
gaagcacttgttgaaaatccaaaaattcaacgaggtgaacataaatgggttgataaaatctacgaattaatggcagcagtagatagttatattccaacacctgaacgtcaaattgacaaacctttctattag
cagtagaatcaactgtatctattacagggcgtggaacagttgctactggtcgtgttgaaagaggtactgtaaaaatcggcgaagttgtggaaatagttggcttaagagaaactaaaagtacaatcgttactg
gtctagagatgttcaaaaaaacgctcgatgaatcagttgccggtgataatgtcggagtactgttacgtggcatccaaaaaaaagatatcgaacgaggcatggtgttagctaagccaaaaagtattcttcct
cattctaaatttgaagctcaagtctatattcttacaaaagaagaaggtggacgacattctgctttttagcaggttatcaacctcaatttttgttcgaacaacagacgtaactggaaaagtaattagttttactcat
attcaaatgcgtaatccttcttctgtagcagaagaacattctaataaaatggcaatgccgggtgaccgtattagtatgttagttgaacttatgtcaccaattgctattgaaaaaggtgttagatttgcaattcgtg
aaggtggacgtactgtaggagctggtgttgttactgcaattattgaatcaaaa

>Chloromonas_clathrata_-_ycf3
atgccaagaactcaaaaaaatgataatttattgacaaaacatttacggtaatcgcagacattttactaaaagttttaccaacttcacaacgagaaaaacaagctttttcatattatcgaaatggtatgtctgccc
aagctgaaggtgaatatgctgaagctcttcaaaattattatgaagcaatgcgtttagaaattgatgcttatgatcgtagttatattctatataatattggtttaattcatacaagtaatggagaacatggtagagca

ttagaatactattatcaagctttagaaagaaatccttctttacctagtgcattaaataatattgctgtgatttatcattatcgaggtgaacaagctattcaagataatcagccagaaatttgtcaacttttatttgaaa
aagcagctgattattggaaagaagctattcgtttagctcctacaaactatattgaagctcaaaattggttaaaaatgactggtcgagaataa

>Chloromonas_clathrata_-_ycf4
atgaacaattctcttttatcgcaagagagctcttcaaaggtcggccctttagagccaaactcaaaacaaacagaactaaatcgtcgttattttattgtcggagaacgtagactgagcaattattggtgggcttc
tgtaatattgttaggtggatttggtttttattaacaggaatttcgtcttatttgaactataatattttagcaaatgcatttaaaatatttaacgttactagcaatttgactccagtaaacgaaaattccgttattgctttttt
tcctcaaggattattaatgtgtttttatggtagcttaggtgttttattaagtatatattggtggcttttaatttattgggatgttggtggtggttttaatgaatttaataaaaaagaaggctttatgagaatttttcgttgg
ggatatcctggaaaaaatcgtcgaattgatatcgtatatccattacaagatattgaagctattcgagttgaattaaaccaaacgcaaggtttattagcttctgaacaaacaatttttgttcgtttactttcaactcc
aaacaacaacggtgttactactcctctttctagtttcgcgccgatttcaaagcaaaaaggggttgctttgccagataggggcccctaccccaagcaaggggaagggaaagcgttaaaagaaaaacgtga
aatccctttaggtggaattggccaaccattaacgttgaaagaaatagaaaaacaagcgtctgaattagctaatttttacaagtagaacttgaaggtttataa

Appendix 23 zDOPE scores generated by Chimera v1.13 (Pettersen *et al*., 2004) for all the Rubisco models. Rows highlighted in yellow are models with the lowest zDOPE scores and used for final Rubisco modelling.

Chlamydomonas augustae

SSU	zDOPE	LSU1	zDOPE	LSU2	zDOPE	LSU3	zDOPE	LSU4	zDOPE
5.1	-0.01	**1.1**	-1.04	**2.1**	-1.13	**3.1**	-1.09	**4.1**	-1.08
5.2	0.01	**1.2**	-1.08	**2.2**	-1.13	**3.2**	-1.01	**4.2**	-1.12
5.3	0	**1.3**	-1.08	**2.3**	-1.12	**3.3**	-1.17	**4.3**	-1.14
5.4	0.07	**1.4**	-1.10	**2.4**	-1.11	**3.4**	-1.09	**4.4**	-1.10
5.5	0.04	**1.5**	-1.14	**2.5**	-1.1	**3.5**	-1.10	**4.5**	-1.13
5.6	0.02	**1.6**	-1.09	**2.6**	-1.1	**3.6**	-1.10	**4.6**	-1.09
5.7	-0.11	**1.7**	-1.08	**2.7**	-1.11	**3.7**	-1.09	**4.7**	-1.09
5.8	-0.07	**1.8**	-1.08	**2.8**	-1.11	**3.8**	-1.10	**4.8**	-1.10
5.9	-0.02	**1.9**	-1.11	**2.9**	-1.14	**3.9**	-1.10	**4.9**	-1.16
5.10	-0.04	**1.10**	-1.08	**2.10**	-1.14	**3.10**	-1.11	**4.10**	-1.18

Chlamydomonas mutabilis

SSU	zDOPE	LSU1	zDOPE	LSU2	zDOPE	LSU3	zDOPE	LSU4	zDOPE
1.1	-0.31	**2.1**	-1.22	**5.1**	-1.23	**3.1**	-1.16	**4.1**	-1.17
1.2	-0.41	**2.2**	-0.12	**5.2**	-1.18	**3.2**	-1.17	**4.2**	-1.14
1.3	-0.50	**2.3**	-1.16	**5.3**	-1.15	**3.3**	-1.22	**4.3**	-1.24
1.4	-0.37	**2.4**	-1.23	**5.4**	-1.19	**3.4**	-1.16	**4.4**	-1.16
1.5	-0.30	**2.5**	-1.21	**5.5**	-1.19	**3.5**	-1.18	**4.5**	-1.17
1.6	-0.36	**2.6**	-1.19	**5.6**	-1.16	**3.6**	-1.22	**4.6**	-1.18
1.7	-0.36	**2.7**	-1.16	**5.7**	-1.2	**3.7**	-1.18	**4.7**	-1.15
1.8	-0.35	**2.8**	-1.17	**5.8**	-1.22	**3.8**	-1.16	**4.8**	-1.16
1.9	-0.34	**2.9**	-1.16	**5.9**	-1.18	**3.9**	-1.21	**4.9**	-1.18
1.10	-0.33	**2.10**	-1.17	**5.10**	-1.19	**3.10**	-1.14	**4.10**	-1.22

Chloromonas serbinowii

SSU	zDOPE	LSU1	zDOPE	LSU2	zDOPE	LSU3	zDOPE	LSU4	zDOPE
5.1	0.19	**4.1**	-0.94	**3.1**	0.11	**1.1**	0.21	**2.1**	0.16
5.2	0.06	**4.2**	-0.99	**3.2**	0.07	**1.2**	0.07	**2.2**	0
5.3	0.13	**4.3**	-0.95	**3.3**	0.08	**1.3**	0.11	**2.3**	0.14
5.4	0.02	**4.4**	-0.97	**3.4**	0.13	**1.4**	0.08	**2.4**	0.2
5.5	0.06	**4.5**	-0.97	**3.5**	0.13	**1.5**	0.01	**2.5**	0.13
5.6	0.11	**4.6**	-0.98	**3.6**	0.08	**1.6**	0.15	**2.6**	0.09
5.7	-0.05	**4.7**	-0.97	**3.7**	0.11	**1.7**	0.2	**2.7**	0.06
5.8	0.11	**4.8**	-0.97	**3.8**	0.13	**1.8**	0	**2.8**	0.15
5.9	0.07	**4.9**	-0.94	**3.9**	0.08	**1.9**	1.6	**2.9**	0.15
5.10	0.05	**4.10**	-0.94	**3.10**	0.11	**1.10**	0.16	**2.10**	0.2

Chloromonas clathrata

SSU	zDOPE	LSU1	zDOPE	LSU2	zDOPE	LSU3	zDOPE	LSU4	zDOPE
5.1	-0.06	**4.1**	-0.92	**3.1**	0.14	**1.1**	0.11	**2.1**	0.14
5.2	0.08	**4.2**	-0.94	**3.2**	0.09	**1.2**	0.14	**2.2**	0.1
5.3	-0.16	**4.3**	-0.96	**3.3**	0.1	**1.3**	0.11	**2.3**	0.14
5.4	0	**4.4**	-0.98	**3.4**	0.18	**1.4**	0.02	**2.4**	0.16
5.5	-0.13	**4.5**	-0.96	**3.5**	0.13	**1.5**	0.1	**2.5**	0.15
5.6	-0.12	**4.6**	-0.9	**3.6**	0.15	**1.6**	0.22	**2.6**	0.19
5.7	0	**4.7**	-0.86	**3.7**	0.23	**1.7**	0.17	**2.7**	0.14
5.8	-0.05	**4.8**	-0.96	**3.8**	0.12	**1.8**	0.23	**2.8**	0.22
5.9	0	**4.9**	-0.91	**3.9**	0.16	**1.9**	0.18	**2.9**	0.17
5.10	0.02	**4.10**	-0.93	**3.10**	0.14	**1.10**	0.26	**2.10**	0.1

Appendix 24 List of all the residues at the interface SSU/LSUs. Amino acids in red are those located either in the two α-helices or in the βA-βB loop.

Chlamydomonas augustae

		Interactions SSU with LSU1		Interactions SSU with LSU2		Interactions SSU with LSU3		Interactions SSU with LSU4
Residues LSU 1 to 4 interacting on SSU		Residues interacting on LSU1		Residues interacting with LSU2		Residues interacting on LSU3		No Residues interacting with SSU
W	4	G		G	173	H	147	
P	6	A		L	174	I	149	
V	7	G		S	175	Q	150	
N	8	F		K	177	R	153	
N	9	W		N	178	D	154	
K	10	G		G	180	K	155	
M	11	L		R	181	L	156	
F	12	T		Y	184	N	157	
E	13	S		E	185	K	158	
T	14	R		R	188	Y	159	
F	15	Y		F	205	G	160	
S	16			R	209	R	161	
Y	17			I	213	R	188	
L	18			F	214	G	189	
P	19			A	216	G	190	
L	21			E	217	Y	220	
Q	25			A	218	Q	223	
A	28			Y	220	A	224	
Q	29			K	221	E	225	
Y	32			Q	223	T	226	
W	38			A	224	G	227	
E	43			E	225	E	228	
A	47			V	250	I	229	
K	49			D	253	K	252	
Y	51			L	254	D	253	
V	52			G	255	G	255	
N	54			H	403	V	256	
S	56			P	404	P	257	
R	59			W	405	N	281	
F	60			G	406	G	282	
G	61			N	407	L	283	
S	62					L	284	
S	62					R	344	
V	63					E	346	
V	63					D	390	
S	64					D	391	
C	65					A	412	
C	65					R	415	
L	66					V	416	
L	66					L	418	
Y	67					E	419	
Y	68					A	420	
Y	68					T	422	
D	69					Q	423	
N	70					R	425	

R	71
R	71
Y	72
Y	72
W	73
T	74
M	75
M	75
K	77
K	77
L	78
L	78
P	79
F	81
G	82
C	83
R	84
D	85
A	86
V	89
R	106
F	110
F	110
D	111
N	112
N	112
Q	113
K	114
Q	115
Q	115
Q	115
V	116
Q	117
Q	117
I	118
M	119
G	120

N	426
E	427
W	445
E	448

Chlamydomonas mutabilis

		Interactions SSU with LSU1		Interactions SSU with LSU2		Interactions SSU with LSU3		Interactions SSU with LSU4
Residues LSU 1 to 4 interacting on SSU		**Residues interacting on LSU1**		**Residues interacting on LSU2**		**Residues interacting on LSU3**		**No Residues interacting with SSU**
M	1	T		G	173	Q	150	
M	2	R		S	175	R	153	
W	4	G		K	177	D	154	
P	6	A		N	178	K	155	
V	7	G		G	180	L	156	
N	8	F		R	181	N	157	
N	9	W		Y	184	K	158	
K	10	G		E	185	Y	159	
M	11	L		R	188	G	160	
E	13	T		F	205	R	161	
T	14	S		V	213	R	188	
F	15	R		F	214	G	189	
S	16	Y		A	216	G	190	
Y	17			E	217	Y	220	
L	18			A	218	Q	223	
P	19			Y	220	A	224	
L	21			K	221	E	225	
Q	25			A	224	T	226	
R	28			E	225	G	227	
Q	29			V	250	E	228	
Y	32			K	252	V	229	
V	34			S	253	K	252	
N	36			L	254	G	255	
G	37			G	255	V	256	
W	38			H	403	P	257	
W	38			P	404	N	281	
I	39			W	405	G	282	
P	40			G	406	L	283	
E	43			N	407	L	284	
K	49					D	346	
Y	51					D	390	
V	52					D	391	
S	56					W	405	
R	59					A	408	
F	60					P	409	
G	61					A	412	
S	62					R	415	
S	62					V	416	
V	63					E	419	
C	65					A	420	
C	65					T	422	
L	66					Q	423	
L	66					R	425	
Y	67					N	426	
Y	68					E	427	
Y	68					G	428	
D	69					W	445	
N	70					S	446	
R	71					P	447	
R	71					E	448	
Y	72					A	451	

W	73
T	74
M	75
M	75
K	77
L	78
L	78
P	79
F	81
G	82
C	83
R	84
F	110
D	111
N	112
Q	113
Q	113
K	114
Q	115
Q	115
Q	115
V	116
Q	117
I	118
M	119
G	120

Chloromonas serbinowii

Interactions SSU with LSU1		Interactions SSU with LSU2		Interactions SSU with LSU3		Interactions SSU with LSU4
Residues LSU 1 to 4 interacting on SSU		Residues interacting on LSU1		Residues interacting on LSU2		Residues interacting on LSU3		No residues interacting
P	6	K	4	G	174	I	150	
N	8	K	5	L	175	E	153	
N	9	R	6	R	178	R	154	
K	10	W	65	Y	180	E	155	
M	11	R	68	V	184	R	156	
E	13	L	69	Y	185	L	157	
T	14	T	70	L	188	D	158	
F	15	A	71	P	205	K	159	
S	16	M	74	F	213	F	160	
Y	17	Y	75	L	214	G	161	
L	18			V	216	G	187	
P	19			M	217	L	188	
L	21			D	218	K	189	
Q	25			V	220	G	190	
A	28			N	221	L	192	
Y	32			S	224	A	223	
W	38			A	225	S	224	
I	39			E	250	A	225	
E	43			K	253	A	226	
K	49			E	254	T	227	
Y	51			L	255	G	228	
V	52			H	403	E	229	
S	56			P	404	V	230	
A	57			Q	405	A	252	
R	59			G	406	K	253	
F	60			I	407	L	255	
G	61			Q	408	G	256	
S	62					S	257	
S	62					N	281	
V	63					D	282	
V	63					M	283	
S	64					V	284	
C	65					R	344	
C	65					K	366	
L	66					K	367	
L	66					D	390	
Y	67					D	391	
Y	68					T	412	
Y	68					R	415	
D	69					V	416	
N	70					L	418	
N	70					E	419	
R	71					A	420	
R	71					V	422	
Y	72					L	423	
Y	72					R	425	
W	73					N	426	
T	74					E	427	
M	75					G	428	
M	75					S	445	
K	77							

L	78
L	78
P	79
M	80
F	81
G	82
C	83
R	84
R	106
V	108
F	110
F	110
D	111
D	111
N	112
N	112
Q	113
K	114
Q	115
Q	115
Q	115
V	116
V	116
Q	117
Q	117
I	118
M	119
G	120
F	121

Chloromonas clathrata

	Interactions SSU with LSU1	Interactions SSU with LSU2	Interactions SSU with LSU3	Interactions SSU with LSU4
Residues LSU 1 to 4 interacting on SSU	**Residues interacting on LSU1**	**Residues interacting on LSU2**	**Residues interacting on LSU3**	**No residues interacting**
W 4	K 5	L 173	I 150	
N 9	R 6	L 175	E 153	
K 10	Y 7	G 177	R 154	
M 11	W 65	R 178	E 155	
F 12	R 68	Y 180	R 156	
E 13	L 69	G 181	L 157	
T 14	T 70	R 182	D 158	
F 15	A 71	V 184	K 159	
S 16	M 74	Y 185	F 160	
Y 17	Y 75	L 188	G 161	
L 18		P 205	L 188	
P 19		F 213	K 189	
L 21		L 214	G 190	
Q 25		V 216	A 223	
S 27		M 217	S 224	
A 28		S 224	A 225	
Q 29		A 225	A 226	
K 31		E 250	T 227	
Y 32		K 253	G 228	
I 39		S 254	E 229	
Y 51		L 255	V 230	
V 52		H 403	A 252	
E 55		P 404	K 253	
S 56		Q 405	L 255	
R 59		G 406	G 256	
F 60		I 407	S 257	
G 61		Q 408	N 281	
G 61			D 282	
S 62			M 283	
S 62			I 284	
V 63			K 366	
V 63			D 390	
S 64			T 412	
C 65			R 415	
C 65			V 416	
L 66			E 419	
L 66			A 420	
Y 67			V 422	
Y 67			L 423	
Y 68			R 425	
Y 68			N 426	
D 69			E 427	
N 70			G 428	
N 70				
R 71				
R 71				
Y 72				
Y 72				
W 73				
T 74				

M	75
M	75
K	77
L	78
L	78
P	79
F	81
G	82
R	106
F	110
F	110
F	110
D	111
D	111
N	112
N	112
Q	113
K	114
K	114
Q	115
Q	115
Q	115
V	116
V	116
Q	117
Q	117
I	118
M	119
G	120
F	121

Appendix 25 Presence/Absence of the 88 essential genes tested with BLAST (Kent, 2002) on the 5 new sequenced strains.

Gene ID	Gene function	*Chlamydomonas augustae*	*Chlamydomonas mutabilis*	*Chloromonas serbinowii*	*Chloromonas rosae*	*Chloromonas clathrata*
		CCM + pyr+	CCM + pyr +	CCM+pyr-	CCM+ pyr-	CCM-pyr-
		28/88	64/88	31/88	57/88	51/88
Cre01.g014350.t1.2	PRX5 - Peroxiredoxin, type II					
Cre01.g027150.t1.1						
Cre01.g030900.t1.1	6.2.1.26 - o-succinylbenzoate--CoA ligase / OSB-CoA synthetase					
Cre01.g045902.t1.1						
Cre01.g051500.t1.2	Uncharacterized thykaloid lumenal polypeptide					
Cre01.g054850.t1.2						
Cre02.g073850.t1.2						
Cre02.g078507.t1.2	PF13326 - Photosystem II Psb27 (PSII_Psb27)					
Cre02.g097800.t1.1						
Cre02.g105650.t1.2						
Cre02.g111550.t1.1	Kinase. From pooled screens					
Cre02.g120100	RbcS1					
Cre02.g120150	RbcS2					
Cre02.g120250.t1.1	STT7- found in pyrenoid proteome and interacts with CAH3 - we would be interested in further characterizing this interaction					
Cre02.g143450.t1.2	PTHR36738:SF1 -expressed protein					
Cre03.g146167.t1.1	TEF10a - predicted protein					
Cre03.g151650.t1.1	SMM					
Cre03.g156600.t1.2	GluTRBP - Glutamyl-tRNA reductase binding protein					
Cre03.g162800.t1.2						
Cre03.g179800.t1.2	LCI 24					
Cre03.g183850.t1.2	FDX6 - Ferredoxin					
Cre03.g185550.t1.2						
Cre03.g188700.t1.2						

Cre10.g439350.t1.2	PTHR17130:SF24 - GAN									
Cre10.g440000.t1.1										
Cre10.g440050.t1.2	CSP41- in Lemaire proteome, binds Rubisco, has two motifs - could fit well with Moritz' story									
Cre10.g444700.t1.1	SBE3 - Starch branching enzyme									
Cre10.g452800.t1.2	LCI B									
Cre11.g467712.t1.1	SAGA 1									
Cre12.g484200.t1.2	GGPS1									
Cre12.g485050.t1.2										
Cre12.g494850.t1.2	ADK3 - Adenylate kinase 3									
Cre12.g497300.t2.1										
Cre12.g507300.t1.2										
Cre12.g509050.t1.1	PSBP3									
Cre12.g519300.t1.2	TEF9 - predicted protein									
Cre12.g524300.t1.2										
Cre12.g524500.t1.2										
Cre12.g531050.t1.1										
Cre12.g560950.t1.2	PSAG									
Cre13.g574000.t1.1	Putative voltage-gated bicarbonate transporter from screens									
Cre13.g577100.t1.2	ACP2 - Acyl carrier protein									
Cre13.g578650.t1.1										
Cre13.g581850.t1.2	Kinase. Identified in Leif's screen and Frieder pooled screen									
Cre14.g616600.t1.2	FZL; mutant has mislocalized pyrenoid; being characterized by Moritz									
Cre14.g626700.t1.2	Fd/FDX1 - Ferrodoxin									
Cre16.g651050.t1.2	CYC6 - cytochrome c6									
Cre16.g652800.t1.2										
Cre16.g658400.t1.2	FDX2 - Ferrodoxine									
Cre16.g659050.t1.1										
Cre16.g662150.t1.2	CCB1/CPLD51 cytochrome b6f complex assembly									

Cre03.g189800.t1.2	CYN38 - Peptidyl-prolyl cis-trans isomerase, cyclophilin-type					
Cre03.g191250.t1.2	LCI 34					
Cre04.g223050.t1.2						
Cre04.g223300.t1	CCP1- binds weakly to Rubisco, found in Zhan/Lemaire proteome					
Cre04.g229300.t1.1	RCA1 - Rubisco activase 1					
Cre05.g248450.t1.2						
Cre06.g259100.t1.1	SAGA like 2					
Cre06.g259900.t1.2	ATPc - ATP synthase gamma chain, chloroplastic					
Cre06.g261750	RBMP1					
Cre06.g273700.t1.2						
Cre06.g283750.t1.2						
Cre06.g295450.t1.2						
Cre06.g298300.t1.1						
Cre06.g307500.t1.1	LCI C					
Cre06.g309000.t1.2						
Cre07.g330250.t1.2	PSAH - Subunit H of photosystem I					
Cre08.g362900.t1.1	PSBP4 - Lumenal PsbP-like					
Cre08.g372450.t1.2	PSBQ - Oxygen evolving enhancer protein 3					
Cre09.g389615.t1.1						
Cre09.g394150.t1.1						
Cre09.g394473.t1.1	LCI 9 - Low Co2 inducible					
Cre09.g394621.t1.1	SAGA like 1					
Cre09.g396950.t1.1	Candidate Na+/HCO3- transporter from screens					
Cre09.g415700.t1.2	CAH3 - Carbonic anhydrase					
Cre09.g416800.t1.2						
Cre09.g416850.t1.2	Kinase? Rubisco physical interactor RBMP2					
Cre10.g423500.t1.2						
Cre10.g430150.t1.2						
Cre10.g436550	EPYC1 /LCI 5					
Cre16.g662600.t1.2	BST1					
Cre16.g663450.t1.2	LCI 11					
Cre17.g721500.t1.2	STA2 - Starch synthase, chloroplastic/amyloplastic					
Cre17.g724300.t1.2	PSAk - Photosystem I reaction center subunit psaK					
Cre17.g725500.t1.2						
Cre17.g740950.t1.2	LHL4 - High intensity light-inducible lhc-like gene					
BST4						
rbcL						

Appendix 26 Paper accepted in New Phytologist

MRS MYRIAM MADELEINE MARTHE GOUDET (Orcid ID : 0000-0001-9522-6659)
DR DOUGLAS J ORR (Orcid ID : 0000-0003-1217-537X)

Article type : Regular Manuscript

Rubisco and carbon concentrating mechanism (CCM) co-evolution across Chlorophyte and Streptophyte green algae

Myriam M. M. Goudet[1], Douglas J. Orr[2], Michael Melkonian[3], Karin H. Müller[4], Moritz T. Meyer[5], Elizabete Carmo-Silva[2] and Howard Griffiths[1]

[1]Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK; [2]Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK; [3]Institute for Plant Sciences, Department of Biological Sciences, University of Cologne, 50674 Cologne, and Central Collection of Algal Cultures, Faculty of Biology, University of Duisburg-Essen, 45141 Essen, Germany; [4]Cambridge Advanced Imaging Centre, University of Cambridge, Cambridge, CB2 3DY, UK; [5]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA

ORCID ID: MMM Goudet (0000-0001-9522-6659); DJ Orr (0000-0003-1217-537X); KH Müller (0000-0003-4693-8558); MT Meyer (0000-0001-8516-2591); E Carmo-Silva (0000-0001-6059-9359) and H Griffiths (0000-0002-3009-6563).

Author for correspondence:
Myriam M. M. Goudet
Tel: +44 (0)1223 330218
Email: mmmg2@cam.ac.uk

And

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/NPH.16577

Prof. Howard Griffiths
Tel: +44 (0)1223 333946
Email: hg230@cam.ac.uk

Received: *19 November 2019*
Accepted: *23 March 2020*

Summary

- Green algae expressing a carbon concentrating mechanism (CCM) are usually associated with a Rubisco-containing micro-compartment, the pyrenoid. A link between the small subunit (SSU) of Rubisco and pyrenoid formation in *Chlamydomonas reinhardtii* has previously suggested that specific *Rbc*S residues could explain pyrenoid occurrence in green algae.
- A phylogeny of *Rbc*S was used to compare the protein sequence and CCM distribution across the green algae and positive selection in *Rbc*S was estimated. For six streptophyte algae, Rubisco catalytic properties, affinity for CO_2 uptake ($K_{0.5}$), carbon isotope discrimination ($\delta^{13}C$) and pyrenoid morphology were compared.
- The length of the βA-βB loop in *Rbc*S provided a phylogenetic marker discriminating chlorophyte from streptophyte green algae. Rubisco kinetic properties in streptophyte algae have responded to the extent of inducible CCM activity, as indicated by changes in inorganic carbon uptake affinity, $\delta^{13}C$ and pyrenoid ultrastructure between high and low CO_2 conditions for growth.
- We conclude that the Rubisco catalytic properties found in streptophyte algae have co-evolved and reflect the strength of any CCM or degree of pyrenoid leakiness, and limitations to inorganic carbon in the aquatic habitat, whereas Rubisco in extant land plants reflects more recent selective pressures associated with improved diffusive supply the terrestrial environment.

Key words: carbon concentrating mechanism (CCM), green algae, photosynthesis, pyrenoid, Rubisco, streptophyte algae,

Introduction

Photoautotrophic organisms globally fix 111-117x10^{15} grams of carbon per year and around half of this global net primary production is aquatic (Behrenfeld *et al*., 2001; Field *et al*., 1998), with green algae a major contributor to this global carbon fixation. Green algae are classified into two major groups: chlorophytes and streptophytes, the latter demonstrating a wide range of ultrastructural and developmental traits closely related to land plants. Despite the existence of terrestrial green algae (Warren *et al.*, 2019), both groups remain subject to key limitations in the aquatic milieu (low CO_2 diffusion and availability, light limitation; Borges & Frankignoulle, 2002; Yamano *et al*., 2015).

Green algal inter-relationships have been resolved through numerous molecular phylogenies, including the chloroplast gene (*rbcL*) encoding the large subunit (LSU) of the primary carboxylase Rubisco (ribulose 1,5-bisphosphate carboxylase/oxygenase). An early split after the primary endosymbiosis saw the diversification of the hypothetical ancestral flagellate into two main lineages (Leliaert *et al*., 2011; 2012). First, the chlorophytes, which diversified early as prasinophytes in marine waters, which then gave rise to the core chlorophytes (chlorophytes without prasinophytes, Fig. S1, Supporting Information) in fresh or marine waters. Second, the streptophyte algae, which diversified in fresh water and some subaerial/terrestrial habitats (Harholt *et al*., 2016). The split between chlorophyte and streptophyte probably occurred during the Neoproterozoic (between 1,000 and 541 million years ago; Becker, 2013; Del Cortona *et al*., 2020). Extant photosynthetic chlorophyte and streptophyte algae (as well as non-algal streptophytes, i.e. land plants) have a form 1B Rubisco. Selection pressures on the Rubisco catalytic properties are driven by the availability and diffusive supply of inorganic carbon, the CO_2:O_2 ratio and the development of any carbon concentrating mechanism (CCM) which improves the operating efficiency of Rubisco in many aquatic photosynthetic microorganisms (Tortell, 2000; Young *et al*., 2012; Meyer & Griffiths, 2013; Griffiths *et al*., 2017; Rickaby & Hubbard, 2019). The origins of the algal CCM could be related to equimolar CO_2:O_2 concentrations in surface waters around 500 million years ago (Griffiths *et al*., 2017).

The challenge for inorganic carbon delivery within aquatic environments is that bicarbonate (HCO_3^-) or carbonate (CO_3^{2-}) are often much more prevalent, and under current conditions, the concentration of CO_2 is often ~2,000 times lower in water than in air, and diffusion is 8,000 times slower (Raven *et al*., 1985; Falkowski & Raven, 2007; Young *et al*., 2012). A CCM is typically

associated with active transport of bicarbonate across membranes, and catalytic conversion to CO_2 within a chloroplast microcompartment, the pyrenoid (Meyer *et al*., 2017). Although the presence of a pyrenoid is a robust marker of the presence of a CCM, not all the eukaryotic algae with a CCM have a pyrenoid (Morita *et al.,* 1999; Raven *et al.*, 2005).

The CCM has been particularly well-defined in the model unicellular chlorophyte *Chlamydomonas reinhardtii*, where the pyrenoid is present with a clearly defined starch sheath, and the associated inner Rubisco matrix transversed by knotted thylakoid tubules, thought to be involved in the delivery of CO_2 within the matrix (Meyer & Griffiths, 2013; Engel *et al*., 2015; Mackinder *et al*., 2017; Meyer *et al*., 2017; Mukherjee *et al*., 2019). The CCM is inducible following transfer from elevated to ambient CO_2, and a key linker protein (EPYC1) has been associated with the recruitment of Rubisco to the pyrenoid (Mackinder *et al*., 2016; Freeman-Rosensweig *et al*., 2017). This recruitment ultimately involves interactions with the Rubisco Small Subunit (SSU) (Wunder *et al*., 2018; Atkinson *et al*., 2019), presumably at the level of surface exposed α-helices (Meyer *et al*., 2012). However, there has been little systematic analysis of the extent to which some form of carbon accumulation mechanism occurs across this chlorophyte clade, or comparative physiological and molecular studies on CCM characteristics or Rubisco kinetic properties, and whether these traits are captured across chlorophyte, prasinophyte and streptophyte algal lineages in *Rbc*S.

Chlamydomonas reinhardtii has also been used as a model organism to explore the interactions between Rubisco LSU, SSU and catalytic properties. The eight identical 55-kDa LSUs assemble as four dimers, while two sets of four 15-kDa SSUs, top and tail the Rubisco holoenzyme. A central 'solvent channel' runs through Rubisco and the width of its aperture is dependent on the length of the βA-βB loop in each set of four SSUs capping the LSU octamer (Spreitzer, 2003) and interacting residues between LSUs and SSUs affect Rubisco operating efficiency and catalytic properties (Spreitzer *et al*., 2005). Natural variation in Rubisco catalytic properties exists among photosynthetic organisms (Jordan & Ogren, 1981), however, a shift in the catalytic parameters towards higher turnover rate per active site (k_{cat}) and higher affinity for CO_2 (K_c) has been observed from cyanobacteria, chlorophyte to land plants (reviewed in Badger *et al*., 1998; Meyer & Griffiths, 2013). However, it has also been suggested that selective pressures on the Rubisco kinetic parameters V_C and K_c could have been relaxed due to the saturating CO_2 environment

provided by a CCM over evolutionary time (Tortell, 2000; Young *et al*., 2012; Meyer & Griffiths, 2013).

The overall aim of the presents study was to address the possible interactions between Rubisco SSU structure and phylogeny, and occurrence of any reported CCM or pyrenoid across the green algae. Additionally, we set out to define key Rubisco catalytic properties for a range of streptophyte algae representing the main streptophyte lineages (Fig. S1), as compared to *C. reinhardtii*. The few Rubisco kinetic measurements available for green algae were performed on chlorophytes (*Coccomyxa* sp., Palmqvist *et al*., 1995; *Scenedesmus obliquus*, Jordan & Ogren, 1981, Badger *et al*., 1998), not streptophyte algae. Surprisingly, there is yet no streptophyte model alga, despite the previous interest in using species with giant cells to characterise carbon uptake mechanisms (Lucas & Berry, 1985) or the recently published genome of *Chara braunii* (Nishiyama *et al.,* 2018).

Specifically, this study sought to *(i)* develop a phylogeny for *Rbc*S sequences in green algae as compared to consensus phylogenies (e.g. Leliart *et al*., 2012; Leebens-Mack *et al.,* 2019), and compare the distribution of pyrenoid and CCM across the algal clades; *(ii)* to identify whether any selection pressure on residues within the SSU were associated with the broader phylogeny or CCM activity and, *(iii)* to determine whether the catalytic properties of Rubisco across contrasting streptophyte algal groups reflected the overall phylogeny or specific activity of a CCM at the whole organism level. Our results reveal that a change in Rubisco SSU secondary structure (namely the βA-βB loop) is a distinctive trait of the division between core chlorophytes and streptophyte algae. We also demonstrate that Rubisco catalytic properties have co-evolved in association with the extent of CCM activity in streptophytes. Finally, this study provides additional insights for selection pressures driving the evolution of green algae and photosynthetic processes, particularly for Rubisco during the transition to terrestrial plant life forms.

Materials and Methods

Collection of protein sequences, phylogenetic analysis, βA-βB loop length and pyrenoid presence/absence mapping

2,674 protein *Rbc*S sequences of green algae were kindly provided by «The 1000 plants project» (1KP; Leebens-Mack *et al*., 2019; Carpenter *et al*., 2019). All the protein sequences were

manually and individually screened. Sequences showing cross-contamination (Carpenter *et al*., 2019), or which were too short or incomplete, were removed. The dataset did now allow to unambiguously identify *Rbc*S isoforms. Although it is generally taken that all photosynthetic members of the Viridiplantae have multiple copies of the *RbcS* gene, conservatively only one sequence was used in the analysis for each species, except when the data was sourced from independently sequenced genomes (e.g. for *Asteromonas*). A total of 187 protein sequences belonging to 113 species (31 streptophyte algae, 10 prasinophytes, 72 chlorophytes) were then aligned with Clustal Omega (Sievers *et al*., 2011). ProTest v2.4 (Abascal *et al.*, 2005) was used to identify the best model of protein evolution. Bayesian phylogenetic analyses were performed using BEAST v2.3.1 (Bouckaert *et al*., 2014) with a LG model of protein evolution (Le & Gascuel, 2008), a gamma distribution model with four categories, a relaxed molecular clock and finally with a Yule model of speciation. Three independent chains were run, each of length $8x10^7$ steps, parameters values and trees were sampled every $10x10^2$ steps. Chain convergences were checked using Tracer v1.6 (Drummond & Rambaut, 2007). Posterior parameters were summarized with Tree Annotator v1.8.2 (Drummond & Rambaut, 2007) using a maximum clade credibility tree (MCC) and a posterior limit of 0.5. Figtree v1.4.2 (Rambaut, 2007) was used for tree visualizations. The length of the βA-βB loop was determined after the analysis of the protein sequences, with the number of residues in the loop (Spreitzer, 2003) mapped on to the phylogeny of *Rbc*S. Finally, the same phylogeny was used to map the pyrenoid presence/absence. The scoring for pyrenoid presence/absence was based on the available literature (Table S1).

Likelihood ratio test for positive selection

To test the importance of two SSU α-helices for pyrenoid formation in *C. reinhardtii* (Meyer *et al*., 2012), the Codon-based package (codeml) implemented in PAML v4.9 (Yang, 2007) was used to detect residues under positive selection across the green algae lineage. In addition, the presence of a CCM is not universal across the green algae so the branch model also implemented in PAML was used to detect branches under positive selection. All the analyses were performed using "user tree" mode. The DNA phylogenetic tree was reconstructed using BEAST v2.3.1 with 135 cDNA *Rbc*S sequences of green algae from the 1KP, with a GTR model of protein evolution (Tavaré, 1986) and the same gamma distribution, molecular clock and model of speciation previously used. Three independent chains were run, each of length $5x10^7$ steps, parameters values and trees were sampled every $10x10^2$ steps. Chain convergences, posterior parameters and tree visualization were

analysed with the same method explained above. Several models of codon evolution that allow for variations in ω (d*N*/dS) among codons were tested (Site model) and evaluated using Likelihood Ratio Tests (LRTs) (Neyman & Pearson, 1928) as described in Kapralov & Filatov (2007). Branch models were used to test for positive selection across branches. The null model allowed for variations in ω among branches (0<d*N*/dS<1 and d*N*/dS=1 for both foreground and background branches) and also included two additional classes of codons with fixed d*N*/dS=1 on foreground branches but restricted as 0< d*N*/dS<1 and d*N*/dS=1 for background branches. The alternative model allowed 0<d*N*/dS<1 and d*N*/dS=1 for both foreground and background branches but also included two additional classes of codons under positive selection with d*N*/dS>1 on foreground branches with restriction as 0<d*N*/dS<1 and d*N*/dS=1 on background branches. Branches leading to species without pyrenoid were labelled as foreground branches (allows positive selection) and the rest of the branches were considered as background branches (with no positive selection). The level of significance was tested as described above.

Streptophyte algae culturing, Rubisco purification and Rubisco catalytic properties

Six streptophyte algae (Table S2-3; Fig. S1) were ordered from the Culture Collection of Algae at Göttingen. These consisted of: *Chlorokybus atmophyticus* (Chlorokybophyceae), *Klebsormidium subtile* (Klebsormidiophyceae), *Cosmarium subtumidum*, *Onychonema laeve*, *Spirogyra* sp. (Zygnematophyceae) and *Coleochaete scutata* (Coleochaetophyceae). The wild type *Chlamydomonas reinhardtii* (strain CC-4533, Li *et al.*, 2016) was used as control to test protocols since the Rubisco catalytic properties are well characterised (Jordan & Ogren, 1981; Genkov & Spreitzer, 2009). Strains were cultured in an incubator shaker (Innova 42, New Brunswick Scientific) under constant agitation (130 RPM) in the recommended medium (Table S2), in 2L conical flasks, under constant light at 20°C and bubbled with ambient air. Due to the low concentration of Rubisco in algae (Losh *et al*., 2013; Valegård *et al*., 2018) a minimum of 30g wet paste per sample was harvested in order to have enough material for the Rubisco extraction and purification.

Algal cells were broken using an Emulsiflex-C5 high pressure homogenizer (Avestin Inc., Ottawa, Canada) kindly loaned by Biopharma Group (Winchester, UK). Cell pastes were re-suspended in *ca.* 200 mL of extraction buffer containing 10 mM $MgCl_2$, 50 mM Bicine, 10 mM $NaHCO_3$, 1 mM DTT, 1 mM ε-aminocaproic acid, 1 mM benzamidine, 0.1 M phenylmethylsulfonyl fluoride,

and 200 μL of protease inhibitor cocktail (Sigma, UK). Total soluble proteins were extracted via centrifugation at 22,000 ***g*** for 12 minutes (min) at 4°C. After this initial centrifugation step, PEG 4000 (60% w/v) and 1 M $MgCl_2$ were added to the supernatant and the rest of the purification carried out as described previously (Orr & Carmo-Silva, 2018). Peak fractions containing Rubisco (based on CABP binding [Sharwood *et al*., 2016]) were concentrated using Amicon Ultracel-15 concentrators (100 kDa MWCO, Merck-Millipore, UK). Aliquots were snap-frozen in liquid nitrogen and stored at -80°C.

Rubisco activity for the six streptophyte algae was determined by incorporation of $H^{14}CO_3$ into acid-stable products at 25°C as described in Prins *et al*. (2016) with some modifications. Rubisco activity was measured at a higher temperature (25°C) than for growth in the natural environment, to allow comparison with the expression of standard Rubisco kinetic properties (Jordan & Ogren, 1984). Purified Rubisco was diluted using desalting buffer (Orr & Carmo-Silva, 2018) and then desalted using a G-25 MidiTrap column (GE Healthcare, UK). Samples were allowed to activate on ice for 45 mins prior to assaying. Carboxylation activity was measured at nine different concentrations of CO_2 (8, 16, 24, 36, 68, 100, 180, 280 and 400 μM) and with O_2 concentrations of 0 and 21% (250 μM). In order to ensure that the activity measured was entirely due to Rubisco, three controls were performed: CO_2 fixation (acid-stable ^{14}C) was measured in reaction solutions lacking RuBP or $NaHCO_3$, and following total inhibition of Rubisco by prior treatment with an excess of the tight-binding inhibitor 2-carboxyarabinitol-1,5-bisphosphate (CABP). Radioactive content of ^{14}C-labelled compounds was measured in 0.4 ml aqueous solutions to which were added 3.6 ml Gold Star Quanta Scintillation cocktail (Meridian Biotechnologies, UK), in a Tri-Carb 2250 CA Liquid Scintillation Analyser (Perkin-Elmer, USA). Turnover number (k_{cat}: mol product mol active site^{-1} s^{-1}) was calculated from the corresponding V_{max} value (V_c : μmol acid-stable ^{14}C mg Rubisco^{-1} min^{-1}).

Rubisco quantification was via [^{14}C]CABP binding assay as described Sharwood *et al*. (2016). Rubisco was incubated for 25 min after adding [^{14}C]CABP. Each quantification was performed in duplicate. Radioactive content of ^{14}C-labelled compounds was measured using scintillation counting as described above.

Photosynthetic affinity for inorganic carbon

Apparent affinity for inorganic carbon (Ci) was determined by oxygen evolution (Badger *et al*., 1980) and as described in Mitchell *et al*. (2014). Five extra concentrations were added in cultures grown in high CO_2 condition in order to reach maximum rate of oxygen evolution (2500, 3000, 4000, 4500 and 5000 μM). Chlorophyll *a* and *b* concentrations were measured for normalization of oxygen evolution measurements as described in Mitchell *et al*. (2014).

Carbon isotope analysis

Algae cultures were grown under low and high CO_2 conditions and were harvested by centrifugation at 3,234 ***g*** for 5 minutes at 20°C (Eppendorf, Centrifuge 5804 R), resuspended in 0.1M HCl to remove inorganic carbon and washed several times with deionized water. Samples were dried in a freeze drier overnight and weighed (0.5 mg) in triplicate into 3mm x 5mm tin capsules (Experimental Microanalysis Ltd., Okehampton, UK). The results were reported with reference to the international standard VPDB with a precision better than +/- 0.08 per mil for $^{12}C/^{13}C$. All the analyses were performed at the Godwin Laboratory for Paleoclimate Research at the University of Cambridge.

Pyrenoid morphologies

Pyrenoid morphologies were examined using blockface imaging by SEM. Sample preparation and imaging were undertaken at the Cambridge Advanced Imaging Centre (CAIC). Cells were cultured as explained above in liquid Tris-phosphate medium and bubbled under ambient air supply (0.04% CO_2). After centrifugation, they were then fixed and embedded as described in Chan (2018). Resin blocks were mounted on aluminium SEM stubs and sputter-coated with 35 nm gold. Blockfaces were obtained with an ultramicrotome (Leica, Wetzlar, Germany) and coated with 30 nm carbon. Finally, blockfaces were imaged using a FEI Verios 460 scanning electron microscope (Thermo Fisher Scientific), running at 4 keV accelerating voltage and 0.2 nA probe current. Images were obtained using the Through-lens detector in immersion and backscatter mode. Automated image acquisition was set up using FEI MAPS software using a pixel resolution of 1536 x 1024, a dwell time of 3 μs, a horizontal field width of 15.9 μm/tile (magnification 8000x), an x-y tile overlap of 15%/20% and the MAPS default stitching profile.

Results

The length of the βA-βB loop drives the phylogeny of *Rbc*S

A protein phylogeny of *Rbc*S was constructed to identify any residues specific to species with a pyrenoid as a determinant of CCM activity. Despite the low number of variable sites, attributable to the brevity of the sequence, *RbcS* recapitulated at the phylum level the green lineage phylogeny (e.g. Leliart *et al*., 2012; Leebens-Mack *et al.,* 2019; the present study: Fig. S2). However, the present study found that species without a pyrenoid were dispersed throughout the whole *RbcS* phylogeny. Therefore, specific residues in the SSU α-helices (Meyer *et al*., 2012) were not sufficient to explain the pyrenoid occurrence across the entire phylum (Fig. 1). A closer examination of the solvent-exposed residues (available for possible interactions with the Rubisco linker EPYC1) of the amino acids and their electrostatic properties in the two α-helices, hypothesised to be the key elements for the formation of a pyrenoid (Meyer *et al*., 2012; Mackinder *et al*., 2016), varied in their distribution (Fig. S3). For example, the two pyrenoid-less species *Spermatozopsis similis* and *Chloromonas oogama* exhibited α-helices identical to *C. reinhardtii* (pyrenoid-positive) (Fig. S3). The absence of any consistent pattern which could differentiate pyrenoid-less from pyrenoid-positive species suggests that the residues in the two α-helices are not sufficient to singlehandedly explain pyrenoid occurrence in green algae, as we had hypothesized.

However, the *Rbc*S phylogeny did systematically differentiate streptophyte algae and core chlorophytes, which were clustered separately into two sister clades (Fig. 1). Prasinophytes clustered with the core chlorophytes, except *Picocystis salinarum*. The phylogenetic differentiation in *Rbc*S clearly coincided with differences in the βA-βB loop length. Core chlorophytes and prasinophytes consistently showed a βA-βB loop length of 25 or more residues, whereas the vast majority of streptophyte algae exhibited a βA-βB loop length of less than 23 residues with 52 of the 58 sequences having a βA-βB loop 21 residues long. The short loop of *P. salinarum* (21 residues) matches that of *Picocystis* sp. (draft genome; Junkins *et al.* 2019). The nested position within streptophyte algae could be due to this singular property, although the overall short length of *RbcS* and low bootstrap values at internal branches were likely additional factors. The difference in loop length between core chlorophytes and streptophyte algae revealed different Rubisco structures between these two groups. With a wider central solvent channel due to the shorter βA-βB loop, streptophyte algae have a Rubisco structure more similar to that in land plants as embryophytes (Spreitzer, 2003).

***Rbc*S is not under positive selection**

As an additional test for residues under positive selection in *Rbc*S, in association with a CCM or at the level of the SSU α-helices, 135 DNA sequences from green algae were used (Fig. S4). One Likelihood Ratio Test (LRT) for d*N*/dS heterogeneity across codons (M0-M3) was successfully performed and was significant, indicating expected heterogeneity in selective pressure across *Rbc*S molecules (2ΔlnL =2312.99, *P*-value<0.0001, df=8) (Table 1). Two LRTs were also performed to test for the presence of codons under positive selection (M7-M8 and M8-M8a) and both comparisons rejected models with positive selection (Table 1). The model M7 (which allows for 10 site classes, each with a ω >1) was selected in favour of the model M8 (11 sites classes with one of which allows for ω>1) and was consequently not significant (2ΔlnL= -0.00049, *P*-value=0.5, df=2). The more stringent comparison between the model M8a (which is similar to M7 but which allows for an extra class of codons with dN/dS=1) and M8 was also not significant (2ΔlnL= -0.07013, *P*-value=0.5, df=1) confirming the absence of codons under positive selection in *Rbc*S. The absence of residues under positive selection suggests that the appearance of new residues would not confer selective advantages in *Rbc*S, and particularly at the level of the α-helices (consistent with observations arising from Fig. 1 and Fig. S3, described above).

Branches under positive selection were successfully tested with the branch-model implemented in PAML. The LRT for heterogeneity across branches (H0-H1) was significant (2ΔlnL=9.358, *P*-value=0.0011, df=1) (Table 2). However, background and foreground omega showed values less than 1, implying positive selection was absent among foreground branches (ωa=0.082; ωb=0.16 <1). These results suggest that the presence of variation in ω across branches in *Rbc*S*,* but not significant enough to show positive selection, or any correlation with pyrenoid occurrence.

Streptophyte algae share Rubisco catalytic properties with both chlorophytes and embryophytes

A more detailed investigation of Rubisco catalytic properties was undertaken in order to explore whether any evolutionary progression towards land plant characteristics was evident in streptophyte algae. The multiple alignment of *Rbc*S in six representative streptophyte algae selected for this component of the study confirmed the deletion of five amino-acids in this group

compared to *C. reinhardtii* (Fig. 2; Spreitzer, 2003). This shortens the loop between the first and the second β-sheets, reducing the constriction at the entry of the holoenzyme's solvent channel. Rubisco catalytic properties at 25°C for the six green algae are shown in Table 3, including *C. reinhardtii* as a control. In *C. reinhardtii*, Rubisco catalytic properties varied slightly from previous measurements (Satagopan & Spreitzer, 2008; Jordan & Ogren, 1981) but remained in the same range. Michaelis-Menten constant for carboxylation (K_c) showed similar values (39.6 and 34 μM) whereas the Rubisco turnover rate (k_{cat}) was somewhat higher in this study compared to the value found in Satagopan & Spreitzer (2008). The streptophyte algae did not show a clear systematic shift from chlorophyte towards land plant catalytic properties despite similar Rubisco SSU structural changes. Of the five streptophyte algae, only *Klebsormidium subtile* and *Onychonema laeve* showed a higher affinity for CO_2 (lower K_c values), closer to land plant values (e.g. *Arabidopsis thaliana*; 10.7 μM) with K_c of 18.7 and 27.3 μM respectively (Table 3). *Cosmarium subtumidum*, *Spirogyra* sp. and *Coleochaete scutata* had a relative low affinity for CO_2 with K_c values in the range of the core chlorophytes or slightly higher (45.3, 49.1 and 43.1 μM respectively).

The catalytic turnover rate (k_{cat}) showed a trend towards lower values. *Onychonema laeve* and *Cosmarium subtumidum*, both members of the Zygnematophyceae, had similar k_{cat} values (2.39 and 2.51 s^{-1} respectively). *Spirogyra* sp. appeared to be an exception with a high k_{cat} value compared to the other streptophyte algae (4.90 s^{-1}), similar to the land plant *A. thaliana* (4.1 s^{-1}, Atkinson *et al*., 2017). *Coleochaete scutata* showed the lowest k_{cat} of all the streptophyte algae (1.67 s^{-1}). Higher K_c is usually correlated to high k_{cat} and lower specificity factor (Badger, 1987; von Caemmerer & Quick, 2000; Tcherkez *et al*., 2006; Savir *et al*., 2010; Tcherkez, 2013). *Klebsormidium subtile* presented the highest value for carboxylation catalytic efficiency (k_{cat}/K_c^{air}) (0.14 s^{-1} μM^{-1}), and whilst this was the highest streptophyte algae value determined, remains well below that of land plants like *A. thaliana* (Atkinson *et al.*, 2017). The remaining streptophyte algae displayed lower efficiency, with *Coleochaete scutata* showing the lowest efficiency (0.032 s^{-1} μM^{-1}).

Streptophyte algae have a CCM, albeit leaky in some species

Oxygen evolution measurements, pyrenoid imaging and $\delta^{13}C$ were used to characterise CCM activity in the different streptophyte algae and to investigate whether CCM activity was associated

with Rubisco catalytic properties. Oxygen evolution was used to determine the whole cell affinity for inorganic carbon and therefore the extent of any inducible carbon concentrating mechanism. The photosynthetic $K_{0.5}$ (Ci) value (Table 4) of the wild-type *C. reinhardtii* under low CO_2 showed a strong affinity for Ci (54 μM Ci), similar to previous values in the literature (Mitchell *et al*., 2014; Wang *et al*., 2014). *Klebsormidium subtile*, *Spirogyra* sp. and *Coleochaete scutata* showed a whole cell affinity for Ci similar to *C. reinhardtii* with $K_{0.5}$ ranging from 45 to 53 μM Ci, consistent with a fully functional CCM, whereas *Chlorokybus atmophyticus*, *Cosmarium subtumidum* and *Onychonema laeve* exhibited a c.20% lower apparent affinity for CO_2 compared to the other species ($K_{0.5}$ 62, 64 and 62 μM Ci respectively), but still suggestive of CCM activity. Photosynthetic $K_{0.5}$ (Ci) values of all the species grown under high CO_2 confirmed the absence of CCM activity under such conditions (Table S4), and thereby the inducible character of the CCM in all species under examination.

Stable carbon isotope composition ($\delta^{13}C$) for organic matter was also used as a second proxy for CCM activity in the different species (Table 4). *Coleochaete scutata, Chlorokybus atmophyticus*, *Spirogyra* sp. and *Cosmarium subtumidum* appeared to be isotopically enriched at -15.8 to -18.8‰ (Table 4), values close to *C. reinhardtii* (-18.9‰) and close to the upper range typically seen in C_4 terrestrial plants and consistent with a fully-functioning CCM (Raven *et al*., 1982). On the other hand, *Klebsormidium subtile* and *Onychonema laeve* were somewhat isotopically depleted compared to the other species, with values intermediate between typical C_3 and C_4 plants ($\delta^{13}C$ of -21.1 and -21.3‰ respectively; O'Leary, 1988) and consistent with a CCM phenotype prone to leakiness (retro-diffusion of CO_2: Meyer *et al*., 2008) or limited carbon accumulation capacity.

Taken together, these observations reveal that Rubisco catalytic properties correlate to some extend with the strength of CCM activity. Similarly to *C. reinhardtii*, the three streptophytes algae *Cosmarium subtumidum*, *Spirogyra* sp. and *Coleochaete scutata* revealed a fully functioning CCM (low whole-cell affinity, $K_{0.5}$, and low carbon isotope discrimination) but lower Rubisco catalytic affinity for inorganic carbon (high K_c values), whereas *Klebsormidium subtile* and *Onychonema laeve* have a less effective CCM but higher affinity for inorganic carbon in terms of Rubisco catalytic properties (low K_c values). Therefore, in the presence of a less-effective CCM, Rubisco catalytic properties for *Klebsormidium subtile* and *Onychonema laeve* show a systematic shift towards values more typically associated with land plants.

Finally, electronic microscopy was used to diagnose the presence/absence of a pyrenoid in the algal material used in the present study, as an additional diagnostic for an active biophysical CCM. The presence of a pyrenoid was confirmed for all the species except for *Coleochaete scutata* for which tissue embedding was unsuccessful. Presence and morphology of a pyrenoid in that species had been previously published (McKay *et al*., 1991). CCM activities were supported by the presence of a pyrenoid in all species (Fig. 3). *Cosmarium subtumidum* (Fig. 3b), *Onychonema laeve* (Fig. 3d)*, Coleochaete scutata* (Fig. 3f) and *Spirogyra* sp. (Fig. 3e) exhibited pyrenoid morphologies similar to *C. reinhardtii* (Fig. 3g) with a spheroidal electron dense matrix traversed by multiple tubules, and a single layered peripheral starch sheath. There were, however, differences in the fine structure (starch sheath thickness and continuity, density of thylakoid tubules network) that perhaps provide clues to the variability in $K_{0.5}$ and $\delta^{13}C$ measurements. *Klebsormidium subtile* lacked a peripheral starch sheath (Fig. 3a), although a starch sheath may occur in *Klebsormidium* dependent on growth stage or light intensity (M. Melkonian, unpublished observations). *Chlorokybus atmophyticus* had multiple layers of short starch plates surrounding the matrix (Fig. 3c). The network of cross-pyrenoidal tubules was regular and dense in *Cosmarium* and *Chlorokybus* (Figs 3 b, c).

Overall, the results show that Rubisco catalytic properties are CCM dependent. However, at this stage, it remains difficult to differentiate limitations in carbon uptake versus leakiness of CO_2 as the selective pressure operating on Rubisco, and more detailed physiological experiments are warranted to fully characterize these contrasting processes.

Discussion

Rubisco SSU residues do not systematically equate to a CCM.

There was no immediately apparent correlation between SSU amino-acid sequence and pyrenoid occurrence/inferred CCM activity across the newly-created phylogeny of *Rbc*S for green algae. Our expectation was based on (*i*) the observations that the *Rbc*S α-helices are important for pyrenoid formation in *C. reinhardtii* (Meyer *et al.,* 2012), as well as (*ii*) recent *in vitro* and *in vivo* experiments showing that the SSU is needed to interact with the *Chlamydomonas* Rubisco linker EPYC1 (Wunder *et al*., 2018; Atkinson *et al*., 2019).Whether streptophyte pyrenoids assemble with an EPYC1 analogue is currently unknown. Based on the primary sequence alone, there are no EPYC1 homologues outside the Chlamydomonadales, so it would seem that other Rubisco

aggregation mechanisms may occur in more distantly related lineages, perhaps through interactions with other elements of the SSU and/or the LSU, which is the *modus operandi* in some cyanobacterial carboxysomes (Long *et al*., 2011; Oltrogge *et al*., 2019; Wang *et al*., 2019). It would be interesting to determine whether the widespread occurrence of some form of pyrenoid across green algae was due to multiple independent origins of the algal CCM (Meyer *et al.,* 2017), as found in C_4 and CAM pathways (Sage *et al*., 2011). However, the absence of a pyrenoid does not always equate to lack of a CCM (Giordano *et al.*, 2005), particularly in *Chloromonas*, which is closely related to *Chlamydomonas* (Morita *et al.,* 1999; Nozaki *et al.,* 2002; Pröschold *et al*., 2001; Meyer *et al.,* 2017), and although the underlying mechanisms of carbon accumulation of such species remain unknown there is also a consistent relationship between carbon isotope composition and CCM activity in those closely related species (M.M.M.Goudet, unpublished observations).

Overall, alignments of the *Rbc*S α-helix residues did not discriminate between pyrenoid-positive and pyrenoid-negative species (Fig. 1; Fig. S3). The two *Chlamydomonas Rbc*S isoforms (Goldschmidt-Clermont & Rahire, 1986) show inverse patterns of gene expression across the day-night cycle (Zones *et al*., 2015). For the present study, it was not possible to establish the functionality of *Rbc*S paralogues in terms of CCM expression (See Materials & Methods). Therefore, determining the exact number of copies, and their sequence specificity, for each of the pyrenoidless species would provide additional confirmation for the absence of specific residues essential for pyrenoid formation in green algae. An extensive evaluation of positive selection also showed no significant shifts in *Rbc*S amino acid residues associated with the CCM across the phylogeny (Table 1) whereas 13 residues under positive selection have been detected in *Rbc*S in angiosperms (Yamada *et al*., 2019). The absence of positive selection along branches leading to a pyrenoid could be an artefact of the small number of species *lacking* a pyrenoid within the green algae (Fig. 1), or indeed those possessing some form of a CCM but lacking a pyrenoid (see above). A possible alternative explanation is that all green algae retained a pyrenoid-competent Rubisco SSU but that the absence of a pyrenoid is rather determined by the lack (ancestral or through secondary loss) of a Rubisco linker, of similar or different ancestry as the *C. reinhardtii* EPYC1 (Mackinder *et al.*, 2016). Here too, future comparative proteomic studies with pyrenoidless algal CCMs will help resolve this question.

Streptophyte algal Rubisco SSU structure is similar to land plants

The phylogeny of *Rbc*S revealed a Rubisco structure in streptophyte algae similar to that of embryophytes, with SSUs possessing a shorter βA-βB loop and therefore a central solvent channel with a similar open structure as that shown for embryophytes (Spreitzer, 2003). Although the shorter loop in land plants has been well described (Spreitzer, 2003) and was probably thought to be a consequence of the transition from the aquatic environment to land, the presence of a similar structure in the streptophyte algae has not been previously reported. The phylogeny of *Rbc*S showed that this loss of amino acids is more ancient, and probably occurred during the split between chlorophytes and streptophyte algae, which occurred somewhere between 736 Mya (Becker, 2013) and 1,000 Mya (early Neoproterozoic; Del Cortona *et al*., 2020). The Rubisco structural change was not an isolated event at this time. The split between chlorophytes and streptophytes coincides with the appearance of multiple new traits (Hori *et al*., 2014; Nishiyama *et al*., 2018) such as lateral flagella, a flagellar peroxidase and also a Gap A/B gene duplication (McCourt *et al*., 2004; Finet *et al*., 2010). Interestingly, the photorespiratory pathway, which would have to contend with CCM inefficiencies, has been shown to differ between chlorophytes and streptophyte algae. Chlorophytes use a mitochondrial glycolate dehydrogenase, which produces NADH and H^+ whereas streptophytes use a peroxisomal glycolate oxidase which produces H_2O_2 for the conversion of glycolate to glyoxylate (Stabenau & Winkler, 2005).

The role of the SSU and of the βA-βB loop in particular is not entirely understood but the central solvent channel may facilitate channelling of substrates and products to and from the active sites (Esquivel *et al*., 2013). Spreitzer (2001; 2002) demonstrated the importance of the loop for holoenzyme assembly and showed that direct mutagenesis within and near the βA-βB loop changed Rubisco catalytic properties. However, these studies did not investigate the relationship to presence or absence of the pyrenoid in green algae and CCM activity. Despite the change in Rubisco SSU structure between chlorophytes and streptophytes, and effect on solvent channel width, the present work showed that there was a continued need for CCMs across the entire phylogeny (Fig. 1), as reflected in the catalytic properties of the streptophyte algae.

Rubisco catalytic properties in green algae depend on CCM efficiency

The above observations led to the investigation of Rubisco catalytic properties within the streptophyte algae and their associated physiological CCM activity. Streptophyte algae are

difficult to investigate physiologically. Oxygen electrode measurements were also extremely challenging (Table 4).

Despite the clear structural change associated with the βA-βB loop length, Rubisco catalytic properties remained generally similar to chlorophytes (Table 3) without systematic shift towards values associated with land plants (Satagopan & Spreitzer, 2008; Kapralov *et al*., 2010; Atkinson *et al*., 2017). Of the six streptophyte algae, only *Klebsormidium subtile* and *Onychonema laeve* showed K_c values in this lower range. Direct mutagenesis has shown the importance of the SSU βA-βB loop in Rubisco catalytic properties (see paragraph above) but the data in the present study suggested that they were more influenced by the effectiveness of the CCM, consistent with systematic changes in carbon isotope composition ($\delta^{13}C$: Table 4). Carbon isotopes have been used to infer leakiness of CCMs found in algae and hornworts (Meyer *et al*., 2008). Although whole cell inorganic carbon (Ci) uptake affinity was broadly similar for all species under ambient growth conditions ($K_{0.5}$, Table 4), the weaker CCM activities (identified through more negative $\delta^{13}C$ values: Table 4) in *Klebsormidium subtile* and *Onychonema laeve*, were associated with the highest affinity of Rubisco for CO_2 (K_c, Table 3). The importance of the CCM in shaping the adaptation within Rubisco catalytic properties has been a long-standing hypothesis (Tortell, 2000; Young *et al*. 2012; Meyer *et al*., 2013, Galmés *et al*., 2014, 2016, 2019; Griffiths *et al*., 2017), consistent with the shifts seen in C_4 Rubisco (Jordan & Ogren, 1981; Sage, 2002; Kubien *et al*., 2008). Our results show that Rubisco catalytic properties for this representative range of streptophyte algae are adapted to the presence of the CCM.

A strong CCM (uptake and conversion of inorganic carbon) or reduced retrodiffusion (leakiness) is partly consistent with pyrenoid presence for these two species (with either a naked pyrenoid or simple starch sheath: Fig. 3a, d, respectively). In addition, *Klebsormidium subtile* has often been reported to be a cosmopolitan species, colonising a great variety of aquatic and terrestrial habitats (Table S3; Hoffmann, 1989; Rindi *et al.,* 2011; Mikhailyuk *et al*., 2015). The Rubisco catalytic properties found in *Klebsormidium subtile* would place this species as an intermediate between obligate aquatic green algae and land plants. The future study of real subaerial algae such as *Klebsormidium flaccidum* or *Mesotaenium endlicherianum* would allow a more complete understanding of the photosynthetic adaptation to life on land. In the absence of the liquid boundary layer impeding CO_2 diffusion on land which could affect Rubisco catalytic properties (Raven *et al.,* 1985; Sáez *et al*., 2017), the naked pyrenoid in *Klebsormidium subtile* would

account for the more land-plant-like Rubisco catalytic properties and a reliance on direct diffusive CO_2 supply.

The co-evolution of Rubisco and CCMs has been demonstrated in multiple non-green photosynthetic organisms (Badger *et al*., 1998). Diatoms and haptophytes, which possess Form 1D Rubisco, are known to carry most of the oceanic photosynthesis (Delwiche & Palmer, 1997; Yoon *et al*., 2002; Falkowski *et al*., 2004). In these groups, Rubisco affinity for CO_2 (K_c) exhibits larger variations, exceeding those of C_4 plant Rubisco suggesting a large diversity of CCM strengths (Young *et al*., 2016; Heureux *et al*., 2017). In addition, the CO_2:O_2 ratio around the active site led to the suggestion that pyrenoids could have an oxygen exclusion function (McKay & Gibbs, 1991; Griffiths *et al.*, 2017). In land plants, Rubisco catalytic properties have been shown to be linked to changes in the atmospheric CO_2:O_2 ratio over time as well as temperature, in addition to leaf architecture, morphology and conductance (Beerling *et al*., 2001; Franks & Beerling, 2009; Haworth *et al.,* 2011; Galmes *et al*., 2014; 2015; Sharwood *et al.,* 2016; Conesa *et al*., 2019). As the atmospheric CO_2:O_2 ratio decreased over time, Galmes *et al*. (2014) showed that land plants developed a Rubisco that was more efficient at carboxylation (higher k_{cat}/K_c ratio) with increased affinity for CO_2 (lower K_c) but slower carboxylation rate (k_{cat}). Alongside these changes in catalytic properties, the proportion of soluble protein present as Rubisco increased, counteracting somewhat the effect of the decrease in carboxylation rate (Galmes *et al.,* 2014). Furthermore, higher temperatures increase maximum carboxylase turnover rate (k_{cat}) of Rubisco and decrease CO_2 affinity (Bernacchi *et al*., 2001; Galmés *et al*., 2015, 2016).

In conclusion, this study has highlighted that Rubisco SSU structure effectively differentiates between streptophytes and core chlorophytes, with a transition occurring in the prasinophyte clade which contains mostly species with a long βA-βB loop. Otherwise, the *Rbc*S phylogeny recaptures the latest consensus green algal phylogenies built from many marker genes, including *rbcL* (Leebens-Mack *et al*., 2019). A more focussed study on Rubisco catalytic properties in streptophyte algae suggests that the activity of any CCM, which may have arisen because of limitations in bulk CO_2 delivery to Rubisco, has permitted the retention of a lower affinity (high K_c) Rubisco. We showed that the extent of adaptation which occurs should either cause CCM activity to be reduced, or indeed lost during the transition to land, as the reliance on gaseous diffusion to deliver CO_2 to Rubisco began to increase. Overall, the observations confirm the widespread occurrence of a CCM across the entire green algal lineage through the means of a

pyrenoid-based CCM to fuel carbon fixation by Rubisco. However, rather than being intransigent and slow, Rubisco catalytic properties adapt to local conditions of CO_2 availability. This is consistent with the changes seen in Rubisco from C_4 (Jordan & Ogren, 1981; Sage, 2002; Kubien *et al*., 2008) and CAM plants (Griffiths *et al*., 2008), which have been associated with operating within a CCM for the past 5-10 million years. Based on this study, the selective pressures driven by local conditions of photosynthetic CO_2 supply are more likely to explain the shifts in Rubisco catalytic properties during life on land, rather than any long term transition seen in land plants.

Acknowledgements

This work was supported by the Natural Environment Research Council (grant number NE/L002507/1 to HG) and resources associated with BBSRC-BB/M007693/1, BB/I024518/1 as part of the Combining Algal and Plant Photosynthesis (CAPP), supported by BBSRC and NSF. We are grateful for a Cambridge Trust Vice Chancellor's award and Lucy Cavendish College, Cambridge, for supporting the PhD scholarship of MMMG. DJO and ECS acknowledge support from the UK Biotechnology and Biological Sciences Research Council (BBSRC; grant number BB/I024488/1). We thank Lyn Carter at the Cambridge Advanced Imaging Centre (CAIC) for her help in the pyrenoid imaging. We also thank James Rolfe for the $\delta^{13}C$ measurements at the Goodwin Laboratory, Department of Earth Sciences, University of Cambridge, UK. Finally, we thank Juan-Carlos Villarreal, Université de Laval, Québec, Canada, for his help with the 1KP data.

Author Contributions

M.M.M.G, H.G and M.T.M planned the research. D.J.O, E.C-S and M.M.M.G designed and performed the experiments on Rubisco kinetics and D.J.O. analysed the data. M.M.M.G performed the phylogenetic analyses, positive selection and physiological data collection and analysis. K.H.M. performed SEM imaging. M.M. provided the *RbcS* sequences. M.M.M.G and H.G. interpreted the data and wrote the manuscript with assistance from all authors.

References

Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104-2105.

Atkinson N, Leitão N, Orr DJ, Meyer MT, Carmo-Silva E, Griffiths H, Smith AM, McCormick AJ. 2017. Rubisco small subunits from the unicellular green alga *Chlamydomonas* complement Rubisco-deficient mutants of Arabidopsis. *New Phytologist* **214**: 655-667.

Atkinson N, Velanis CN, Wunder T, Clarke DJ, Mueller-Cajar O, McCormick AJ. 2019. The pyrenoidal linker protein EPYC1 phase separates with hybrid *Arabidopsis–Chlamydomonas* Rubisco through interactions with the algal Rubisco small subunit. *Journal of experimental botany* **70**: 5271-5285.

Badger MR, Kaplan A, Berry JA. 1980. Internal inorganic carbon pool of *Chlamydomonas reinhardtii*: evidence for a carbon dioxide-concentrating mechanism. *Plant Physiology* **66**: 407-413.

Badger MR. 1987. The CO_2-concentrating mechanism in aquatic phototrophs. *Photosynthesis*. Academic press.

Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD. 1998. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2 concentrating mechanisms in algae. *Canadian Journal of Botany* **76**: 1052-1071.

Becker B. 2013. Snow ball earth and the split of Streptophyta and Chlorophyta. *Trends in Plant science* **18**: 180-183.

Beerling DJ, Osborne CP, Chaloner WG. 2001. Evolution of leaf-form in land plants linked to atmospheric CO_2 decline in the Late Palaeozoic era. *Nature* **410**: 352.

Behrenfeld MJ, Randerson JT, McClain CR, Feldman GC, Los SO, Tucker CJ, Falkowski PG, Field CB, Frouin R, Esaias WE *et al.* 2001. Biospheric primary production during an ENSO transition. *Science* **291**: 2594-2597.

Bernacchi CJ, Singsaas EL, Pimentel C, Portis Jr AR, Long SP. 2001. Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell & Environment* **24**: 253-259.

Borges AV, Frankignoulle M. 2002. Distribution and air-water exchange of carbon dioxide in the Scheldt plume off the Belgian coast. *Biogeochemistry* **59**: 41-67.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology* **10**: e1003537.

Carpenter EJ, Matasci N, Ayyampalayam S, Wu S, Sun J, Yu J, Jimenez-Vieira FB, Bowler C, Dorrel RG, Gitzendanner MA *et al.* 2019. Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). *GigaScience* **8**: 1-7.

Chan KX. 2018. Morphological and physiological studies of the carbon concentrating mechanism in *Chlamydomonas reinhardtii*. PhD thesis, University of Cambridge, UK.

Conesa MÀ, Muir CD, Molins A, Galmés J. 2019. Stomatal anatomy coordinates leaf size with Rubisco kinetics in the Balearic *Limonium*. *AoB PLANTS* **12**: plz050

Del Cortona A, Jackson CJ, Bucchini F, Van Bel M, D'hondt S, Škaloud P, Delwiche CF, Knoll AH, Raven JA, Verbruggen H *et al.* 2020. Neoproterozoic origin and multiple transitions to macroscopic growth in green seaweeds. *Proceedings of the National Academy of Sciences, USA* **117**: 2551-2559.

Delwiche CF, Palmer JD. 1997. The origin of plastids and their spread via secondary symbiosis. In: Bhattacharya D, eds. *Origins of algae and their plastids.* Springer, Vienna, 53-86

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology* **7**: 214.

Engel BD, Schaffer M, Cuellar LK, Villa E, Plitzko JM, Baumeister W. 2015. Native architecture of the *Chlamydomonas* chloroplast revealed by in situ cryo-electron tomography. *Elife* **4**: e04889.

Esquivel MG, Genkov T, Nogueira AS, Salvucci ME, Spreitzer RJ. 2013. Substitutions at the opening of the Rubisco central solvent channel affect holoenzyme stability and CO_2/O_2 specificity but not activation by Rubisco activase. *Photosynthesis research* **118**: 209-218.

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR. 2004. The evolution of modern eukaryotic phytoplankton. *Science* **305**: 354-360.

Falkowski PG, Raven JA. 2007. Photosynthesis and primary production in nature. In: Falkowski PG, Raven JA, 2nd ed. *Aquatic photosynthesis*. Princeton University Press, Princeton. 1-43

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**: 237-240.

Finet C, Timme RE, Delwiche CF, Marlétaz F. 2010. Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. *Current Biology* **20**: 2217-2222.

Franks PJ, Beerling DJ. 2009. Maximum leaf conductance driven by CO_2 effects on stomatal size and density over geologic time. *Proceedings of the National Academy of Sciences* **106**: 10343-10347.

Galmés J, Kapralov MV, Andralojc PJ, Conesa MÀ, Keys AJ, Parry MA, Flexas J. 2014. Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. *Plant, Cell & Environment* **37**: 1989-2001.

Galmés J, Kapralov MV, Copolovici LO, Hermida-Carrera C, Niinemets Ü. 2015. Temperature responses of the Rubisco maximum carboxylase activity across domains of life: phylogenetic signals, trade-offs, and importance for carbon gain. *Photosynthesis research* **123**: 183-201.

Galmés J, Hermida-Carrera C, Laanisto L, Niinemets Ü. 2016. A compendium of temperature responses of Rubisco kinetic traits: variability among and within photosynthetic groups and impacts on photosynthesis modeling. *Journal of Experimental Botany* **67**: 5067-5091.

Galmés J, Capó-Bauçà S, Niinemets Ü, Iñiguez C. 2019. Potential improvement of photosynthetic CO_2 assimilation in crops by exploiting the natural variation in the temperature response of Rubisco catalytic traits. *Current opinion in plant biology* **49**: 60-67.

Genkov T, Spreitzer RJ. 2009. Highly conserved small subunit residues influence RuBisCO large subunit catalysis. *Journal of Biological Chemistry* **284**: 30105-30112.

Giordano M, Beardall J, Raven JA. 2005. CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annual Review of Plant Biology* **56**: 99-131.

Goldschmidt-Clermont M, Rahire M. 1986. Sequence, evolution and differential expression of the two genes encoding variant small subunits of ribulose bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *Journal of Molecular Biology* **191**: 421-432.

Griffiths H, Robe WE, Girnus J, Maxwell K. 2008. Leaf succulence determines the interplay between carboxylase systems and light use during Crassulacean acid metabolism in Kalanchoë species. *Journal of Experimental Botany* **59**: 1851-1861.

Griffiths H, Meyer MT, Rickaby REM. 2017. Overcoming adversity through diversity: aquatic carbon concentrating mechanisms. 3689-3695.

Harholt J, Moestrup Ø, Ulvskov P. 2016. Why plants were terrestrial from the beginning. *Trends in Plant Science* **21**: 96-101.

Haworth M, Elliott-Kingston C, McElwain JC. 2011. Stomatal control as a driver of plant evolution. *Journal of Experimental Botany* **62**: 2419-2423.

Heureux AM, Young JN, Whitney SM, Eason-Hubbard MR, Lee RB, Sharwood RE, Rickaby RE. 2017. The role of Rubisco kinetics and pyrenoid morphology in shaping the CCM of haptophyte microalgae. *Journal of experimental botany* **68**: 3959-3969.

Hoffmann L. 1989. Algae of terrestrial habitats. *The botanical review* **55**: 77-105.

Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajiima N *et al.* 2014. *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nature communications* **5**: 3978.

Jordan DB, Ogren WL. 1981. Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase. *Nature* **291**: 513-515.

Jordan DB, Ogren WL. 1984. The CO_2/O_2 specificity of ribulose 1, 5-bisphosphate carboxylase/oxygenase. *Planta* **161**: 308-313.

Junkins EN, Stamps BW, Corsetti FA, Oremland RS, Spear JR, Stevenson BS. 2019. Draft Genome Sequence of *Picocystis sp*. Strain ML, Cultivated from Mono Lake, California. *Microbiol Resour Announc* **8**: e01353-18.

Kapralov MV, Filatov DA. 2007. Widespread positive selection in the photosynthetic Rubisco enzyme. *BMC Evolutionary Biology* **7**: 73.

Kapralov MV, Kubien DS, Andersson I, Filatov DA. 2010. Changes in Rubisco kinetics during the evolution of C_4 photosynthesis in *Flaveria* (Asteraceae) are associated with positive selection on genes encoding the enzyme. *Molecular Biology and Evolution* **28**: 1491-1503.

Kubien DS, Whitney SM, Moore PV, Jesson LK. 2008. The biochemistry of Rubisco in *Flaveria*. *Journal of Experimental Botany* **59**: 1767-1777.

Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Molecular biology and evolution* **25**: 1307-1320.

Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, Yue R, Robertson JM, Lefebvre PA, Fitz-Gibbon ST, Grossman AR, Jonikas MC. 2016. An indexed, mapped mutant library enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. *Plant Cell* **28**: 367-387.

Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, Grosse I, Li Z, Melkonian M, Mirarab S *et al.* 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679-685.

Leliaert F, Verbruggen H, Zechman FW. 2011. Into the deep: new discoveries at the base of the green plant phylogeny. *BioEssays* **33**: 683-692.

Leliaert F, Smith DR, Moreau H, Herron MD, Verbruggen H, Delwiche CF, De Clerck, O. 2012. Phylogeny and molecular evolution of the green algae. *Critical reviews in plant sciences* **31**: 1-46.

Long BM, Rae BD, Badger MR., Price GD. 2011. Over-expression of the β-carboxysomal CcmM protein in Synechococcus PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. *Photosynthesis research* **109**: 33-45.

Losh JL, Young JN, Morel FM. 2013. Rubisco is a small fraction of total protein in marine phytoplankton. *New Phytologist* **198**: 52-58.

Lucas WJ, Berry JA. 1985. Inorganic carbon transport in aquatic photosynthetic organisms. *Physiologia plantarum* **65**: 539-543.

Mackinder LC, Meyer MT, Mettler-Altmann T, Chen VK, Mitchell MC, Caspari O, Freeman Rosenzweig ES, Pallesen L, Reeves G, Itakura A *et al*. 2016. A repeat protein links RuBisCO to form the eukaryotic carbon- concentrating organelle. *Proceedings of the National Academy of Sciences, USA* **113**: 5958-5963.

Mackinder LC, Chen C, Leib RD, Patena W, Blum SR, Rodman M, Ramundo S, Adams CM, Jonikas MC. 2017. A Spatial Interactome Reveals the Protein Organization of the Algal CO_2-Concentrating Mechanism. *Cell* **171**: 133-147.

McCourt RM, Delwiche CF, Karol KG. 2004. Charophyte algae and land plant origins. *Trends in Ecology & Evolution* **19**: 661-666.

McKay RML, Gibbs SP. 1991. Composition and function of pyrenoids: cytochemical and immunocytochemical approaches. *Canadian Journal of Botany* **69**: 1040-1052.

McKay RML, Gibbs SP. & Vaughn KC. 1991. RuBisCo activase is present in the pyrenoid of green algae. *Protoplasma* **162**: 38-45.

Meyer M, Seibt U, Griffiths H. 2008. To concentrate or ventilate? Carbon acquisition, isotope discrimination and physiological ecology of early land plant life forms. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**: 2767-2778.

Meyer MT, Genkov T, Skepper JN, Jouhet J, Mitchell MC, Spreitzer RJ, Griffiths H. 2012. RuBisCO small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. *Proceedings of the National Academy of Sciences, USA* **109**: 19474-19479.

Meyer MT, Griffiths H. 2013. Origins and diversity of eukaryotic CO_2-concentrating mechanisms: lessons for the future. *Journal of experimental botany* **64**: 769-786.

Meyer MT, Whittaker C, Griffiths H. 2017. The algal pyrenoid: key unanswered questions. *Journal of experimental botany* **68**: 3739-3749.

Mikhailyuk T, Glaser K, Holzinger A, Karsten U. 2015. Biodiversity of Klebsormidium (Streptophyta) from alpine biological soil crusts (Alps, Tyrol, Austria, and Italy). *Journal of phycology* **51**: 750-767.

Mitchell MC, Meyer MT, Griffiths H. 2014. Dynamics of carbon-concentrating mechanism induction and protein relocalization during the dark-to-light transition in synchronized *Chlamydomonas reinhardtii*. *Plant Physiology* **166**: 1073-1082.

Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K, Nozaki H. 1999. Role of pyrenoids in the CO_2-concentrating mechanism: comparative morphology, physiology and molecular phylogenetic analysis of closely related strains of *Chlamydomonas* and *Chloromonas* (Volvocales). *Planta* **208**: 365-372.

Mukherjee A, Lau CS, Walker CE, Rai AK, Prejean CI, Yates G, Emrich-Mills T, Lemoine SG, Vinyard DJ, Mackinder LCM, Moroney JV. 2019. Thylakoid localized bestrophin-like proteins are essential for the CO_2 concentrating mechanism in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **116**: 16915-16920.

Neyman J, Pearson ES. 1928. On the use and interpretation of certain test criteria for purposes of statistical inference: Part II. *Biometrika* **20A**: 263-294.

Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK, Haas FB, Vanderstraeten L, Becker D, Lang D *et al.* 2018. The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell* **174**: 448-464.

Nozaki H, Onishi K, Morita E. 2002. Differences in pyrenoid morphology are correlated with differences in the rbcL genes of members of the *Chloromonas* lineage (Volvocales, Chlorophyceae). *Journal of Molecular Evolution* **55**: 414-430.

O'Leary MH. 1988. Carbon isotopes in photosynthesis. *Bioscience* **38**: 328-336.

Oltrogge LM, Chaijarasphong T, Chen AW, Bolin ER, Marqusee S, Savage DF. 2019. α-carboxysome formation is mediated by the multivalent and disordered protein CsoS2. *bioRxiv*. doi: 10.1101/708164.

Orr DJ, Carmo-Silva E. 2018. Extraction of Rubisco to determine catalytic constants. In: Covshoff S, ed. *Photosynthesis: methods and protocols, Methods in molecular biology, vol 1770*. Humana Press, New York, NY, 229-238

Palmqvist K, Sültemeyer D, Baldet P, Andrews TJ, Badger MR. 1995. Characterisation of inorganic carbon fluxes, carbonic anhydrase(s) and ribulose-1, 5-biphosphate carboxylase-oxygenase in the green unicellular alga *Coccomyxa*. *Planta* **197**: 352-361.

Prins A, Orr DJ, Andralojc PJ, Reynolds MP, Carmo-Silva E, Parry MA. 2016. Rubisco catalytic properties of wild and domesticated relatives provide scope for improving wheat photosynthesis. *Journal of Experimental Botany* **67**: 1827-1838.

Pröschold T, Marin B, Schlösser UG, Melkonian, M. 2001. Molecular phylogeny and taxonomic revision of *Chlamydomonas* (Chlorophyta). I. Emendation of *Chlamydomonas* Ehrenberg and *Chloromonas* Gobi, and description of *Oogamochlamys* gen. nov. and *Lobochlamys* gen. nov. *Protist* **152**: 265-300.

Rambaut A. 2007. FigTree, a graphical viewer of phylogenetic trees. URL *http://tree. bio. ed. ac. uk/software/figtree*.

Raven J, Beardall J, Griffiths H. 1982. Inorganic C-sources for *Lemanea, Cladophora* and *Ranunculus* in a fast-flowing stream: measurements of gas exchange and of carbon isotope ratio and their ecological implications. *Oecologia* **53**: 68-78.

Raven JA, Osborne BA, Johnston AM. 1985. Uptake of CO_2 by aquatic vegetation. *Plant, Cell & Environment* **8**: 417-425.

Raven JA, Ball LA, Beardall J, Giordano M, Maberly SC. 2005. Algae lacking carbon-concentrating mechanisms. *Canadian Journal of Botany* **83**: 879-890.

Rickaby REM, Hubbard MRE. 2019. Upper ocean oxygenation, evolution of RuBisCO and the Phanerozoic succession of phytoplankton. *Free Radical Biology and Medicine* **140**: 295-304.

Rindi F, Mikhailyuk TI, Sluiman HJ, Friedl T, López-Bautista JM. 2011. Phylogenetic relationships in *Interfilum* and *Klebsormidium* (Klebsormidiophyceae, Streptophyta). *Molecular phylogenetics and evolution* **58**: 218-231.

Rosenzweig ESF, Xu B, Cuellar LK, Martinez-Sanchez A, Schaffer M, Strauss M, Cartwright HN, Ronceray P, Plitzko JM, Förster F *et al.* 2017. The eukaryotic CO_2 concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**: 148-162.

Sáez PL, Bravo LA, Cavieres LA, Vallejos V, Sanhueza C, Font-Carrascosa M, Gil-Pelegrin E, Peguero-Pina JJ, Galmés J. 2017. Photosynthetic limitations in two Antarctic vascular plants: importance of leaf anatomical traits and Rubisco kinetic parameters. *Journal of experimental botany* **68**: 2871-2883.

Sage RF. 2002. Variation in the kcat of Rubisco in C_3 and C_4 plants and some implications for photosynthetic performance at high and low temperature. *Journal of Experimental Botany* **53**: 609-620.

Sage RF, Christin PA, Edwards EJ. 2011. The C_4 plant lineages of planet Earth. *Journal of Experimental botany* **62**: 3155-3169.

Satagopan S, Spreitzer RJ. 2008. Plant-like substitutions in the large-subunit carboxy terminus of *Chlamydomonas* Rubisco increase CO_2/O_2 Specificity. *BMC plant biology* **8**: 85.

Savir Y, Noor E, Milo R, Tlusty T. 2010. Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proceedings of the National Academy of Sciences, USA* **107**: 3475–3480.

Sharwood RE, Sonawane BV, Ghannoum O, Whitney SM. 2016. Improved analysis of C_4 and C_3 photosynthesis via refined in vitro assays of their carbon fixation biochemistry. *Journal of Experimental Botany* **67**: 3137-3148.

Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**: 539.

Spreitzer RJ, Esquivel MG, Du YC, McLaughlin PD. 2001. Alanine-Scanning Mutagenesis of the Small-Subunit βA−βB Loop of Chloroplast Ribulose-1, 5-Bisphosphate Carboxylase/Oxygenase: Substitution at Arg-71 Affects Thermal Stability and CO_2/O_2 Specificity. *Biochemistry* **40**: 5615-5621.

Spreitzer RJ, Salvucci ME. 2002. RuBisCO: structure, regulatory interactions, and possibilities for a better enzyme. *Annual review of plant biology* **53**: 449-475.

Spreitzer RJ. 2003. Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Archives of Biochemistry and Biophysics* **414**: 41-149.

Spreitzer RJ, Peddi SR, Satagopan S. 2005. Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal RuBisCO. *Proceedings of the National Academy of Sciences, USA* **102**: 17225-17230.

Stabenau H, Winkler U. 2005. Glycolate metabolism in green algae. *Physiologia Plantarum* **123**: 235-245.

Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on mathematics in the life sciences* **17**: 57-86.

Tcherkez, G, Farquhar GD, Andrews TJ. 2006. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proceedings of the National Academy of Sciences, USA* **103**: 7246–7251.

Tcherkez G. 2013. Modelling the reaction mechanism of ribulose-1, 5-bisphosphate carboxylase/oxygenase and consequences for kinetic parameters. *Plant, Cell & Environment* **36**: 1586-1596.

Tortell PD. 2000. Evolutionary and ecological perspectives on carbon acquisition in phytoplankton. *Limnology and Oceanography* **45**: 744–750.

Valegård K, Andralojc PJ, Haslam RP, Pearce FG, Eriksen GK, Madgwick PJ, Kristoffersen AK, van Lun M, Klein U, Eilertsen HC *et al.* 2018. Structural and functional analyses of Rubisco from arctic diatom species reveal unusual posttranslational modifications. *Journal of Biological Chemistry* **293**: 13033-13043.

Von Caemmerer S, Quick WP. 2000. Rubisco: physiology in vivo. In: Leegood RC, Sharkey TD, von Caemmerer S, eds. *Photosynthesis. Advances in Photosynthesis and Respiration, vol 9*. Springer, Dordrecht, 85-113.

Wang H, Yan X, Aigner H, Bracher A, Nguyen ND, Hee WY, Long BM, Price GD, Hartl FU, Hayer-Hartl M. 2019. Rubisco condensate formation by CcmM in β-carboxysome biogenesis. *Nature* **566**: 131-135.

Wang L, Yamano T, Kajikawa M, Hirono M, Fukuzawa H. 2014. Isolation and characterization of novel high-CO_2-requiring mutants of *Chlamydomonas reinhardtii*. *Photosynthesis research* **121**: 175-184.

Warren SD, Clair LLS, Stark LR, Lewis LA, Pombubpa N, Kurbessoian T, Stajich JE, Aanderud ZT. 2019. Reproduction and dispersal of biological soil crust organisms. *Frontiers In Ecology Evolution* **7**: 344.

Wunder T, Cheng SLH, Lai SK, Li HY, Mueller-Cajar O. 2018. The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. *Nature communications* **9**: 5076.

Yamada K, Davydov II, Besnard G, Salamin N. 2019. Duplication history and molecular evolution of the *rbcS* multigene family in angiosperms. *Journal of experimental botany* **70**: 6127-6139.

Yamano T, Sato E, Iguchi H, Fukuda Y, Fukuzawa H. 2015. Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **112**: 7315-7320.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution* **24**: 1586-1591.

Yoon HS, Hackett JD, Pinto G, Bhattacharya D. 2002. The single, ancient origin of chromist plastids. *Journal of phycology* **38**: 40-40.

Young JN, Rickaby REM, Kapralov MV, Filatov DA. 2012. Adaptive signals in algal Rubisco reveal a history of ancient atmospheric carbon dioxide. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**: 483-492.

Young JN, Heureux AM, Sharwood RE, Rickaby RE, Morel FM, Whitney SM. 2016. Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *Journal of Experimental Botany* **67**: 3445-3456.

Zones JM, Blaby IK, Merchant SS, Umen JG 2015. High-resolution profiling of a synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. *Plant Cell* **27**: 2743-2769.

Figure legends:

Fig. 1 Protein phylogeny of the small subunit of Rubisco (*RbcS*) in green algae built with BEAST 2 (Bouckaert *et al*., 2014). Branches were colored according to the different phylum [chlorophytes: green (with prasinophytes in blue); streptophyte algae: orange]. Species lacking pyrenoids are indicated in bold red font. Length of the βA-βB loop was mapped onto each species and highlighted by the colour chart in the top left corner (species with a βA-βB loop length superior or equal to 25 residues are highlighted in the different shade of orange whereas species with a loop length inferior to 25 are highlighted in the different shade of blue). The phylogeny is clustered in two main clades. The first includes all the chlorophytes (green branches) and some prasinophytes (blue branches) and shows a loop length greater than, or equal to 25 residues. The second cluster includes all the streptophyte algae (orange branches) and the remaining prasinophytes (blue branches) with a loop length lower than 25 residues. Species without a pyrenoid (red font) are distributed across the phylogeny and not clustered together.

Fig. 2 Alignment of Rubisco small subunit (*RbcS*) sequences sampled from the 1KP, representative of streptophyte algae. The two isoforms of *RbcS* in *Chlamydomonas reinhardtii* (Chlorophytes, *Cr1* and *Cr2*) and *Arabidopsis thaliana* (*At*, land plants) are shown for comparison. *Ca* (*Chlorokybus atmophyticus*), *Ks* (*Klebsormidium subtile*), *Cs* (*Cosmarium subtumidum*), *Ol* (*Onychonema laeve*), *Ci* (*Coleochaete irregularis*) and *Ss* (*Spirogyra* sp.). Red boxes indicate residues of the two α-helices, green boxes indicate residues of the four β sheets and the blue box includes all the residues of the βA-βB loop. The alignment clearly shows the absence of five amino acids, between sites 61 to 66 (*Chlamydomonas reinhardtii* amino acid position is taken as reference).

Fig. 3 Electron micrographs of the six representative streptophyte algae and of *Chlamydomonas reinhardtii* (a: *Klebsormidium subtile*, b: *Cosmarium subtumidum*, c: *Chlorokybus atmophyticus*, d: *Onychonema laeve*, e: *Spirogyra* sp., f: *Coleochaete scutata* (from McKay *et al*., 1991), g: *Chlamydomonas reinhardtii*). Three distinct pyrenoid morphologies can be observed: matrix enclosed by one layer of starch plates (b, d, e, f and g); matrix enclosed by multiple starch grains (c); and pyrenoid without observable starch sheath (a). Bars: 2 μm (a to e) and 0.5 μm (f and g).

Table 1 Results of the three Likelihood Ratio Tests (LRTs) for positive selection using the site-models (M0-M8) (codeml) implemented in PAML (Yang, 2007) and their associated parameters.

	Number of classes (ω)	N[a]	Length (bp)[b]	LRT (2ΔlnL)	*P*-value (P<0.05)	df[c]
M0	1	135	462	2312.99077	<0.0001	8
M3	5	135	462			
M7	10	135	462	0	0.5	2
M8	11	135	462			
M8a	11	135	462	0	0.5	1
M8	11	135	462			

a: number of sequences analysed

b: length of *RbcS* sequences analysed

c: degrees of freedom

Table 2: Results of the three LRTs for positive selection using the branch-models (H0-H1) (codeml) implemented in PAML (Yang, 2007) and their associated parameters.

	dN/dS	**LRT (2ΔlnL)**	***P*-value (P<0.05)**	**df**
H0	ω=0.08445			
H1	ω^a=0.08262	9.358	0.0011	1
	ω^b=0.16371			

a: omega for background branches

b: omega for foreground branches

Table 3 Kinetic parameters of Rubisco at 25 °C in streptophyte algae in comparison to *Chlamydomonas reinhardtii* (Chlorophytes) and *Arabidopsis thaliana* (land plant) previously measured using the same protocol (Atkinson *et al*., 2017).

Species name	**n[a]**	**k_{cat} (s-1)**	**K_c (µM)**	**K_c^{air} (µM)**	**k_{cat} /K_c**	**k_{cat} /K_c^{air}**
Chlamydomonas reinhardtii	**3**	3.25 ± 0.18	39.6 ± 5.1	50.9 ± 7.0	0.086 ± 0.015	0.067 ± 0.011
Klebsormidium subtile	**6**	3.79 ± 0.67	18.7 ± 1.4	28.8 ± 2.1	0.228 ± 0.070	0.144 ± 0.040
Cosmarium subtumidum	**4**	2.51 ± 0.45	45.3 ± 13.1	55.6 ± 12.7	0.061 ± 0.008	0.040 ± 0.006
Onychonema laeve	**4**	2.39 ± 0.44	27.3 ± 5.5	40.9 ± 1.6	0.088 ± 0.003	0.052 ± 0.010
Spirogyra sp.	**5**	4.90 ± 0.32	49.1 ± 8.0	56.9 ± 4.3	0.108 ± 0.015	0.086 ± 0.010
Coleochaete scutata	**4**	1.67 ± 0.29	43.1 ± 9.8	48.2 ± 3.9	0.047 ± 0.013	0.032 ± 0.009
Arabidopsis thaliana (Atkinson et al., 2017)		4.10 ± 0.10	10.7 ± 0.7	15.8 ± 1.0	–	0.25 ± 0.01

a: number of replicates

Species are ordered from the furthest species (*Chlamydomonas reinhardtii*, Chlorophytes, Chlorophyceae) away from land plants to the closest (*Coleochaete scutata*, Coleochaetophyceae, Streptophytes). Values are means ± SEM.

Table 4 Whole cell affinity for inorganic carbon in the six streptophyte algae representative species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under low CO_2 conditions (9 μM) and their associated δ13C for organic matter.

Species name	**$K_{0.5}$(Ci) (μM)**	**$\delta^{13}C$ (‰)**
Chlamydomonas reinhardtii (n=3)	54 ± 23	-18.86 ± 0.01
Chlorokybus atmophyticus (n=3)	62 ± 26	-18.36 ± 0.02
Klebsormidium subtile (n=3)	53 ± 20	-21.18 ± 0.02
Cosmarium subtumidum (n=3)	64 ± 32	-15.80 ± 0.03
Onychonema laeve (n=3)	62 ± 40	-21.31 ± 0.03
Spirogyra sp. (n=3)	48 ± 38	-17.85 ± 0.04
Coleochaete scutata (n=3)	45 ± 23	-18.50 ± 0.09

Species are ordered from the furthest species away from land plants (*Chlamydomonas reinhardtii*, Chlorophytes, Chlorophyceae) to the closest (*Coleochaete scutata*, Coleochaetophyceae, Charophytes). Values are means ± SEM. n=number of replicates.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Evolutionary relationship of algae issued of the primary endosymbiosis and the major glaciation events which occurred during the diversification of the green algal lineages.

Fig. S2 Phylogenetic tree of 64 green algae species based on the nucleotide alignment of 44 chloroplastic genes.

Fig. S3 Comparison of the amino acid composition of the two Rubisco SSU α-helices for species without pyrenoid, compared to *Chlamydomonas reinhardtii* (pyrenoid positive).

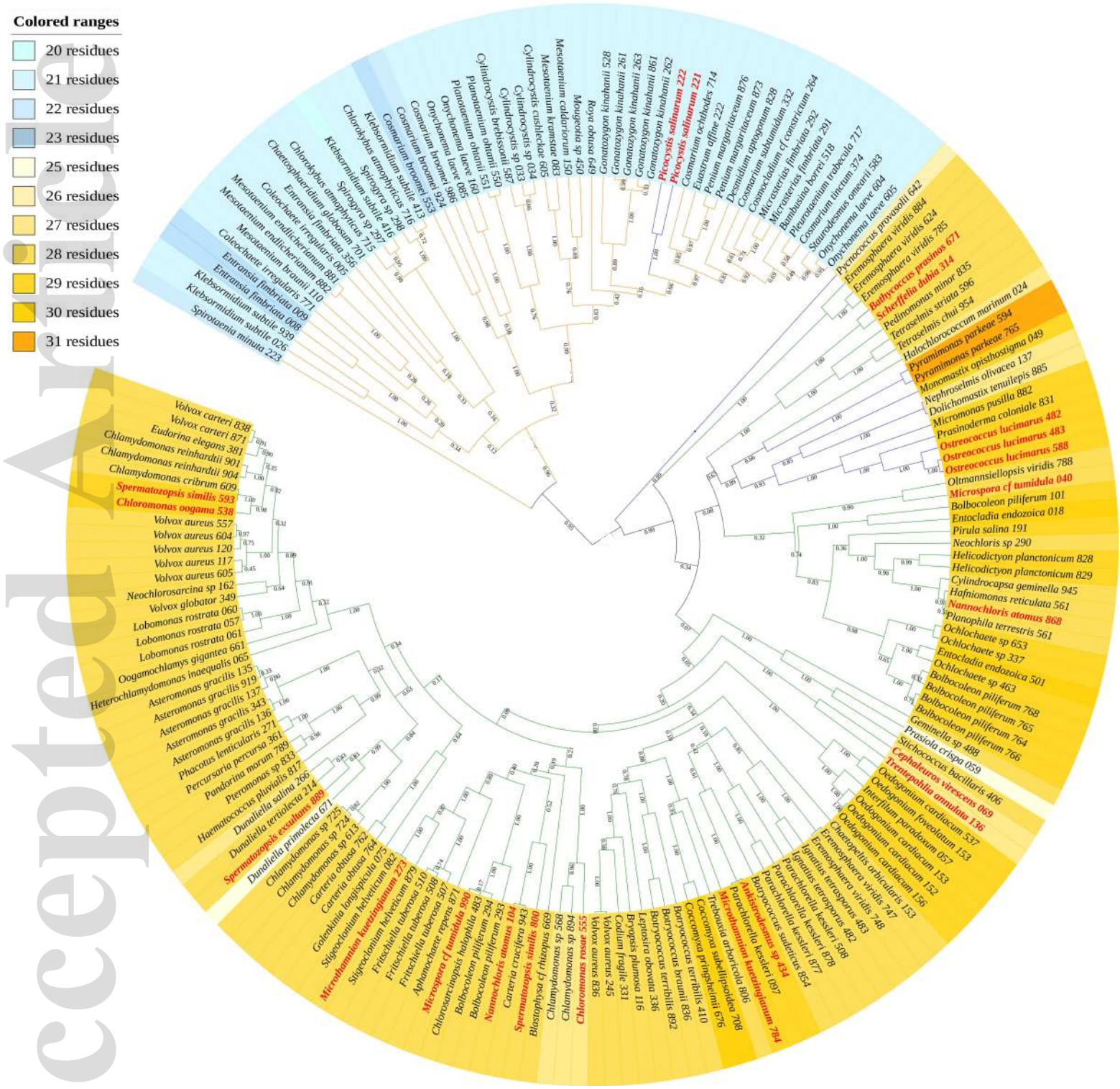
Fig. S4 DNA phylogeny of *RbcS* used for the PAML analysis and built with BEAST v2.3.1.

Table S1 Pyrenoid diagnostic for all the species present in the phylogeny of *RbcS* and the associated references.

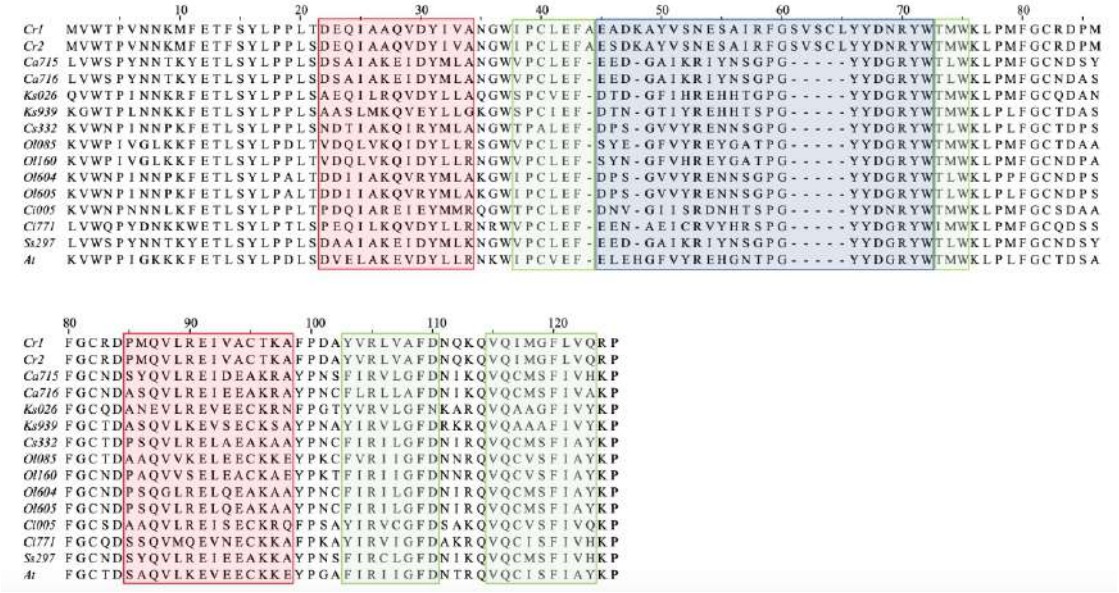
Table S2 Growth media and accession number of the six streptophyte algae.

Table S3 Classification and habitat description of the six streptophyte algae**.**

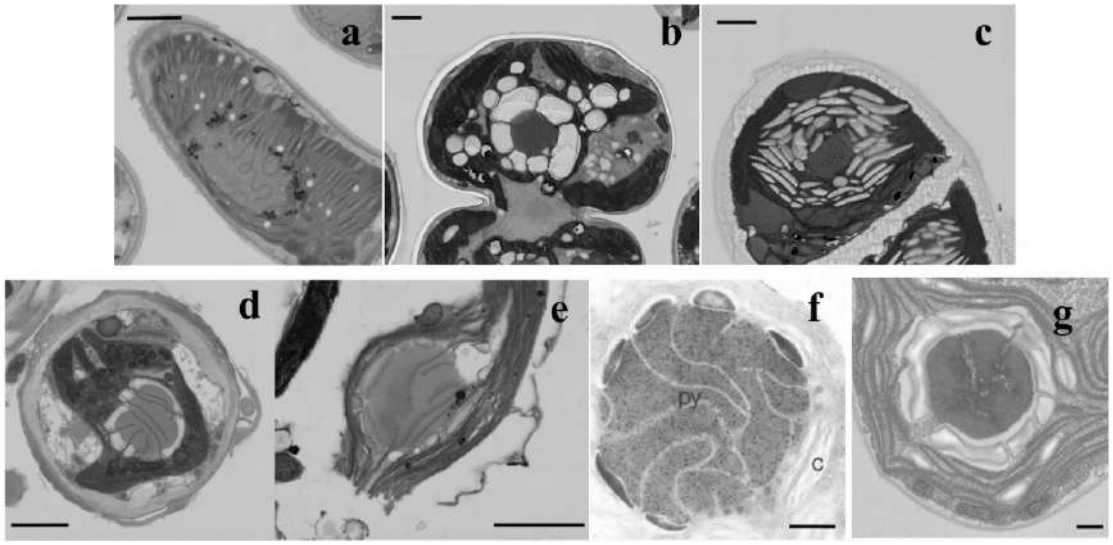
Table S4 Whole-cell affinity for inorganic carbon in the six streptophyte algae representative species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under high CO_2 conditions (5% CO_2) and their associated $\delta^{13}C$ for organic matter.



nph_16577_f1.eps



nph_16577_f2.jpg



nph_16577_f3.jpg