1 **Rubisco and carbon concentrating mechanism (CCM) co-evolution across Chlorophyte**

2 **and Streptophyte green algae**

3 Myriam M. M. Goudet[1], Douglas J. Orr[2], Michael Melkonian[3], Karin H. Müller[4], Moritz T.

4 Meyer[5], Elizabete Carmo-Silva[2] and Howard Griffiths[1]

5 [1]Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK;

6 [2]Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK; [3]Institute for

7 Plant Sciences, Department of Biological Sciences, University of Cologne, 50674 Cologne,

8 and Central Collection of Algal Cultures, Faculty of Biology, University of Duisburg-Essen,

9 45141 Essen, Germany; [4]Cambridge Advanced Imaging Centre, University of Cambridge,

10 Cambridge, CB2 3DY, UK; [5]Department of Molecular Biology, Princeton University,

11 Princeton, NJ 08544, USA

12

13 Author for correspondence:

14 *Myriam M. M. Goudet*

15 *Tel: +44 (0)1223 330218*

16 *Email: mmmg2@cam.ac.uk*

17

18 And

19

20 *Prof. Howard Griffiths*

21 *Tel: +44 (0)1223 333946*

22 *Email: hg230@cam.ac.uk*

Total word count (excluding summary, references and legends):	6472	No. of figures	3 (Fig. 1- 2 in colour)
Summary:	199	No. of Tables:	4
Introduction	1128	No. of Supporting Information files:	7 (Fig. S1-4; Table S1-S4)
Materials and Methods	1574		
Results:	1853		
Discussion:	1917		
Acknowledgements:	132		

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 $111\text{-}117\text{x}10^{15}$ 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).

54   Green algal inter-relationships have been resolved through numerous molecular phylogenies,

55   including the chloroplast gene (*rbcL*) encoding the large subunit (LSU) of the primary

56   carboxylase Rubisco (ribulose 1,5-bisphosphate carboxylase/oxygenase). An early split after

57   the primary endosymbiosis saw the diversification of the hypothetical ancestral flagellate into

58   two main lineages (Leliaert *et al*., 2011; 2012). First, the chlorophytes, which diversified early

59   as prasinophytes in marine waters, which then gave rise to the core chlorophytes (chlorophytes

60   without prasinophytes, Fig. S1, Supporting Information) in fresh or marine waters. Second, the

61   streptophyte algae, which diversified in fresh water and some subaerial/terrestrial habitats

62   (Harholt *et al*., 2016). The split between chlorophyte and streptophyte probably occurred

63   during the Neoproterozoic (between 1,000 and 541 million years ago; Becker, 2013; Del

64   Cortona *et al*., 2020). Extant photosynthetic chlorophyte and streptophyte algae (as well as

65   non-algal streptophytes, i.e. land plants) have a form 1B Rubisco. Selection pressures on the

66   Rubisco catalytic properties are driven by the availability and diffusive supply of inorganic

67   carbon, the CO_2:O_2 ratio and the development of any carbon concentrating mechanism (CCM)

68   which improves the operating efficiency of Rubisco in many aquatic photosynthetic

69   microorganisms (Tortell, 2000; Young *et al*., 2012; Meyer & Griffiths, 2013; Griffiths *et al*.,

70   2017; Rickaby & Hubbard, 2019). The origins of the algal CCM could be related to equimolar

71   CO_2:O_2 concentrations in surface waters around 500 million years ago (Griffiths *et al*., 2017).

72   The challenge for inorganic carbon delivery within aquatic environments is that bicarbonate

73   (HCO_3^-) or carbonate (CO_3^{2-}) are often much more prevalent, and under current conditions, the

74   concentration of CO_2 is often ~2,000 times lower in water than in air, and diffusion is 8,000

75   times slower (Raven *et al*., 1985; Falkowski & Raven, 2007; Young *et al*., 2012). A CCM is

76   typically associated with active transport of bicarbonate across membranes, and catalytic

77   conversion to CO_2 within a chloroplast microcompartment, the pyrenoid (Meyer *et al*., 2017).

78   Although the presence of a pyrenoid is a robust marker of the presence of a CCM, not all the

79   eukaryotic algae with a CCM have a pyrenoid (Morita *et al.,* 1999; Raven *et al*., 2005).

80   The CCM has been particularly well-defined in the model unicellular chlorophyte

81   *Chlamydomonas reinhardtii*, where the pyrenoid is present with a clearly defined starch sheath,

82   and the associated inner Rubisco matrix transversed by knotted thylakoid tubules, thought to

83   be involved in the delivery of CO_2 within the matrix (Meyer & Griffiths, 2013; Engel *et al*.,

84   2015; Mackinder *et al*., 2017; Meyer *et al*., 2017; Mukherjee *et al*., 2019). The CCM is

85   inducible following transfer from elevated to ambient CO_2, and a key linker protein (EPYC1)

86   has been associated with the recruitment of Rubisco to the pyrenoid (Mackinder *et al.*, 2016;

87   Freeman-Rosensweig *et al.*, 2017). This recruitment ultimately involves interactions with the

88   Rubisco Small Subunit (SSU) (Wunder *et al.*, 2018; Atkinson *et al.*, 2019), presumably at the

89   level of surface exposed α-helices (Meyer *et al.*, 2012). However, there has been little

90   systematic analysis of the extent to which some form of carbon accumulation mechanism

91   occurs across this chlorophyte clade, or comparative physiological and molecular studies on

92   CCM characteristics or Rubisco kinetic properties, and whether these traits are captured across

93   chlorophyte, prasinophyte and streptophyte algal lineages in *Rbc*S.

94   *Chlamydomonas reinhardtii* has also been used as a model organism to explore the interactions

95   between Rubisco LSU, SSU and catalytic properties. The eight identical 55-kDa LSUs

96   assemble as four dimers, while two sets of four 15-kDa SSUs, top and tail the Rubisco

97   holoenzyme. A central 'solvent channel' runs through Rubisco and the width of its aperture is

98   dependent on the length of the βA-βB loop in each set of four SSUs capping the LSU octamer

99   (Spreitzer, 2003) and interacting residues between LSUs and SSUs affect Rubisco operating

100   efficiency and catalytic properties (Spreitzer *et al.*, 2005). Natural variation in Rubisco

101   catalytic properties exists among photosynthetic organisms (Jordan & Ogren, 1981), however,

102   a shift in the catalytic parameters towards higher turnover rate per active site (k_{cat}) and higher

103   affinity for CO_2 (K_c) has been observed from cyanobacteria, chlorophyte to land plants

104   (reviewed in Badger *et al.*, 1998; Meyer & Griffiths, 2013). However, it has also been

105   suggested that selective pressures on the Rubisco kinetic parameters Vc and K_c could have been

106   relaxed due to the saturating CO_2 environment provided by a CCM over evolutionary time

107   (Tortell, 2000; Young *et al.*, 2012; Meyer & Griffiths, 2013).

108   The overall aim of the presents study was to address the possible interactions between Rubisco

109   SSU structure and phylogeny, and occurrence of any reported CCM or pyrenoid across the

110   green algae. Additionally, we set out to define key Rubisco catalytic properties for a range of

111   streptophyte algae representing the main streptophyte lineages (Fig. S1), as compared to *C.*

112   *reinhardtii*. The few Rubisco kinetic measurements available for green algae were performed

113   on chlorophytes (*Coccomyxa* sp., Palmqvist *et al.*, 1995; *Scenedesmus obliquus*, Jordan &

114   Ogren, 1981, Badger *et al.*, 1998), not streptophyte algae. Surprisingly, there is yet no

115   streptophyte model alga, despite the previous interest in using species with giant cells to

116   characterise carbon uptake mechanisms (Lucas & Berry, 1985) or the recently published

117   genome of *Chara braunii* (Nishiyama *et al.,* 2018).

118 Specifically, this study sought to *(i)* develop a phylogeny for *Rbc*S sequences in green algae as
119 compared to consensus phylogenies (e.g. Leliart *et al*., 2012; Leebens-Mack *et al.,* 2019), and
120 compare the distribution of pyrenoid and CCM across the algal clades; *(ii)* to identify whether
121 any selection pressure on residues within the SSU were associated with the broader phylogeny
122 or CCM activity and, *(iii)* to determine whether the catalytic properties of Rubisco across
123 contrasting streptophyte algal groups reflected the overall phylogeny or specific activity of a
124 CCM at the whole organism level. Our results reveal that a change in Rubisco SSU secondary
125 structure (namely the βA-βB loop) is a distinctive trait of the division between core
126 chlorophytes and streptophyte algae. We also demonstrate that Rubisco catalytic properties
127 have co-evolved in association with the extent of CCM activity in streptophytes. Finally, this
128 study provides additional insights for selection pressures driving the evolution of green algae
129 and photosynthetic processes, particularly for Rubisco during the transition to terrestrial plant
130 life forms.

131 **Materials and Methods**

132

133 **Collection of protein sequences, phylogenetic analysis, βA-βB loop length and pyrenoid**
134 **presence/absence mapping**

135 2,674 protein *Rbc*S sequences of green algae were kindly provided by «The 1000 plants
136 project» (1KP; Leebens-Mack *et al*., 2019). All the protein sequences were manually and
137 individually screened. Sequences showing cross-contamination (Carpenter *et al*., 2019), or
138 which were too short or incomplete, were removed. The dataset did now allow to
139 unambiguously identify *Rbc*S isoforms. Although it is generally taken that all photosynthetic
140 members of the Viridiplantae have multiple copies of the *Rbc*S gene, conservatively only one
141 sequence was used in the analysis for each species, except when the data was sourced from
142 independently sequenced genomes (e.g. for *Asteromonas*). A total of 187 protein sequences
143 belonging to 113 species (31 streptophyte algae, 10 prasinophytes, 72 chlorophytes) were then
144 aligned with Clustal Omega (Sievers *et al*., 2011). ProTest v2.4 (Abascal *et al.*, 2005) was used
145 to identify the best model of protein evolution. Bayesian phylogenetic analyses were performed
146 using BEAST v2.3.1 (Bouckaert *et al*., 2014) with a LG model of protein evolution (Le &
147 Gascuel, 2008), a gamma distribution model with four categories, a relaxed molecular clock
148 and finally with a Yule model of speciation. Three independent chains were run, each of length
149 8×10^7 steps, parameters values and trees were sampled every 10×10^2 steps. Chain convergences
150 were checked using Tracer v1.6 (Drummond & Rambaut, 2007). Posterior parameters were

151 summarized with Tree Annotator v1.8.2 (Drummond & Rambaut, 2007) using a maximum
152 clade credibility tree (MCC) and a posterior limit of 0.5. Figtree v1.4.2 (Rambaut, 2007) was
153 used for tree visualizations. The length of the βA-βB loop was determined after the analysis of
154 the protein sequences, with the number of residues in the loop (Spreitzer, 2003) mapped on to
155 the phylogeny of *Rbc*S. Finally, the same phylogeny was used to map the pyrenoid
156 presence/absence. The scoring for pyrenoid presence/absence was based on the available
157 literature (Table S1).

158 **Likelihood ratio test for positive selection**
159 To test the importance of two SSU α-helices for pyrenoid formation in *C. reinhardtii* (Meyer
160 *et al.*, 2012), the Codon-based package (codeml) implemented in PAML v4.9 (Yang, 2007)
161 was used to detect residues under positive selection across the green algae lineage. In addition,
162 the presence of a CCM is not universal across the green algae so the branch model also
163 implemented in PAML was used to detect branches under positive selection. All the analyses
164 were performed using "user tree" mode. The DNA phylogenetic tree was reconstructed using
165 BEAST v2.3.1 with 135 cDNA *Rbc*S sequences of green algae from the 1KP, with a GTR
166 model of protein evolution (Tavaré, 1986) and the same gamma distribution, molecular clock
167 and model of speciation previously used. Three independent chains were run, each of length
168 5×10^7 steps, parameters values and trees were sampled every 10×10^2 steps. Chain
169 convergences, posterior parameters and tree visualization were analysed with the same method
170 explained above. Several models of codon evolution that allow for variations in ω (dN/dS)
171 among codons were tested (Site model) and evaluated using Likelihood Ratio Tests (LRTs)
172 (Neyman & Pearson, 1928) as described in Kapralov & Filatov (2007). Branch models were
173 used to test for positive selection across branches. The null model allowed for variations in ω
174 among branches (0<dN/dS<1 and dN/dS=1 for both foreground and background branches) and
175 also included two additional classes of codons with fixed dN/dS=1 on foreground branches but
176 restricted as 0< dN/dS<1 and dN/dS=1 for background branches. The alternative model allowed
177 0<dN/dS<1 and dN/dS=1 for both foreground and background branches but also included two
178 additional classes of codons under positive selection with dN/dS>1 on foreground branches
179 with restriction as 0<dN/dS<1 and dN/dS=1 on background branches. Branches leading to
180 species without pyrenoid were labelled as foreground branches (allows positive selection) and
181 the rest of the branches were considered as background branches (with no positive selection).
182 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 4,200 rpm 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 *Rbc*S 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 dN/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).

309 Branches under positive selection were successfully tested with the branch-model implemented

310 in PAML. The LRT for heterogeneity across branches (H0-H1) was significant ($2\Delta lnL$=9.358,

311 *P*-value=0.0011, df=1) (Table 2). However, background and foreground omega showed values

312 less than 1, implying positive selection was absent among foreground branches

313 ($\omega\alpha$=0.082; $\omega\beta$=0.16 <1). These results suggest that the presence of variation in ω across

314 branches in *Rbc*S, but not significant enough to show positive selection, or any correlation with

315 pyrenoid occurrence.

316

317 Streptophyte algae share Rubisco catalytic properties with both chlorophytes and

318 embryophytes

319 A more detailed investigation of Rubisco catalytic properties was undertaken in order to

320 explore whether any evolutionary progression towards land plant characteristics was evident

321 in streptophyte algae. The multiple alignment of *Rbc*S in six representative streptophyte algae

322 selected for this component of the study confirmed the deletion of five amino-acids in this

323 group compared to *C. reinhardtii* (Fig. 2; Spreitzer, 2003). This shortens the loop between the

324 first and the second β-sheets, reducing the constriction at the entry of the holoenzyme's solvent

325 channel. Rubisco catalytic properties at 25°C for the six green algae are shown in Table 3,

326 including *C. reinhardtii* as a control. In *C. reinhardtii*, Rubisco catalytic properties varied

327 slightly from previous measurements (Satagopan & Spreitzer, 2008; Jordan & Ogren, 1981)

328 but remained in the same range. Michaelis-Menten constant for carboxylation (K_c) showed

329 similar values (39.6 and 34 μM) whereas the Rubisco turnover rate (k_{cat}) was somewhat higher

330 in this study compared to the value found in Satagopan & Spreitzer (2008). The streptophyte

331 algae did not show a clear systematic shift from chlorophyte towards land plant catalytic

332 properties despite similar Rubisco SSU structural changes. Of the five streptophyte algae, only

333 *Klebsormidium subtile* and *Onychonema laeve* showed a higher affinity for CO_2 (lower K_c

334 values), closer to land plant values (e.g. *Arabidopsis thaliana*; 10.7 μM) with K_c of 18.7 and

335 27.3 μM respectively (Table 3). *Cosmarium subtumidum*, *Spirogyra* sp. and *Coleochaete*

336 *scutata* had a relative low affinity for CO_2 with K_c values in the range of the core chlorophytes

337 or slightly higher (45.3, 49.1 and 43.1 μM respectively).

338 The catalytic turnover rate (k_{cat}) showed a trend towards lower values. *Onychonema laeve* and

339 *Cosmarium subtumidum*, both members of the Zygnematophyceae, had similar k_{cat} values (2.39

340 and 2.51 s^{-1} respectively). *Spirogyra* sp. appeared to be an exception with a high k_{cat} value

341 compared to the other streptophyte algae (4.90 s^{-1}), similar to the land plant *A. thaliana* (4.1 s^{-}

342   [1], Atkinson *et al.*, 2017). *Coleochaete scutata* showed the lowest k_{cat} of all the streptophyte

343   algae (1.67 s^{-1}). Higher K_c is usually correlated to high k_{cat} and lower specificity factor (Badger,

344   1987; von Caemmerer & Quick, 2000; Tcherkez *et al.*, 2006; Savir *et al.*, 2010; Tcherkez,

345   2013). *Klebsormidium subtile* presented the highest value for carboxylation catalytic efficiency

346   ($k_{cat}/K_c{}^{air}$) (0.14 s^{-1} μM^{-1}), and whilst this was the highest streptophyte algae value determined,

347   remains well below that of land plants like *A. thaliana* (Atkinson *et al.*, 2017). The remaining

348   streptophyte algae displayed lower efficiency, with *Coleochaete scutata* showing the lowest

349   efficiency (0.032 s^{-1} μM^{-1}).

350

351   **Streptophyte algae have a CCM, albeit leaky in some species**

352   Oxygen evolution measurements, pyrenoid imaging and δ^{13}C were used to characterise CCM

353   activity in the different streptophyte algae and to investigate whether CCM activity was

354   associated with Rubisco catalytic properties. Oxygen evolution was used to determine the

355   whole cell affinity for inorganic carbon and therefore the extent of any inducible carbon

356   concentrating mechanism. The photosynthetic $K_{0.5}$ (Ci) value (Table 4) of the wild-type *C.*

357   *reinhardtii* under low CO_2 showed a strong affinity for Ci (54 μM Ci), similar to previous

358   values in the literature (Mitchell *et al.*, 2014; Wang *et al.*, 2014). *Klebsormidium subtile*,

359   *Spirogyra* sp. and *Coleochaete scutata* showed a whole cell affinity for Ci similar to *C.*

360   *reinhardtii* with $K_{0.5}$ ranging from 45 to 53 μM Ci, consistent with a fully functional CCM,

361   whereas *Chlorokybus atmophyticus*, *Cosmarium subtumidum* and *Onychonema laeve* exhibited

362   a c.20% lower apparent affinity for CO_2 compared to the other species ($K_{0.5}$ 62, 64 and 62 μM

363   Ci respectively), but still suggestive of CCM activity. Photosynthetic $K_{0.5}$ (Ci) values of all the

364   species grown under high CO_2 confirmed the absence of CCM activity under such conditions

365   (Table S4), and thereby the inducible character of the CCM in all species under examination.

366   Stable carbon isotope composition (δ^{13}C) for organic matter was also used as a second proxy

367   for CCM activity in the different species (Table 4). *Coleochaete scutata, Chlorokybus*

368   *atmophyticus*, *Spirogyra* sp. and *Cosmarium subtumidum* appeared to be isotopically enriched

369   at -15.8 to -18.8‰ (Table 4), values close to *C. reinhardtii* (-18.9‰) and close to the upper

370   range typically seen in C_4 terrestrial plants and consistent with a fully-functioning CCM (Raven

371   *et al.*, 1982). On the other hand, *Klebsormidium subtile* and *Onychonema laeve* were somewhat

372   isotopically depleted compared to the other species, with values intermediate between typical

373   C_3 and C_4 plants (δ^{13}C of -21.1 and -21.3‰ respectively; O'Leary, 1988) and consistent with

374   a CCM phenotype prone to leakiness (retro-diffusion of CO_2: Meyer *et al.*, 2008) or limited

375   carbon accumulation capacity.

376   Taken together, these observations reveal that Rubisco catalytic properties correlate to some

377   extend with the strength of CCM activity. Similarly to *C. reinhardtii*, the three streptophytes

378   algae *Cosmarium subtumidum*, *Spirogyra* sp. and *Coleochaete scutata* revealed a fully

379   functioning CCM (low whole-cell affinity, $K_{0.5}$, and low carbon isotope discrimination) but

380   lower Rubisco catalytic affinity for inorganic carbon (high K_c values), whereas *Klebsormidium*

381   *subtile* and *Onychonema laeve* have a less effective CCM but higher affinity for inorganic

382   carbon in terms of Rubisco catalytic properties (low K_c values). Therefore, in the presence of

383   a less-effective CCM, Rubisco catalytic properties for *Klebsormidium subtile* and *Onychonema*

384   *laeve* show a systematic shift towards values more typically associated with land plants.

385   Finally, electronic microscopy was used to diagnose the presence/absence of a pyrenoid in the

386   algal material used in the present study, as an additional diagnostic for an active biophysical

387   CCM. The presence of a pyrenoid was confirmed for all the species except for *Coleochaete*

388   *scutata* for which tissue embedding was unsuccessful. Presence and morphology of a pyrenoid

389   in that species had been previously published (McKay *et al.*, 1991). CCM activities were

390   supported by the presence of a pyrenoid in all species (Fig. 3). *Cosmarium subtumidum* (Fig.

391   3b), *Onychonema laeve* (Fig. 3d)*, Coleochaete scutata* (Fig. 3f) and *Spirogyra* sp. (Fig. 3e)

392   exhibited pyrenoid morphologies similar to *C. reinhardtii* (Fig. 3g) with a spheroidal electron

393   dense matrix traversed by multiple tubules, and a single layered peripheral starch sheath. There

394   were, however, differences in the fine structure (starch sheath thickness and continuity, density

395   of thylakoid tubules network) that perhaps provide clues to the variability in $K_{0.5}$ and $\delta^{13}C$

396   measurements. *Klebsormidium subtile* lacked a peripheral starch sheath (Fig. 3a), although a

397   starch sheath may occur in *Klebsormidium* dependent on growth stage or light intensity (M.

398   Melkonian, unpublished observations). *Chlorokybus atmophyticus* had multiple layers of short

399   starch plates surrounding the matrix (Fig. 3c). The network of cross-pyrenoidal tubules was

400   regular and dense in *Cosmarium* and *Chlorokybus* (Figs 3 b, c).

401   Overall, the results show that Rubisco catalytic properties are CCM dependent. However, at

402   this stage, it remains difficult to differentiate limitations in carbon uptake versus leakiness of

403   CO_2 as the selective pressure operating on Rubisco, and more detailed physiological

404   experiments are warranted to fully characterize these contrasting processes.

405

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

439   been detected in *Rbc*S in angiosperms (Yamada *et al*., 2019). The absence of positive selection

440   along branches leading to a pyrenoid could be an artefact of the small number of species *lacking*

441   a pyrenoid within the green algae (Fig. 1), or indeed those possessing some form of a CCM but

442   lacking a pyrenoid (see above). A possible alternative explanation is that all green algae

443   retained a pyrenoid-competent Rubisco SSU but that the absence of a pyrenoid is rather

444   determined by the lack (ancestral or through secondary loss) of a Rubisco linker, of similar or

445   different ancestry as the *C. reinhardtii* EPYC1 (Mackinder *et al*., 2016). Here too, future

446   comparative proteomic studies with pyrenoidless algal CCMs will help resolve this question.

447

Streptophyte algal Rubisco SSU structure is similar to land plants

449   The phylogeny of *Rbc*S revealed a Rubisco structure in streptophyte algae similar to that of

450   embryophytes, with SSUs possessing a shorter βA-βB loop and therefore a central solvent

451   channel with a similar open structure as that shown for embryophytes (Spreitzer, 2003).

452   Although the shorter loop in land plants has been well described (Spreitzer, 2003) and was

453   probably thought to be a consequence of the transition from the aquatic environment to land,

454   the presence of a similar structure in the streptophyte algae has not been previously reported.

455   The phylogeny of *Rbc*S showed that this loss of amino acids is more ancient, and probably

456   occurred during the split between chlorophytes and streptophyte algae, which occurred

457   somewhere between 736 Mya (Becker, 2013) and 1,000 Mya (early Neoproterozoic; Del

458   Cortona *et al*., 2020). The Rubisco structural change was not an isolated event at this time. The

459   split between chlorophytes and streptophytes coincides with the appearance of multiple new

460   traits (Hori *et al*., 2014; Nishiyama *et al*., 2018) such as lateral flagella, a flagellar peroxidase

461   and also a Gap A/B gene duplication (McCourt *et al*., 2004; Finet *et al*., 2010). Interestingly,

462   the photorespiratory pathway, which would have to contend with CCM inefficiencies, has been

463   shown to differ between chlorophytes and streptophyte algae. Chlorophytes use a

464   mitochondrial glycolate dehydrogenase, which produces NADH and H^+ whereas streptophytes

465   use a peroxisomal glycolate oxidase which produces H_2O_2 for the conversion of glycolate to

466   glyoxylate (Stabenau & Winkler, 2005).

467

468   The role of the SSU and of the βA-βB loop in particular is not entirely understood but the

469   central solvent channel may facilitate channelling of substrates and products to and from the

470   active sites (Esquivel *et al*., 2013). Spreitzer (2001; 2002) demonstrated the importance of the

471   loop for holoenzyme assembly and showed that direct mutagenesis within and near the βA-βB

472 loop changed Rubisco catalytic properties. However, these studies did not investigate the
473 relationship to presence or absence of the pyrenoid in green algae and CCM activity. Despite
474 the change in Rubisco SSU structure between chlorophytes and streptophytes, and effect on
475 solvent channel width, the present work showed that there was a continued need for CCMs
476 across the entire phylogeny (Fig. 1), as reflected in the catalytic properties of the streptophyte
477 algae.

478

479 **Rubisco catalytic properties in green algae depend on CCM efficiency**

480 The above observations led to the investigation of Rubisco catalytic properties within the
481 streptophyte algae and their associated physiological CCM activity. Streptophyte algae are
482 difficult to investigate physiologically. Oxygen electrode measurements were also extremely
483 challenging (Table 4).

484 Despite the clear structural change associated with the βA-βB loop length, Rubisco catalytic
485 properties remained generally similar to chlorophytes (Table 3) without systematic shift
486 towards values associated with land plants (Satagopan & Spreitzer, 2008; Kapralov *et al.*, 2010;
487 Atkinson *et al.*, 2017). Of the six streptophyte algae, only *Klebsormidium subtile* and
488 *Onychonema laeve* showed K_c values in this lower range. Direct mutagenesis has shown the
489 importance of the SSU βA-βB loop in Rubisco catalytic properties (see paragraph above) but
490 the data in the present study suggested that they were more influenced by the effectiveness of
491 the CCM, consistent with systematic changes in carbon isotope composition ($\delta^{13}C$: Table 4).
492 Carbon isotopes have been used to infer leakiness of CCMs found in algae and hornworts
493 (Meyer *et al.*, 2008). Although whole cell inorganic carbon (Ci) uptake affinity was broadly
494 similar for all species under ambient growth conditions ($K_{0.5}$, Table 4), the weaker CCM
495 activities (identified through more negative $\delta^{13}C$ values: Table 4) in *Klebsormidium subtile* and
496 *Onychonema laeve*, were associated with the highest affinity of Rubisco for CO_2 (K_c, Table 3).
497 The importance of the CCM in shaping the adaptation within Rubisco catalytic properties has
498 been a long-standing hypothesis (Tortell, 2000; Young *et al.* 2012; Meyer *et al.*, 2013, Galmés
499 *et al.*, 2014, 2016, 2019; Griffiths *et al.*, 2017), consistent with the shifts seen in C_4 Rubisco
500 (Jordan & Ogren, 1981; Sage, 2002; Kubien *et al.*, 2008). Our results show that Rubisco
501 catalytic properties for this representative range of streptophyte algae are adapted to the
502 presence of the CCM.

503 A strong CCM (uptake and conversion of inorganic carbon) or reduced retrodiffusion
504 (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

539   recaptures the latest consensus green algal phylogenies built from many marker genes,

540   including *rbcL* (Leebens-Mack *et al.*, 2019). A more focussed study on Rubisco catalytic

541   properties in streptophyte algae suggests that the activity of any CCM, which may have arisen

542   because of limitations in bulk CO_2 delivery to Rubisco, has permitted the retention of a lower

543   affinity (high K_c) Rubisco. We showed that the extent of adaptation which occurs should either

544   cause CCM activity to be reduced, or indeed lost during the transition to land, as the reliance

545   on gaseous diffusion to deliver CO_2 to Rubisco began to increase. Overall, the observations

546   confirm the widespread occurrence of a CCM across the entire green algal lineage through the

547   means of a pyrenoid-based CCM to fuel carbon fixation by Rubisco. However, rather than

548   being intransigent and slow, Rubisco catalytic properties adapt to local conditions of CO_2

549   availability. This is consistent with the changes seen in Rubisco from C_4 (Jordan & Ogren,

550   1981; Sage, 2002; Kubien *et al.*, 2008) and CAM plants (Griffiths *et al.*, 2008), which have

551   been associated with operating within a CCM for the past 5-10 million years. Based on this

552   study, the selective pressures driven by local conditions of photosynthetic CO_2 supply are more

553   likely to explain the shifts in Rubisco catalytic properties during life on land, rather than any

554   long term transition seen in land plants.

555

568

569   **Author Contributions**

570   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

571   performed the experiments on Rubisco kinetics and D.J.O. analysed the data. M.M.M.G

572 performed the phylogenetic analyses, positive selection and physiological data collection and

573 analysis. K.H.M. performed SEM imaging. M.M. provided the *RbcS* sequences. M.M.M.G and

574 H.G. interpreted the data and wrote the manuscript with assistance from all authors.

575

576

577

578

579 **References**

580 **Abascal F, Zardoya R, Posada D. 2005.** ProtTest: selection of best-fit models of protein

581 evolution. *Bioinformatics* **21**: 2104-2105.

582 **Atkinson N, Leitão N, Orr DJ, Meyer MT, Carmo-Silva E, Griffiths H, Smith AM,**

583 **McCormick AJ. 2017.** Rubisco small subunits from the unicellular green alga

584 *Chlamydomonas* complement Rubisco-deficient mutants of Arabidopsis. *New Phytologist* **214**:

585 655-667.

586 **Atkinson N, Velanis CN, Wunder T, Clarke DJ, Mueller-Cajar O, McCormick AJ. 2019.**

587 The pyrenoidal linker protein EPYC1 phase separates with hybrid Arabidopsis–

588 Chlamydomonas Rubisco through interactions with the algal Rubisco small subunit. *Journal*

589 *of experimental botany*.

590 **Badger MR, Kaplan A, Berry JA. 1980.** Internal inorganic carbon pool of *Chlamydomonas*

591 *reinhardtii*: evidence for a carbon dioxide-concentrating mechanism. *Plant Physiology* **66**:

592 407-413.

593 **Badger MR. 1987.** The CO_2-concentrating mechanism in aquatic phototrophs. *Photosynthesis*.

594 Academic press.

595 **Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD.**

596 **1998.** The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based

597 CO_2 concentrating mechanisms in algae. *Canadian Journal of Botany* **76**: 1052-1071.

598 **Becker B. 2013.** Snow ball earth and the split of Streptophyta and Chlorophyta. *Trends in*

599 *Plant science* **18**: 180-183.

600 **Beerling DJ, Osborne CP, Chaloner WG. 2001.** Evolution of leaf-form in land plants linked

601 to atmospheric CO_2 decline in the Late Palaeozoic era. *Nature* **410**: 352.

602 **Behrenfeld MJ, Randerson JT, McClain CR, Feldman GC, Los SO, Tucker CJ,**

603 **Falkowski PG, Field CB, Frouin R, Esaias WE *et al*. 2001.** Biospheric primary production

604 during an ENSO transition. *Science* **291**: 2594-2597.

605 **Bernacchi CJ, Singsaas EL, Pimentel C, Portis Jr AR, Long SP. 2001.** Improved

606 temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell &*

607 *Environment* **24**: 253-259.

608 **Borges AV, Frankignoulle M. 2002.** Distribution and air-water exchange of carbon dioxide

609 in the Scheldt plume off the Belgian coast. *Biogeochemistry* **59**: 41-67.

610 **Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut**

611 **A, Drummond AJ. 2014.** BEAST 2: a software platform for Bayesian evolutionary analysis.

612 *PLoS computational biology* **10**: e1003537.

613 **Carpenter EJ, Matasci N, Ayyampalayam S, Wu S, Sun J, Yu J, Jimenez-Vieira FB,**

614 **Bowler C, Dorrel RG, Gitzendanner MA *et al*. 2019.** Access to RNA-sequencing data from

615 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). *GigaScience* **8**: 1-7.

616 **Chan KX. 2018.** Morphological and physiological studies of the carbon concentrating

617 mechanism in *Chlamydomonas reinhardtii*. PhD thesis, University of Cambridge, UK.

618 **Conesa MÀ, Muir CD, Molins A, Galmés J. 2019.** Stomatal anatomy coordinates leaf size

619 with Rubisco kinetics in the Balearic Limonium. *AoB PLANTS*.

620 **Del Cortona A, Jackson CJ, Bucchini F, Van Bel M, D'hondt S, Škaloud P, Delwiche CF,**

621 **Knoll AH, Raven JA, Verbruggen H *et al*. 2020.** Neoproterozoic origin and multiple

622 transitions to macroscopic growth in green seaweeds. *Proceedings of the National Academy of*

623 *Sciences, USA* **117**: 2551-2559.

624 **Delwiche CF, Palmer JD. 1997.** The origin of plastids and their spread via secondary

625 symbiosis. *Origins of algae and their plastids.* Springer, Vienna.

626 **Drummond AJ, Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling

627 trees. *BMC evolutionary biology* **7**: 214.

628 **Engel BD, Schaffer M, Cuellar LK, Villa E, Plitzko JM, Baumeister W. 2015.** Native

629 architecture of the *Chlamydomonas* chloroplast revealed by in situ cryo-electron tomography.

630 *Elife* **4**: e04889.

631 **Esquivel MG, Genkov T, Nogueira AS, Salvucci ME, Spreitzer RJ. 2013.** Substitutions at

632 the opening of the Rubisco central solvent channel affect holoenzyme stability and CO_2/O_2

633 specificity but not activation by Rubisco activase. *Photosynthesis research* **118**: 209-218.

634 **Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR. 2004.**

635 The evolution of modern eukaryotic phytoplankton. *Science* **305**: 354-360.

636 **Falkowski PG, Raven JA. 2007.** Photosynthesis and primary production in nature. *Aquatic*

637 *photosynthesis 2nd ed*. Princeton University Press, Princeton.

638 **Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. 1998.** Primary production of the

639 biosphere: integrating terrestrial and oceanic components. *Science* **281**: 237-240.

640 **Finet C, Timme RE, Delwiche CF, Marlétaz F. 2010.** Multigene phylogeny of the green

641 lineage reveals the origin and diversification of land plants. *Current Biology* **20**: 2217-2222.

642 **Franks PJ, Beerling DJ. 2009.** Maximum leaf conductance driven by CO_2 effects on stomatal

643 size and density over geologic time. *Proceedings of the National Academy of Sciences* **106**:

644 10343-10347.

645 **Galmés J, Kapralov MV, Andralojc PJ, Conesa MÀ, Keys AJ, Parry MA, Flexas J. 2014.**

646 Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and

647 evolutionary trends. *Plant, Cell & Environment* **37**: 1989-2001.

648 **Galmés J, Kapralov MV, Copolovici LO, Hermida-Carrera C, Niinemets Ü. 2015.**

649 Temperature responses of the Rubisco maximum carboxylase activity across domains of life:

650 phylogenetic signals, trade-offs, and importance for carbon gain. *Photosynthesis research* **123**:

651 183-201.

652 **Galmés J, Hermida-Carrera C, Laanisto L, Niinemets Ü. 2016.** A compendium of

653 temperature responses of Rubisco kinetic traits: variability among and within photosynthetic

654 groups and impacts on photosynthesis modeling. *Journal of Experimental Botany* **67**: 5067-

655 5091.

656 **Galmés J, Capó-Bauçà S, Niinemets Ü, Iñiguez C. 2019.** Potential improvement of

657 photosynthetic CO_2 assimilation in crops by exploiting the natural variation in the temperature

658 response of Rubisco catalytic traits. *Current opinion in plant biology* **49**: 60-67.

659 **Genkov T, Spreitzer RJ. 2009.** Highly conserved small subunit residues influence RuBisCO

660 large subunit catalysis. *Journal of Biological Chemistry* **284**: 30105-30112.

661 **Giordano M, Beardall J, Raven JA. 2005.** CO_2 concentrating mechanisms in algae:

662 mechanisms, environmental modulation, and evolution. *Annual Review of Plant Biology* **56**:

663 99-131.

664 **Goldschmidt-Clermont M, Rahire M. 1986.** Sequence, evolution and differential expression

665 of the two genes encoding variant small subunits of ribulose bisphosphate

666 carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *Journal of Molecular Biology* **191**:

667 421-432.

668 **Griffiths H, Robe WE, Girnus J, Maxwell K. 2008.** Leaf succulence determines the interplay

669 between carboxylase systems and light use during Crassulacean acid metabolism in Kalanchoë

670 species. *Journal of Experimental Botany* **59**: 1851-1861.

671 **Griffiths H, Meyer MT, Rickaby REM. 2017.** Overcoming adversity through diversity:
672 aquatic carbon concentrating mechanisms. 3689-3695.

673 **Harholt J, Moestrup Ø, Ulvskov P. 2016.** Why plants were terrestrial from the beginning.
674 *Trends in Plant Science* **21**: 96-101.

675 **Haworth M, Elliott-Kingston C, McElwain JC. 2011.** Stomatal control as a driver of plant
676 evolution. *Journal of Experimental Botany* **62**: 2419-2423.

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

680 **Hoffmann L. 1989.** Algae of terrestrial habitats. *The botanical review* **55**: 77-105.

681 **Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T,**
682 **Mori H, Tajiima N *et al*. 2014.** *Klebsormidium flaccidum* genome reveals primary factors for
683 plant terrestrial adaptation. *Nature communications* **5**: 3978.

684 **Jordan DB, Ogren WL. 1981.** Species variation in the specificity of ribulose biphosphate
685 carboxylase/oxygenase. *Nature* **291**: 513-515.

686 **Jordan DB, Ogren WL. 1984.** The CO_2/O_2 specificity of ribulose 1, 5-bisphosphate
687 carboxylase/oxygenase. *Planta* **161**: 308-313.

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

691 **Kapralov MV, Filatov DA. 2007.** Widespread positive selection in the photosynthetic
692 Rubisco enzyme. *BMC Evolutionary Biology* **7**: 73.

693 **Kapralov MV, Kubien DS, Andersson I, Filatov DA. 2010.** Changes in Rubisco kinetics
694 during the evolution of C4 photosynthesis in *Flaveria* (Asteraceae) are associated with positive
695 selection on genes encoding the enzyme. *Molecular Biology and Evolution* **28**: 1491-1503.

696 **Kubien DS, Whitney SM, Moore PV, Jesson LK. 2008.** The biochemistry of Rubisco in
697 *Flaveria. Journal of Experimental Botany* **59**: 1767-1777.

698 **Le SQ, Gascuel O. 2008.** An improved general amino acid replacement matrix. *Molecular*
699 *biology and evolution* **25**: 1307-1320.

700 **Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, Yue R, Robertson JM, Lefebvre**
701 **PA, Fitz-Gibbon ST, Grossman AR, Jonikas MC. 2016.** An indexed, mapped mutant library
702 enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii. Plant*
703 *Cell* **28**: 367-387.

704 **Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA,**
705 **Graham SW, Grosse I, Li Z, Melkonian M, Mirarab S *et al.* 2019.** One thousand plant
706 transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679-685.

707 **Leliaert F, Verbruggen H, Zechman FW. 2011.** Into the deep: new discoveries at the base
708 of the green plant phylogeny. *BioEssays* **33**: 683-692.

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

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

715 **Losh JL, Young JN, Morel FM. 2013.** Rubisco is a small fraction of total protein in marine
716 phytoplankton. *New Phytologist* **198**: 52-58.

717 **Lucas WJ, Berry JA. 1985.** Inorganic carbon transport in aquatic photosynthetic organisms.
718 *Physiologia plantarum* **65**: 539-543.

719 **Mackinder LC, Meyer MT, Mettler-Altmann T, Chen VK, Mitchell MC, Caspari O,**
720 **Freeman Rosenzweig ES, Pallesen L, Reeves G, Itakura A *et al*. 2016.** A repeat protein
721 links RuBisCO to form the eukaryotic carbon- concentrating organelle. *Proceedings of the*
722 *National Academy of Sciences, USA* **113**: 5958-5963.

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

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

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

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

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

735 **Meyer MT, Genkov T, Skepper JN, Jouhet J, Mitchell MC, Spreitzer RJ, Griffiths H.**
736 **2012.** RuBisCO small-subunit α-helices control pyrenoid formation in *Chlamydomonas*.
737 *Proceedings of the National Academy of Sciences, USA* **109**: 19474-19479.

738 **Meyer MT, Griffiths H. 2013.** Origins and diversity of eukaryotic CO_2-concentrating
739 mechanisms: lessons for the future. *Journal of experimental botany* **64**: 769-786.

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

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

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

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

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

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

758 **Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK,**
759 **Haas FB, Vanderstraeten L, Becker D, Lang D *et al*. 2018.** The *Chara* genome: secondary
760 complexity and implications for plant terrestrialization. *Cell* **174**: 448-464.

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

764 **O'Leary MH. 1988.** Carbon isotopes in photosynthesis. *Bioscience* **38**: 328-336.

765 **Oltrogge LM, Chaijarasphong T, Chen AW, Bolin ER, Marqusee S, Savage DF. 2019.** α-
766 carboxysome formation is mediated by the multivalent and disordered protein CsoS2.
767 **doi:** https://doi.org/10.1101/708164.

768 **Orr DJ, Carmo-Silva E. 2018.** Extraction of Rubisco to determine catalytic constants.
769 *Covshoff S (ed) Photosynthesis: methods and protocols, Methods in molecular biology, vol*
770 *1770*. Springer, New York.

771 **Palmqvist K, Sültemeyer D, Baldet P, Andrews TJ, Badger MR. 1995.** Characterisation of

772 inorganic carbon fluxes, carbonic anhydrase(s) and ribulose-1, 5-biphosphate carboxylase-

773 oxygenase in the green unicellular alga Coccomyxa. *Planta* **197**: 352-361.

774 **Prins A, Orr DJ, Andralojc PJ, Reynolds MP, Carmo-Silva E, Parry MA. 2016.** Rubisco

775 catalytic properties of wild and domesticated relatives provide scope for improving wheat

776 photosynthesis. *Journal of Experimental Botany* **67**: 1827-1838.

777 **Pröschold T, Marin B, Schlösser UG, Melkonian, M. 2001.** Molecular phylogeny and

778 taxonomic revision of *Chlamydomonas* (Chlorophyta). I. Emendation of *Chlamydomonas*

779 Ehrenberg and *Chloromonas* Gobi, and description of *Oogamochlamys* gen. nov. and

780 *Lobochlamys* gen. nov. *Protist* **152**: 265-300.

781 **Rambaut A. 2007.** FigTree, a graphical viewer of phylogenetic trees. URL *http://tree. bio.*

782 *ed. ac. uk/software/figtree.*

783 **Raven J, Beardall J, Griffiths H. 1982.** Inorganic C-sources for *Lemanea, Cladophora* and

784 *Ranunculus* in a fast-flowing stream: measurements of gas exchange and of carbon isotope

785 ratio and their ecological implications. *Oecologia* **53**: 68-78.

786 **Raven JA, Osborne BA, Johnston AM. 1985.** Uptake of CO_2 by aquatic vegetation. *Plant,*

787 *Cell & Environment* **8**: 417-425.

788 **Raven JA, Ball LA, Beardall J, Giordano M, Maberly SC. 2005.** Algae lacking carbon-

789 concentrating mechanisms. *Canadian Journal of Botany* **83**: 879-890.

790 **Rickaby REM, Hubbard MRE. 2019.** Upper ocean oxygenation, evolution of RuBisCO and

791 the Phanerozoic succession of phytoplankton. *Free Radical Biology and Medicine* **140**: 295-

792 304.

793 **Rindi F, Mikhailyuk TI, Sluiman HJ, Friedl T, López-Bautista JM. 2011.** Phylogenetic

794 relationships in *Interfilum* and *Klebsormidium* (Klebsormidiophyceae, Streptophyta).

795 *Molecular phylogenetics and evolution* **58**: 218-231.

796 **Rosenzweig ESF, Xu B, Cuellar LK, Martinez-Sanchez A, Schaffer M, Strauss M,**

797 **Cartwright HN, Ronceray P, Plitzko JM, Förster F *et al*. 2017.** The eukaryotic CO_2

798 concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**: 148-162.

799 **Sáez PL, Bravo LA, Cavieres LA, Vallejos V, Sanhueza C, Font-Carrascosa M, Gil-**

800 **Pelegrin E, Peguero-Pina JJ, Galmés J. 2017.** Photosynthetic limitations in two Antarctic

801 vascular plants: importance of leaf anatomical traits and Rubisco kinetic parameters. *Journal*

802 *of experimental botany* **68**: 2871-2883.

803 **Sage RF. 2002.** Variation in the kcat of Rubisco in C_3 and C_4 plants and some implications for
804 photosynthetic performance at high and low temperature. *Journal of Experimental Botany* **53**:
805 609-620.

806 **Sage RF, Christin PA, Edwards EJ. 2011.** The C_4 plant lineages of planet Earth. *Journal of*
807 *Experimental botany* **62**: 3155-3169.

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

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

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

816 **Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H,**
817 **Remmert M, Söding J, Thompson JD, Higgins DG. 2011.** Fast, scalable generation of high-
818 quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*
819 **7**: 539.

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

824 **Spreitzer RJ, Salvucci ME. 2002.** RuBisCO: structure, regulatory interactions, and
825 possibilities for a better enzyme. *Annual review of plant biology* **53**: 449-475.

826 **Spreitzer RJ. 2003.** Role of the small subunit in ribulose-1,5-bisphosphate
827 carboxylase/oxygenase. *Archives of Biochemistry and Biophysics* **414**: 41-149.

828 **Spreitzer RJ, Peddi SR, Satagopan S. 2005.** Phylogenetic engineering at an interface
829 between large and small subunits imparts land-plant kinetic properties to algal RuBisCO.
830 *Proceedings of the National Academy of Sciences, USA* **102**: 17225-17230.

831 **Stabenau H, Winkler U. 2005.** Glycolate metabolism in green algae. *Physiologia Plantarum*
832 **123**: 235-245.

833 **Tavaré S. 1986.** Some probabilistic and statistical problems in the analysis of DNA sequences.
834 *Lectures on mathematics in the life sciences* **17**: 57-86.

835 **Tcherkez, G, Farquhar GD, Andrews TJ. 2006.** Despite slow catalysis and confused

836 substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized.

837 *Proceedings of the National Academy of Sciences, USA* **103**: 7246–7251.

838 **Tcherkez G. 2013.** Modelling the reaction mechanism of ribulose-1, 5-bisphosphate

839 carboxylase/oxygenase and consequences for kinetic parameters. *Plant, Cell & Environment*

840 **36**: 1586-1596.

841 **Tortell PD. 2000.** Evolutionary and ecological perspectives on carbon acquisition in

842 phytoplankton. *Limnology and Oceanography* **45**: 744–750.

843 **Valegård K, Andralojc PJ, Haslam RP, Pearce FG, Eriksen GK, Madgwick PJ,**

844 **Kristoffersen AK, van Lun M, Klein U, Eilertsen HC *et al*. 2018.** Structural and functional

845 analyses of Rubisco from arctic diatom species reveal unusual posttranslational modifications.

846 *Journal of Biological Chemistry* **293**: 13033-13043.

847 **Von Caemmerer S, Quick WP. 2000.** Rubisco: physiology in vivo. *Photosynthesis*. Springer,

848 Dordrecht.

849 **Wang H, Yan X, Aigner H, Bracher A, Nguyen ND, Hee WY, Long BM, Price GD, Hartl**

850 **FU, Hayer-Hartl M. 2019.** Rubisco condensate formation by CcmM in β-carboxysome

851 biogenesis. *Nature* **566**: 131-135.

852 **Wang L, Yamano T, Kajikawa M, Hirono M, Fukuzawa H. 2014.** Isolation and

853 characterization of novel high-CO_2-requiring mutants of *Chlamydomonas reinhardtii*.

854 *Photosynthesis research* **121**: 175-184.

855 **Warren SD, Clair LLS, Stark LR, Lewis LA, Pombubpa N, Kurbessoian T, Stajich JE,**

856 **Aanderud ZT. 2019.** Reproduction and dispersal of biological soil crust organisms. *Frontiers*

857 *In Ecology Evolution* **7**: 344.

858 **Wunder T, Cheng SLH, Lai SK, Li HY, Mueller-Cajar O. 2018.** The phase separation

859 underlying the pyrenoid-based microalgal Rubisco supercharger. *Nature communications* **9**:

860 5076.

861 **Yamada K, Davydov II, Besnard G, Salamin N. 2019.** Duplication history and molecular

862 evolution of the *rbcS* multigene family in angiosperms. *Journal of experimental botany*.

863 **Yamano T, Sato E, Iguchi H, Fukuda Y, Fukuzawa H. 2015.** Characterization of

864 cooperative bicarbonate uptake into chloroplast stroma in the green alga *Chlamydomonas*

865 *reinhardtii*. *Proceedings of the National Academy of Sciences, USA* **112**: 7315-7320.

866 **Yang Z. 2007.** PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology*

867 *and evolution* **24**: 1586-1591.

868 **Yoon HS, Hackett JD, Pinto G, Bhattacharya D. 2002.** The single, ancient origin of chromist
869 plastids. *Journal of phycology* **38**: 40-40.

870 **Young JN, Rickaby REM, Kapralov MV, Filatov DA. 2012.** Adaptive signals in algal
871 Rubisco reveal a history of ancient atmospheric carbon dioxide. *Philosophical Transactions of*
872 *the Royal Society B: Biological Sciences* **367**: 483-492.

873 **Young JN, Heureux AM, Sharwood RE, Rickaby RE, Morel FM, Whitney SM. 2016.**
874 Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-
875 concentrating mechanisms. *Journal of Experimental Botany* **67**: 3445-3456.

876 **Zones JM, Blaby IK, Merchant SS, Umen JG 2015.** High-resolution profiling of a
877 synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell
878 and metabolic differentiation. *Plant Cell* **27**: 2743-2769.

879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901

902

903

904

905

906

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* 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*

936 (from McKay *et al*., 1991), g: *Chlamydomonas reinhardtii*). Three distinct pyrenoid

937 morphologies can be observed: matrix enclosed by one layer of starch plates (b, d, e, f and g);

938 matrix enclosed by multiple starch grains (c); and pyrenoid without observable starch sheath

939 (a). Bars: 2 μm (a to e) and 0.5 μm (f and g).

940

941 **Table 1** Results of the three Likelihood Ratio Tests (LRTs) for positive selection using the

942 site-models (M0-M8) (codeml) implemented in PAML (Yang, 2007) and their associated

943 parameters.

944

945

	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			

946
947 a: number of sequences analysed
948 b: length of *RbcS* sequences analysed
949 c: degrees of freedom
950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

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.

972

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

973

a: omega for background branches
b omega for foreground branches

976

977

978

979

980

981

982

983

984

985
986
987
988
989

990
991
992
993
994
995
996

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 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.

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

997 **Table 4** Whole cell affinity for inorganic carbon in the six streptophyte algae representative

998 species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under low CO_2 conditions (10

999 μM) and their associated δ13C for organic matter. Species are ordered from the furthest species

1000 away from land plants (*Chlamydomonas reinhardtii*, Chlorophytes, Chlorophyceae) to the

1001 closest (*Coleochaete scutata*, Coleochaetophyceae, Charophytes). Values are means ± SEM.

1002 n=number of replicates

1003

Species name	$K_{0.5}(Ci)$ (μM)	$δ^{13}C$ (‰)
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

1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023

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, 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 endosymbiosis is indicated by a red arrow and dashed lines represent uncertain relationships.

Fig. S2 Phylogenetic tree of 64 green algae species. 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. Streptophyte algae were labelled in orange, prasinophytes in blue and chlorophyte in green.

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). Acid polar residues are in yellow, basic 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.

Fig. S4 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. Cluster A 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.

Table S1 Pyrenoid diagnostic for all the species present in the phylogeny of *RbcS* and the associated references. Species without pyrenoid are highlighted in light grey.

Table S2 Growth media and accession number of the six streptophyte algae.

1057 **Table S3** Classification and habitat description of the six streptophyte algae**.**

1058

1059 **Table S4** Whole-cell affinity for inorganic carbon in the six streptophyte algae representative

1060 species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under high CO_2 conditions (5%

1061 CO_2) and their associated $\delta^{13}C$ for organic matter. Species are ordered from the furthest species

1062 (*Chlamydomonas reinhardtii*, Chlorophytes, Chlorophyceae) away from land plants to the

1063 closest (*Coleochaete scutata*, Coleochaetophyceae, Streptophytes).Values are means ± SEM.

1064

1065
1066

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* amino acid position is taken as reference).
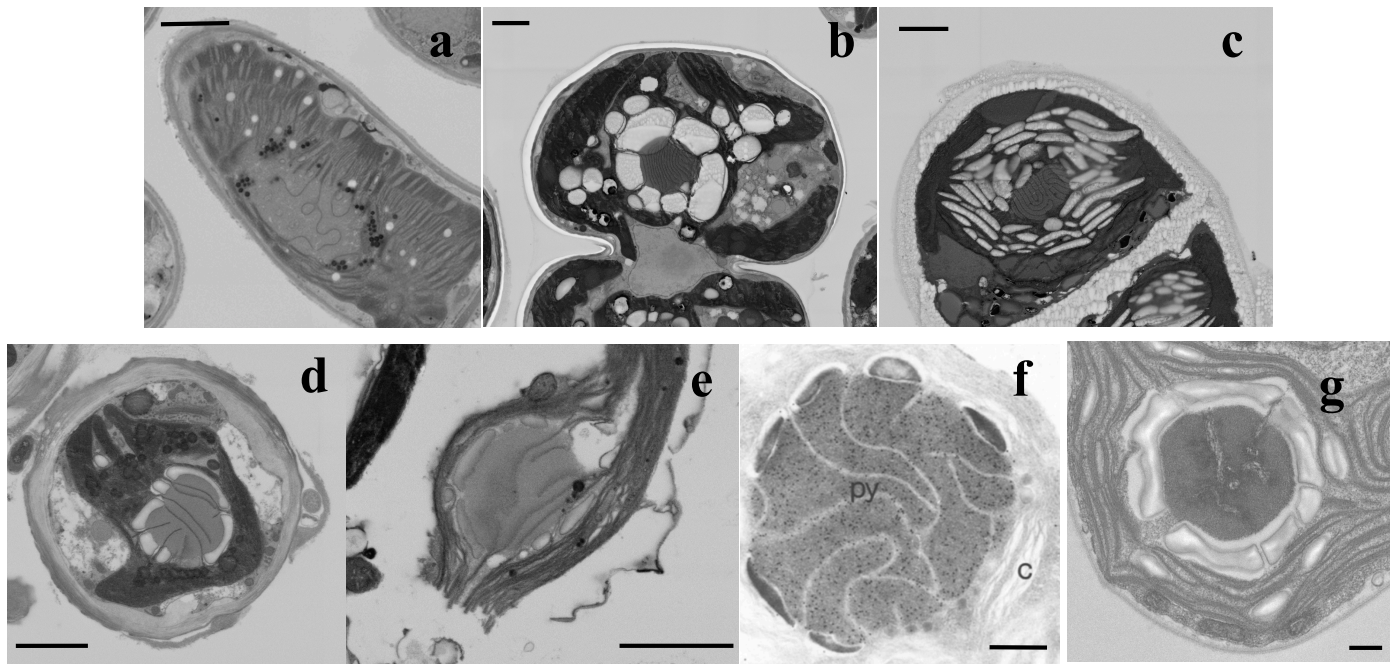
Fig. 3 Electron micrographs of the six representative streptophyte algae and of *Chlamydomonas reinhardtii* (a: *Klebsormidium subtile*, b: *Cosmarium subtumidum*, c: *Chlorokybus atmophyticus*, d: *Onychonema laeve*, e: *Spirogyra* sp., f: *Coleochaete scutata* (from McKay *et al*., 1991), g: *Chlamydomonas 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).