1 **TECHNICAL ADVANCE**

2 **A procedure to introduce point mutations into the Rubisco large subunit gene in wild-type**
3 **plants**

4

5 Myat T. Lin[1+], Douglas J. Orr[2+], Dawn Worrall[2], Martin A. J. Parry[2*], Elizabete Carmo-Silva[2],
6 Maureen R. Hanson[1]
7 [1] Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14850, USA.
8 [2] Lancaster Environment Centre, Lancaster University, Library Avenue, Lancaster, LA1 4YQ, UK.

9

10 **Running title:** Introducing point mutations into chloroplast genes

11

12 **Keywords:** chloroplast transformation, food security, homologous recombination, Rubisco,
13 photosynthesis, site-directed mutagenesis, *Nicotiana tabacum*

14

15

16 + M.T. Lin and D.J. Orr should be considered joint first authors.
17 *Correspondence: m.parry@lancaster.ac.uk; tel +44 1524 595084

18

19 **SUMMARY**

20 Photosynthetic inefficiencies limit the productivity and sustainability of crop production, and
21 the resilience of agriculture to future societal and environmental challenges. Rubisco is a key
22 target for improvement as it plays a central role in carbon fixation during photosynthesis and is
23 remarkably inefficient. Introduction of mutations to the chloroplast-encoded Rubisco large
24 subunit *rbc*L is of particular interest to improve the catalytic activity and efficiency of the
25 enzyme. However, manipulation of *rbc*L is hampered by its location in the plastome, with many
26 species recalcitrant to plastome transformation, and by the plastid's efficient repair system,
27 which can prevent effective maintenance of mutations introduced with homologous
28 recombination. Here we present a system where the introduction of a number of silent
29 mutations into *rbc*L within the model plant *Nicotiana tabacum* facilitates simplified screening
30 via additional restriction enzyme sites. This system was used to successfully generate a range of
31 transplastomic lines from wild-type *N. tabacum* with stable point mutations within *rbc*L in 40%
32 of the transformants, allowing assessment of the effect of these mutations on Rubisco assembly
33 and activity. With further optimization, the approach offers a viable way forward for mutagenic
34 testing of Rubisco function in planta within tobacco and modifying *rbc*L in other crops where
35 chloroplast transformation is feasible. The transformation strategy could also be applied to
36 introduce point mutations in other chloroplast-encoded genes.

37

38 **Significance Statement**

39 A simplified transformation strategy was developed to perform site-directed mutagenesis in the
40 chloroplast-encoded Rubisco large subunit of tobacco. This approach reduces unwanted
41 mismatch repairevents, enables rapid screening of transformed lines that possess desired
42 mutations and can be applied to modify any chloroplast-encoded gene in plant species
43 amenable to plastid genome transformation.

44

45 **INTRODUCTION**

46 The transformation of chloroplast or plastid genomes in higher plants represents a promising
47 technology in multiple biotechnological applications such as introducing agronomically
48 important traits, metabolic engineering, recombinant protein expression and production of
49 high-value therapeutic compounds (Maliga and Bock 2011). Chloroplast transformation offers
50 several benefits: precise manipulation of its genome via homologous recombination, no
51 transgene silencing and better control over the transgene escape into the environment. In
52 addition to the introduction and expression of foreign genes, chloroplast transformation can
53 also be used to delete or mutate plastid-encoded protein subunits for functional studies. Since
54 many of these proteins are involved in photosynthesis, the technology has great potential to
55 improve photosynthesis and productivity in crops (Hanson *et al*. 2013, Bock 2015, Martin-Avila
56 *et al*. 2020).

57 One of the most commonly studied plastid genes is *rbc*L, which encodes the large
58 subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). In plants, Rubisco
59 catalyses two competing reactions in the stroma of chloroplasts: carboxylation or oxygenation
60 of ribulose-1,5-bisphosphate (RuBP) (Ogren and Bowes 1971, Tcherkez *et al.* 2006, Andersson
61 and Backlund 2008). The carboxylation of RuBP is an essential step for photosynthesis in plants,
62 while its oxygenation generates 2-phosphoglycolate, which is recycled through the
63 photorespiratory pathway, spanning multiple organelles (Ogren and Bowes 1971, Ogren 1984,
64 Keys 1986, Busch 2020). Rubisco in C_3 plants is a relatively slow enzyme with k_{cat} about 3-4 CO_2
65 s^{-1} (Orr *et al.* 2016, Flamholz *et al.* 2019). As a result, C_3 plants generally express a large amount
66 of Rubisco within the chloroplast stroma in order to achieve sufficient carbon fixation. Since the
67 reactions of Rubisco control major metabolic fluxes, manipulating its kinetics has been an
68 important target to improve photosynthesis in plants (Sharwood 2017, Zhu *et al.* 2020). Two
69 general goals of engineering Rubisco are to improve its carboxylation efficiency under ambient
70 O_2, which is defined as $k_{\mathrm{cat}}^{\mathrm{C}}/K_{\mathrm{M,air}}^{\mathrm{C}}$, and to increase its CO_2/O_2 specificity factor, which is the
71 ratio of its carboxylation efficiency to oxygenation efficiency or $k_{\mathrm{cat}}^{\mathrm{C}}K_{\mathrm{M}}^{\mathrm{O}}/k_{\mathrm{cat}}^{\mathrm{O}}K_{\mathrm{M}}^{\mathrm{C}}$, where k_{cat} and
72 K_M represent the catalytic turnover number and Michaelis-Menten constant, respectively
73 (Whitney *et al.* 2011a). Increasing Rubisco content in maize and rice led to higher plant biomass,

74 supporting the hypothesis that Rubisco represents a bottleneck in photosynthesis, and crop
75 yields can be improved with a more efficient Rubisco (Salesse-Smith *et al.* 2018, Yoon *et al.*
76 2020). However, to date attempts to engineer vascular plants with such a Rubisco enzyme have
77 been unsuccessful.

78 Rubisco enzymes with different kinetic properties exist in nature despite its well-
79 characterized catalytic constraints (Galmés *et al.* 2015, Galmés *et al.* 2016, Orr *et al.* 2016,
80 Sharwood *et al.* 2016a, Flamholz *et al.* 2019). As organisms adapt to different environments,
81 they have evolved Rubisco enzymes that are optimized to their immediate surroundings
82 (Tcherkez *et al.* 2006, Savir *et al.* 2010). For example, Rubisco in C_4 plants is generally associated
83 with a lower affinity for CO_2 or higher $K_{\mathrm{M}}^{\mathrm{C}}$ and a higher $k_{\mathrm{cat}}^{\mathrm{C}}$ compared to those in C_3 plants
84 (Whitney *et al.* 2011a). Modeling studies indicated that a typical C_3 Rubisco is optimized for 220
85 ppm of atmospheric CO_2 and Rubisco from several C_4 plants would improve carbon fixation in C_3
86 plants (Zhu *et al.* 2004, Sharwood *et al.* 2016b). Rubisco from red algae such as *Griffithsia*
87 *monilis* were shown to have the highest known specificity factors and generally assumed to be
88 promising candidates for improving photosynthesis in C_3 plants (Whitney *et al.* 2011a).

89 Plants possess form I Rubisco, which is a hexadecameric complex made up of eight
90 chloroplast-encoded large subunits (LSu) and eight small subunits (SSu) encoded by a family of
91 *Rbc*S nuclear genes and imported to the chloroplast stroma (Whitney *et al.* 2011a). Each
92 holoenzyme complex consists of four LSu dimers with two active sites located inside each dimer
93 at the interface between the two LSu (Andersson and Backlund 2008), and capped with 4 SSu
94 monomers at each end. Plants produce specific chaperonins to prevent irreversible aggregation
95 of LSu as well as multiple chaperones for step-by-step assembly of functional L_8S_8 complexes
96 (Bracher *et al.* 2017, Wilson and Hayer-Hartl 2018). Thus, engineering plants with a more
97 efficient Rubisco that can accumulate the enzyme at sufficiently high levels (Carmo-Silva *et al.*
98 2015) has been a major challenge.

99 Engineering Rubisco in higher plants has historically been carried out exclusively in
100 tobacco (*Nicotiana tabacum*), where well-established procedures for chloroplast transformation
101 allowed precise modification of its *rbc*L gene through homologous recombination and, recently,
102 co-engineering of both Rubisco subunits (Sharwood 2017, Martin-Avila *et al*. 2020). Although

103 LSu from sunflower, Arabidopsis, and C_3 and C_4 Flaveria species were able to assemble with
104 native tobacco SSu to form functional enzymes in tobacco transformant plants, the
105 accumulation of such hybrid enzymes in the leaves was significantly lower than the normal
106 amount, likely due to suboptimal interactions between the foreign LSu and native chaperones
107 or SSu (Kanevski *et al.* 1999, Whitney *et al.* 2011b, Whitney *et al.* 2015). Attempts to replace the
108 *rbc*L gene in tobacco with red algal Rubisco genes produced transformants without functional
109 Rubisco due to incompatibility with the chaperonin machinery and chaperones in tobacco
110 chloroplast stroma (Whitney *et al.* 2001, Lin and Hanson 2018). A recent study demonstrated
111 that a red-type Rubisco from *Rhodobacter sphaeroides* was able to assemble and function in
112 tobacco chloroplasts, but its poor compatibility with native Rubisco activases led to low
113 activation levels in the absence of its cognate Rubisco activase from *Rhodobacter sphaeroides*
114 (Gunn *et al*. 2020). Thus, successfully expressing Rubisco sufficiently distant phylogenetically
115 from the host plant will likely require manipulation of ancillary proteins such as Rubisco activase
116 and/or assembly related factors.
117 Unless chaperones and SSu can also be optimized to work with a foreign LSu, an
118 alternative approach could be to introduce carefully selected site-directed mutations in the
119 native LSu via chloroplast transformation. However, generating such transformants from wild-
120 type tobacco has been inefficient because the desired mutations introduced by recombination
121 between the mutated *rbc*L and the native version of the gene can be removed by the plastid's
122 repair system before transformants reached homoplasmy (Whitney *et al.* 1999, Kanevski *et al*.
123 1999). Thus, a tobacco master line has been created where the native *rbc*L gene was replaced
124 with the codon-modified *rbc*L gene from *Rhodospirillum rubrum* with low sequenece homology
125 followed by removal of the selectable marker gene to facilitate the introduction of mutated and
126 foreign *rbc*L genes and polycistrons into tobacco (Whitney and Sharwood 2008). This tobacco
127 master line has been successfully used to study the residues critical for the catalytic properties
128 of C_4 Rubisco enzymes in Flaveria species as well as introduction of Arabidopsis LSu into tobacco
129 (Whitney *et al.* 2011b, Whitney *et al.* 2015). In a recent study, inhibiting the expression of native
130 SSu in the tobacco master line with RNA interference allowed the simultaneous transformation
131 of both Rubisco subunits and investigation of the effects of novel Rubisco complexes in tobacco

132 (Martin-Avila *et al.* 2020). This tobacco master line has been a great resource to carry out
133 Rubisco engineering in a model species, though the generation of such a master line in other
134 plants is a lengthy process, and an efficient procedure to modify chloroplast genes including
135 *rbc*L using a wild-type line would therefore be helpful.

136 In this study, we developed an approach to effectively introduce specific mutations into
137 chloroplast-encoded genes. As a proof of concept, we synthesized a tobacco *rbc*L gene with
138 synonymous or silent mutations resulting in unique restriction sites. Using this modified *rbc*L
139 gene as a template, we introduced single or double residue substitutions that were predicted in
140 previous phylogenetic and biochemical studies to be potentially important for the enzyme's
141 kinetic properties (Kapralov *et al.* 2012, Galmés *et al.* 2014, Studer *et al.* 2014, Orr *et al.* 2016).
142 This approach allowed us to successfully replace the *rbc*L gene with the modified *rbc*L genes
143 directly within wild-type tobacco plants with high retention of the mutations. A procedure to
144 screen for transformants that possess the mutant *rbc*L genes is described, as well as preliminary
145 analyses of Rubisco activities in ten different transplastomic lines.
146
147 **RESULTS**
148
149 **Modified restriction sites in the *rbc*L gene allowed efficient screening of the transformants**
150 We synthesized part of a modified *rbc*L gene (*Nt-rbc*L^{m}) by introducing 26 silent mutations such
151 that four restriction sites were removed while ten were added (Figure 1a, S1). The *Nt-rbc*L^{m}
152 gene was then seamlessly joined with the native *rbc*L promoter from tobacco and inserted into
153 a chloroplast transformation plasmid, pCT-rbcL, described previously (Lin *et al.* 2014). In the
154 resultant chloroplast transformation vector pCT-Nt-rbcL^{m}, the *Nt-rbc*L^{m} gene followed by the
155 native *rbc*L terminator, Nt-TrbcL, and a selectable marker operon expressing the *aadA* gene
156 driven by the tobacco *psbA* promoter are flanked between a 980 bp upstream homologous
157 region or Flank 1, which contains the native *rbc*L promoter, and a 1kbp downstream
158 homologous region or Flank 2 (Figure 1b). The homologous recombination between the plastid
159 genome and pCT-Nt-rbcL^{m} plasmid through the two flanking regions should introduce the *aadA*
160 marker gene into the plastid genome and facilitate the selection of transformants with

161 spectinomycin on regeneration medium. However, it is expected that the intended mutations in
162 *Nt-rbc*L^m will be incorporated into only a portion of the transformants with the *aadA* gene. If the
163 cross-over site upstream of the aadA gene is located within *Nt-rbc*L^m, the resulting
164 transformants will not possess those mutations upstream of the cross-over site (Figure 1c). Our
165 strategy is to use the unique restriction sites in *Nt-rbc*L^m to screen for the transformants with
166 the intended mutations.

167 Phylogenetic studies of Rubisco had previously suggested that several amino acid
168 substitutions in LSu were positively selected during the C_3 to C_4 transition (Kapralov *et al.* 2011,
169 Kapralov *et al.* 2012, Studer *et al.* 2014). As detailed in Table 1, we selected five of those
170 mutations (V101I, V255A, L270I, A281S and H282N) and introduced each into the *Nt-rbc*L^m gene
171 as well as a double mutant with both A281S and H282N. We also included the *Nt-rbc*L^m gene
172 with a C449G mutation that was suggested to be associated with improved catalytic efficiency
173 by a previous wide survey of Rubisco kinetic properties from 75 plant species (Orr *et al.* 2016).
174 In addition, we introduced the K429Q mutation to obtain the enzyme from tobacco's paternal
175 parent, *N. tomentosiformis* (accession YP_398871.1). We also created the *Nt-rbc*L^m gene that
176 encodes both K429Q and L225I mutations as the inverse I225L change was found to be selected
177 in evolution of C_3 branches, and ancestors of tobacco Rubisco potentially possessed an L225I
178 mutation (Studer *et al.* 2014).

179 We introduced each of the pCT-Nt-rbcL^m vectors with these mutations into tobacco
180 seedlings of either Samsun or Petit Havana tobacco cultivar with biolistics and performed
181 restriction digestion of the PCR-amplified *rbc*L gene from the transformed shoots arising from
182 the first selection round (Figure 2a). We found that typically 40 percent of the shoots at this
183 stage possessed the restriction sites corresponding to the *Nt-rbc*L^m gene. Our result was
184 comparable to that in a previous study, which introduced a different set of silent mutations to
185 remove commonly used restriction sites in the tobacco *rbcL* gene and found that 6 out of 12 or
186 50% of the transformants had the silent mutations (Sinagawa *et al.* 2009). After the shoots we
187 produced with the modified restriction sites at the *rbc*L locus were subjected to a second round
188 of selection, we analyzed restriction fragment length polymorphisms (RFLPs) in the
189 transformants with DNA blotting (Figure 2b,c). Those transformants that had achieved

190 homoplasmy with the *Nt-rbc*L^m gene were transferred to rooting medium and subsequently to
191 soil until they set seeds. We obtained multiple independent transformants for six out of nine
192 LSu mutations and one each for the remaining three mutations (Table 1). Sequencing of the *rbc*L
193 locus in the final transformants confirmed that all but one of the transformants possessed the
194 entire set of silent mutations in the *Nt-rbc*L^m gene along with the intended non-synonymous
195 mutations. The NtLwt transformant inherited all the silent mutations in the *Nt-rbc*L^m gene minus
196 the single nucleotide change for the new HindIII site (Table 1). Our results are consistent with
197 the previous study which found that the majority, five out of six transformants with silent
198 mutations in the *rbc*L gene, possessed the entire set of mutations and only one originated from
199 a cross-over event within the *rbc*L gene (Sinagawa *et al.* 2009).
200
201 **Mutations in the Rubisco large subunit did not prevent assembly of Rubisco holoenzyme**
202 We further investigated one tobacco transformant for each Rubisco LSu mutant to determine
203 the effects of the introduced mutations. DNA blotting with a probe hybridized to a region
204 upstream of the *rbc*L gene locus confirmed that all transformants were homoplasmic and
205 possessed a restriction site from the *Nt-rbc*L^m gene (Figure 3). Likewise, an RNA blot of the same
206 samples with a probe to detect the *Nt-rbc*L gene showed that all transformants had an extra
207 dicistronic transcript with both the *Nt-rbc*L^m and downstream *aadA* genes, in addition to the
208 monocistronic *Nt-rbc*L^m transcript (Figure 4a). Both of the *rbc*L transcripts present in the
209 transformants were much less abundunt than the single monocistronic transcript from the wild-
210 type sample, however, the total *rbc*L transcript levels in the transformants were comparable to
211 that from the single transcript in the wild-type (Figure 4b). The proportion of discistronic mRNA
212 amongst the transgenic lines did not significantly vary (ANOVA, P = 0.848). Soluble protein
213 samples from the leaf tissues of the wild-type plant and all transformants displayed a similar
214 band for L_8S_8 holoenzyme on blue native PAGE, indicating these mutations did not prevent
215 Rubisco assembly, though measurements indicate an effect on enzyme abundance in the leaf
216 (Table 2).
217
218 **Introduced mutations often affected Rubisco carboxylation**

219 To compare the potential impact of these introduced mutations on Rubisco activity an analysis
220 was conducted using glasshouse-grown plants of each transplastomic line and wild-type
221 controls. Leaf discs were taken from 26 day old plants and analysed for maximum Rubisco
222 carboxylation rate (k_{cat}) and other related parameters including total soluble protein and
223 chlorophyll content (Table 2, Table S1). There was significant variation amongst genotypes in the
224 cv. Samsun background (Table 2). Two mutations, V101I and V255A, showed significant
225 negative impact on Rubisco maximum carboxylation rate. Amongst cv. Petit-Havana genotypes
226 there was some evidence for the negative impacts of the C449G mutation, with variation among
227 these genotypes approaching statistical signifiance (P = 0.054, ANOVA). In cv. Petit-Havana
228 genotypes there was also some support for the effect of mutations on Rubisco content and
229 reflected in total soluble protein content (Table 2). Plants of each transplastomic line grew to
230 comparable size and displayed no obvious phenotype under the conditions used (Figure 5).
231 There were no significant differences in chlorophyll content (Table S1). Growth under additional
232 environment conditions will be needed to further explore the effect of the mutations on plant
233 phenotype and enzymatic activity.
234
235
236

237 **DISCUSSION**

238 Engineering the chloroplast genome, or plastome, to utilise the unique characteristics of this
239 organelle is of increasing interest for goals such as improving photosynthesis and the use of
240 plants as bio-factories. The central CO_2 fixing enzyme Rubisco, due its inefficiencies, has long
241 been considered a promising target to improve photosynthesis and increase biomass and yield
242 in crops. Manipulation of Rubisco subunit genes presents contrasting complications. Chloroplast
243 transformation allows precise site-directed mutagenesis of the large subunit (*rbc*L) gene in the
244 chloroplast genome, and although this has long been successfully used to investigate Rubisco
245 biogenesis and biochemistry, the technology is currently limited to a relatively small set of
246 species. Understanding of thepervasive role, in some cases, of the small subunit (encoded by
247 the nuclear RbcS family) on catalysis is rapidly increasing, and despite the complexity of
248 manipulating highly similar gene families, often involving a large number of homologues,
249 advances are being made in this area (e.g. Martin-Avila *et al.* 2020, Khumsupan *et al*. 2020). In
250 this study, we introduced ten restriction sites with 26 silent mutations in a modified tobacco
251 *rbc*L gene (*Nt-rbc*L^m), which allowed PCR-RFLP screening of transformants after the first round
252 of selection. The elimination of null transformants without the intended mutations after the
253 first selection round also improved the efficiency of workflow to obtain the final stable
254 transformants. Compared to a previous report where only one in eight transformants had the
255 desired mutation in the *rbc*L gene (Whitney *et al.* 1999), about 40 percent of our transformants
256 obtained after the first selection had *Nt-rbc*L^m, with the majority of these giving rise to stable
257 transplastomic plants with the entire set of mutations in the *Nt-rbc*L^m gene. Our results are
258 generally consistent with a previous study to remove common restriction sites in the *rbc*L gene
259 with silent mutations (Sinagawa *et al.* 2009).

260 It proved useful to generate multiple shoots or subclones from each transformant after
261 the first selection round since more than half of the shoots tested after the second selection
262 round had lost the *Nt-rbc*L^m. This is not surprising since *Nt-rbc*L^m is over 98% identical to the
263 wild-type sequence and can be removed before the transformants reach homoplasmy (Whitney
264 *et al.* 1999, Kanevski *et al*. 1999). There may be potential to further improve the transformation
265 efficiency by increasing the number of silent mutations in *Nt-rbc*L^m such that there is no

266 sequence homology to the wild-type gene. However, having too many silent mutations could
267 possibly interfere with translation efficiency and other underlying sequence-specific regulatory
268 processes, which are not yet fully understood. One such regulatory element that is widely
269 conserved among plants is a major translation pause site within *rbc*L transcripts caused by an
270 internal ribosome-binding site and mRNA structure (Gawroński *et al.* 2018).

271 Our analyses indicate that many of the transformants with the *Nt-rbc*L^m gene had lower
272 Rubisco content than the wild-type plants although in some cases the difference was not
273 statistically significant. One reason for reduced Rubisco levels could be the 23 modified codons
274 in the Nt-*rbc*L^m gene, with the majority of these changes resulting in the incorporation of less
275 frequently used codons that could negatively impact the protein's translation efficiency.
276 Although translation efficiencies in chloroplasts cannot always be predicted from codon usage
277 and chloroplasts do not possess rare codons similar to those found in *Escherichia coli*, different
278 codons have varying translation effienciencies so that it may still be desirable to avoid
279 unnecessary codon changes (Nakamura and Sugiura 2011). For example, strategically targetting
280 a single restriction site closest to each non-synonymous mutation instead of introducing the
281 entire set of silent mutations in the *Nt-rbc*L^m gene should minimise unintentional influence on
282 translation efficiencies.

283 The processing of the *rbcL* mRNA 3' end in our transformants was not efficient, giving
284 rise to a dicistronic transcript with *rbcL* and *aadA* genes. This is likely due to insufficient length
285 of the *rbcL* 3'-UTR, which was 205 nucleotides in the transformants. In addition to the
286 dicistronic transcipt, we also observed two sizes for monocistronic mRNAs, indicating a second
287 transcript processing site that is likely in the *psbA* promoter drownstream of the *rbcL* 3'-UTR. In
288 a previous study where an *rbcL* 3'-UTR that was 269 nucleotides in length was incorporated
289 downstream of the *rbcL* gene, the dicistronic transcripts were much less abundant, probably
290 due to more efficient processing of the transcripts (Whitney and Sharwood 2008). It was
291 previously shown that 410 nucleotides following *rbcL* was necessary for proper maturation of
292 the *rbcL* transcript (Sinagawa *et al.* 2009). Thus, future work should consider incorporation of a
293 complete *rbcL* 3'-UTR that is at least 410 nucleotides long so that the transformants can
294 produce *rbcL* transcripts that are similar in size and abundance to those in the wild-type. In

295 addition, stem-loop structures at the *loxP* sites flanking the *aadA* operon were suggested to
296 interfere with the processing of both *rbcL* and *aadA* transcripts (Sinagawa *et al.* 2009). Thus, it
297 may be preferable to replace *loxP* sites with long direct repeats that can spontaneously trigger
298 removal of the marker gene through homologous recombination (Iamtham and Day 2000).
299 Alternatively, the marker operon can be flanked with *attB* and *attP* sequences and subsequently
300 removed with the expression of PhiC31 phage integrase (Kittiwongwattana *et al.* 2007).

301 Previously, the development of a tobacco master line, where the *rbc*L gene had been
302 replaced with a homolog from *Rhodospirillum rubrum* encoding a form II Rubisco, allowed
303 modification of the *rbc*L gene and rapid characterization of the subsequent mutant Rubisco
304 enzymes within 6-9 weeks of transformation, although this master line required a high CO_2
305 environment to grow in soil (Whitney and Sharwood 2008). Recently, functional Rubisco
306 enzymes from Arabidopsis and tobacco were successfully assembled in *Escherichia coli* with the
307 co-expression of at least five chaperones (Aigner *et al.* 2017, Wilson *et al.* 2019, Lin *et al*. 2020).
308 Thus, modified Rubisco enzymes from plants can now be readily produced in *E. coli*, and once
309 modifications that lead to superior carboxylation kinetics are identified, they can be introduced
310 into host plants for further characterization of their effects on photosynthesis and plant growth.
311 These new tools are complemented by directed evolution approaches with cyanobacterial
312 Rubisco, which have shown promising improvements to catalytic efficiency and specificity
313 (Wilson *et al*. 2018).

314 As a proof of concept for this approach to mutating chloroplast genes, we generated
315 seven tobacco transformants each with one residue substitution and two transformants each
316 with two residue substitutions in the Rubisco LSu. Most of these residue substitutions were
317 predicted to be selected during the C_3 to C_4 transitions or associated with a higher catalytic
318 efficiency, while L225I and K429Q were potentially present in ancestors of tobacco Rubisco
319 (Studer *et al.* 2014). Mutations that are *bona fide* kinetic switches between C_3 and C_4 Rubisco
320 should result in enzymes with higher k_{cat} (Whitney *et al.* 2011b). However, our analyses
321 indicated a lack of significant improvement in the carboxylation rate of Rubisco from these
322 transformants, indeed a number of the mutant Rubiscos displayed lower maximum
323 carboxylation activities than the control plants. Importantly, the transformants exhibited

324 Rubisco with similar mobility on native PAGE as the wild-type plants, suggesting no impairment
325 of Rubisco biogenesis as a result of the mutations introduced in *rbc*L.

326 A previous study identified a residue substitution in the Rubisco LSu of *Flaveria* species
327 as a C_3-C_4 kinetic switch, but the same residue substitution which is native to the tobacco
328 enzyme did not convey the same kinetic effect (Whitney *et al.* 2011b). Thus, kinetic switches
329 likely require accessorial components, which may be one or more additional complementary
330 mutations in the large or even the small subunit of Rubisco. Indeed, evidence that the residues
331 at the interface between the two subunits play important roles in determining the kinetic
332 properties was previously reported by studies of the Rubisco from *Chlamydomonas* (Spreitzer *et*
333 *al.* 2005, Genkov and Spreitzer 2009) and cyanobacteria (Wilson *et al*. 2018). Understanding the
334 role of the small subunit in plant Rubisco catalysis is an expanding area, and will be an
335 important complement to engineering the large subunit via chloroplast transformation (e.g
336 Ishikawa *et al.* 2011, Morita *et al*. 2014, Atkinson *et al.* 2017, Khumosupan *et al.* 2020),
337 alongside newly developed tools for rbcL-rbcS co-engineering in the plastome (Martin-Avila *et*
338 *al*. 2020).

339 Reliable plastid transformation procedures for many agriculturally important crops are
340 not yet available. A recent study showed that spectinomycin-resistant Arabidopsis plastid
341 transformants could be selected with reasonably high efficiency once the nuclear-encoded
342 ACC2, an enzyme subunit involved in fatty acid biosynthesis pathway inside plastids, was
343 knocked out, rendering the plants hypersensitive to spectinomycin (Yu *et al.* 2017). Fertile
344 Arabidopsis plastid transformants were then successfully generated when root-derived
345 microcalli of *acc2* knockout lines were used for transformation and regeneration (Ruf *et al.*
346 2019). These latest developments have potential to inform future attempts to extend plastid
347 transformation technology to other species. Once the plastid genome in crops can be readily
348 transformed, an important goal would be to employ changes within the Rubisco large subunit to
349 improve its properties in key crops such as rice and wheat. Since our approach does not require
350 initial generation of master lines in target plant species, which can be time-intensive, it can
351 greatly expedite future work in introducing targeted changes not only in the Rubisco large
352 subunit, but also in other chloroplast-encoded protein subunits.

353

354 **EXPERIMENTAL PROCEDURES**

355

356 **Construction of the chloroplast transformation vectors with modified *Nt-rbc*L genes**

357 All primers used were obtained from Integrated DNA Technologies and listed in Table S2.

358 Phusion^{TM} high-fidelity DNA polymerase, FastDigest restriction enzymes and T4 DNA ligase from

359 Thermo Scientific were used to generate amplicons, restriction digests and ligation products

360 respectively. The partial *Nt-rbc*L^{m} gene (bp172-1362) with silent mutations was synthesized by

361 GenScript and amplified with NtLm-154f and NtLm-1343r primers (Figure 1a, S1). The upstream

362 and downstream fragments were amplified from tobacco DNA with LSU-FL1f+NtLm-171r and

363 NtLm-1363f+NtLrev primers respectively. The three amplicons were joined with overlapping

364 PCR with LSU-FL1f and NtLrev primers, digested with ClaI and MauBI and ligated into the

365 similarly digested pCT-rbcL vector (described in Lin *et al*. 2014) to obtain pCT-Nt-rbcL^{m} vector

366 (Figure 1b).

367 To introduce V101I, L225I, A281S, H282N or A281S/H282N amino-acid substitutions,

368 amplicons obtained with the respective forward primers and NtLrev were digested with *Mlu*I-

369 *Age*I, *Age*I-*Hind*III or *Hind*III-*Not*I and ligated into the similarly digested pCT-Nt-rbcL^{m} vector.

370 V255A, L270I and C449G mutations were introduced by ligating the amplicons from NtL-97SF

371 and the respective reverse primers into the *Mlu*I-*Sac*I, *Mlu*I-*Hind*III or *Hind*III-*Not*I sites of pCT-

372 Nt-rbcL^{m} vector. To generate a *Nt-rbc*L^{m} gene with the K429Q mutation, the two amplicons

373 obtained with NtL-783SF+K429Qr and K429Qf+NtLrev primers were joined with overlapping

374 PCR, digested with *Kpn*I-*Mau*BI and ligated into the similarly digested pCT-Nt-rbcL^{m} vector. To

375 obtain pCT-Nt-rbcL^{m} vector with L225I/K429Q mutations, *Nt-rbc*L^{m} with K429Q was digested

376 with *Kpn*I-*Not*I and ligated into the similarly digestd pCT-Nt-rbcL^{m}-L225I vector. All modified *Nt-*

377 *rbc*L^{m} genes were confirmed by Sanger sequencing with NtL-219SR, NtL-97SF and NtL-783SF

378 primers.

379

380 **Generation of tobacco transformants with modified *Nt-rbc*L^{m} genes**

381 Each MS agar plate with 25 two-week-old seedlings of either the Samsun or the Petit Havana

382 cultivars of *Nicotiana tabacum* occupying 1 inch^{2} area was bombarded twice with 0.6 µm gold

383 nanoparticles coated with pCT-Nt-rbcL^{m} vectors in a PDS-1000/He Particle Delivery System (Bio-
384 Rad Laboratories) and the leaf tissues were placed on RMOP agar medium with 500 µg/mL
385 spectinomycin as described previously (Maliga and Tungsuchat-Huang 2014). Five plates of
386 seedlings were bombarded for each transformation vector. The *rbc*L locus was amplified from
387 DNA extracted from the shoots arising from the first selection round with NtL-97SF and NtLrev
388 primers. The amplicons were purified, digested with MluI, AgeI or HindIII and analyzed on an
389 agarose gel. The shoots possessing the restriction sites in the *rbc*L gene were cut into 25 mm^2
390 pieces and placed on RMOP agar medium with 500 µg/mL spectinomycin. After shoots
391 appeared 4-6 weeks later, DNA was extracted from each shoot with CTAB solution (Allen *et al.*
392 2006), 1 µg each was digested with NheI and XhoI and the RFLP was analyzed on a DNA blot
393 with a DIG-labeled probe as described previously (Orr *et al.* 2020). The shoots homoplasmic
394 with *Nt-rbc*L^{m} genes were then placed on MS agar medium with 500 µg/mL spectinomycin for
395 rooting and subsequently transferred to soil. The *Nt-rbc*L^{m} genes in those plants were confirmed
396 by Sanger sequencing.
397

398 **Analyses of *Nt-rbc*L^{m} transcripts in the tobacco transformants**

399 RNA was extracted and purified from young leaves of five-week-old plants with PureLink RNA
400 mini kit from Life Technologies, and 200 ng of each sample was separated in a denaturing gel
401 with 1.3% agarose gel and 2% formaldehyde, transferred to a Nylon membrane and detected
402 with a DIG-labeled RNA probe as described previously (Occhialini *et al.* 2016). The
403 chemiluminescence detection was performed with a ChemiDoc MP imaging system (Bio-Rad)
404 and the band intensities were quantified with Gel Analyzer options in ImageJ software
405 (https://imagej.nih.gov/ij/). The experiment was performed for two sets of plants. Five different
406 RNA concentrations (25, 50, 100, 200 and 400 ng) from a wild-type cv. Samsun plant were
407 applied along with each set of samples, and a quadratic function obtained from their band
408 intensities was used to estimate the relative *rbcL* transcript quantity in each band. Each *rbcL*
409 transcript quantity was then normalized with the corresponding 25S rRNA band intensity
410 obtained from the UV exposure of the agarose gel before transfer.
411

412 **Analyses of Rubisco holoenzyme on blue native PAGE**

413 Leaf tissues were homogenized in an extraction buffer consisting of 100 mM Bicine-NaHO pH
414 7.9, 5 mM $MgCl_2$, 1 mM EDTA, 5 mM ε-aminocaproic acid, 50 mM 2-mercaptoethanol, 5% (w/v)
415 PEG 4000, 10 mM $NaHCO_3$, 10 mM DTT and Pierce™ protease inhibitor mini tablets (Thermo
416 Scientific part # A32955) and insoluble materials were removed with centrifugation at 16,000 x
417 g at 4°C for 5 min. The protein concentrations were estimated with the Bradford method using a
418 protein assay dye reagent concentrate (Bio-Rad part # 5000006), and 3 µg of each total soluble
419 extract mixed with a sample buffer consisting of 50 mM BisTris, 50 mM NaCl, 10% w/v glycerol
420 and 0.001% w/v Ponceau S pH 7.2 was loaded to a NativePAGE™ 3-12% Bis-Tris protein gel
421 (Thermo Scientific part # BN1003BOX). The electrophoresis was carried out in an XCell™
422 *SureLock*™ Mini-Cell with an anode buffer consisting of 50 mM Bis-Tris and 50 mM Tricine pH
423 6.8 and a cathode buffer consisting of additional 0.002% (w/v) Coomassie G-250 dye at 4°C
424 150V for 30 min and 250V for ~ 60 min. The gel was fixed with 100 mL of 40% methanol and
425 10% acetic acid solution for 15-30 min, stained with 100 mL 0.02% Coomassie R-250 in 30%
426 methanol and 10% acetic acid solution for 15-30 min and destained with 100 mL 8% acetic acid
427 solution for 2-5 hours.

428

429 **Plant material**

430 To obtain plant material for protein analyses, seeds of wild type and transgenic *N. tabacum* cv.
431 Petit Havana and cv. Samsun were sown into planter trays of a commercial potting mix
432 (Petersfield Products, UK) with a slow-release fertiliser (Osmocote, Scotts UK Professional, UK).
433 All lines germinated at similar time post sowing (6 d), and seedlings were thinned out after *ca.*
434 two weeks, with individuals transferred to 1 L pots after *ca.* three weeks. Plants were grown
435 inside a heated glasshouse at Lancaster University, UK, during June-July with minimum
436 day/night temperatures of 25/18 ± 2°C and a 16 h photoperiod. Supplemental lighting was
437 supplied by sodium lamps when light levels fell below 200 µmol m^{-2} s^{-1}. Plants were kept well-
438 watered.

439

440 **Protein extraction and analysis**

441 Leaf samples for Rubisco analyses were collected from the youngest fully expanded leaf 32 days
442 after sowing. Samples were collected 3 h after the beginning of the photoperiod, and plants
443 positioned to avoid shading for at least 1 h prior to sampling. Three 0.5 cm^2 leaf discs were
444 rapidly collected using a cork borer and immediately snap frozen, then stored at -80 °C prior to
445 analysis.

446 Fozen leaf samples were homogenised in an ice-cold mortar and pestle in 0.6 mL of
447 extraction buffer, the soluble proteins collected after centrifugation for 1 min at 4 °C and 14,700
448 × *g*, followed immediately by assays of initial and total Rubisco activity, as described by Carmo-
449 Silva *et al*. (2017). Rubisco activation state was calculated as the ratio of initial/total activity. An
450 aliquot (100 µL) of the same soluble protein extract was incubated at RT for 30 mins with 100 µL
451 of CABP binding buffer (Carmo-Silva *et al.* 2017) including [^{14}C]CABP (carboxyarabintol-1,5-
452 bisphosphate), for determination of Rubisco content via [^{14}C]CABP binding (Sharwood *et al.*
453 2016c).

454 The same soluble protein extract was used to determine total soluble protein (TSP) via
455 Bradford assay (Bradford 1976). The method of Wintermans and de Mots (1965) was used to
456 determine chlorophyll content, using 20 µL of the homogenate taken in duplicate prior to
457 centrifugation. This was added to 480 µL ethanol, inverted to mix, and kept in the dark for 2-3 h.
458 Absorbance was measured using a SPECTROstar Nano (BMG LabTech, UK) to determine
459 concentrations of chlorophyll a and b.

460

461 **Statistical analysis**

462 Statistical differences between biochemical trait means were assessed via ANOVA. Where a
463 genotype effect was observed ($P < 0.05$), a post-hoc Tukey's honest significance difference test
464 was used to conduct multiple pairwise comparisons. Analyses were performed using RStudio
465 (version 1.2.5033 (R Studio Team 2019)) and R (version 3.6.2, (R Core Development Team
466 2013)). Plots were prepared with ggplot2 (Wickham 2016). Outliers were detected using the
467 Tukey's fences method, where outliers are defined as extreme values that are 1.5 times the
468 interquartile range (1.5 IQR) below the first quartile or 1.5 IQR above the third quartile.

469

Data availability

The data generated in this study can be obtained from the corresponding author upon request.

Accession numbers

Tobacco RbcL sequence data can be found in the GenBank data library under accession number NP_054507.1.

ACKNOWLEDGMENTS

This work was funded by the UK Biotechnology and Biological Sciences Research Council (BBSRC) under grant number BB/I024488/1 to M.A.J.P., the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy under awards numbers DE-SC0014339 and DE-SC0020142 to M.R.H. and the US National Science Foundation (NSF) under grant number MCB-1642386 to M.R.H.

List of author contributions:

MTL, MRH, DJO, ECS & MAJP conceived research. MTL, MRH, DJO & ECS designed experiments. MTL, DJO & DW performed the experiments and analyzed data. All authors contributed to writing the manuscript.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

493 **SUPPORTING INFORMATION**

494

495

496 **Supplemental Figure S1.** Sequence alignment of native *rbc*L and modified *rbc*L (*Nt-rbc*L^m)

497 **Supplemental Figure S2.** Blue native PAGE analyses of soluble proteins in the transformants

498

499 **Supplemental Table S1.** Chlorophyll data from analysis of tobacco transplastomic lines.

500 **Supplemental Table S2.** Oligonucleotide sequences used in the construction of transformation

501 vectors

502 **REFERENCES**

503

504 **Aigner, H., Wilson, R.H., Bracher, A., Calisse, L., Bhat, J.Y., Hartl, F.U. and Hayer-Hartl, M.**
505 (2017) Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including
506 BSD2. *Science*, 358, 1272-1278. https://doi.org/10.1126/science.aap9221

507 **Allen, G.C., Flores-Vergara, M.A., Krasynanski, S., Kumar, S. and Thompson, W.F.** (2006) A
508 modified protocol for rapid DNA isolation from plant tissues using
509 cetyltrimethylammonium bromide. *Nat. Protoc.*, 1, 2320-2325.
510 https://doi.org/10.1038/nprot.2006.384

511 **Andersson, I. and Backlund, A.** (2008) Structure and function of Rubisco. *Plant Physiol.*
512 *Biochem.*, 46, 275-291. https://doi.org/10.1016/j.plaphy.2008.01.001

513 **Atkinson, N., Leitão, N., Orr, D.J., Meyer, M.T., Carmo-Silva, E., Griffiths, H., Smith, A.M., and**
514 **McCormick, A.J.** (2017). Rubisco small subunits from the unicellular green alga
515 Chlamydomonas complement Rubisco-deficient mutants of Arabidopsis. *New Phytol*.
516 214:655–667.

517 **Bock, R.** (2015). Engineering Plastid Genomes: Methods, Tools, and Applications in Basic
518 Research and Biotechnology. *Ann. Rev. Plant. Biol*. 66, 211-241.

519 **Bracher, A., Whitney, S.M., Hartl, F.U. and Hayer-Hartl, M.** (2017) Biogenesis and metabolic
520 maintenance of Rubisco. *Annu. Rev. Plant. Biol.*, 68, 29-60.
521 https://doi.org/10.1146/annurev-arplant-043015-111633

522 **Bradford, M.M.** (1976) A rapid and sensitive method for the quantitation of microgram
523 quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, 72,
524 248-254. https://doi.org/10.1006/abio.1976.9999

525 **Busch, F.A.** (2020) Photorespiration in the context of Rubisco biochemistry, CO2 diffusion and
526 metabolism. *Plant J.*, 101, 919-939. https://doi.org/10.1111/tpj.14674

527 **Carmo-Silva, E., Andralojc, P.J., Scales, J.C., Driever, S.M., Mead, A., Lawson, T., Raines, C.A.**
528 **and Parry, M.A.J.** (2017) Phenotyping of field-grown wheat in the UK highlights
529 contribution of light response of photosynthesis and flag leaf longevity to grain yield. *J.*
530 *Exp. Bot.*, 68, 3473-3486. https://doi.org/10.1093/jxb/erx169

531 **Carmo-Silva, E., Scales, J.C., Madgwick, P.J. and Parry, M.A.** (2015) Optimizing Rubisco and its
532 regulation for greater resource use efficiency. *Plant Cell Environ.*, 38, 1817-1832.
533 https://doi.org/10.1111/pce.12425
534 **Christin, P.-A., Salamin, N., Muasya, A.M., Roalson, E.H., Russier, F. and Besnard, G.** (2008)
535 Evolutionary switch and genetic convergence on *rbcL* following the evolution of C_4
536 photosynthesis. *Mol. Biol. Evol.*, 25, 2361-2368.
537 https://doi.org/10.1093/molbev/msn178
538 **Flamholz, A.I., Prywes, N., Moran, U., Davidi, D., Bar-On, Y.M., Oltrogge, L.M., Alves, R.,**
539 **Savage, D. and Milo, R.** (2019) Revisiting trade-offs between Rubisco kinetic
540 parameters. *Biochemistry*, 58, 3365-3376. https://doi.org/10.1021/acs.biochem.9b00237
541 **Galmés, J., Hermida-Carrera, C., Laanisto, L. and Niinemets, Ü.** (2016) A compendium of
542 temperature responses of Rubisco kinetic traits: variability among and within
543 photosynthetic groups and impacts on photosynthesis modeling. *J. Exp. Bot.*, 67, 5067-
544 5091. https://doi.org/10.1093/jxb/erw267
545 **Galmés, J., Kapralov, M.V., Andralojc, P.J., Conesa, M.A., Keys, A.J., Parry, M.A. and Flexas, J.**
546 (2014) Expanding knowledge of the Rubisco kinetics variability in plant species:
547 environmental and evolutionary trends. *Plant Cell Environ.*, 37, 1989-2001.
548 https://doi.org/10.1111/pce.12335
549 **Galmés, J., Kapralov, M.V., Copolovici, L.O., Hermida-Carrera, C. and Niinemets, Ü.** (2015)
550 Temperature responses of the Rubisco maximum carboxylase activity across domains of
551 life: phylogenetic signals, trade-offs, and importance for carbon gain. *Photosynth. Res.*,
552 123, 183-201. https://doi.org/10.1007/s11120-014-0067-8
553 **Gawroński, P., Jensen, P.E., Karpiński, S., Leister, D. and Scharff, L.B.** (2018) Pausing of
554 chloroplast ribosomes Is induced by multiple features and is linked to the assembly of
555 photosynthetic complexes. *Plant Physiol.*, 176, 2557-2569.
556 https://doi.org/10.1104/pp.17.01564
557 **Genkov, T. and Spreitzer, R.J.** (2009) Highly conserved small subunit residues influence rubisco
558 large subunit catalysis. *J. Biol. Chem.*, 284, 30105-30112.
559 https://doi.org/10.1074/jbc.M109.044081
560 **Gunn, L.H., Martin-Avila, E., Birch, R. and Whitney, S.M.** (2020). The dependency of red
561 Rubisco on its cognate activase for enhancing plant photosynthesis and growth. *Proc*
562 *Nat Acad Sci*, 117, 25890-25896. doi: 10.1073/pnas.2011641117

563 **Hanson, M.R., Gray, B.N. and Ahner, B.A.** (2013). Chloroplast transformation for engineering of
564 photosynthesis. J. Exp. Bot., 64, 731–742. https://doi.org/10.1093/jxb/ers325

565 **Iamtham, S. and Day, A.** (2000) Removal of antibiotic resistance genes from transgenic tobacco
566 plastids. *Nat. Biotechnol.*, 18, 1172-1176. https://doi.org/10.1038/81161
567 **Ishikawa, C., Hatanaka, T., Misoo, S., Miyake, C., and Fukayama, H.** (2011). Functional
568 incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco
569 in transgenic rice. *Plant Physiol.,* 156:1603–11.
570 **Kanevski, I., Maliga, P., Rhoades, D.F. and Gutteridge, S.** (1999) Plastome engineering of
571 ribulose-1,5-bisphosphate carboxylase/oxygenase in tobacco to form a sunflower large

572 subunit and tobacco small subunit hybrid. *Plant Physiol.*, 119, 133-142.
573 https://doi.org/10.1104/pp.119.1.133
574 **Kapralov, M.V. and Filatov, D.A.** (2007) Widespread positive selection in the photosynthetic
575 Rubisco enzyme. *BMC Evol. Biol.*, 7, 73. https://doi.org/10.1186/1471-2148-7-73
576 **Kapralov, M.V., Kubien, D.S., Andersson, I. and Filatov, D.A.** (2011) Changes in Rubisco kinetics
577 during the evolution of C_4 photosynthesis in *Flaveria* (Asteraceae) are associated with
578 positive selection on genes encoding the enzyme. *Mol. Biol. Evol.*, 28, 1491-1503.
579 https://doi.org/10.1093/molbev/msq335
580 **Kapralov, M.V., Smith, J.A.C. and Filatov, D.A.** (2012) Rubisco evolution in C_4 eudicots: An
581 analysis of Amaranthaceae *Sensu Lato*. *PLoS One*, 7, e52974.
582 https://doi.org/10.1371/journal.pone.0052974
583 **Keys, A.J.** (1986) Rubisco: its role in photorespiration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*,
584 313, 325-336. https://doi.org/10.1098/rstb.1986.0040
585 **Khumsupan, P., Kozlowska, M.A., Orr, D.J., Andreou, A.I., Nakayama, N., Patron, N., Carmo-**
586 **Silva, E. McCormick, A.J.** (2020) Generating and characterising single- and multi-gene
587 mutants of the Rubisco small subunit family in Arabidopsis, *J. Exp. Bot.*,
588 https://doi.org/10.1093/jxb/eraa316
589 **Kittiwongwattana, C., Lutz, K., Clark, M. and Maliga, P.** (2007) Plastid marker gene excision by
590 the phiC31 phage site-specific recombinase. *Plant Mol. Biol.*, 64, 137-143.
591 https://doi.org/10.1007/s11103-007-9140-4
592 **Morita, K., Hatanaka, T., Misoo, S., and Fukayama, H.** (2014). Unusual small subunit that is not
593 expressed in photosynthetic cells alters the catalytic properties of Rubisco in rice. *Plant*
594 *Physiol.,* 164:69–79.
595 **Lin, M.T. and Hanson, M.R.** (2018) Red algal Rubisco fails to accumulate in transplastomic
596 tobacco expressing *Griffithsia monilis RbcL* and *RbcS* genes. *Plant direct*, 2, e00045.
597 https://doi.org/10.1002/pld3.45
598 **Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A. and Hanson, M.R.** (2014) A faster Rubisco
599 with potential to increase photosynthesis in crops. *Nature*, 513, 547-550.
600 https://doi.org/10.1038/nature13776
601 **Lin, M.T., Stone, W.D., Chaudhari, V. and Hanson, M.R.** (2020) Small subunits can determine
602 enzyme kinetics of tobacco Rubisco expressed in Escherichia coli. *Nat. Plants* 6, 1289–
603 1299. https://doi.org/10.1038/s41477-020-00761-5
604 **Maliga, P. and Bock, R.** (2011) Plastid biotechnology: food, fuel, and medicine for the 21st
605 century. *Plant Physiol.*, 155, 1501-1510. https://doi.org/10.1104/pp.110.170969
606 **Maliga, P. and Tungsuchat-Huang, T.** (2014) Plastid transformation in *Nicotiana tabacum* and
607 *Nicotiana sylvestris* by biolistic DNA delivery to leaves. *Methods Mol. Biol.*, 1132, 147-
608 163. https://doi.org/10.1007/978-1-62703-995-6_8
609 **Martin-Avila, E., Lim, Y.-L., Birch, R., Dirk, L.M.A., Buck, S., Rhodes, T., Sharwood, R.E.,**
610 **Kapralov, M.V. and Whitney, S.M.** (2020) Modifying plant photosynthesis and growth
611 via simultaneous chloroplast transformation of Rubisco large and small subunits. *Plant*
612 *Cell*, 32, 2898-2916. https://doi.org/10.1105/tpc.20.00288
613 **Nakamura, M. and Sugiura, M.** (2011) Translation efficiencies of synonymous codons for
614 arginine differ dramatically and are not correlated with codon usage in chloroplasts.
615 *Gene*, 472, 50-54. https://doi.org/10.1016/j.gene.2010.09.008

616 **Occhialini, A., Lin, M.T., Andralojc, P.J., Hanson, M.R. and Parry, M.A.** (2016) Transgenic
617 tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly
618 factors grow at near wild-type rates if provided with elevated CO_2. *Plant J.*, 85, 148-160.
619 https://doi.org/10.1111/tpj.13098
620 **Ogren, W.L.** (1984) Photorespiration: pathways, regulation, and modification. *Ann. Rev. Plant*
621 *Physiol.*, 35, 415-442. https://doi.org/10.1146/annurev.pp.35.060184.002215
622 **Ogren, W.L. and Bowes, G.** (1971) Ribulose diphosphate carboxylase regulates soybean
623 photorespiration. *Nat. New Biol.*, 230, 159-160.
624 https://doi.org/10.1038/newbio230159a0
625 **Orr, D.J., Alcantara, A., Kapralov, M.V., Andralojc, P.J., Carmo-Silva, E. and Parry, M.A.** (2016)
626 Surveying Rubisco diversity and temperature response to improve crop photosynthetic
627 efficiency. *Plant Physiol.*, 172, 707-717. https://doi.org/10.1104/pp.16.00750
628 **Orr, D.J., Worrall, D., Lin, M.T., Carmo-Silva, E., Hanson, M.R. and Parry, M.A.J.** (2020) Hybrid
629 cyanobacterial-tobacco Rubisco supports autotrophic growth and procarboxysomal
630 aggregation. *Plant Physiol.*, 182, 807-818. https://doi.org/10.1104/pp.19.01193
631 **R Core Development Team** (2013) *A language and environment for statistical computing.*
632 http://www.r-project.org/.
633 **R Studio Team** (2019) *RStudio Cloud: Integrated Development for R.*
634 https://www.rstudio.com/.
635 **Ruf, S., Forner, J., Hasse, C., Kroop, X., Seeger, S., Schollbach, L., Schadach, A. and Bock, R.**
636 (2019) High-efficiency generation of fertile transplastomic Arabidopsis plants. *Nat.*
637 *Plants*, 5, 282-289. https://doi.org/10.1038/s41477-019-0359-2
638 **Salesse-Smith, C.E., Sharwood, R.E., Busch, F.A., Kromdijk, J., Bardal, V. and Stern, D.B.** (2018)
639 Overexpression of Rubisco subunits with RAF1 increases Rubisco content in maize. *Nat.*
640 *Plants*, 4, 802-810. https://doi.org/10.1038/s41477-018-0252-4
641 **Savir, Y., Noor, E., Milo, R. and Tlusty, T.** (2010) Cross-species analysis traces adaptation of
642 Rubisco toward optimality in a low-dimensional landscape. *Proc. Natl. Acad. Sci. U. S. A.*,
643 107, 3475-3480. https://doi.org/10.1073/pnas.0911663107
644 **Sharwood, R.E.** (2017) Engineering chloroplasts to improve Rubisco catalysis: prospects for
645 translating improvements into food and fiber crops. *New Phytol.*, 213, 494-510.
646 https://doi.org/10.1111/nph.14351
647 **Sharwood, R.E., Ghannoum, O., Kapralov, M.V., Gunn, L.H. and Whitney, S.M.** (2016a)
648 Temperature responses of Rubisco from Paniceae grasses provide opportunities for
649 improving C_3 photosynthesis. *Nat. Plants*, 2, 16186.
650 https://doi.org/10.1038/nplants.2016.186
651 **Sharwood, R.E., Ghannoum, O. and Whitney, S.M.** (2016b) Prospects for improving CO_2
652 fixation in C_3-crops through understanding C_4-Rubisco biogenesis and catalytic diversity.
653 *Curr. Opin. Plant Biol.*, 31, 135-142. https://doi.org/10.1016/j.pbi.2016.04.002
654 **Sharwood, R.E., Sonawane, B.V., Ghannoum, O. and Whitney, S.M.** (2016c) Improved analysis
655 of C_4 and C_3 photosynthesis via refined *in vitro* assays of their carbon fixation
656 biochemistry. *J. Exp. Bot.*, 67, 3137-3148. https://doi.org/10.1093/jxb/erw154
657 **Sinagawa-García, S.R., Tungsuchat-Huang, T., Paredes-López, O. and Maliga, P.** (2009) Next
658 generation synthetic vectors for transformation of the plastid genome of higher plants.
659 *Plant Mol. Biol.*, 70, 487-498. https://doi.org/10.1007/s11103-009-9486-x

660 **Spreitzer, R.J., Peddi, S.R. and Satagopan, S.** (2005) Phylogenetic engineering at an interface
661 between large and small subunits imparts land-plant kinetic properties to algal Rubisco.
662 *Proc. Natl. Acad. Sci. U. S. A.*, 102, 17225-17230.
663 https://doi.org/10.1073/pnas.0508042102
664 **Studer, R.A., Christin, P.A., Williams, M.A. and Orengo, C.A.** (2014) Stability-activity tradeoffs
665 constrain the adaptive evolution of RubisCO. *Proc. Natl. Acad. Sci. U. S. A.*, 111, 2223-
666 2228. https://doi.org/10.1073/pnas.1310811111
667 **Tcherkez, G.G.B., Farquhar, G.D. and Andrews, T.J.** (2006) Despite slow catalysis and confused
668 substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly
669 optimized. *Proc. Natl. Acad. Sci. U. S. A.*, 103, 7246-7251.
670 https://doi.org/10.1073/pnas.0600605103
671 **Whitney, S.M., Baldet, P., Hudson, G.S. and Andrews, T.J.** (2001) Form I Rubiscos from non-
672 green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant*
673 *J.*, 26, 535-547. https://doi.org/10.1046/j.1365-313x.2001.01056.x
674 **Whitney, S.M., Birch, R., Kelso, C., Beck, J.L. and Kapralov, M.V.** (2015) Improving recombinant
675 Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1
676 chaperone. *Proc. Natl. Acad. Sci. U. S. A.*, 112, 3564-3569.
677 https://doi.org/10.1073/pnas.1420536112
678 **Whitney, S.M., Houtz, R.L. and Alonso, H.** (2011a) Advancing our understanding and capacity
679 to engineer nature's CO_2-sequestering enzyme, Rubisco. *Plant Physiol.*, 155, 27-35.
680 https://doi.org/10.1104/pp.110.164814
681 **Whitney, S.M. and Sharwood, R.E.** (2008) Construction of a tobacco master line to improve
682 Rubisco engineering in chloroplasts. *J. Exp. Bot.*, 59, 1909-1921.
683 https://doi.org/10.1093/jxb/erm311
684 **Whitney, S.M., Sharwood, R.E., Orr, D., White, S.J., Alonso, H. and Galmes, J.** (2011b)
685 Isoleucine 309 acts as a C_4 catalytic switch that increases ribulose-1,5-bisphosphate
686 carboxylase/oxygenase (rubisco) carboxylation rate in *Flaveria*. *Proceedings of the*
687 *National Academy of Sciences*, 108, 14688-14693.
688 https://doi.org/10.1073/pnas.1109503108
689 **Whitney, S.M., von Caemmerer, S., Hudson, G.S. and Andrews, T.J.** (1999) Directed mutation
690 of the Rubisco large subunit of tobacco influences photorespiration and growth. *Plant*
691 *Physiol.*, 121, 579-588. https://doi.org/10.1104/pp.121.2.579
692 **Wickham, H.** (2016) *ggplot2: Elegant Graphics for Data Analysis Using the Grammar of Graphics*
693 New York: Springer-Verlag.
694 **Wilson, R.H. and Hayer-Hartl, M.** (2018) Complex chaperone dependence of Rubisco
695 biogenesis. *Biochemistry*, 57, 3210-3216. https://doi.org/10.1021/acs.biochem.8b00132
696 **Wilson, R. H., Martin-Avila, E., Conlan, C. and Whitney, S.M**. (2018) An improved Escherichia
697 coli screen for Rubisco identifies a protein-protein interface that can enhance CO2-
698 fixation kinetics. *Journal of Biological Chemistry*. 293, 18-27. doi:
699 10.1074/jbc.M117.810861.
700 **Wilson, R.H., Thieulin-Pardo, G., Hartl, F.U. and Hayer-Hartl, M.** (2019) Improved recombinant
701 expression and purification of functional plant Rubisco. *FEBS Lett.*, 593, 611-621.
702 https://doi.org/10.1002/1873-3468.13352

703 **Wintermans, J.F. and de Mots, A.** (1965) Spectrophotometric characteristics of chlorophylls *a*
704 and *b* and their pheophytins in ethanol. *Biochim. Biophys. Acta*, 109, 448-453.
705 https://doi.org/10.1016/0926-6585(65)90170-6
706 **Yoon, D.-K., Ishiyama, K., Suganami, M., Tazoe, Y., Watanabe, M., Imaruoka, S., Ogura, M.,**
707 **Ishida, H., Suzuki, Y., Obara, M., Mae, T. and Makino, A.** (2020) Transgenic rice
708 overproducing Rubisco exhibits increased yields with improved nitrogen-use efficiency
709 in an experimental paddy field. *Nat. Food*, 1, 134-139. https://doi.org/10.1038/s43016-
710 020-0033-x
711 **Yu, Q., Lutz, K.A. and Maliga, P.** (2017) Efficient plastid transformation in Arabidopsis. *Plant*
712 *Physiol.*, 175, 186-193. https://doi.org/10.1104/pp.17.00857
713 **Zhu, X.G., Ort, D.R., Parry, M. and von Caemmerer, S.** (2020) A wish list for synthetic biology in
714 photosynthesis research. *J. Exp. Bot.* https://doi.org/10.1093/jxb/eraa075
715 **Zhu, X.G., Portis, A.R. and Long, S.P.** (2004) Would transformation of C_3 crop plants with foreign
716 Rubisco increase productivity? A computational analysis extrapolating from kinetic
717 properties to canopy photosynthesis. *Plant Cell Environ.*, 27, 155-165.
718 https://doi.org/10.1046/j.1365-3040.2004.01142.x

Table 1. Summary on the generation of tobacco chloroplast transformants with point mutations in the Rubisco large subunit and the number of independent shoots obtained at each step of selection.

Cultivar	Mutations	Predicted roles and references	**Number of independent shoots obtained at each stage of selection**			
			1st round	1st round with expected restriction fragments	homoplasmic shoots after the 2nd selection round	Final independent transformants
Samsun	None (NtLwt)	Unmodified Rubisco	8	2	1	1*
Samsun	V101I	C_3 to C_4 transition (Christin *et al.* 2008, Studer *et al.* 2014)	2	1	1	1
Samsun	V255A	C_3 to C_4 transition (Studer *et al.* 2014, Christin *et al.* 2008, Kapralov and Filatov 2007)	4	1	1	1
Samsun	A281S	C_3 to C_4 transition (Christin *et al.* 2008, Studer *et al.* 2014)	8	3	3	3
Samsun	H282N	C_3 to C_4 transition (Studer *et al.* 2014)	5	3	3	3
Samsun	A281S,H282N	See A281S and H282N above	4	2	2	2
Petit Havana	L270I	C_3 to C_4 transition (Christin *et al.* 2008, Studer *et al.* 2014)	7	4	3	3
Petit Havana	L225I, K429Q	C_3 branch for L225I (Studer *et al.* 2014, Kapralov and Filatov 2007)	5	2	2	2
Petit Havana	K429Q	Found in paternal parent (*N. tomentosiformis*)	4	3	3	2
Petit Havana	C449G	Improved catalytic efficiency (Orr *et al.* 2016)	3	1	1	1

719 * The NtLwt transformant possesses all the silent mutations in the *Nt-rbc*L^m gene except for a single mutation necessary for the new HindIII site.
720 All the other transformants have the entire set of silent mutations in the *Nt-rbc*L^m gene as well as the intended non-synonymous mutations.
721

Table 2. Rubisco carboxylation rate (k_{cat}), content and total soluble protein (TSP) of transplastomic tobacco producing Rubisco with mutations in the large subunit. Transplastomic lines were compared statistically with the respective wild-type background. Values represent mean ± S.E. (n = 3-6 replicate plants), letters denote significant differences ($p < 0.05$) as determined by Tukey's pairwise comparisons following ANOVA.

722

Genotype	Rubisco k_{cat} (s^{-1})	Rubisco (µmol sites m^{-2})	TSP (g m^{-2})	TSP (% Rubisco)	Activation state (%)
WT-SS	2.6^{a} ± 0.2	24.6 ± 2.8	4.8 ± 0.5	36.2 ± 2.7	82.4 ± 2.6
NtLwt	2.4^{ab} ± 0.2	18.6 ± 2.5	4.0 ± 0.7	36.0 ± 4.6	78.0 ± 2.9
V101I	1.6^{b} ± 0.1	19.5 ± 1.1	4.7 ± 0.1	29.1 ± 1.3	71.2 ± 6.2
V255A	1.6^{b} ± 0.0	18.2 ± 2.9	3.7 ± 0.2	34.1 ± 3.1	69.4 ± 5.9
A281S	1.8^{ab} ± 0.2	17.8 ± 2.7	3.5 ± 0.3	35.4 ± 4.3	74.4 ± 5.1
H282N	2.2^{ab} ± 0.3	15.3 ± 2.8	3.2 ± 0.2	33.6 ± 4.4	71.1 ± 4.5
A281S/H282N	2.1^{ab} ± 0.2	18.9 ± 3.3	3.5 ± 0.3	37.2 ± 3.8	72.4 ± 2.1
	P = **0.007**	= 0.337	= 0.249	= 0.842	= 0.216
Genotype	**Rubisco k_{cat} (s^{-1})**	**Rubisco (µmol sites m^{-2})**	**TSP (g m^{-2})**	**TSP (% Rubisco)**	**Activation state (%)**
WT-PH	2.5 ± 0.2	26.6^{a} ± 1.3	5.6^{a} ± 0.4	33.9 ± 2.5	71.1 ± 2.5
C449G	1.5 ± 0.2	22.1^{ab} ± 4.2	5.2^{ab} ± 0.5	29.3 ± 3.2	68.3 ± 1.9
K429Q	1.9 ± 0.1	19.7^{ab} ± 2.2	5.0^{ab} ± 0.6	28.0 ± 2.6	68.4 ± 2.8
L225I/K429Q	2.4 ± 0.3	15.1^{b} ± 0.7	3.2^{b} ± 0.6	35.6 ± 6.7	83.6 ± 4.5
L270I	1.9 ± 0.2	19.3^{ab} ± 2.4	4.4^{ab} ± 0.3	30.9 ± 3.3	74.7 ± 3.9
	P = 0.054	= **0.018**	= **0.025**	= 0.612	= 0.165

723
724
725
726
727
728
729 **FIGURE LEGENDS**
730

731 **Figure 1.** Modification of the tobacco *rbc*L gene for chloroplast transformation. (a) Schematic
732 diagrams of the synthesized *Nt-rbc*L^{m} gene with modified restriction sites. (b) The chloroplast
733 transformation vector pCT-Nt-rbcL^{m} for replacing the tobacco *rbc*L gene with *Nt-rbc*L^{m}. (c)
734 Comparison of two hypothetical scenarios of cross-over events between the *rbcL* locus of the
735 plastid genome and pCT-Nt-rbcL^{m} plasmid upstream of the selectable marker aadA gene. The
736 location of a hypothetical point mutation in *Nt*-rbcL^{m} gene on pCT-Nt-rbcL^{m} plasmid is indicated

737 with an 'X'. The two cross-over events are indicated with dashed lines between the plastid
738 genome and pCT-Nt-rbcL^{m} plasmid. If the cross-over takes place in Flank 1 or inside Nt-rbcL
739 upstream of 'X' as in event 1, the point mutation will be introduced into the transformant. On
740 the other hand, the cross-over taking place downstream of 'X' will fail to introduce the point
741 mutation into the transformant. Note that although the segment upstream of *Nt-rbcL* has been
742 marked as "Flank 1", *Nt-rbcL* should be considered as part of Flank 1 due to high homology
743 between *Nt-rbcL* and *Nt-rbcL*^{m}.
744
745 **Figure 2.** Selection of tobacco transformants with the *Nt-rbc*L^{m} gene. (a) Restriction analyses of
746 the amplified *rbc*L locus from tobacco transformants after the first selection round. The
747 examples shown include *MluI*, *AgeI* and *HindIII* digests of transformants with V101I, H282N and
748 A281S/H282N mutations respectively. The samples with digestion fragments expected for *Nt-*
749 *rbc*L^{m} are indicated with asterisks (*). (b) Schematics of the *rbc*L loci of the wild-type and *Nt-*
750 *rbc*L^{m} transformants with the restriction sites used in DNA blot analyses. The green bars
751 represent the region where the probe binds. (c) DNA blot analyses of tobacco transformants
752 with *Nt-rbc*L^{m}. The samples with successful replacement of the native *rbc*L gene with *Nt-rbc*L^{m}
753 are indicated with asterisks (*).
754
755 **Figure 3**. DNA blot analyses of the tobacco chloroplast transformants with *Nt-rbcL*^{m} gene
756 encoding Rubisco large subunit mutants. (a) Schematics of the the *rbcL* loci of the wild-type (wt)
757 and *Nt-rbcL*^{m} transformants with the restriction sites used in DNA blot analyses. The green bars
758 represent the region where the probe binds. (b) DNA blot indicating successful replacement of
759 the native *rbcL* gene with *Nt-rbcL*^{m} in the transformants.
760
761 **Figure 4.** RNA blot analyses of the *rbcL* transcripts in tobacco transformants with *Nt-rbcL*^{m}
762 encoding Rubisco large subunit mutants. (a) The RNA blot for *rbcL* transcripts in the wt (left six
763 lanes and transformants (right ten lanes). In addition to the monocistronic *rbcL* transcript, all
764 transformants display an additional dicistronic transcript with *rbcL* and *aadA* genes. The bottom
765 panel shows ethidium bromide staining of the agarose gel. 25S rRNA band was used to estimate

766 the relative amounts of total RNA loaded for each sample. The transformants are in the same
767 order as in (b). (b) The relative quantities of *rbcL* transcripts in the transformants compared to
768 that in the wild-type plants obtained from two sets of plants. *P* values were determined by
769 Tukey's pairwise comparisons following ANOVA.
770
771 **Figure 5.** Visual comparison of plant development. (a) wild-type tobacco cv. Samsun (WT-SS)
772 and *rbc*L transplastomic lines, and (b) wild-type tobacco cv. Petite Havana (WT-PH) and *rbcL*
773 transplastomic lines. Photos were taken of plants at the same age (35 days after sowing) and
774 growth stage.
775