Rubisco and carbon concentrating mechanism (CCM) co-evolution across Chlorophyte and Streptophyte green algae

SCHOLARONE™
Manuscripts

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]Botanical

7   Institute, Department of Biological Sciences, Universität zu Köln, Köln D-50674, Germany;

8   [4]Cambridge Advanced Imaging Centre, University of Cambridge, Cambridge, CB2 3DY, UK;

9   [5]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA

10

11   Author for correspondence:

12   *Myriam M.M. Goudet*

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

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

15

16   And

17

18   *Prof. Howard Griffiths*

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

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

21

Total word count (excluding summary, references and legends):	6499	No. of figures	3 (Fig. 1- 2 in colour)
Summary:	198	No. of Tables:	4
Introduction	1088	No. of Supporting Information files:	7 (Fig. S1-3; Table S1-S4)
Materials and Methods	1552		
Results:	1913		
Discussion:	1946		
Acknowledgements:	132		

22

23 Summary

24 • Green algae expressing a Carbon Concentrating Mechanism (CCM) are usually
25 associated with a Rubisco-containing micro-compartment, the pyrenoid. A link
26 between the small subunit (SSU) of Rubisco and pyrenoid formation in
27 *Chlamydomonas reinhardtii* has previously suggested that specific *Rbc*S residues could
28 explain pyrenoid occurrence in green algae.

29 • A phylogeny of *Rbc*S was used to compare the protein sequence and CCM distribution
30 across the green algae and positive selection in *Rbc*S was estimated. Six streptophyte
31 algae, Rubisco catalytic properties, affinity for CO_2 uptake ($K_{0.5}$), carbon isotope
32 discrimination ($\delta^{13}C$) and pyrenoid morphology were compared.

33 • The *Rbc*S sequence did not correlate with CCM occurrence, but the length of the βA-
34 βB loop discriminated chlorophyte from streptophyte green algae, with prasinophytes
35 representing an intermediate group. Rubisco catalytic traits in streptophyte algae ranged
36 between values typical for algae to those of embryophytes and correlated well with
37 CCM activity, $\delta^{13}C$ and pyrenoid ultrastructure.

38 • We conclude that the Rubisco catalytic properties found in streptophyte algae reflect
39 the strength of any CCM and pyrenoid leakiness, with selective pressures associated
40 with the availability of inorganic carbon in the aquatic habitat, whereas Rubisco in
41 extant land plants reflects more recent selective pressures associated with the terrestrial
42 environment.

43 Key words: carbon concentrating mechanism (CCM), green algae, photosynthesis, pyrenoid,
44 Rubisco, streptophyte algae,

45 **Introduction**

46 Photoautotrophic organisms globally fix $111\text{-}117\times10^{15}$ grams of carbon per year and around
47 half of this global net primary production is aquatic (Behrenfeld *et al*., 2001; Field *et al*., 1998),
48 with green algae a major contributor to this global carbon fixation. Among green algae, the
49 streptophytes demonstrate a wide range of ultrastructural and developmental traits closely
50 related to land plants. However streptophyte algae and chlorophytes remain subject to key
51 limitations in the aquatic milieu (low CO_2 diffusion and availability, light limitation; Borges &
52 Frankignoulle, 2002; Yamano *et al*., 2015).

53   The chloroplast gene (*rbcL*) encoding the large subunit (LSU) of the primary carboxylase

54   Rubisco (ribulose 1,5-bisphosphate carboxylase/oxygenase; Spreitzer & Salvucci, 2002;) as

55   well as transcriptome data have helped to resolve green algal inter-relationships: an early split

56   after the primary endosymbiosis saw the diversification of the hypothetical ancestral flagellate

57   into two main lineages (Leliaert *et al.*, 2011). On one side, the chlorophytes diversified early

58   as prasinophytes in marine waters, which then gave rise to the Core chlorophytes (chlorophytes

59   without prasinophytes, Fig. S2, Supplementary Materials) in both fresh and marine waters. The

60   streptophytes (which include both embryophytes and streptophyte algae) diversified in fresh

61   water but also in some subaerial/terrestrial habitats (Harholt *et al.*, 2016). The split between

62   Chlorophyte and Streptophyte probably occurred during the Neoproterozoic (between 1,000

63   and 541 million years ago; Becker, 2013; Del Cortona *et al.*, 2019). Selection pressures on the

64   Rubisco holoenzyme catalytic properties are driven by the availability and diffusive supply of

65   inorganic carbon, the CO_2:O_2 ratio and the development of any carbon concentrating

66   mechanism (CCM) which improves the operating efficiency of Rubisco in many algae (Meyer

67   & Griffiths 2013). The origins of the algal CCM could be related to equimolar CO_2:O_2

68   concentrations in surface waters around 500 million years ago (Griffiths *et al.*, 2017).

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

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

71   conditions, the concentration of CO_2 is often 2,200 times lower in water than in air, and

72   diffusion is also 8,000 times slower (Raven *et al.*, 1985; Falkowski & Raven, 2007; Young *et*

73   *al.*, 2012). A CCM is typically associated with transmembrane inorganic carbon transporters,

74   and a specific carbonic anhydrase (CA) for conversion of HCO_3^- to elevated CO_2

75   concentrations within the Rubisco matrix forming the pyrenoid. The latter microcompartment

76   is often traversed by thylakoid tubules, and in some green algae the pyrenoid is demarcated by

77   a starch sheath (Meyer *et al.*, 2017).

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

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

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

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

82   2015; Mackinder *et al.*, 2017; Meyer *et al.*, 2017; Mukherjee *et al.*, 2019). The CCM is

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

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

85   Freeman-Rosensweig *et al*., 2017), primarily via an interaction with Rubisco Small Subunit

86   (SSU) (Wunder *et al*., 2018; Atkinson *et al*., 2019), presumably situated at the level of surface

87   exposed α-helices (Meyer *et al*., 2012). However, there has been little systematic analysis of

88   the extent to which some form of carbon accumulation mechanism occurs across this

89   chlorophyte clade, or comparative physiological and molecular studies on CCM characteristics

90   or Rubisco kinetic properties, and whether these traits are captured across chlorophyte species,

91   prasinophyte and streptophyte algal lineages in *Rbc*S.

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

93   between Rubisco LSU, SSU and catalytic properties. The eight identical 55-kDa large subunits

94   assemble as four dimers, while two sets of four 15-kDa small subunits, top and tail the Rubisco

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

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

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

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

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

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

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

102   *et al*., 1998; Meyer & Griffiths, 2013). However, Meyer & Griffiths (2013) suggested that

103   selective pressures on V_C and K_c could have been relaxed due to the saturating CO_2

104   environment provided by a CCM over evolutionary time.

105   The overall aim of this study was to address the possible interactions between Rubisco SSU

106   structure and phylogeny, and occurrence of any reported CCM or pyrenoid across the green

107   algae. Additionally, we set out to define key Rubisco catalytic properties for selected

108   streptophyte algae, as compared to *Chlamydomonas reinhardtii*. Surprisingly, no model

109   organisms for physiological studies have been identified in streptophyte algae, despite the

110   previous interest in using species with giant algal cells to characterise carbon uptake

111   mechanisms (Lucas & Berry, 985). In addition, only few Rubisco catalytic properties are

112   available for green alga species including *Euglena gracilis* (Yokota *et al*., 1989), *Coccomyxa*

113   *sp.* (Pamlqvist *et al*., 1995) or *Scenedesmus obliquus* (Jordan & Ogren, 1981; Badger *et al*.,

114   1998) but none of them are streptophyte alga. Recent measurements have largely focussed on

115   embryophytes (Kapralov *et al*., 2010; Galmes *et al.,* 2014, 2015, 2016; Hermida-Carrera *et al*.,

116   2016; Orr *et al.*, 2016; Prins *et al.*, 2016) or Core chlorophytes (Jordan & Ogren, 1981;

117   Spreitzer, 2003; Spreitzer *et al.*, 2005).

118   Specifically, this study sought to *(i)* establish a phylogeny of *Rbc*S sequences in green algae,

119   and compare the distribution of pyrenoid and CCM across the algal clades; *(ii)* to identify

120   whether any selection pressure on residues within the SSU were associated with the broader

121   phylogeny or were related to CCM activity and, *(iii)* to determine whether the catalytic

122   properties of Rubisco across contrasting streptophyte algal groups reflected the overall

123   phylogeny or specific activity of a CCM at the whole organism level. Our results reveal that

124   the division between Core chlorophytes and streptophyte algae in *Rbc*S is defined by a change

125   in SSU secondary structure but also highlight a more complex relationship between Rubisco

126   catalytic properties and CCM activity. This study also provides additional insights for selection

127   pressures driving the evolution of green algae and photosynthetic processes, particularly during

128   the transition to terrestrial plant life forms.

129   **Materials and Methods**

130

131   **Collection of protein sequences, phylogenetic analysis, βA-βB loop length and pyrenoid**

132   **presence/absence mapping**

133   2,674 protein *Rbc*S sequences of green algae were kindly provided by «The 1000 plants

134   project» (1KP; Leebens-Mack *et al.*, 2019). All the protein sequences were manually and

135   individually screened. Sequences showing cross-contamination (Carpenter *et al.*, 2019), or

136   which were too short or incomplete, were removed. The dataset did now allow to

137   unambiguously identify *Rbc*S isoforms. Although it is generally taken that all photosynthetic

138   members of the Viridiplantae have multiple copies of the *Rbc*S gene, conservatively only one

139   sequence was used in the analysis for each species, except when the data was sourced from

140   independently sequenced genomes (e.g. for *Asteromonas*). A total of 187 protein sequences

141   belonging to 113 species (31 streptophyte algae, 10 prasinophytes, 72 chlorophytes) were then

142   aligned with Clustal Omega (Sievers *et al.*, 2011). ProTest v2.4 (Abascal *et al*, 2005) was used

143   to identify the best model of protein evolution. Bayesian phylogenetic analyses were performed

144   using BEAST v2.3.1 (Boukaert *et al.*, 2014) with a LG model of protein evolution (Le &

145   Gascuel, 2008), a gamma distribution model with four categories, a relaxed molecular clock

146   and finally with a Yule model of speciation. Three independent chains were run, each of length

147   8×10^7 steps, parameters values and trees were sampled every 10×10^2 steps. Chain convergences

148   were checked using Tracer v1.6 (Drummond & Rambaut, 2007). Posterior parameters were

149   summarized with Tree Annotator v1.8.2 (Drummond & Rambaut, 2007) using a maximum

150   clade credibility tree (MCC) and a posterior limit of 0.5. Figtree v1.4.2 (Rambaut, 2007) was

151   used for tree visualizations. The length of the βA-βB loop was determined after the analysis of

152   the protein sequences, with the number of residues in the loop (Spreitzer, 2003) mapped on to

153   the phylogeny of *Rbc*S. Finally, the same phylogeny was used to map the pyrenoid

154   presence/absence. The scoring for pyrenoid presence/absence was based on the available

155   literature (Table S4, Supplementary Materials).

Likelihood ratio test for positive selection

157   To test the importance of two SSU α-helices for pyrenoid formation in *C. reinhardtii* (Meyer

158   *et al.*, 2012), the Codon-based package (codeml) implemented in PAML v4.9 (Yang, 2007)

159   was used to detect residues under positive selection across the green algae lineage. In addition,

160   the presence of a CCM is not universal across the green algae so the branch model also

161   implemented in PAML was used to detect branches under positive selection. All the analyses

162   were performed using "user tree" mode. The DNA phylogenetic tree was reconstructed using

163   BEAST v2.3.1 with 135 cDNA *Rbc*S sequences of green algae from the 1KP, with a GTR

164   model of protein evolution (Tavaré, 1986) and the same gamma distribution, molecular clock

165   and model of speciation previously used. Three independent chains were run, each of length

166   5×10^7 steps, parameters values and trees were sampled every 10×10^2 steps. Chain

167   convergences, posterior parameters and tree visualization were analysed with the same method

168   explained above. Several models of codon evolution that allow for variations in ω (dN/dS)

169   among codons were tested (Site model) and evaluated using Likelihood Ratio Tests (LRTs)

170   (Neyman & Pearson, 1928) as described in Kapralov & Filatov (2007). Branch models were

171   used to test for positive selection across branches. The null model allowed for variations in ω

172   among branches (0<dN/dS<1 and dN/dS=1 for both foreground and background branches) and

173   also included two additional classes of codons with fixed dN/dS=1 on foreground branches but

174   restricted as 0< dN/dS<1 and dN/dS=1 for background branches. The alternative model allowed

175   0<dN/dS<1 and dN/dS=1 for both foreground and background branches but also included two

176   additional classes of codons under positive selection with dN/dS>1 on foreground branches

177   with restriction as 0<dN/dS<1 and dN/dS=1 on background branches. Branches leading to

178   species without pyrenoid were labelled as foreground branches (allows positive selection) and

179   the rest of the branches were considered as background branches (with no positive selection).

180   The level of significance was tested as described above.

181 **Streptophyte algae culturing, Rubisco purification and Rubisco catalytic properties**

182 Six streptophyte algae (from the Chlorokybophyceae to Coleochaetophyceae; Table S1-3; Fig.

183 S2, Supplementary Materials) were ordered from the Culture Collection of Algae at Göttingen.

184 These consisted of: *Chlorokybus atmophyticus* (Chlorokybophyceae), *Klebsormidium subtile*

185 (Klebsormidiophyceae), *Cosmarium subtumidum*, *Onychonema laeve*, *Spirogyra sp.*

186 (Zygnematophyceae) and *Coleochaete scutata* (Coleochaetophyceae). The wild type

187 *Chlamydomonas reinhardtii* (strain CC-4533, Li *et al.*, 2016) was used as control to test

188 protocols since the Rubisco catalytic properties are well characterised (Jordan & Ogren, 1981;

189 Genkov & Spreitzer, 2009). Strains were cultured in an incubator shaker (Innova 42, New

190 Brunswick Scientific) under constant agitation (130 RPM) in the recommended medium (Table

191 S1, Supplementary Materials), in 2L conical flasks, under constant light at 20°C and bubbled

192 with ambient air. Due to the low concentration of Rubisco in algae (Losh *et al.*, 2013; Valegård

193 *et al.*, 2018) a minimum of 30g wet paste per sample was harvested in order to have enough

194 material for the Rubisco extraction and purification.

195 Algal cells were broken using an Emulsiflex-C5 high pressure homogenizer (Avestin Inc.,

196 Ottawa, Canada) kindly loaned by Biopharma Group (Winchester, UK). Cell pastes were re-

197 suspended in *ca.* 200 mL of extraction buffer containing 10 mM $MgCl_2$, 50 mM Bicine, 10

198 mM $NaHCO_3$, 1 mM DTT, 1 mM ε-aminocaproic acid, 1 mM benzamidine, 0.1 M

199 phenylmethylsulfonyl fluoride, and 200 µL of protease inhibitor cocktail (Sigma, UK). Total

200 soluble proteins were extracted via centrifugation at 22,000 ×*g* for 12 minutes (min) at 4°C.

201 After this initial centrifugation step, PEG 4000 (60% w/v) and 1 M $MgCl_2$ were added to the

202 supernatant and the rest of the purification carried out as described previously (Orr & Carmo-

203 Silva, 2018). Peak fractions containing Rubisco (based on CABP binding [Sharwood *et al.*,

204 2016]) were concentrated using Amicon Ultracel-15 concentrators (100 kDa MWCO, Merck-

205 Millipore, UK). Aliquots were snap-frozen in liquid nitrogen and stored at -80°C.

206 Rubisco activity for the six streptophyte algae was determined by incorporation of $H^{14}CO_3$ into

207 acid-stable products at 25°C as described in Prins *et al.* (2016) with some modifications.

208 Purified Rubisco was diluted using desalting buffer (Orr & Carmo-Silva, 2018) and then

209 desalted using a G-25 MidiTrap column (GE Healthcare, UK). Samples were allowed to

210 activate on ice for 45 mins prior to assaying. Carboxylation activity was measured at nine

211 different concentrations of CO_2 (8, 16, 24, 36, 68, 100, 180, 280 and 400 µM) and with O_2

212 concentrations of 0 and 21%. In order to ensure that the activity measured was entirely due to

213   Rubisco, three controls were performed: CO_2 fixation (acid-stable ^{14}C) was measured in

214   reaction solutions lacking RuBP or $NaHCO_3$, and following total inhibition of Rubisco by prior

215   treatment with an excess of the tight-binding inhibitor 2-carboxyarabinitol-1,5-bisphosphate

216   (CABP). Radioactive content of ^{14}C-labelled compounds was measured in 0.4 ml aqueous

217   solutions to which were added 3.6 ml Gold Star Quanta Scintillation cocktail (Meridian

218   Biotechnologies, UK), in a Tri-Carb 2250 CA Liquid Scintillation Analyser (Perkin-Elmer,

219   USA). Turnover number (k_{cat}: mol product mol active site^{-1} s^{-1}) was calculated from the

220   corresponding V_{max} value (V_c: μmol acid-stable ^{14}C mg Rubisco^{-1} min^{-1}).

221   Rubisco quantification was via [^{14}C]CABP binding assay as described Sharwood *et al*. (2016).

222   Rubisco was incubated for 25 min after adding [^{14}C]CABP. Each quantification was performed

223   in duplicate. Radioactive content of ^{14}C-labelled compounds was measured using scintillation

224   counting as described above.

225   **Photosynthetic affinity for inorganic carbon**

226   Apparent affinity for inorganic carbon (Ci) was determined by oxygen evolution (Badger *et*

227   *al*., 1980) and as described in Mitchell *et al*. (2014). Five extra concentrations were added in

228   cultures grown in high CO_2 condition in order to reach maximum rate of oxygen evolution

229   (2500, 3000, 4000, 4500 and 5000 μM). Chlorophyll *a* and *b* concentrations were measured

230   for normalization of oxygen evolution measurements as described in Mitchell *et al*. (2014).

231   **Carbon isotope analysis**

232   Algae cultures were grown under low and high CO_2 conditions and were harvested by

233   centrifugation at 4,200 rpm for 5 minutes at 20°C (Eppendorf, Centrifuge 5804 R), resuspended

234   in 0.1M HCl to remove inorganic carbon and washed several times with deionized water.

235   Samples were dried in a freeze drier overnight and weighed (0.5 mg) in triplicate into 3mm x

236   5mm tin capsules (Experimental Microanalysis Ltd., Okehampton, UK). The results were

237   reported with reference to the international standard VPDB with a precision better than +/- 0.08

238   per mil for $^{12}C/^{13}C$. All the analyses were performed at the Godwin Laboratory for Paleoclimate

239   Research at the University of Cambridge.

240

241   **Pyrenoid morphologies**

242   Pyrenoid morphologies were examined using blockface imaging by SEM. Sample preparation

243   and imaging were undertaken at the Cambridge Advanced Imaging Centre (CAIC). Cells were

244   cultured as explained above in liquid Tris-phosphate medium and bubbled under ambient air

245   supply (0.04% CO_2). After centrifugation, they were then fixed and embedded as described in

246   Chan (2018). Resin blocks were mounted on aluminium SEM stubs and sputter-coated with 35

247   nm gold. Blockfaces were obtained with an ultramicrotome (Leica, Wetzlar, Germany) and

248   coated with 30 nm carbon. Finally, blockfaces were imaged using a FEI Verios 460 scanning

249   electron microscope (Thermo Fisher Scientific), running at 4 keV accelerating voltage and 0.2

250   nA probe current. Images were obtained using the Through-lens detector in immersion and

251   backscatter mode. Automated image acquisition was set up using FEI MAPS software using a

252   pixel resolution of 1536 x 1024, a dwell time of 3 μs, a horizontal field width of 15.9 μm/tile

253   (magnification 8000x), an x-y tile overlap of 15%/20% and the MAPS default stitching profile.

254   **Results**

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

256   The protein phylogeny of *Rbc*S was originally constructed to identify any residues specific to

257   species with a pyrenoid as a determinant of CCM activity. The present study found that species

258   without a pyrenoid were dispersed throughout the whole *Rbc*S phylogeny. Therefore, specific

259   residues in the SSU α-helices (Meyer *et al*., 2012) were not sufficient to explain the pyrenoid

260   occurrence across the entire phylum (Fig. 1). A direct comparison of the solvent-exposed

261   residues (available for possible interactions with EPYC1) of the amino acids and their

262   electrostatic properties in the two α-helices, hypothesised to be the key elements for the

263   formation of a pyrenoid (Meyer *et al*., 2012; Mackinder *et al*., 2016), varied in their distribution

264   (Fig. S1, Supplementary Material). For example, *Spermatozopsis similis* (pyrenoid-less)

265   exhibited α-helices identical to *C. reinhardtii* (pyrenoid-positive), and *Chloromonas oogama*

266   (pyrenoid-less) differed by only one residue (Fig. S1, Supplementary Material). The absence

267   of any consistent pattern which could differentiate pyrenoid-less from pyrenoid-positive

268   species suggests that neither the specific residues in the two α-helices and their properties nor

269   the solvent-exposed residues, can singlehandedly explain pyrenoid occurrence in green algae,

270   as we had hypothesized.

271

272   However, the *Rbc*S phylogeny did systematically differentiate streptophyte algae and Core

273   chlorophytes, which were clustered separately into two sister clades (Fig. 1) with nine

274   prasinophyte species clustered with the Core chlorophytes, and one with the streptophyte algae

275   (*Picocystis salinarum*). The phylogenetic differentiation in *Rbc*S clearly coincided with

276 differences in the βA-βB loop length. Core chlorophytes and prasinophytes consistently
277 showed a βA-βB loop length of 25 or more residues, whereas the vast majority of streptophyte
278 algae exhibited a βA-βB loop length of less than 23 residues with 52 of the 58 sequences having
279 a βA-βB loop 21 residues long. *Picocystis salinarum* (prasinophytes) appeared to be an
280 exception with a loop only 21 residues long and clustered with the streptophyte algae. The draft
281 genome of *Picocystis sp.* (Junkins *et al.* 2019) confirmed the *RbcS* short loop for this species
282 and therefore explain why it is clustered with the streptophyte algae. However, its position in
283 this clade is probably due to the short length overall of *RbcS* and the lack of confidence in
284 determining the internal branches. The difference in loop length between Core chlorophytes
285 and streptophyte algae revealed different Rubisco structures between these two groups. With a
286 wider central solvent channel due to the shorter βA-βB loop, streptophyte algae have a Rubisco
287 structure more similar to that in land plants as embryophytes (Spreitzer, 2003).

288

*Rbc*S is not under positive selection
290 As an additional test for residues under positive selection in *Rbc*S, in association with a CCM
291 or at the level of the SSU α-helices, 135 DNA sequences from green algae were used. One
292 Likelihood Ratio Test (LRT) for dN/dS heterogeneity across codons (M0-M3) was successfully
293 performed and was significant, indicating expected heterogeneity in selective pressure across
294 *Rbc*S molecules (2ΔlnL =2312.99 > χ^2 =15.507, df=8) (Table 1). Two LRTs were also
295 performed to test for the presence of codons under positive selection (M7-M8 and M8-M8a)
296 and both comparisons rejected models with positive selection (Table 1). The model M7 (which
297 allows for 10 site classes, each with a ω >1) was selected in favour of the model M8 (11 sites
298 classes with one of which allows for ω>1) and was consequently not significant (2ΔlnL= -
299 0.00049 < χ^2 =5.99, df=2). The more stringent comparison between the model M8a (which is
300 similar to M7 but which allows for an extra class of codons with dN/dS=1) and M8 was also
301 not significant (2ΔlnL= -0.07013 < χ^2=3.84, df=1) confirming the absence of codons under
302 positive selection in *Rbc*S. The absence of residues under positive selection suggests that the
303 appearance of new residues would not confer selective advantages in *Rbc*S, and particularly at
304 the level of the α-helices (consistent with observations arising from Fig. 1 and Fig. S1,
305 described above).
306 Branches under positive selection were successfully tested with the branch-model implemented
307 in PAML. The LRT for heterogeneity across branches (H0-H1) was significant (2ΔlnL=9.358
308 < χ^2=3.84, df=1) (Table 2). However, background and foreground omega showed values less

309   than 1, implying positive selection was absent among foreground branches
310   (ωα=0.082; ωβ=0.16 <1). These results suggest that the presence of variation in ω across
311   branches in *Rbc*S, but not significant enough to show positive selection, or any correlation with
312   CCM occurrence.

313

314   **Streptophyte algae share Rubisco catalytic properties with both chlorophytes and**
315   **embryophytes**
316   A more detailed investigation of Rubisco catalytic properties was undertaken in order to
317   explore whether any evolutionary progression towards land plant characteristics was evident
318   in streptophyte algae. The multiple alignment of *Rbc*S in six representative streptophyte algae
319   selected for this component of the study confirmed the deletion of five amino-acids in this
320   group compared to *Chlamydomonas reinhardtii* (Fig. 2; Spreitzer, 2003). This shortens the
321   loop between the first and the second β-sheets, reducing the constriction at the entry of the
322   holoenzyme's solvent channel. Rubisco catalytic properties at 25°C for the six green algae are
323   shown in Table 3, including *Chlamydomonas reinhardtii* which was used as a control, and to
324   compare this analytical system with previous measurements for this species, albeit of different
325   genetic parentage (Jordan & Ogren, 1981; Satagopan & Spreitzer, 2008). The absence of
326   measurements for *Chlorokybus atmophyticus* was due to many unsuccessful attempts at
327   Rubisco extraction. In *Chlamydomonas reinhardtii*, Rubisco catalytic properties varied slightly
328   from previous measurements (Satagopan & Spreitzer, 2008; Jordan & Ogren, 1981) but
329   remained in the same range. Michaelis-Menten constant for carboxylation (K_c) showed similar
330   values (39.6 and 34 μM) whereas the Rubisco turnover rate (k_{cat}) was somewhat higher in this
331   study compared to the value found in Satagopan & Spreitzer (2008). The streptophyte algae
332   did not show a clear systematic shift from chlorophyte towards land plant catalytic properties
333   despite similar Rubisco SSU structural changes. Of the five streptophyte algae, only
334   *Klebsormidium subtile* and *Onychonema laeve* showed a higher affinity for CO_2 (lower K_c
335   values), similar to land plants (e.g. *Arabidopsis thaliana*; 10.7 μM) with K_c values of 18.7 and
336   27.3 μM respectively (Table 3). *Cosmarium subtumidum*, *Spirogyra sp.* and *Coleochaete*
337   *scutata* had a relative low affinity for CO_2 with K_c values in the range of the Core chlorophytes
338   or slightly higher (45.3, 49.1 and 43.1 μM respectively).
339   The catalytic turnover rate (k_{cat}) showed a trend towards lower k_{cat} values. *Onychonema laeve*
340   and *Cosmarium subtumidum*, both members of the Zygnematophyceae, had similar k_{cat} values
341   (2.39 and 2.51 s^{-1} respectively). *Spirogyra sp* appeared to be an exception with a high k_{cat} value

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

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

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

345   1987; von Caemmerer & Quick, 2000; Tcherkez *et al*., 2006; Savir *et al*., 2010; Tcherkez,

346   2013). *Klebsormidium subtile* presented the highest value for carboxylation catalytic efficiency

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

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

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

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

351

352   **Rubisco catalytic properties are CCM dependent**

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

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

355   associated with Rubisco catalytic properties. The rate of photosynthetic oxygen evolution

356   under different concentrations of inorganic carbon was used to determine the whole cell affinity

357   for inorganic carbon and therefore the extent of any inducible carbon concentrating

358   mechanism. The photosynthetic K$_{0.5}$ (Ci) value (Table 4) of the wild-type *C. reinhardtii* under

359   low CO$_2$ showed a strong affinity for Ci (54 μM Ci), similar to previous values in the literature

360   and in the range of photosynthetic responses of cells expressing a CCM of 10-100 μM Ci

361   (Mitchell *et al*., 2014; Wang *et al*., 2014). *Klebsormidium subtile*, *Chlorokybus atmophyticus*,

362   *Spirogyra sp.* and *Coleochaete scutata* showed a whole cell affinity for Ci in the range of *C.*

363   *reinhardtii* with K$_{0.5}$ ranging from 45 to 54 μM Ci, consistent with a fully functional CCM,

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

365   lower K$_{0.5}$ compared to the other species (62, 64 and 62 μM Ci respectively) suggestive of

366   some CCM activity. Photosynthetic K$_{0.5}$ (Ci) values of all the species grown under high CO$_2$

367   confirmed the absence of CCM activity under such conditions (Table S3, Supplementary

368   Materials), and thereby the inducible character of the CCM in all species under examination.

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

370   for CCM activity in the different species (Meyer *et al*., 2008) (Table 4). *Chlamydomonas*

371   *reinhardtii, Coleochaete scutata, Chlorokybus atmophyticus*, *Spirogyra sp.* and *Cosmarium*

372   *subtumidum* appeared to be isotopically enriched -15.8 to -18.8‰ (Table 4), with values close

373   to the upper range typically seen in C$_4$ terrestrial plants and consistent with a fully-functioning

374   CCM (Raven *et al.*, 1982). On the other hand, *Klebsormidium subtile* and *Onychonema laeve*

375   were somewhat isotopically depleted compared to the other species, with values intermediate

376   between typical C$_3$ and C$_4$ plants (δ^{13}C of -21.1 and -21.3‰ respectively; O'Leary, 1988) and

377   consistent with a CCM phenotype prone to leakiness (retro-diffusion of CO$_2$: Meyer *et al.*,

378   2008) or limited carbon accumulation capacity.

379   These observations reveal that Rubisco catalytic properties correlate with the strength of CCM

380   activity. *C. reinhardtii*, *Cosmarium subtumidum*, *Spirogyra sp.* and *Coleochaete scutata*

381   revealed a fully functioning CCM (low whole-cell affinity, K$_{0.5}$, and low carbon isotope

382   discrimination) but lower Rubisco catalytic affinity for inorganic carbon (high K_c values),

383   whereas *Klebsormidium subtile* and *Onychonema laeve* have a less effective CCM but higher

384   affinity for inorganic carbon in terms of Rubisco catalytic properties (low K_c values).

385   Therefore, in the presence of a less-effective CCM, Rubisco catalytic properties for

386   *Klebsormidium subtile* and *Onychonema laeve* show a systematic shift towards values more

387   typically associated with land plants.

388   Finally, scanning electronic microscopy (SEM) was used to confirm the presence of a pyrenoid

389   in all the streptophyte algae, as an additional diagnostic for an active biophysical CCM. The

390   presence of a pyrenoid was successfully confirmed for all the species except for *Coleochaete*

391   *scutata* for which tissue embedding was unsuccessful. However, presence and morphology of

392   the pyrenoid were confirmed based on McBride *et al.* (1974) for this species, and through

393   carbon isotope discrimination traits (Meyer *et al.*, 2008). CCM activities were successfully

394   linked to presence of a pyrenoid in all the species. *Cosmarium subtumidum*, *Onychonema*

395   *laeve, Coleochaete scutata* and *Spirogyra sp* exhibited pyrenoid morphologies similar to *C.*

396   *reinhardtii* with a typical single layered starch sheath (Fig. 3). A naked pyrenoid was observed

397   in *Klebsormidium subtile* with a total absence of any starch sheath (Fig. 3a) although starch

398   sheath may occur dependent on growth stage or light intensity (M. Melkonian, unpublished

399   observations). The pyrenoid of *Chlorokybus atmophyticus* consisted of multiple layers of short

400   starch plates around the Rubisco microcompartment (Fig. 3c). The network of cross-pyrenoidal

401   tubules was clearly visible in all the species. Finally, it is intriguing that the pyrenoid

402   surrounded by multiple starch plates seems to support strong CCM activity (*Chlorokybus*

403   *atmophyticus*, Fig. 3c) based on carbon isotope composition.

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

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

406   CO₂ as the selective pressure operating on Rubisco, and more detailed physiological
407   experiments are warranted to fully characterize these contrasting processes.

408

409   **Discussion**

410

411   **Rubisco SSU residues do not systematically equate to a CCM.**

412   There was no immediately apparent correlation between SSU amino-acid sequence and
413   pyrenoid occurrence/inferred CCM activity across the newly-created phylogeny of *Rbc*S for
414   green algae. Our expectation was based on (*i*) the observations that the *Rbc*S α-helices are
415   important for pyrenoid formation in *Chlamydomonas reinhardtii* (Meyer *et al.,* 2012), as well
416   as (*ii*) recent *in vitro* and *in vivo* experiments showing that both SSU α-helices are necessary
417   and sufficient to interact with the *Chlamydomonas* Rubisco linker EPYC1 when expressed in
418   heterologous systems (Atkinson *et al.*, 2019). Based on the primary sequence alone, there are
419   however no EPYC1 homologues outside the Chlamydomonadales, so it would seem that other
420   Rubisco aggregation mechanisms may occur in more distantly related lineages, perhaps
421   through interactions with the LSU, which is the *modus operandi* in cyanobacterial
422   carboxysomes (Oltrogge *et al.*, 2019; Wang *et* al., 2019). It would be interesting to determine
423   whether the widespread occurrence of some form of pyrenoid across green algae was due to
424   multiple independent origins of the algal CCM (Meyer *et al.,* 2017), as found in C₄ and CAM
425   pathways (Sage *et al.,* 2011). However, the absence of a pyrenoid does not always equate to
426   lack of a CCM (Giordano *et al.*, 2005), particularly in *Chloromonas*, which is closely related
427   to *Chlamydomonas* (Morita *et al.,* 1999; Nozaki *et al.,* 2002; Pröschold *et al.*, 2001; Meyer *et*
428   *al.,* 2017) but the underlying mechanisms of carbon accumulation of such species remain
429   unknown.

430

431   Overall, detailed alignments of the *Rbc*S α-helix residues did not discriminate between
432   pyrenoid-positive and pyrenoid-negative species (Fig. 1; Fig. S1). In *Chlamydomonas*
433   *reinhardtii*, for instance, two *Rbc*S copies (Goldschmidt-Clermont & Rahire, 1986) show
434   inverse patterns of gene expression across the day-night cycle (Zones *et al.*, 2015). For the
435   present study, it was not possible to establish the functionality of *Rbc*S paralogues in terms of
436   CCM expression (See Materials & Methods). Therefore, determining the exact number of
437   copies, and their sequence specificity, for each of the pyrenoidless species would provide
438   additional confirmation for the absence of specific residues essential for pyrenoid formation in

439   green algae. An extensive evaluation of positive selection also showed no significant shifts in

440   *Rbc*S amino acid residues associated with the CCM across the phylogeny (Table 1) whereas

441   13 residues under positive selection have been detected in *Rbc*S in angiosperms (Yamada *et*

442   *al.*, 2019). The absence of positive selection along branches leading to a pyrenoid could be an

443   artefact of the small number of species *lacking* a pyrenoid within the green algae (Fig. 1), or

444   indeed those possessing some form of a CCM but lacking a pyrenoid structure (see above). A

445   possible alternative explanation is that all green algae retained a pyrenoid-competent Rubisco

446   SSU (as also supported by *in vitro* assays; Wunder *et al.*, 2018; Atkinson *et al.*, 2019) but that

447   the absence of a pyrenoid is rather determined by the lack (ancestral or through secondary loss)

448   of a Rubisco linker, of similar or different ancestry as the *C. reinhardtii* EPYC1 (Mackinder *et*

449   *al.*, 2016). Here too, future comparative proteomic studies with pyrenoidless algal CCMs will

450   help resolve this question .

451

452   **Streptophyte algal Rubisco SSU structure is similar to land plants**

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

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

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

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

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

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

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

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

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

462   Cortona *et al.*, 2019). The Rubisco structural change was not an isolated event at this time. The

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

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

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

466   the photorespiratory pathway has been shown to differ between chlorophytes and streptophyte

467   algae. Chlorophytes use a mitochondrial glycolate dehydrogenase, which produces NADH and

468   H^+ whereas streptophytes use a peroxisomal glycolate oxidase which produces H_2O_2 for the

469   conversion of glycolate to glyoxylate (Stabenau & Winkler, 2005).

470

471 The role of the SSU and of the βA-βB loop in particular is not entirely understood but the
472 central solvent channel may facilitate channelling of substrates and products to and from the
473 active sites (Esquivel *et al*., 2013). Spreitzer (2001; 2002) demonstrated the importance of the
474 loop for holoenzyme assembly and direct mutagenesis at the level of the βA-βB loop changed
475 Rubisco catalytic properties. Direct substitution of a non-surface exposed residue, distant from
476 the solvent channel, R71A, decreased Rubisco specificity and increased K_c and K_o values in *C.*
477 *reinhardtii* (Spreitzer *et al*., 2001) whereas suppressor substitutions of two SSU residues nearer
478 the solvent channel, N54V and A57V, increased Vc, the specificity and the thermal stability of
479 the large subunit L290F mutant enzyme (Du *et al.,* 2000). In addition, Spreitzer *et al*. (2005)
480 demonstrated that the interface between SSU/LSU, far from the actives sites, contributes to
481 different catalytic properties between *C. reinhardtii* and *Spinacia oleracea*. Despite the change
482 in Rubisco SSU structure between chlorophytes and streptophytes, and effect on solvent
483 channel width and possible "suppressor" interactions between LSU and SSU (Spreitzer *et al*.,
484 2001, 2005), there was a continued need for CCMs across the entire phylogeny (Fig. 1) which
485 is reflected in the catalytic properties of the streptophyte algae.

486

487 **Rubisco catalytic properties in green algae depend on CCM efficiency**
488 The above observations led to the investigation of Rubisco catalytic properties within the
489 streptophyte algae and their associated physiological CCM activity. Streptophyte algae are
490 difficult to investigate physiologically. Oxygen electrode measurements were also extremely
491 challenging (Table 4).
492 Despite the clear structural change associated with the βA-βB loop length, Rubisco catalytic
493 properties remained generally similar to chlorophytes (Table 3) without systematic shift
494 towards values associated with land plants (Satagopan & Spreitzer, 2008; Kapralov *et al*., 2010;
495 Atkinson *et al*., 2017). Over the six streptophyte algae, only two species (*Klebsormidium*
496 *subtile* and *Onychonema laeve*) showed K_c values in this lower range. Direct mutagenesis has
497 shown the importance of the SSU βA-βB loop in Rubisco catalytic properties (see paragraph
498 above) but the data in the present study suggested that they were more influenced by the
499 effectiveness of the CCM, consistent with systematic changes in carbon isotope composition
500 (δ^{13}C: Table 4). Carbon isotopes have been used to infer leakiness of CCMs found in algae and
501 hornworts (Meyer *et al*., 2008), although whole cell inorganic carbon (Ci) uptake affinity was
502 similar for all species under ambient growth conditions ($K_{0.5}$, Table 4). *Klebsormidium subtile*
503 and *Onychonema laeve*, the weaker CCM activities (identified through more negative δ^{13}C

504 values: Table 4), were associated with the highest affinity of Rubisco for CO_2 (K_c, Table 3).

505 The importance of the CCM in shaping the adaptation within Rubisco catalytic properties has

506 been a long-standing hypothesis (Meyer *et al.*, 2013, Galmes *et al.*, 2014, 2016, 2019; Griffiths

507 *et al.*, 2017), consistent with the shifts seen in C_4 Rubisco (Jordan & Ogren, 1981; Sage, 2002;

508 Kubien *et al.*, 2008). Our results show that Rubisco catalytic properties for this range of

509 representative streptophyte algae are adapted to the presence of the CCM.

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

511 (leakiness) is partly consistent with pyrenoid presence for these two species (with either a

512 naked pyrenoid or simple starch sheath: Fig. 3a,d, respectively). In addition, *Klebsormidium*

513 *subtile* has often been reported to be a cosmopolitan species, colonising a great variety of

514 aquatic and terrestrial habitats (Table S2; Supplementary Materials; Hoffmann, 1989; Rindi *et*

515 *al.,* 2011; Mikhailyuk *et al.*, 2015). The Rubisco catalytic properties found in *Klebsormidium*

516 *subtile* would place this species as an intermediate between obligate aquatic green algae and

517 land plants, but only the study of real subaerial algae such as *Klebsormidium flaccidum* or

518 *Mesotaenium endlicherianum* would help us to fully understand the photosynthetic adaptation

519 for life on land. In the absence of the liquid boundary layer impeding CO_2 diffusion on land

520 which could affect Rubisco catalytic properties (Raven *et al.,* 1985; Sáez *et al.*, 2017), the

521 naked pyrenoid in *Klebsormidium subtile* would account for the more land-plant-like Rubisco

522 catalytic properties and a reliance on direct diffusive CO_2 supply.

523

524 The co-evolution of Rubisco and CCMs has been demonstrated in multiple organisms (Badger

525 *et al.*, 1998). In diatoms (Young *et al.*, 2016) and haptophytes (Heureux *et al.,* 2017), which

526 are known to carry out most of the oceanic photosynthesis but which possess Form 1D Rubisco

527 (Delwiche & Palmer, 1997; Yoon *et al.,* 2002; Falkowski *et al.*, 2004), Rubisco affinity for

528 CO_2 (K_c) exhibits larger variations, exceeding those of C_4 plant Rubisco suggesting a large

529 diversity of CCM strengths in this group. In addition, the CO_2:O_2 ratio around the active site

530 led to the suggestion that pyrenoids could have an oxygen exclusion function (McKay & Gibbs,

531 1991; Griffiths *et al.*, 2017). In land plants, Rubisco catalytic properties have been shown to

532 be linked to changes in the atmospheric CO_2:O_2 ratio over time as well as temperature, in

533 addition to leaf architecture, morphology and conductance (Beerling *et al.*, 2001; Franks &

534 Beerling, 2009; Haworth *et al.,* 2011; Galmes *et al.*, 2014; 2015; Sharwood *et al.,* 2016; Conesa

535 *et al.*, 2019). In a decreasing atmospheric CO_2:O_2 ratio environment, Rubisco exhibits a higher

536 affinity for CO_2, a fall in K_c and $k_{cat}{}^c$ values, which then has been improved by a higher

537 proportion of leaf protein accounted for by Rubisco (Galmes *et al.,* 2014). Furthermore, higher

538 temperatures increase maximum carboxylase turnover rate (k_{cat}^c) of Rubisco and decrease CO_2

539 affinity (Bernacchi *et al.*, 2001; Galmes *et al.*, 2015, 2016).

540

541 In conclusion, this study has highlighted that Rubisco SSU structure effectively differentiates

542 between streptophytes and Core chlorophytes, with a transition occurring in the intermediate

543 prasinophyte clade which contains mostly species with a long βA-βB loop. Otherwise, the

544 *Rbc*S phylogeny recaptures the latest consensus green algal phylogenies built from many

545 marker genes, including *rbcL* (Leebens-Mack *et al.*, 2019). A more focussed study on Rubisco

546 catalytic properties in streptophyte algae suggests that the activity of any CCM, which may

547 have arisen because of limitations in bulk CO_2 delivery to Rubisco, has permitted the retention

548 of a lower affinity (high K_c) Rubisco. We demonstrated that the extent of adaptation which

549 occurs should either cause CCM activity to be reduced, or indeed lost during the transition to

550 land, as the reliance on gaseous diffusion to deliver CO_2 to Rubisco began to increase. Overall,

551 the observations confirm the widespread occurrence of a CCM across the entire green algal

552 lineage, and the need for active bicarbonate uptake and conversion within some form of

553 pyrenoid to fuel carbon fixation by Rubisco. However, rather than being intransigent and slow,

554 Rubisco catalytic properties adapt to local conditions of CO_2 availability. This is consistent

555 with the changes seen in Rubisco from C_4 (Jordan & Ogren, 1981; Sage, 2002; Kubien *et al.*,

556 2008) and CAM plants (Griffiths *et al.*, 2008), which have been associated with operating

557 within a CCM for the past 5-10 million years. Based on this study, the selective pressures

558 driven by local conditions of photosynthetic CO_2 supply are more likely to explain the shifts

559 in Rubisco catalytic properties during life on land, rather than any long term transition seen in

560 land plants.

561

562 **Acknowledgements**

571 the $\delta^{13}C$ measurements at the Goodwin Laboratory, Department of Earth Sciences, University
572 of Cambridge, UK. Finally, we thank Juan-Carlos Villareal, Université de Laval, Québec,
573 Canada, for his help with the 1KP data.
574

575 **Author Contributions**
576 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
577 performed the experiments on Rubisco kinetics and D.J.O. analysed the data. ~~M.M provided~~
578 ~~the 1Kp data.~~ M.M.M.G performed the phylogenetic analyses, positive selection and
579 physiological data collection and analysis. K.H.M. performed SEM imaging. M.M. provided
580 the *RbcS* sequences. M.M.M.G and H.G. interpreted the data and wrote the manuscript with
581 assistance from all authors.
582

583 **References**
584 **Abascal F, Zardoya R, Posada D. 2005.** ProtTest: selection of best-fit models of protein
585 evolution. *Bioinformatics* **21**: 2104-2105.
586 **Atkinson N, Leitão N, Orr DJ, Meyer MT, Carmo-Silva E, Griffiths H, Smith AM,**
587 **McCormick AJ. 2017.** Rubisco small subunits from the unicellular green alga
588 *Chlamydomonas* complement Rubisco-deficient mutants of Arabidopsis. *New Phytologist* **214**:
589 655-667.
590 **Atkinson N, Velanis CN, Wunder T, Clarke DJ, Mueller-Cajar O, McCormick AJ. 2019.**
591 The pyrenoidal linker protein EPYC1 phase separates with hybrid Arabidopsis–
592 Chlamydomonas Rubisco through interactions with the algal Rubisco small subunit. *Journal*
593 *of experimental botany*.
594 **Badger MR, Kaplan A, Berry JA. 1980.** Internal inorganic carbon pool of *Chlamydomonas*
595 *reinhardtii*: evidence for a carbon dioxide-concentrating mechanism. *Plant Physiology* **66**:
596 407-413.
597 **Badger MR. 1987.** The CO_2-concentrating mechanism in aquatic phototrophs. *Photosynthesis*.
598 Academic press.
599 **Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD.**
600 **1998.** The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based
601 CO_2 concentrating mechanisms in algae. *Canadian Journal of Botany* **76**: 1052-1071.
602 **Becker B. 2013.** Snow ball earth and the split of Streptophyta and Chlorophyta. *Trends in*
603 *Plant science* **18**: 180-183.

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

606 **Behrenfeld MJ, Randerson JT, McClain CR, Feldman GC, Los SO, Tucker CJ,**
607 **Falkowski PG, Field CB, Frouin R, Esaias WE *et al*. 2001.** Biospheric primary production
608 during an ENSO transition. *Science* **291**: 2594-2597.

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

612 **Borges AV, Frankignoulle M. 2002.** Distribution and air-water exchange of carbon dioxide
613 in the Scheldt plume off the Belgian coast. *Biogeochemistry* **59**: 41-67.

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

617 **Carpenter EJ, Matasci N, Ayyampalayam S, Wu S, Sun J, Yu J, Jimenez-Vieira FB,**
618 **Bowler C, Dorrel RG, Gitzendanner MA *et al*. 2019.** Access to RNA-sequencing data from
619 1,173 plant species: The 1000 Plant transcriptomes initiative (1KP). *GigaScience* **8**: 1-7

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

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

624 **Del Cortona A, Jackson CJ, Bucchini F, Van Bel M, D'hondt S, Škaloud P, Delwiche CF,**
625 **Knoll AH, Raven JA, Verbruggen H *et al*. 2019.** Neoproterozoic origin and multiple
626 transitions to macroscopic growth in green seaweeds. *bioRxiv*, 668475.

627 **Delwiche CF, Palmer JD. 1997.** The origin of plastids and their spread via secondary
628 symbiosis. *Origins of algae and their plastids.* Springer, Vienna.

629 **Drummond AJ, Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling
630 trees. *BMC evolutionary biology* **7**: 214.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

650   10343-10347.

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

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

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

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

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

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

657   183-201.

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

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

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

661   5091.

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

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

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

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

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

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

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

669   99-131.

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

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

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

673 421-432.

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

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

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

677 **Griffiths H, Meyer MT, Rickaby REM. 2017.** Overcoming adversity through diversity:

678 aquatic carbon concentrating mechanisms. 3689-3695.

679 **Harholt J, Moestrup Ø, Ulvskov P. 2016.** Why plants were terrestrial from the beginning.

680 *Trends in Plant Science* **21**: 96-101.

681 **Haworth M, Elliott-Kingston C, McElwain JC. 2011.** Stomatal control as a driver of plant

682 evolution. *Journal of Experimental Botany* **62**: 2419-2423.

683 **Hermida-Carrera C, Kapralov MV, Galmés J. 2016.** Rubisco catalytic properties and

684 temperature response in crops. *Plant Physiology* **171**: 2549-2561.

685 **Heureux AM, Young JN, Whitney SM, Eason-Hubbard MR, Lee RB, Sharwood RE,**

686 **Rickaby RE. 2017.** The role of Rubisco kinetics and pyrenoid morphology in shaping the

687 CCM of haptophyte microalgae. *Journal of experimental botany* **68**: 3959-3969.

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

689 **Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T,**

690 **Mori H, Tajiima N *et al*. 2014.** *Klebsormidium flaccidum* genome reveals primary factors for

691 plant terrestrial adaptation. *Nature communications* **5**: 3978.

692 **Jordan DB, Ogren WL. 1981.** Species variation in the specificity of ribulose biphosphate

693 carboxylase/oxygenase. *Nature* **291**: 513-515.

694 **Junkins EN, Stamps BW, Corsetti FA, Oremland RS, Spear JR, Stevenson BS. 2019.** Draft

695 Genome Sequence of Picocystis sp. Strain ML, Cultivated from Mono Lake, California.

696 *Microbiol Resour Announc* **8**: e01353-18.

697 **Kapralov MV, Filatov DA. 2007.** Widespread positive selection in the photosynthetic

698 Rubisco enzyme. *BMC Evolutionary Biology* **7**: 73

699 **Kapralov MV, Kubien DS, Andersson I, Filatov DA. 2010.** Changes in Rubisco kinetics

700 during the evolution of C4 photosynthesis in *Flaveria* (Asteraceae) are associated with positive

701 selection on genes encoding the enzyme. *Molecular Biology and Evolution* **28**: 1491-1503.

702 **Kubien DS, Whitney SM, Moore PV, Jesson LK. 2008.** The biochemistry of Rubisco in

703 Flaveria. *Journal of Experimental Botany* **59**: 1767-1777.

704   **Le SQ, Gascuel O. 2008.** An improved general amino acid replacement matrix. *Molecular*

705   *biology and evolution* **25**: 1307-1320.

706   **Li X, Zhang R, Patena W, Gang SS, Blum SR, Ivanova N, Yue R, Robertson JM, Lefebvre**

707   **PA, Fitz-Gibbon ST, Grossman AR, Jonikas MC. 2016.** An indexed, mapped mutant library

708   enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. *Plant*

709   *Cell* **28**: 367-387.

710   **Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA,**

711   **Graham SW, Grosse I, Li Z, Melkonian M, Mirarab S** *et al.* **2019.** One thousand plant

712   transcriptomes and the phylogenomics of green plants. *Nature* **574**: 679-685

713   **Leliaert F, Verbruggen H, Zechman FW. 2011.** Into the deep: new discoveries at the base

714   of the green plant phylogeny. *BioEssays* **33**: 683-692.

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   **McBride G, LaBounty J, Adams J, Berns M. 1974.** The totipotency and relationship of seta-

727   bearing cells to thallus development in the green alga *Coleochaete scutata*. A laser microbeam

728   study. *Developmental biology* **37** 90-99.

729   **McCourt RM, Delwiche CF, Karol KG. 2004.** Charophyte algae and land plant origins.

730   *Trends in Ecology & Evolution* **19**: 661-666.

731   **McKay RML, Gibbs SP. 1991.** Composition and function of pyrenoids: cytochemical and

732   immunocytochemical approaches. *Canadian Journal of Botany* **69**: 1040-1052.

733   **Meyer M, Seibt U, Griffiths H. 2008.** To concentrate or ventilate? Carbon acquisition, isotope

734   discrimination and physiological ecology of early land plant life forms. *Philosophical*

735   *Transactions of the Royal Society B: Biological Sciences* **363**: 2767-2778.

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

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

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

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

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

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

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

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

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

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

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

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

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

772 **Orr DJ, Carmo-Silva E. 2018.** Extraction of Rubisco to determine catalytic constants.
773 *Covshoff S (ed) Photosynthesis: methods and protocols, Methods in molecular biology, vol*
774 *1770*. Springer, New York.

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

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

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

785 **Rambaut A. 2007.** FigTree, a graphical viewer of phylogenetic trees. [WWW.document]
786 URL *http://tree. bio. ed. ac. uk/software/figtree*.

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

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

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

795 **Rosenzweig ESF, Xu B, Cuellar LK, Martinez-Sanchez A, Schaffer M, Strauss M,**
796 **Cartwright HN, Ronceray P, Plitzko JM, Förster F *et al*. 2017.** The eukaryotic CO_2
797 concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**: 148-162.

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

802 **Sage RF. 2002.** Variation in the kcat of Rubisco in C3 and C4 plants and some implications
803 for photosynthetic performance at high and low temperature. *Journal of Experimental Botany*
804 **53**: 609-620.

805 **Sage RF, Christin PA, Edwards EJ. 2011.** The C$_4$ plant lineages of planet Earth. *Journal of*
806 *Experimental botany* **62**: 3155-3169.

807 **Satagopan S, Spreitzer RJ. 2008.** Plant-like substitutions in the large-subunit carboxy
808 terminus of *Chlamydomonas* Rubisco increase CO$_2$/O$_2$ Specificity. *BMC plant biology* **8**: 85.

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

812 **Sharwood RE, Sonawane BV, Ghannoum O, Whitney SM. 2016.** Improved analysis of C$_4$
813 and C$_3$ photosynthesis via refined in vitro assays of their carbon fixation biochemistry. *Journal*
814 *of Experimental Botany* **67**: 3137-3148.

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

819 **Spreitzer RJ, Esquivel MG, Du YC, McLaughlin PD. 2001.** Alanine-Scanning Mutagenesis
820 of the Small-Subunit βA−βB Loop of Chloroplast Ribulose-1, 5-Bisphosphate
821 Carboxylase/Oxygenase: Substitution at Arg-71 Affects Thermal Stability and CO$_2$/O$_2$
822 Specificity. *Biochemistry* **40**: 5615-5621.

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

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

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

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

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

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

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

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

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

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

839 **36**: 1586-1596.

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

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

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

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

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

845 Dordrecht.

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

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

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

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

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

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

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

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

854 5076.

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

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

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

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

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

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

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

862 **Yokota A, Harada A, Kitaoka, S. 1989.** Characterization of ribulose 1, 5-bisphosphate

863 carboxylase/oxygenase from Euglena gracilis Z. *The Journal of Biochemistry* **105**: 400-405.

864 **Yoon HS, Hackett JD, Pinto G, Bhattacharya D. 2002.** The single, ancient origin of chromist

865 plastids. *Journal of phycology* **38**: 40-40.

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

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

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

875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899

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], and species lacking pyrenoids are indicated in red font. The βA-βB loop length 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: 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*.

Fig. 3: Scanning Electron Microscopy (SEM) 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, g: *Chlamydomonas reinhardtii*). Three distinct pyrenoid morphologies can be observed: Pyrenoid enclosed by one layer of starch plates (b, d and e); pyrenoid enclosed by multiple starch grains (c); and pyrenoid without observable starch sheaths (k). Bars: 2 μm (a to e) and 0.5 μm (f and g).

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

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

935 parameters.

936
937

	Number of classes (ω)	N^a	Length (bp)[b]	LRT (2ΔlnL)	critical values (P<0.05)	df[c]
M0	1	135	462	2312.99077	15.507	8
M3	5	135	462			
M7	10	135	462	-0.000494	5.9915	2
M8	11	135	462			
M8a	11	135	462	-0.07013	3.8415	1
M8	11	135	462			

938
939 a: Number of sequences analysed

940 b: length of *RbcS* sequences analysed

941 c: degrees of freedom

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962 **Table 2:** Results of the three LRTs for positive selection using the branch-models (H0-H1)

963 (codeml) implemented in PAML (Yang, 2007) and their associated parameters.

964

	dN/dS	LRT (2ΔlnL)	critical values (P<0.05)	df
H0	$\omega=0.08445$			
H1	$\omega^a=0.08262$	9.358	3.8415	1
	$\omega^b=0.16371$			

965

966 a: omega for background branches

967 b omega for foreground branches

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

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	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

1003 **Table 4:** Whole cell affinity for inorganic carbon in the six streptophyte algae representative

1004 species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under low CO_2 conditions

1005 (0.04% CO_2) and their associated δ13C for organic matter. Species are ordered from the

1006 furthest species away from land plants (*Chlamydomonas reinhardtii*, Chlorophytes,

1007 Chlorophyceae) to the closest (*Coleochaete scutata*, Coleochaetophyceae, Charophytes).

1008 Values are means ± SEM.

1009
1010

Species name	$K_{0.5}(Ci)$ (μM)	$δ^{13}C$ (‰)
Chlamydomonas reinhardtii	54 ± 23	-18.86
Chlorokybus atmophyticus	62 ± 26	-18.36
Klebsormidium subtile	53 ± 2	-21.18
Cosmarium subtumidum	64 ± 32	-15.80
Onychonema laeve	62 ± 40	-21.31
Spirogyra sp	48 ± 38	-17.85
Coleochaete scutata	45 ± 23	-18.50

1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030

1031 **Supporting Information**

1032

1033 Additional supporting information may be found in the online version of this article.

1034

1035 **Fig. S1:** Comparison of the chemical properties of the two α-helices for species without

1036 pyrenoid and compared to *Chlamydomonas reinhardtii* (pyrenoid positive).

1037

1038 **Fig. S2:** Evolutionary relationship of algae issued of the primary endosymbiosis and the major

1039 glaciation events which occurred during the diversification of the green algae lineages modified

1040 from Leliaert *et al*. (2012) and Becker (2013).

1041

1042 **Fig. S3:** DNA phylogeny of *RbcS* used for the PAML analysis and built with BEAST v2.3.1.

1043

1044 **Table S1:** Growth media and accession number of the six streptophyte algae

1045

1046 **Table S2:** Systematic classification and habitat description of the six streptophyte algae

1047

1048 **Table S3:** Whole cell affinity for inorganic carbon in the six streptophyte algae representative

1049 species and *Chlamydomonas reinhardtii* (Chlorophytes) grown under high CO_2 conditions (5%

1050 CO_2) and their associated δ13C for organic matter.

1051

1052 **Table S4:** Pyrenoid diagnostic for all the species present in the phylogeny of *RbcS* and the

1053 associated references. Species without pyrenoid are highlighted in light grey.
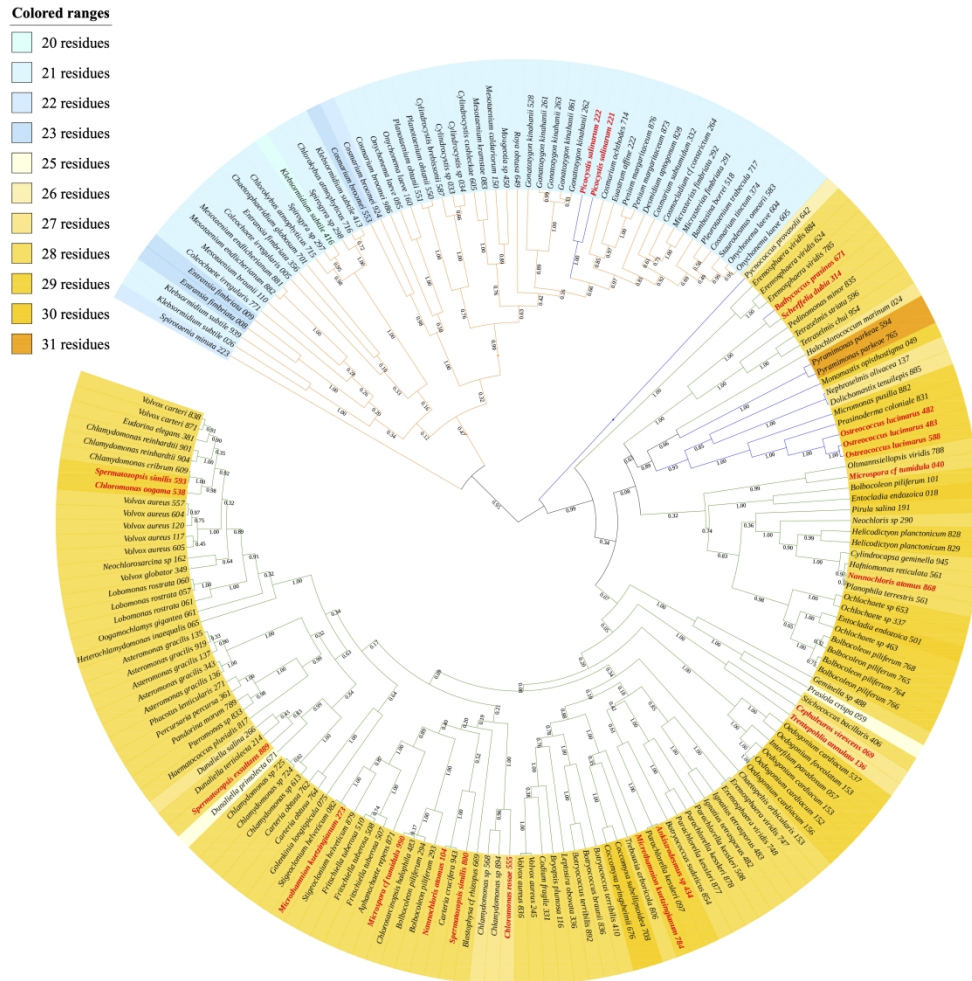
1054

1055
1056

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], and species lacking pyrenoids are indicated in red font. The βA-βB loop length 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.
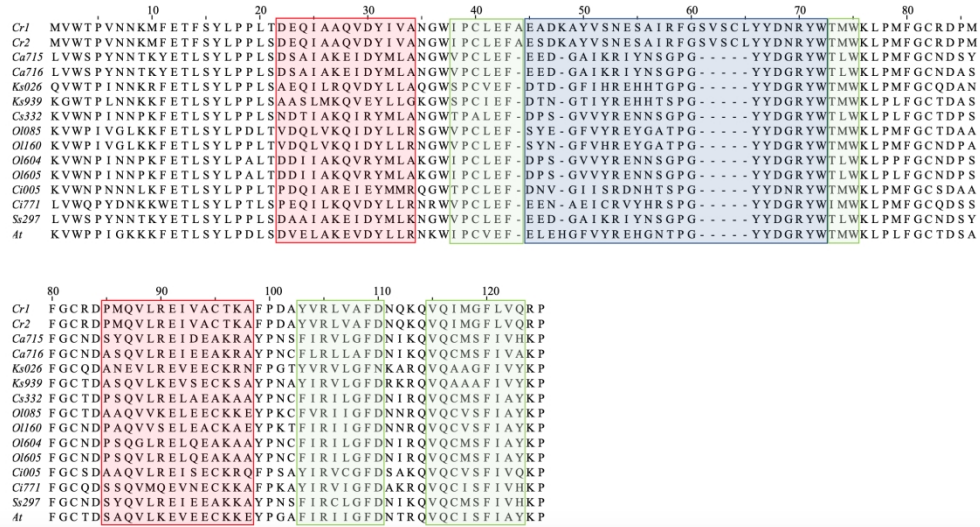
Fig. 2: Subset alignment of sequences from the 1KP of the representative streptophyte algae Rubisco small subunit (RbcS) and their primary structures compared to the two copies of RbcS in Chlamydomonas reinhardtii (Chlorophytes, Cr1 and Cr2) and Arabidopsis thaliana (At, land plants). Ca (Chlorokybus 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.
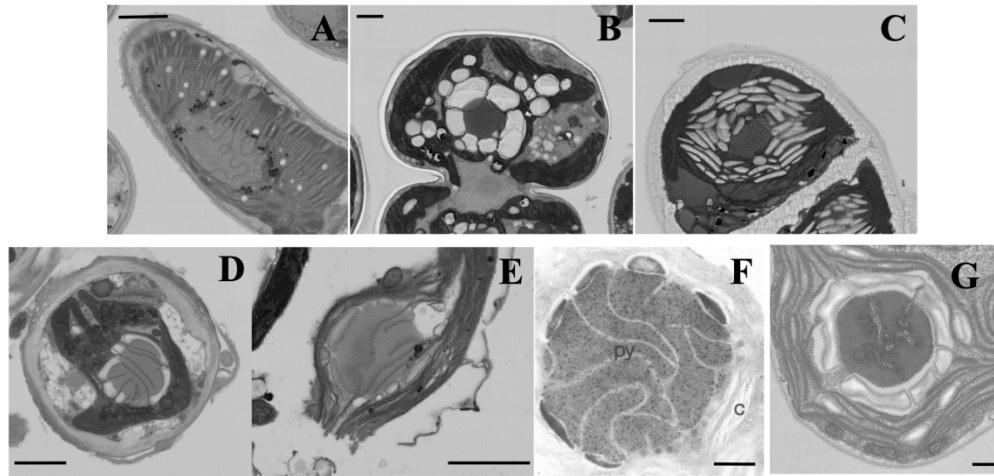
816x436mm (72 x 72 DPI)

Fig. 3: Scanning Electron Microscopy (SEM) 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, g: Chlamydomonas reinhardtii). Three distinct pyrenoid morphologies can be observed: Pyrenoid enclosed by one layer of starch plates (b, d and e); pyrenoid enclosed by multiple starch grains (c); and pyrenoid without observable starch sheaths (k). Bars: 2 µm (a to e) and 0.5 µm (f and g).

262x125mm (250 x 250 DPI)