1 **Mutation of the *ALBOSTRIANS* Ohnologous Gene *HvCMF3***

2 **Impairs Chloroplast Development and Thylakoid Architecture**

3 **in Barley due to Reduced Plastid Translation**

4 Mingjiu Li[a], Goetz Hensel[b], Michael Melzer[c], Astrid Junker[d], Henning Tschiersch[e],

5 Daniel Arend[f], Jochen Kumlehn[b], Thomas Börner[g,1] and Nils Stein[a,h,1]

6 [a]Genomics of Genetic Resources Group, Department of Genebank, Leibniz Institute of

7 Plant Genetics and Crop Plant Research (IPK), 06466 Seeland, Germany

8 [b]Plant Reproductive Biology Group, Department of Physiology and Cell Biology, IPK,

9 06466 Seeland, Germany

10 [c]Structural Cell Biology Group, Department of Physiology and Cell Biology, IPK, 06466

11 Seeland, Germany

12 [d]Acclimation Dynamics and Phenotyping Group, Department of Molecular Genetics,

13 IPK, 06466 Seeland, Germany

14 [e]Heterosis Group, Department of Molecular Genetics, IPK, 06466 Seeland, Germany

15 [f]Bioinformatics and Information Technology Group, Department of Breeding Research,

16 IPK, 06466 Seeland, Germany

17 [g]Molecular Genetics Group, Institute of Biology, Humboldt University, 10115 Berlin,

18 Germany

19 [h]Department of Crop Sciences, Center for Integrated Breeding Research (CiBreed),

20 Georg-August-University, Göttingen, Germany

21 **[1]Corresponding authors:** Thomas Börner: thomas.boerner@rz.hu-berlin.de; Nils

22 Stein; stein@ipk-gatersleben.de

23 **Short title:** HvASL/HvCMF3 needed for chloroplast development

24 **One-sentence summary:** Phylogenetic and mutant analyses of the barley protein HvCMF3
25 (ALBOSTRIANS-LIKE) identified, in higher plants, a subfamily of CCT domain proteins with
26 essential function in chloroplast development.

27

28 The author responsible for distribution of materials integral to the findings presented in
29 this article in accordance with the policy described in the Instructions for Authors
30 (www.plantcell.org) is: Nils Stein (stein@ipk-gatersleben.de)

ABSTRACT

Gene pairs resulting from whole genome duplication (WGD), so-called ohnologous genes, are retained only if at least one gene of the pair undergoes neo- or subfunctionalization. Sequence-based phylogenetic analyses of the ohnologous genes *ALBOSTRIANS* (*HvAST/HvCMF7*) and *ALBOSTRIANS-LIKE* (*HvASL*/*HvCMF3*) of barley (*Hordeum vulgare*) revealed that they belong to a newly identified subfamily of genes encoding CCT domain proteins with putative N-terminal chloroplast transit peptides. Recently, we showed that HvCMF7 is needed for chloroplast ribosome biogenesis. Here we demonstrate that mutations in *HvCMF3* lead to seedlings delayed in development. They exhibit a *xantha* phenotype and successively develop pale green leaves. Compared to the wild type, plastids of the mutant seedlings show decreased PSII efficiency and lower amounts of ribosomal RNAs; they contain less thylakoids and grana with a higher number of more loosely stacked thylakoid membranes. Site-directed mutagenesis of *HvCMF3* identified a previously unknown functional region, which is highly conserved within this subfamily of CCT domain containing proteins. HvCMF3:GFP fusion constructs localized to plastids. *Hvcmf3Hvcmf7* double mutants indicated epistatic activity of *HvCMF7* over *HvCMF3.* The chloroplast ribosome deficiency is discussed as the primary defect of the *Hvcmf3* mutants. Our data suggests that HvCMF3 and HvCMF7 have similar but not identical functions.

INTRODUCTION

Chloroplasts are the photosynthetic active type of plastids. Functional chloroplasts normally exhibit an ellipsoidal shape and contain stroma and thylakoid membranes. The thylakoid membranes are the site of light-dependent photosynthesis reactions as mediated by four protein complexes – photosystem I (PSI), photosystem II (PSII), cytochrome b_6f and ATPase (Dekker and Boekema, 2005). Thylakoid membranes appear either in stacks of thylakoid discs, termed grana, or they exist as stroma lamellae, sheets of lipid-bilayers interconnecting the grana. While PSII is mainly found in the grana membranes, PSI and the ATPase complex are enriched in the lamellae, and the cytochrome b_6f complex is distributed evenly between the two structures (Dekker and Boekema, 2005).

Chloroplasts originated from photosynthetic cyanobacteria (Gould et al., 2008). They contain their own genome with a core set of approximately 100 genes inherited from the cyanobacterial ancestor and possess their own machinery for gene expression, i.e., for transcription, transcript processing and translation (Pogson and Albrecht, 2011; Börner et al., 2014; Pogson et al., 2015). Extensive studies have demonstrated that chloroplast development and function require the import of nucleus-encoded proteins; actually, more than 95% of the chloroplast proteins are encoded by the nuclear genome and subsequently targeted to the chloroplasts, in most cases with help of an N-terminal transit peptide, cTP (Leister, 2003; Lee and Hwang, 2018).

The extant land plants exhibit very high species diversity, which is the outcome of a long lasting evolutionary process, during which polyploidisation is considered as having provided one of the major driving forces (De Bodt et al., 2005; Soltis et al., 2009; Lafon-Placette et al., 2016; Van de Peer et al., 2017; Vamosi et al., 2018). Whole genome duplication (WGD) is widespread across land plants as revealed by genome sequencing of an increasing number of species (Muhlhausen and Kollmar, 2013). After WGD genomes tend to return - through a process called diploidisation - to the more stable and less redundant diploid stage. Thus, one copy of all the duplicated genes will be lost in a more or less random fashion. There are three possibilities for the evolutionary fate of duplicated genes (Lynch and Conery, 2000). In most of the cases, the function of one

4

80 copy is lost either by complete deletion of the gene or through accumulating non-sense
81 or deleterious mutations. In maize, a recent auto-polyploid, nearly half of the duplicated
82 genes were lost during evolution (Lai et al., 2004). If both ohnologous genes are
83 retained, one copy typically acquires a novel, beneficial function (neo-functionalization),
84 conserved during natural selection (Lynch and Conery, 2000). The second scenario to
85 maintain duplicated gene pairs is by sub-functionalization; each gene of an ohnologous
86 pair partially retain the original function, but only together providing the complete
87 functional capacity of the ancestral gene (Force et al., 1999).

88 A common ancestor of the family of the *Poaceae*, comprising all extant cereal crops,
89 underwent WGD at around 70 million years ago (Paterson et al., 2004). Traces of this
90 WGD are conserved in the barley (*Hordeum vulgare*) genome (Thiel et al., 2009) and
91 were detected, e.g. as pairs of ohnologs among genes coding for CCT domain
92 containing proteins in the genomes of cereal crops (Cockram et al., 2012). The CCT
93 domain [from the three Arabidopsis (*Arabidopsis thaliana*) proteins CONSTANS,
94 CONSTANS-LIKE and TIMING OF CAB1] comprises 43 amino acids and is found near
95 the C-terminus of numerous proteins. As far as a function could be assigned, CCT
96 domain proteins are transcription (co-) factors typically involved in modulating flowering
97 time, light-induced signaling and circadian rhythms.

98 Among the genes with a proposed ohnologous relationship, the genes *HvCMF7*
99 (*ALBOSRIANS*) and its paralog *HvCMF3* (*ALBOSTRIANS-LIKE*) are representing
100 ohnologs within the CCT domain gene family of barley (Cockram et al., 2012; Li et al.,
101 2019). A mutation in *HvCMF7* confers the variegated "*albostrians*" phenotype (Li et al.,
102 2019). Besides incomplete penetrance of its variegation phenotype (Hagemann and
103 Scholz, 1962) one of the most prominent characteristics of the *albostrians* mutant are
104 the ribosome-free plastids leading to albino leaves and albino sectors of striped leaves
105 (Hess et al., 1993). The mutant served as a model to study the cross-talk between
106 nucleus and the other DNA-containing organelles and greatly extended the field of
107 chloroplast biology (Bradbeer et al., 1979; Hess et al., 1993; Zhelyazkova et al., 2012).
108 The lack of plastid ribosomes and the albino phenotype of the *albostrians* mutant
109 indicate that the presence of the wild-type allele of the ohnologous gene *HvCMF3*

110 cannot rescue the effects of the mutation in *HvCMF7* suggesting that the two ohnologs

111 do not act at redundancy. Strikingly, the ALBOSTRIANS protein HvCMF7 was localized

112 to chloroplasts and the phenotype of the *albostrians* mutant implies that HvCMF7 plays

113 a role in the biogenesis and/or stability of chloroplast ribosomes, i.e., has a function and

114 location entirely different from all previously investigated CCT domain proteins (Li et al.,

115 2019). In contrast, the Arabidopsis homolog of *HvCMF7* and *HvCMF3*, *AtCIA2,* codes

116 for a nuclear transcription factor regulating genes for the transport of nuclear encoded

117 proteins into chloroplasts and for the biogenesis of chloroplast ribosomes (Sun et al.,

118 2009). This function and localization is more similar to the published functions of

119 previously investigated CCT domain proteins. Intriguingly, the *Atcia2* mutant exhibited a

120 pale green phenotype and no indication of leaf variegation (Sun et al., 2001).

121 Here we report on a phylogenetic analysis of CMF genes related to the *albostrians* gene

122 *HvCMF7* and its homolog *HvCMF3* supporting the ohnologous relationship of *HvCMF7*

123 and *HvCMF3* and of their Arabidopsis homologs, *AtCIA2* and *AtCIL*. To find out if

124 HvCMF3 might have a function more similar than HvCMF7 to their Arabidopsis homolog

125 AtCIA2, we analyzed a series of *Hvcmf3* mutants. Mutants of the gene *HvCMF3* were

126 obtained by chemical or Cas9 endonuclease-triggered site-directed mutagenesis to

127 determine the phenotype conferred by a non-functional gene. Site-directed mutagenesis

128 led to the identification of a highly conserved, previously unknown protein domain,

129 which supposedly plays a key role in the determination of phenotype severity. The

130 observed chlorophyll-deficient phenotype was correlated with impaired photosynthesis,

131 distinctly decreased chloroplast rRNA levels, altered stacking of thylakoids and reduced

132 numbers of grana in overall smaller chloroplasts. HvCMF3:GFP fusions localized to

133 plastids; a feature that is shared by the protein encoded by its ohnologous gene

134 *HvCMF7* and which is distinct from the behavior of the Arabidopsis homolog AtCIA2.

135 **RESULTS**

136 **Phylogenetic Relationships of *HvASL* Homologs in Monocots and Dicots**

137 The sequence of the barley genome (Mascher et al., 2017) predicts the gene model

138 *HORVU6Hr1G021460.2* as the closest homolog of *HvCMF7*. We confirmed the

6

139 predicted gene structure by cDNA sequencing. The gene contains three exons
140 separated by two introns, and encodes a protein of 490 amino acids (AA) in length.
141 Sequence comparison of *HvCMF7* and *HORVU6Hr1G021460.2* revealed that both
142 homologs share 50.5% identity at protein level. The gene *HORVU6Hr1G021460.2* was
143 previously designated as *HvCMF3* in a study on the evolution of the CCT domain-
144 containing gene family (CMF) in *Poaceae* (Cockram et al., 2012). Homology searches
145 for *HvCMF3* and *HvCMF7* against Phytozome v12.1.6 (Goodstein et al., 2012) identified
146 131 homologous genes in 66 angiosperm species, while in 14 species with an earlier
147 evolutionary history (Supplemental Dataset 1) no genes with clear homology to
148 *HvCMF3*/*HvCMF7* could be determined. As we found a homolog also in *Amborella*
149 representing the most basal lineage in the clade of angiosperms (Drew et al., 2014), we
150 used in a further search for homologs the *Amborella* sequence as query which lead to
151 the identification of homologous sequences in the genomes of gymnosperms. The
152 homologous genes were filtered by integrity and correctness of their coding sequence;
153 as a result, 91 genes from 48 species were included in an evolutionary analysis
154 (Supplemental Dataset 1). The maximum likelihood tree shows that *Amborella*
155 *trichopoda* forms a sister clade to all the remaining angiosperm plants in accordance
156 with previous reports (Drew et al., 2014). The monocot and dicot species separate from
157 the main branch and form independent clades (Figure 1). Paralogous genes of all grass
158 species in the *Poaceae* family are divided and grouped together forming two subclades.
159 Similarly, we observed this pattern also for the dicot families *Salicaceae*, *Fabaceae*,
160 *Crassulaceae* and *Brassicaceae*, respectively. The conserved presence of paralogous
161 gene pairs in grass species indicates their origin from the ancient whole-genome
162 duplication shared among grass species (Paterson et al., 2004; Thiel et al., 2009;
163 Cockram et al., 2012), i.e., they represent ohnologous genes (ohnologs). Interestingly,
164 tetraploid species in the mono- and dicots, like *Panicum virgatum* and *Brassica rapa*,
165 consistently contain two pairs of paralogs. Evidently, all ohnologs of *HvCMF3* and
166 *HvCMF7* have been retained in the genomes of all analyzed monocot and dicot plant
167 families, strongly suggesting that all ohnologs fulfil important functions in angiosperm
168 plants and have non-redundant functions.

169 Protein alignments based on 131 HvCMF3/HvCMF7 homologs from 66 monocot and

170 dicot species showed that the C-terminal CCT domain is conserved across all analyzed

171 plant species. These proteins have also a putative N-terminal chloroplast transit peptide

172 (cTP) as predicted by ChloroP (Emanuelsson et al., 1999) (Supplemental Dataset 2)

173 suggesting a role of all or most of these proteins (including the ancestor at the origin of

174 all angiosperms) in chloroplast development and function. In the present study we

175 aimed to make first steps in the elucidation of the biological function of the barley gene

176 *HvCMF3 (ALBOSTRIANS-LIKE)*.

Hvcmf3 Mutant Exhibits a *xantha*-to-green Phenotype

178 We screened for mutants of *HvCMF3* by TILLING of an EMS-induced mutant population

179 consisting of more than 7,500 M_2 plants (Gottwald et al., 2009). Fifty-four M_2 mutant

180 families were identified representing 28 non-synonymous, 24 synonymous and 2 pre-

181 stop mutations (Figure 2A and 2B, Supplemental Table 1 and 2) and all mutant families

182 were assigned to phenotypic and genotypic analyses. Owing to the ohnologous

183 relationship of *HvCMF3* and *HvCMF7,* we screened for leaf colour variation in all

184 *HvCMF3* TILLING families. We could not observe any chlorophyll-deficient phenotype in

185 mutant families representing induced non-synonymous or synonymous single

186 nucleotide polymorphisms. In contrast, all homozygous mutants identified at M_3 stage of

187 the pre-stop TILLING family 4383-1 (carries a guanine to adenine transition at

188 nucleotide position +861 leading to a premature stop codon) exhibited a chlorophyll-

189 deficient phenotype; while the segregating wild type and heterozygous plants of this

190 family produced green seedlings (Figure 2C, Supplemental Figure 1 and Supplemental

191 Table 3). The linkage was confirmed by analysis of 245 M_4 individuals derived from nine

192 heterozygous M_3 plants. The phenotype of the homozygous *Hvcmf3* mutant in TILLING

193 family 4383-1 resembles previously identified *xantha* mutants of barley (Henningsen et

194 al., 1993). The *xantha* leaves gradually turn into green along with plant growth

195 (Supplemental Figure 1 and Figure 6F). Therefore, we describe the *Hvcmf3* mutant

196 phenotype as *xantha*-to-green. Homozygous mutants of the second pre-stop TILLING

197 family 13082-1 (carries a transversion from adenine to thymine at nucleotide position

198 +1135 leading to a premature stop codon) were identified only after propagating to the

199 M_5 generation. Also, M_5 homozygous mutants of family 13082-1 exhibit a *xantha*-to-
200 green phenotype but in comparison to the pre-stop line 4383-1 requires a shorter time-
201 span for recovery to fully green (Figure 2C). The two TILLING mutant alleles of 4383-1
202 and 13082-1 were designated as *Hvcmf3-1* and *Hvcmf3-2*, respectively. F_1 hybrids
203 formed between both mutants (*Hvcmf3-1*/*Hvcmf3-2*) displayed consistently a *xantha*-to-
204 green phenotype, thus demonstrating the allelic state of both mutations (Figure 2C),
205 which was further confirmed by analyzing an additional 50 F_2 plants (*Hvcmf3-1*/*Hvcmf3-*
206 *2*) derived from the four F_1 hybrids.

207 Based on these results we concluded that *HvCMF3*, similar to *HvCMF7*, plays a
208 fundamental role in chloroplast development.

Functional Validation of *HvCMF3* by Site-directed Mutagenesis Using Cas9
Endonuclease

211 Remarkably, the recovery rate of *xantha*-to-green phenotype of the *Hvcmf3-1* mutant
212 was much slower than that of the *Hvcmf3-2* mutant. To test whether this was an effect
213 of the different positions in the coding region of the gene of the two mutations
214 (Supplemental Figure 3C), we adopted RNA-guided Cas9 endonuclease mediated site-
215 directed mutagenesis in order to reproduce the position effect of phenotype severity.
216 Two guide RNAs (gRNAs) were designed surrounding the position of the non-sense
217 mutation of TILLING mutant 4383-1 (Figure 3A). In total, 36 primary regenerants were
218 derived from *Agrobacterium*-mediated co-transformation of both gRNAs. Thirty-four of
219 the 36 T_0 plantlets carried integral T-DNA, i.e., they were PCR positive for the presence
220 of *cas9* and the gRNA-driving *OsU3* promoter in combination with at least one gRNA
221 (Supplemental Table 4). Among them, four plants carried both gRNAs, providing the
222 potential of generating insertion/deletion (INDEL) mutations at the target region
223 (Supplemental Tables 1 and 4). Analysis of T_0 plants (Supplemental Table 5) revealed
224 short INDELs as the most frequent result of site-directed mutagenesis, however, larger
225 deletions were also detected (e.g. BG677E1A, BG677E1B and BG677E9B)
226 (Supplemental Figure 2 and Supplemental Table 5). Sequencing of cloned PCR
227 products revealed the chimeric state for most of the T_0 plants; except BG677E1B,
228 representing a homozygous mutant carrying a 316 bp deletion in the collected leaf

9

229 sample, which showed a phenotype resembling the pre-stop TILLING mutants.

230 Additionally, individual leaves from three independent chimeric T_0 mutants BG677E1E,

231 2B and 2D, with *xantha* phenotype were confirmed to harbor frame-shift mutations and

232 to lack the wild-type allele (Supplemental Figure 2 and Supplemental Table 5). We

233 screened eight T_1 plants each from all of the 14 T_0 mutant families to see transmission

234 of mutations through the germline. As expected, all homozygous and homogeneously

235 biallelic mutant plants with frameshift mutations exhibited the *xantha*-to-green

236 phenotype (Figure 3D & 3E and Supplemental Figure 2). It is worth noting that mutants

237 with a lesion at target motif 1 showed a more severe phenotype than with lesions further

238 downstream. This is not only manifested by the *xantha* leaf colour variation at early

239 developmental stage (3 DAG), but also by a slower leaf development at later stages

240 (e.g.10 DAG, Figure 3E). We named the mutant alleles in BG677E18A_6 and

241 BG677E5A_21 *Hvcmf3-3* and *Hvcmf3-4*, respectively. The site-directed mutagenesis

242 experiment consolidated our previous findings by TILLING that mutations in *HvCMF3*

243 are causal for the *xantha*-to-green mutant phenotype. Furthermore, the observed

244 position effect of the induced mutations implies that *HvCMF3* possesses (a) further

245 essential functional region(s) in addition to the C-terminal CCT domain, which is

246 expected to be removed or disrupted in the proteins of all respective induced mutants.

Identification of a Conserved Sequence Essential for HvCMF3 Function

248 Protein alignments of 131 HvCMF3/HvCMF7 homologs from 66 angiosperm species

249 revealed three highly conserved regions as well as further highly conserved AA

250 residues in addition to the CCT domain near the C-terminus and the putative N-terminal

251 cTP (Figure 4A). The *Hvcmf3-3* and *Hvcmf3-4* alleles differ potentially at protein level by

252 a truncation of 17 AA, leading to a more severe phenotype in case of *Hvcmf3-3* (Figure

253 3E & Supplemental Figure 3C). The missing peptide represents a conserved region with

254 a postulated essential functional role in the protein (conserved region 2 in Figure 4A). In

255 an attempt to test this hypothesis, we screened T_1 regenerants carrying both gRNAs

256 with the expectation to observe large deletions extending over the identified conserved

257 region. We identified four homozygous plants with in-frame deletion from mutant family

258 BG677E9B, all exhibiting the *xantha*-to-green mutant phenotype; among them, one with

259 57 bp and another three with 51 bp deletions. Since none of the deletions affected the

260 splicing site they are expected to result in 19 and 17 AA deletions, respectively, at

261 protein level (Supplemental Figures 3 and 4). The mutant allele with a 51 bp deletion is

262 designated as *Hvcmf3-5*. Two homozygous mutants (new allele *Hvcmf3-6*), carrying a

263 19 bp deletion combined with a 34 bp insertion, were identified in family BG677E2C

264 (Supplemental Figure 3). This mutation led to the substitution of seven AA at position

265 290-296 (PAVPVKD) by 12 AA (HSTDATARTGSG) (Supplemental Figure 3D). The

266 *Hvcmf3-6* mutant showed a green (wild-type) phenotype indicating that replacement of

267 the seven original AA (PAVPVKD) did not affect HvCMF3 protein function. We

268 performed conservation analysis for the *Hvcmf3-5* deleted region by comparing 116

269 homologous sequences from 59 angiosperm species as described in Material and

270 Methods. The first AA 'R' (i.e. arginine) is 100% conserved among all 116 sequences

271 (Figure 4B). As revealed by the substitution mutant *Hvcmf3-6* in family BG677E2C, the

272 C-terminal six AA (Figure 4B, positions 12-17) have no effect on HvCMF3 protein

273 function. Therefore, the peptide of AA 279-289 (Figure 4B, positions 1-11) represents a

274 previously unknown conserved functional region within the conserved domain 2. Neither

275 the identified novel functional region nor the entire conserved domain 2 of HvCMF3 is

276 reported in the NCBI's Conserved Domain Database (Marchler-Bauer et al., 2017).

277 **Reduced Chloroplast Ribosome Accumulation in *Hvcmf3* Mutants**

278 One of the most prominent characteristics of the *albostrians* mutant is the lack of

279 ribosomes in plastids of albino leaves and albino sections of striped leaves (Hess et al.,

280 1993; Li et al., 2019). We checked therefore whether mutation of *HvCMF3* has also an

281 effect on plastid ribosomes. The accumulation of rRNA levels can be used as a proxy

282 for ribosomal subunit accumulation (Walter et al., 2010). Thus, we quantified chloroplast

283 and cytosolic rRNA fractions in light- and dark-grown seedlings of *Hvcmf3* mutants. Due

284 to the *xantha*-to-green phenotype of young *Hvcmf3*, we compared with known barley

285 *xantha* mutants, *xan-g44* and *xan-f68*, which contain only trace amounts of chlorophyll

286 in their leaves due to defects in the magnesium chelatase (EC 6.6.1.1) subunits D and

287 H, respectively (Olsson et al., 2004; Axelsson et al., 2006). This enzyme catalyzes the

288 insertion of magnesium into protoporphyrin IX, the first unique step of the chlorophyll

biosynthetic pathway (Figures 5A & 5B). The relative abundance of chloroplast to cytosolic ribosomal subunits was determined by their ratios. Under light condition, *Hvcmf3* mutants as well as *xan-g44* and *xan-f68* have reduced amounts of both large (50S) and small subunits (30S) of the plastidal ribosome, as indicated by the lower 23S:25S and 16S:25S ratios, respectively (Figures 5B & 5C). It should be noted that the 23S rRNA contains so-called hidden breaks and is represented by two smaller RNAs, one of them at the position of the 18S rRNA and one below the 16S rRNA (Figure 5B) resulting in apparently higher amounts of 16S vs. 23S rRNAs (Figures 5B & 5E). The lower level of plastid rRNAs in light-grown *xan-g44* and *xan-f68* is a secondary effect of the low chlorophyll content and accumulation of chlorophyll precursors. Under these conditions, light leads to the production of ROS (Reactive Oxygen Species) in the plastids and consequently to the degradation of plastid rRNAs and low levels of plastid ribosomes (Willi et al., 2018). Interestingly, the dark-grown *Hvcmf3* mutant [only tested: *Hvcmf3-7* (Supplemental Figure 5D), exhibiting an albino-like phenotype] has very low plastid rRNA levels after growth in darkness indicating that the low content of plastid rRNA is not caused by light-induced degradation but a primary rather than a secondary effect of the mutation (Figures 5B to 5D). Consistent with the reduced amount of plastid rRNA, the chlorophyll content in the *Hvcmf3* mutants is significantly decreased compared to the wild type (Figure 5F and 5G). Mutant *Hvcmf3-1*, which exhibits the most severe phenotype, shows a higher chlorophyll *a:b* ratio than wild-type barley (Figure 5H). As PSII is enriched in chlorophyll *b* as compared to PSI, the higher chlorophyll *a:b* ratio may indicate that PSII is more severely affected than PSI in mutant *Hvcmf3-1* (Figure 5H). Nevertheless, the higher chlorophyll *a:b* ratio ameliorates during the greening process as evidenced by mutants *Hvcmf3-7* and *Hvcmf3-2*, suggesting that deficits in biogenesis of the photosynthetic complex can be compensated over time.

Mutation of *HvCMF3* Affects Photosynthesis

Because of plastid ribosome deficiency (Figure 5) the *Hvcmf3* mutants potentially suffer from insufficient levels of RNA translation in chloroplasts. Since proteins of all components of the photosynthetic apparatus are being synthesized on plastid ribosomes, the efficiency of photosynthetic electron transport can serve as a highly

319   sensitive indicator of plastid translational capacity (Rogalski et al., 2008). PSII is known

320   to require a particularly high translation capacity due to the constant requirement for

321   repair synthesis of the D1 protein (Takahashi and Badger, 2011). To test this, we

322   quantified photosynthesis-related traits in a series of *Hvcmf3* mutants with different

323   severity of their pigment-deficiency phenotype by using a chlorophyll fluorescence

324   imaging-based method integrated into an automated, conveyor-based phenotyping

325   platform (Junker et al., 2014). Initially, 96 plants from 12 families, each with 8 replicates,

326   were sown (Supplemental Figure 5, Supplemental Table 6). After filtering the non- or

327   badly-germinated seeds and the chimeric seedlings, 60 plants were left for analysis

328   including seven mutant and two wild-type families, respectively, each with four to eight

329   replicates (Supplemental Table 6). Based on the severity of phenotype, the nine plant

330   families were classified into three groups: Group I: wild type (Barke and Golden

331   Promise); Group II: mutant families 4383-1 (*Hvcmf3-1*), BG677E2A_2 (*Hvcmf3-7*) and

332   BG677E5A_21 (*Hvcmf3-4*); and Group III: BG677E5A_19 (*Hvcmf3-8*), BG677E9B_1

333   (*Hvcmf3-9*), BG677E9B_6 (*Hvcmf3-5*) and 13082-1 (*Hvcmf3-2*) (Supplemental Figure

334   5D). Consistent with the reduced amount of plastid rRNAs in the *Hvcmf3* mutants, the

335   PSII electron transport rate (ETR) is lower in the mutants compared to wild type.

336   Moreover, the ETR of Group II mutants is significantly lower than of Group III (Figure

337   6A). The quantification of PSII operating efficiency (ΦPSII) of light-adapted plants

338   revealed a lower PSII yield of the mutants compared to the wild type during early

339   developmental stages (i.e. 6-14 DAS). Moreover, PSII operating efficiency of the two

340   mutant groups also showed significant difference to each other (Figure 6B). Another

341   parameter, qP, which represents the proportion of PSII reaction centers that are open,

342   was significantly lower in the *Hvcmf3* mutants than in the wild type (Figure 6C). In line

343   with the decreased ΦPSII, the maximum quantum efficiency of PSII (F_v/F_m) was also

344   significantly reduced in the *Hvcmf3* mutants (Figure 6D). In contrast to the lower PSII

345   yield, a higher proportion of excitation energy in the *Hvcmf3* mutants was released as

346   thermal dissipation compared to the wild type (Figure 6E). Group II mutants showed

347   higher levels of non-photochemical quenching (NPQ) compared to Group III mutants

348   (Figure 6E). The distinct PSII electron transport rate and PSII operating efficiency levels

349   were also reflected by the different severity of the phenotype (Figure 6F). In line with the

350 reduced chlorophyll contents of the *Hvcmf3* mutants (Figure 5F-H), quantification of the

351 plant coloration revealed that Group II mutants have higher yellow/green pixel ratio

352 compared to Group III mutants. Meanwhile, Group II mutants exhibited smaller overall

353 projected leaf area than Group III mutants as well as the wild type due to slower

354 development (Figure 6F). Taken together, this data demonstrates that mutants of the

355 gene *Hvcmf3* show a lower PSII activity which correlates with the reduced levels of

356 plastid rRNA. Hence, mutants with the lowest plastid rRNA levels showed also the

357 lowest PSII efficiency and the lowest PSII electron transport rate. This data supports our

358 hypothesis of *Hvcmf3* mutants suffering from impaired chloroplast translation and that

359 the observed impact on PSII (PS I has not been tested) is most likely a consequence of

360 the plastid ribosome deficiency and not a direct effect of the mutations.

361 **Mutation of *HvCMF3* Affects Chloroplast Development and Grana Organization**

362 To clarify if the *HvCMF3* mutant related *xantha*-to-green phenotype is only manifested

363 in physiological or also in anatomical changes, we analyzed leaf samples of the pre-

364 stop TILLING mutants *Hvcmf3-1* and *Hvcmf3-2* at two developmental stages (3 and 10

365 days after germination, DAG) by transmission electron microscopy (TEM)

366 (Supplemental Figures 6 and 7). Cells of mutant *Hvcmf3-1* contained smaller

367 chloroplasts than both the wild type and mutant *Hvcmf3-2* at 3 DAG and 10 DAG

368 (Supplemental Figure 7). At 3 DAG chloroplast size of *Hvcmf3-2* was also reduced in

369 comparison to wild type (Supplemental Figure 7A-F). At 10 DAG, chloroplast size in

370 *Hvcmf3-2* was indistinguishable from wild type, while *Hvcmf3-1* still contained smaller

371 chloroplasts (Supplemental Figure 7G-L). Compared with wild-type chloroplasts, both

372 mutants showed a distinct difference in the structure of their grana, which (at least

373 partially) were build up by a higher number of thylakoids with less condensed stacking

374 at both developmental stages (Supplemental Figure 7). Based on quantitative

375 assessments of chloroplast length, width and surface area, as well as grana number,

376 the extent of grana stacking and distance between thylakoid membranes within the

377 grana (Figure 7A & 7B), in both mutants, chloroplasts are smaller than wild-type leaves

378 at 3 DAG as determined by the parameter 'surface area' (Figure 7C-E). Chloroplast size

379 was also significantly different (Student's *t*-test, $p = 6.4 \times 10^{-15}$) between *Hvcmf3-1* and

14

380 *Hvcmf3-2*, which correlates well with the more severe phenotype of *Hvcmf3-1 vs.*

381 *Hvcmf3-2* at 3 DAG (Figure 7C-E and Figure 2C). At 10 DAG, the development of

382 chloroplast shape and morphology of mutant *Hvcmf3-1* remained delayed. In contrast,

383 although chloroplast length of mutant *Hvcmf3-2* was still reduced if compared to the wild

384 type, width and chloroplast surface area approached to wild-type level (Figure 7C-E).

385 *Hvcmf3* mutations influenced also grana organization. At 3 DAG, chloroplasts of both

386 TILLING mutants contained lower numbers of grana stacks (Figure 7F). In contrast to

387 *Hvcmf3-2*, the number of grana was significantly reduced (Student's *t*-test, $p = 6.8 \times 10^{-15}$

388) in chloroplasts of *Hvcmf3-1,* also at 10 DAG (Figure 7F). The observed increased

389 grana stacking in both mutants is a result of a higher number of thylakoids and of

390 enhanced distances between thylakoid membranes within the stacks (Figure 7G & 7H,

391 Supplemental Figure 8). In summary, the analyzed *Hvcmf3* mutants are affected in the

392 physiological parameters of PSII efficiency and electron transport rate, which is

393 underpinned by severe anatomical changes like smaller than wild-type chloroplasts

394 containing a lower number of thylakoids and larger but loosely stacked grana.

395 **HvCMF3 Is Localized to the Chloroplast**

396 Similar to its ohnolog HvCMF7, which is allocated to barley chloroplasts (Li et al., 2019),

397 *in silico* analysis by PredSL (Petsalaki et al., 2006) predicted the presence of a 95 AA

398 chloroplast transit peptide at the HvCMF3 N-terminus (Supplemental Table 7). To test

399 its function, we performed transient subcellular localization in barley epidermis of green

400 fluorescent protein (GFP) fusion constructs with either the complete wild-type *HvCMF3*

401 allele (HvCMF3:GFP) or the putative cTP of HvCMF3 only (cTP_95AA_HvCMF3:GFP)

402 (Figure 8A and Supplemental Table 7). GFP fused to wild-type HvCMF3 accumulated in

403 the plastids, co-localizing with the mCherry- labelled chloroplast allocation control

404 (Figure 8D). GFP fluorescence was also observed in the nucleus (Figure 8D), however,

405 at the same low level as was also observed for the GFP-only control (Figure 8B). A

406 plastid allocation was also observed for the cTP_95AA_HvCMF3:GFP construct

407 confirming the functionality of the predicted cTP at the N-terminal domain of HvCMF3

408 (Figure 8E). We conclude that HvCMF3, similar to its ohnolog HvCMF7, is targeted to

409 plastids.

15

Hvcmf3/Hvcmf7 Double Mutant Exhibits a Mixed *xantha*-albino Variegation Phenotype

Our results revealed that mutation of either of the ohnologs *HvCMF3* and *HvCMF7* is causing a chlorophyll-deficient phenotype. While *Hvcmf3* mutants exhibit a *xantha*-to-green recovery phenotype, *Hvcmf7* mutants show either a green-white variegation or a complete albino phenotype (Li et al., 2019). Both genes are essential for chloroplast development. *HvCMF3* mutants affect the amounts of plastid ribosomes, chloroplast size and the morphology of grana stacks while *HvCMF7* mutants do not show any development of chloroplasts and possess only proplastid-like ribosome-free plastids in their mesophyll cells. Homozygous *Hvcmf3-1/Hvcmf7-1* double mutants derived from crossing *Hvcmf7-1* x *Hvcmf3-1* showed a *xantha*-albino striped phenotype (Supplemental Figure 9D). If the more severe *Hvcmf7-2* mutant 6460-1 was used as a crossing parent (Li et al., 2019), the resulting homozygous double mutant *Hvcmf3-1/Hvcmf7-2* exhibited always the complete albino phenotype of *Hvcmf7-2* (Supplemental Figure 9E) indicating that the *HvCMF7* mutation has an epistatic effect on *HvCMF3*.

DISCUSSION

Plastid-encoded proteins are mainly involved in plastid gene transcription and translation or are playing a role in photosynthesis. Most of the genes needed for plastid functions and in particular for the development of chloroplasts and their photosynthetic apparatus are, however, encoded in the nuclear genome and are targeted to the plastid/chloroplast; including genes involved in chloroplast transcription, RNA processing, RNA stability, and translation (Börner et al., 2014; Pogson et al., 2015). Here we studied through induced mutagenesis the function of the gene *HvCMF3.* Similar to its ohnolog *HvCMF7* (Li et al., 2019), the gene *HvCMF3* codes for a nuclear protein that is involved in the biogenesis and/or stability of chloroplast ribosomes. Both *HvCMF3* and *HvCMF7* belong to the large family of genes coding for CCT proteins. While most of the intensively studied CCT domain proteins are involved in the regulation of nuclear gene transcription (Wenkel et al., 2006; Jang et al., 2008), *HvCMF7* and *HvCMF3* encoded proteins are allocated to the plastid. Mutations in both genes affect

16

440 plastid ribosomes. They either lead to the complete loss of chloroplast ribosomes
441 resulting in an albino or green-white variegated phenotype in case of *HvCMF7* (Li et al.,
442 2019), or, as in case of *HvCMF3,* show different degrees of chlorophyll and chloroplast
443 ribosome deficiency, altered thylakoid morphology and reduced photosynthetic activity.

HvCMF3 Belongs to a Small Subfamily of CCT Domain Proteins

445 *HvCMF3,* like *HvCMF7,* belongs to the gene family of CCT domain proteins. Numerous
446 CCT-containing genes represent transcription factors that regulate gene expression in
447 the nucleus through DNA-binding or by integration into DNA-binding protein complexes
448 (Wenkel et al., 2006; Jang et al., 2008). Based on their domain structure, CCT proteins
449 may be classified into COL (CONSTANS-LIKE) proteins having one or two zinc-finger
450 B-Box domains, PRR (PSEUDO RESPONSE REGULATOR) proteins with a pseudo
451 response regulator domain, and CMF (CCT MOTIF FAMILY) proteins containing only
452 the CCT domain and lacking other known functional domains (Cockram et al., 2012).
453 Both, HvCMF3 and HvCMF7, carry only a single CCT domain and thus are assigned to
454 the CMF family, which comprises nine genes in barley (Cockram et al., 2012). CMF
455 genes are found likewise in gymnosperms and angiosperms, including the Arabidopsis
456 homologs *AtCIA2* and *AtCIL* and are characterized by the presence of a putative N-
457 terminal chloroplast transit peptide (cTP) but otherwise carry a single CCT motif (AA
458 436-479 in HvCMF3) as the only annotated protein domain. In the present study as well
459 as in our previous work on the characterization of HvCMF7 (Li et al., 2019), we
460 demonstrate that HvCMF3 and HvCMF7 share the conserved cTP and CCT regions,
461 but, in contrast to other CMF domain proteins, carry additional, previously
462 uncharacterized conserved regions, one of them proved to be essential for wild-type
463 gene function in the present study. Based on the three more intensively studied
464 genes/proteins of this CMF gene sub-family, we propose to differentiate them from other
465 CMF genes by assigning them to a new CMF sub-family, the AAC proteins [for:
466 ALBOSTRIANS/HvCMF7 (Li et al., 2019), ALBOSTRIANS-LIKE/HvCMF3,
467 CHLOROPLAST IMPORT APPARATUS 2/AtCIA2 (Sun et al., 2009)]. According to the
468 phylogenetic tree of CCT domain proteins (Cockram et al., 2012), these genes form a
469 branch in a subclade of clade 2. Clade 2 comprises CMF genes/proteins characterized

470 by a specific position of an intron within the gene region coding for the CCT domain
471 (Cockram et al., 2012). We postulate that AAC proteins have evolved to support the
472 biogenesis and/or maintenance of chloroplast ribosomes in land plant species.

473 **HvCMF3 Potentially Plays a Role in Chloroplast Ribosome**
474 **Formation/Maintenance**

475 We observed a very low amount of chloroplast rRNA in leaves with low chlorophyll
476 content in *Hvcmf3* mutants at early developmental stages. Both chlorophyll and
477 chloroplast rRNA content improved with further development, however, without reaching
478 wild-type level. A further striking feature of *HvCMF3* mutants are the drastic changes in
479 the internal structures of chloroplasts with a decreased number of thylakoids and at the
480 same time larger and more loosely stacked grana. Although we cannot rule out other
481 functions of HvCMF3 yet, we regard the observed chloroplast rRNA deficiency as the
482 most likely primary effect of the studied *Hvcmf3* mutants and all other observed effects
483 of the mutations as being caused by the chloroplast translation deficiency. One reason
484 for this conclusion is that similar phenotypes have previously been described for many
485 mutants with reduced chloroplast translation. Although the phenotypes are different in
486 details and highly variable depending on the type of mutated gene (there are many
487 possibilities, e.g. genes for ribosomal proteins, tRNAs, rRNAs, translation factors, RNA
488 processing factors and others), on the severity of the translation deficiency, on the
489 phase of chloroplast development, when the translation deficiency starts to become
490 effective, all mutants with impaired chloroplast translation show pigment deficiencies,
491 lower performance of photosynthesis and altered thylakoid organization, often combined
492 with retarded growth and delayed greening (Albrecht et al., 2006; Delannoy et al., 2009;
493 Tiller and Bock, 2014; Liu et al., 2015; Kohler et al., 2016; Aryamanesh et al., 2017;
494 Zhang et al., 2017). Another reason for proposing the ribosome deficiency as primary
495 effect is that pigment deficiency, altered thylakoid organization or impaired
496 photosynthesis does not cause chloroplast ribosome deficiencies, while the opposite
497 occurs and can be explained by the function of chloroplast translation. Chloroplast
498 genes encode essential components of the photosynthetic apparatus including subunits
499 of PSI, PSII, Cytb$_6$f, ATP synthase and NDH, i.e., these proteins are synthesized on

500   chloroplast ribosomes. Thus, a reduced amount of chloroplast ribosomes, as observed

501   in *HvCMF3* mutants, will negatively affect photosynthesis and will also have effects on

502   thylakoid architecture. In this context it is interesting to note that the formation of large

503   grana was observed in a barley mutant lacking PSII reaction centers (J. Simpson et al.,

504   1989) and in Arabidopsis plants treated with the chloroplast translation inhibitor

505   lincomycin (Belgio et al., 2015). It is well established that large grana are formed under

506   low or red light *vs.* high light or blue light (Mostowska, 1986). An interruption of the

507   electron transport between photosystem II and photosystem I triggers also the formation

508   of large thylakoid stacks (Meier and K. Lichtenthaler, 1981; Jia et al., 2012). Large

509   grana have been further described in mutants with impaired starch formation (Hausler et

510   al., 2009). Recent studies point to phosphorylation levels of PSII and LHCII and/or the

511   degree of oligomerization of the thylakoid curvature protein family as regulators of the

512   dynamic changes in thylakoid stacking (Armbruster et al., 2013; Pietrzykowska et al.,

513   2014; Puthiyaveetil et al., 2017; Wood et al., 2018; Wood et al., 2019). *HvCMF3*

514   encodes a chloroplast-localized CMF protein. To our knowledge, no evidence has yet

515   been reported for any interaction of CMF proteins with kinases or phosphatases

516   involved in phosphorylation/dephosphorylation of thylakoid components, such as LHCII-

517   specific phosphatase PPH1, PSII-specific phosphatase PBCP, protein kinases STN7

518   and STN8 (Fristedt et al., 2009; Pribil et al., 2010; Samol et al., 2012), or thylakoid

519   curvator proteins (Armbruster et al., 2013). Also nuclear genes coding for proteins with

520   roles in photosynthesis, thylakoid formation, and pigment synthesis are likely affected in

521   their expression in chloroplast ribosome deficient mutants, due to plastid-to-nucleus

522   retrograde signaling (Kleine and Leister, 2016; Borner, 2017; de Souza et al., 2017;

523   Hernandez-Verdeja and Strand, 2018). Chloroplast ribosome deficiency as the reason

524   of our phenotypic observations *Hvcmf3* plants is also supported by the fact that severity

525   of ribosome deficiency is correlated with increasingly drastic effects on chlorophyll

526   content, PSII efficiency, and grana morphology.

527   We conclude that HvCMF3 plays a role in the biogenesis and/or maintenance of plastid

528   ribosomes. Its localization to plastids fits to the proposed role. Thus, HvCMF3 might

529   have a similar function as HvCMF7. The functions of HvCMF3 and HvCMF7 are,

530   however, not identical as can expected when two ohnologs have been retained in the

531   genome for a period of about 70 million years since the WGD they originate from. We

532   deduce non-identical functions for *HvCMF3* and *HvCMF7* from our observation that the

533   genes cannot substitute for each other in mutants. Moreover, the mutants of *HvCMF3*

534   and *HvCMF7* have clearly different phenotypes. While mutation of *HvCMF7* results in

535   an albino phenotype, the lack of ribosomes and, consequently, to a complete stop of

536   chloroplast development, *Hvcmf3* mutants show a *xantha*-to-green phenotype, possess

537   plastid ribosomes, although distinctly reduced in their number, and show a retarded

538   chloroplast development. Crossing the two mutants revealed an epistatic effect of

539   *HvCMF7* on *HvCMF3*. This, however, is not surprising, if the function of both proteins is

540   needed to reach the normal number of ribosomes and the malfunction of one alone

541   (HvCMF7) is already sufficient to cause the complete loss of ribosomes and the

542   complete stop of chloroplast development, that is, more effect is not possible.

543   With this initial characterization of HvCMF3, it is possible to compare the function of

544   three AAC proteins, ALBOSTRIANS, ALBOSTRIANS-LIKE and CHLOROPLAST

545   IMPORT APPARATUS 2. All three share a very similar structure with a putative N-

546   terminal cTP, several conserved domains of unknown function (the functional

547   importance of one conserved region has been demonstrated in the present study),

548   additional conserved amino acids and the CCT domain near the C terminus. The exact

549   roles of those conserved regions including the CCT domain have still to be determined.

550   The function of the cTP domain as mediator of the transport of the protein into plastids

551   has been confirmed for HvCMF3 (this report) and HvCMF7 (Li et al., 2019). Since a

552   putative cTP domain is also present in the *Amborella* homolog and in the homologous

553   proteins of gymnosperms, one might speculate about a chloroplast localization of the

554   ancestor of the AAC proteins. However, the example of the Arabidopsis protein AtCIA2,

555   reported to be a nuclear transcription factor (Sun et al., 2001) (Sun et al., 2009) shows

556   that one has to be cautious about speculations. Even though AtCIA2 is a nuclear protein

557   and HvCMF3 and HvCMF7 are chloroplast proteins, all three play roles in chloroplast

558   development. Since AtCIA2 is reported to be involved in the regulation of transcription

559   of nuclear genes coding for chloroplast ribosomal proteins and for proteins of the

560   chloroplast protein import machinery, all three AAC proteins are essential to provide

20

561 chloroplasts with an adequate number of ribosomes. Thus, the AAC family might be a

562 new source of proteins with essential functions in chloroplast development.

MATERIALS AND METHODS

Plant Material and Growth Conditions

565 M_3 TILLING families carrying single nucleotide polymorphisms (SNP) causing non-

566 synonymous or pre-stop mutations were selected for phenotyping. For each family 16

567 plants were characterized phenotypically and further genotyped for the respective

568 *HvCMF3* alleles via either Sanger sequencing or CAPS assay. The barley cultivar

569 'Golden Promise' was used for generation of the transgenic lines. The primary T_0

570 plantlets were grown in a climate chamber with long day condition (16h light/8h dark;

571 constant temperature 22°C) until reaching the third-leaf stage and then transferred to a

572 greenhouse with the same photoperiod regime but variable day/night temperature

573 20°C/15°C. Supplemental light (300 µmol photons m^{-2} s^{-1}) was used to extend the

574 natural light with incandescent lamps (SON-T Agro 400; MASSIVE-GROW, Bochum,

575 Germany). All TILLING mutants and *xantha* mutants were grown under the same

576 greenhouse condition as the transgenic lines. For dark treatment, grains were

577 germinated within a carton box wrapped with aluminum foil under the greenhouse

578 condition.

579 For automated phenotyping, after 24 hours imbibition on water-soaked filter paper,

580 germinated grains were transferred to 10 cm pot (diameter) filled with a mixture of 85%

581 (v) red substrate 1 (Klasmann-Deilmann GmbH, Geeste, Germany) and 15% (v) sand.

582 All the plants were grown under controlled conditions at 20/16°C under a circadian

583 rhythm 16-h light/8-h darkness, 70% relative humidity, photosynthetic active radiation

584 (PAR) of 300 µmol photons m^{-2} s^{-1} in the growth chamber. In total, 96 plants including

585 12 genotypes each with 8 replicates were phenotypically evaluated under the

586 LemnaTec Scanalyzer system (LemnaTec AG, Aachen, Germany) at the IPK

587 Gatersleben. The 12 genotypes consist of two TILLING mutant lines 4383-1 (*Hvcmf3-1*;

588 M5 lines) and 13082-1 (*Hvcmf3-2*; M6 lines); eight *Cas9*-induced T2 mutant lines

589 BG677E1B_3, BG677E2A_2 (*Hvcmf3-7*), BG677E5A_2, BG677E5A_21 (*Hvcmf3-4*),

21

590   BG677E5A_19 (*Hvcmf3-8*), BG677E9B_1 (*Hvcmf3-9*), BG677E9B_6 (*Hvcmf3-5*) and

591   BG677E18A_6 (Supplemental Figure 5), and the two wild-type cultivars 'Barke' and

592   'Golden Promise', which represent the genetic background of the TILLING and *Cas9*-

593   induced mutants, respectively.

Phylogenetic Analysis

595   The barley ALBOSTRIANS protein sequence was used as BLASTP query to retrieve

596   homologs from other species on NCBI and phytozome (Goodstein et al., 2012)

597   databases. Phylogenetic analysis was performed using MEGA6 (Tamura et al., 2013)

598   following the protocol of Hall (Hall, 2013). The alignment method MUSCLE was chosen

599   to build the alignment. Next, phylogenetic tree construction was performed based on the

600   Maximum Likelihood (ML) statistical method. The Bootstrap method with 1,000

601   Bootstrap Replications was set to estimate reliability of the phylogenetic tree. The

602   Jones-Taylor-Thomton (JTT) model and Gamma Distributed (G) were selected for

603   options Model/Method and Rates among Sites, respectively. The gaps were treated with

604   partial deletion option i.e., all positions containing gaps and missing data less than 95%

605   coverage were eliminated. There were a total of 264 positions in the final dataset. The

606   phylogenetic tree was visualized with iTOL (Letunic and Bork, 2016).

TILLING Screening

608   In an effort to identify *HvCMF3* mutated alleles, an EMS-induced TILLING population

609   (Gottwald et al., 2009) was screened by placing three primer pairs to cover the coding

610   regions of the *HvCMF3* gene (Supplemental Tables 1 and 2) and mutations were

611   detected as described previously (Li et al., 2019). Phenotypic and genotypic analyses

612   were performed with the M_3 progeny of the identified M_2 families, which carried non-

613   synonymous or pre-stop mutations. The two pre-stop TILLING families, 4383-1 and

614   13082-1, were further propagated and analyzed in M_4 and M_5 generations to confirm the

615   linkage between the genotype of the *HvCMF3* locus and the observed phenotype.

Site-directed Mutagenesis Using Cas9 Endonuclease

22

617 Targeted mutagenesis using Cas9 endonuclease was adopted to generate mutations in
618 the *HvCMF3* gene. In the first step, the 'KNOCKIN' tool on Deskgen Cloud was chosen
619 for guide RNA (gRNA) design (https://www.deskgen.com/landing/cloud.html). The
620 coding sequence of *HvCMF3* was used as query and two proper gRNA target motifs
621 were selected surrounding the position of the pre-stop mutation of TILLING mutant
622 4383-1. The predicted gRNA activity scored 50 and 58 for target motif 1 (3'-
623 <u>GGG</u>AGTTCGCCGCCCTGCTGCTG-5') and target motif 2 (3'-
624 <u>GGC</u>CACTTCCTGTAGTGCCAGTG-5'), respectively. Both target motifs were located at
625 the antisense strand and the underlined nucleotides represent the protospacer adjacent
626 motif (PAM). Next, the *HvCMF3*-specific protospacer sequences were synthesized by
627 introducing proper overhangs to facilitate downstream cloning steps (gRNA1 forward:
628 5'-<u>GGCG</u>TCGTCGTCCCGCCGCTTGA-3' and reverse: 5'-
629 <u>AAAC</u>TCAAGCGGCGGGACGACGAC-3'; gRNA2 forward: 5'-
630 <u>GGCG</u>TGACCGTGATGTCCTTCAC-3' and reverse: 5'-
631 <u>AAAC</u>GTGAAGGACATCACGGTCAC-3'). The protospacer sequence (i.e., annealed
632 oligonucleotides) was then cloned into vector pSH91 (Budhagatapalli et al., 2016). The
633 derived vector was designated as pGH379-7 for gRNA1 and pGH380-12 for gRNA2.
634 Subsequently, the expression cassette of pGH379-7 and pGH380-12 was transferred
635 into the binary vector p6i-d35S-TE9 (DNA-Cloning-Service, Hamburg, Germany)
636 through *Sfi*I cloning sites. The resulting plasmids pGH449-2 and pGH450-6 were co-
637 transformed into barley cv. 'Golden Promise' following a previously established protocol
638 (Hensel et al., 2009). To check for T-DNA integration in regenerated T_0 plantlets, PCR
639 primers targeting the *hpt* or *cas9* gene and the *OsU3* promoter were used in PCR
640 reactions (Supplemental Table 1). Besides, presence/absence of gRNA1 and/or gRNA2
641 of each plant were verified by protospacer-specific primers (Supplemental Table 1).
642 Primer pair HvCMF3_F2/R2 was employed to detect mutations for the pre-selected
643 target regions of *HvCMF3*. Mutations carried by the chimeric T_0 plants were further
644 characterized by sub-cloning PCR products using the CloneJET PCR cloning Kit
645 (Thermo Scientific, Wilmington, USA); at least eight colonies were sequenced. T_0 plants
646 with mutations were further propagated to T_1 generation. In analogy to analysis of the T_0
647 plants, inheritance of the mutations was checked for T_1 progenies. Additionally, T_1

23

648 plants were phenotyped in terms of its leaf colour variation during developmental stages

649 of the initial three leaves.

HvCMF3 Gene Structure Analysis

651 The structure of the *HvCMF3* gene was determined by analysis of its cDNA. Total RNA

652 was extracted from leaf material of a 3-day-old barley seedling (cv. Barke) using the

653 Trizol reagent (Thermo Scientific, Wilmington, USA) following the manufacturer's

654 instructions. Concentration of the RNA is measured by help of a NanoDrop 1000

655 spectrophotometer (Thermo Scientific, Wilmington, USA) and further diluted to 1 µg/µL

656 for downstream application. The prepared RNA was first treated with RNase-free

657 DNase I (Fermentas, St. Leon-Rot, Germany) to remove potential DNA contamination;

658 then used for cDNA synthesis applying the SuperScript™ III First-Strand Synthesis

659 System Kit (Thermo Scientific, Wilmington, USA) following the manufacturer's

660 instructions. Next, RT-PCR was performed using primers that cover the *HvCMF3* coding

661 regions (Supplemental Table 1) as previously described (Li et al., 2019). RT-PCR

662 products were purified using the NucleoFast® 96 PCR Kit (Macherey-Nagel, Düren,

663 Germany) and Sanger sequenced on an ABI 3730 XL platform (Life Technologies

664 GmbH, Darmstadt, Germany). The *HvCMF3* exon-intron-structure was revealed by

665 alignment of the coding sequence to the corresponding genomic region.

CAPS Assay

667 One CAPS (Cleaved Amplified Polymorphic Sequences) marker was developed for

668 genotyping the two *HvCMF3* pre-stop TILLING mutants, respectively. Briefly, PCR

669 reactions were performed as described earlier (Li et al., 2019) with minor changes, i.e.,

670 the annealing temperature for the touch-down profile was 62°C to 57°C instead of 65°C

671 to 60°C. The SNP carrying by the PCR amplicon was converted into a CAPS marker by

672 help of the SNP2CAPS software (Thiel et al., 2004) for the selection of the proper

673 restriction enzyme (Supplemental Table 3). Differentiation of the genotypes was

674 achieved by the distinct digestion patterns resolved on 1.5% (w/v) agarose gels

675 (Invitrogen GmbH, Darmstadt, Germany).

Identification of Conserved Sequence Regions

24

677   For conservation analysis, all identified 131 *HvCMF3*-homologous sequences were
678   aligned using MEGA6 with the MUSCEL method (Tamura et al., 2013). During the
679   subsequent sequence validation process, the aligned sequences were manually edited
680   by removing wrongly predicted sequence regions and filling gaps. Conservation of the
681   resulting 675 aligned positions was displayed by the online tool WebLogo (Crooks et al.,
682   2004).

683   For conservation analysis of the novel functional region identified in this study, the
684   conserved region 2 was extracted from the above aligned file and then re-aligned in
685   MEGA6 with the MUSCEL method (Tamura et al., 2013). Next, sequences with unequal
686   length compared to the prominent motif (17 AA in length) were eliminated. Finally, 116
687   sequences from 59 species with a consistent 17 AA length were obtained. Peptide
688   conservation was visualized using the online tool MEME (Bailey et al., 2009).

Ribosomal RNA Analysis

690   RNA isolation and determination of RNA concentration were performed as previously
691   described (Li et al., 2019). In short, an Agilent 4200 TapeStation System (Agilent, Santa
692   Clara, USA) was adopted for analysis of rRNA. Initially, the concentration of the RNA
693   was determined by help of a Qubit$^®$ 2.0 Fluorometer (Life Technologies GmbH,
694   Darmstadt, Germany) according to manufacturer's instructions. RNA samples were
695   further diluted within a quantitative range of 1 - 10 ng/μL. RNA quality and quantity was
696   then measured using an Agilent High Sensitivity RNA ScreenTape following the
697   manufacturer's manual (Agilent, Santa Clara, USA).

Chlorophyll Content Measurement

699   Leaf material was collected from primary leaves of 10-day-old seedlings. Samples were
700   weighted and then frozen in liquid nitrogen. After homogenization using Mixer Mill
701   MM400 (Retsch GmbH, Haan, Germany), 1.5 mL of *N,N*-Dimethylformamide (DMF)
702   was added to each sample, followed by mixing on an overhead shaker (Keison
703   Products, Chelmsford, England) for 30 min. Subsequently, the supernatant obtained
704   after centrifugation (14,000x g for 10 min, room temperature) was transferred to a new 2
705   mL Eppendorf tube. Chlorophyll content measurement and calculation were performed

706 according to (J.W.A Porra et al., 1989). In brief, cuvette-based measurement (cuvette

707 with 1 mm path length) was conducted by help of the Spectramax Plus

708 spectrophotometer (GENEO BioTechProducts GmbH, Germany). Chlorophyll content of

709 a and b was calculated by the following equation: chlorophyll $a = 13.43(A^{663.8} - A^{750}) -$

710 $3.47(A^{646.8} - A^{750})$; chlorophyll $b = 22.90(A^{646.8} - A^{750}) - 5.38(A^{663.8} - A^{750})$.

711 **High-throughput Automated, Imaging-based Phenotyping**

712 Phenotyping by RGB (Red Green Blue, i.e., visible light) and static fluorescence

713 imaging as described in (Junker et al., 2014) started at 5 DAS and was thereafter

714 performed daily until 14 DAS. Kinetic chlorophyll fluorescence measurements were

715 performed using the integrated FluorCam imaging fluorimeter (Photon Systems

716 Instruments, Brno, Czech Republic). Chlorophyll fluorescence kinetics was measured

717 following a protocol optimized for the automated high throughput imaging system

718 (Tschiersch et al., 2017). Measurement of PSII operating efficiency (Φ_{PSII}) and electron

719 transport rate (ETR) were performed with light adapted plants. For adaptation, plants

720 were incubated in the adaptation tunnel for 5 min followed by 1 min illumination after

721 moving into the chlorophyll fluorescence imaging (CFI) chamber with equal light

722 intensity of 300 μmol photons m^{-2} s^{-1}. Subsequently, a saturating flash with PAR

723 (photosynthetic active radiation) intensity 4100 μmol photons m^{-2} s^{-1} for a period of 800

724 ms was applied to induce maximal chlorophyll fluorescence (F_m'). The steady state

725 fluorescence emission (F') and F_m' were recorded by the FluorCam imaging module.

726 The formula $\Phi_{PSII} = (F_m'-F')/F_m'$ was used to calculate effective quantum yield of

727 photochemical energy conversion in PSII. The electron transport rate (ETR) was

728 calculated as ETR = ΦPSII x PAR x 0.5 x ABS where PAR equals 300 in this study, 0.5

729 is a factor that accounts for the fraction of excitation energy distributed to PSII, and the

730 factor ABS (Absorbance) represents the leaf absorbance as determined by the near-

731 infrared (NIR) and red light (RED) sources. It is calculated by the equation ABS = (NIR-

732 RED)/(NIR+RED). The PSII operating efficiency was measured at the time points 6, 7,

733 8, 9, 12, and 14 DAS.

734 Quenching parameters were determined during the night when plants were dark-

735 adapted in the growth chamber for at least 2 hours. The minimal chlorophyll

26

736   fluorescence intensity (F_0) was measured after moving into the CFI chamber and the

737   maximal chlorophyll fluorescence intensity (F_m) was induced by application of a

738   saturating flash (4100 µmol photons m^{-2} s^{-1}) for 800 ms. After 10 s in darkness, plants

739   were illuminated with actinic light (300 µmol photons m^{-2} s^{-1}) for 4 min. During the

740   quenching procedure, a saturating flash was applied for 9 s after application of the

741   actinic light and repeated 6 times with an interval of 46 s. The values of maximal

742   chlorophyll fluorescence intensity F_m' and steady state fluorescence emission F' were

743   collected from the last saturating flash when the plants were light-adapted. Non-

744   photochemical quenching (NPQ) was calculated using the equation NPQ = (F_m/F_m')-1;

745   and photochemical quenching (qP) using the equation qP = (F_m'-F')/(F_m'-F_0')

746   (Supplemental Dataset 3). The distance between the FluorCam panels and plants was

747   set to 27 cm. The quenching experiment was performed at 6, 8, 9, 12, and 14 DAS.

748   From daily RGB and static fluorescence imaging, amongst others the traits 'projected

749   leaf area' and yellow to green pixel ratio were extracted after automated image pre-

750   processing and segmentation using the Integrated Analysis Platform (Klukas et al.,

751   2014). Both parameters were measured based on images acquired from the side view.

752   These traits are a proxy for plant growth dynamics during the phenotyping experiment

753   and the dynamics of plant coloration and the *xantha*-to-green phenotype during early

754   seedling development, respectively. To comply with the FAIR principles of data

755   management, the phenotyping procedures and dataset have been described using

756   standardized metadata formats (Rocca-Serra et al., 2010) following the

757   recommendations of the Minimum Information About a Plant Phenotyping Experiment

758   version 1.1 (MIAPPE v1.1) recommendations (Cwiek-Kupczynska et al., 2016) and the

759   entire dataset comprising raw and result image data as well as derived phenotypic trait

760   tables and metadata descriptions was uploaded to the Plant Genomics and Phenomics

761   repository (Arend et al., 2016) using the e!DAL data publication pipeline (Arend et al.,

762   2014). The exact value for all the measured traits at different time points is summarized

763   in Supplemental Dataset 3, and *p* values of the *Student's t-test* is summarized in

764   Supplemental Dataset 4.

765   **Chloroplast Ultrastructural Analysis**

766 Primary leaves of two developmental stages (3 and 10 days after germination) were
767 collected from wild type Barke, mutant 4383-1 and mutant 13082-1. For comparative
768 ultrastructural analysis, leaf cuttings of a size of 1x2 mm from corresponding regions
769 (Supplemental Figure 6) of three biological replicates were used for combined
770 conventional and microwave assisted chemical fixation, substitution and resin
771 embedding as defined in the given protocol (Supplemental Table 8). Sectioning and
772 transmission electron microscopy analysis was performed as described (Daghma et al.,
773 2011).

774 **Subcellular Localization**

775 Two constructs, HvCMF3:GFP and cTP_95AA_HvCMF3:GFP, were used to investigate
776 the subcellular localization of HvCMF3. For HvCMF3:GFP, the coding sequence of cv.
777 'Barke' was amplified using cDNA as a template employing the manually designed
778 primer pair HvCMF3_SC_F/HvCMF3_SC_R with *Spe*I and *Hind*III restriction sites
779 introduced at the 5' and 3' end, respectively. Similarly, primer pair
780 HvCMF3_cTP_95AA_F/HvCMF3_cTP_95AA_R with restriction sites as mentioned
781 above was used to amplify the HvCMF3 cTP predicted by the online tool PredSL
782 (Petsalaki et al., 2006) (Supplemental Tables 1 and 7). The derived PCR fragments
783 were separately inserted into vector pSB179 (Li et al., 2019). The resulting vectors
784 HvCMF3:GFP and cTP_95AA_HvCMF3:GFP were investigated for transient expression
785 in barley epidermal cells via biolistic assay by using the PDS-1000/He Hepta^{TM} device
786 (Bio-Rad, Munich, Germany). A plastid marker pt-rk CD3-999 containing the *mCherry*
787 gene driven by the doubled enhanced *CaMV 35S* promoter was adopted for particle co-
788 bombardment with the *HvCMF3* constructs (Plastid marker TAIR link:
789 https://www.arabidopsis.org/servlets/TairObject?type=stock&id=3001623338). Four to
790 six primary leaves were harvested from 7-day-old seedlings and placed on 1% Agar
791 supplemented with 20 μg/mL benzimidazol and 10 μg/mL chloramphenicol. Gold
792 suspension was prepared by suspending 30 mg gold particles (diameter = 1.0 μm, Bio-
793 Rad, Munich, Germany) in 1 mL 100% ethanol. For each shooting, 50 μL of gold
794 suspension was taken and washed three times with 100 μL ddH$_2$O followed by
795 suspension in 25 μL ddH$_2$O. Then, gold particles were coated with 5 μL of plasmids (2.5

796 µL each of HvCMF3 construct and plastid marker; both with a concentration of 1 µg/µL)

797 in the presence of 25 µL 25 mM $CaCl_2$ and 10 µL 0.1 M spermidine under vortexing for

798 2 minutes. After centrifugation, the plasmid-gold-pellet was washed twice with 100%

799 ethanol and suspended in 60 µL 100% ethanol. A total of 5 µL of plasmids-coated gold

800 suspension was loaded onto each of seven macro-carriers pre-washed with 100%

801 ethanol and dried under a fume hood. Plasmids pSB179 and pt-rk CD3-999 were

802 bombarded individually with 1100 psi acceleration pressure and 27 inch Hg vacuum

803 pressure in controls for distribution pattern of GFP and mCherry fluorescence,

804 respectively. The biolistically transformed leaves were incubated at room temperature

805 for 24 hours followed by detection of the fluorescent signals by help of a Zeiss LSM780

806 confocal laser scanning microscope (Carl Zeiss, Jena, Germany). Green fluorescence

807 of GFP was visualized by using the 488 nm excitation laser line with a manually defined

808 490-530 bandpass; mCherry signals were detected by the 561 nm excitation laser in

809 combination with a 580-620 nm bandpass.

Crossing Experiments

811 Allelism tests between *Hvcmf3-1* and *Hvcmf3-2* were performed by crossing TILLING

812 mutant 4383-1 (maternal parent) with TILLING mutant 13082-1 (pollen donor). F_1

813 hybrids carrying both mutant alleles were phenotypically characterized during the first to

814 three leaf stages. Generation of *Hvcmf3/Hvcmf7* double mutants was achieved by

815 crossing TILLING mutant 4383-1 as pollen donor with heterozygous *albostrians*

816 TILLING mutant 6460-1 and the original *albostrians* mutant M4205, respectively. F_1

817 plants heterozygous for both *HvCMF3* and *HvCMF7* loci were kept and *Hvcmf3/Hvcmf7*

818 double mutants were further selected in F_2 generation.

SUPPLEMENTAL DATA

820 **Supplemental Figure 1:** Phenotype of TILLING mutant *Hvcmf3-1* during development.

821 **Supplemental Figure 2:** Summary of Cas9-induced mutations.

822 **Supplemental Figure 3:** Identification of novel functional region of HvCMF3.

823 **Supplemental Figure 4:** *HvCMF3* cDNA analysis of T_1 homozygous mutants of family

824 BG677E9B.

29

825 **Supplemental Figure 5:** Phenotypes of selected *Hvcmf3* mutants and respective wild-
826 type plants.

827 **Supplemental Figure 6:** Sample collection for ultrastructural analysis.

828 **Supplemental Figure 7**: Chloroplast ultrastructural analysis for *Hvcmf3* mutants and
829 wild-type plants.

830 **Supplemental Figure 8:** Quantification of thylakoid numbers.

831 **Supplemental Figure 9.** Phenotype of double mutant *Hvcmf3/Hvcmf7*.

832 **Supplemental Table 1.** Primers used in this study.

833 **Supplemental Table 2.** Summary of identified TILLING mutations of *HvCMF3*.

834 **Supplemental Table 3.** Markers used for analysis of *Hvcmf3* pre-stop TILLING
835 mutants.

836 **Supplemental Table 4.** PCR screening of T_0 plants for presence and integrity of T-
837 DNA.

838 **Supplemental Table 5.** Genotyping of T_0 regenerants.

839 **Supplemental Table 6.** List of genotypes used for automated phenotyping.

840 **Supplemental Table 7.** *In silico* prediction of subcellular localization of HvCMF3.

841 **Supplemental Table 8.** Sample preparation for transmission electron microscopy.

842 **Supplemental Dataset 1.** Orthologs of *HvASL* and *HvAST* in monocots and dicots.

843 **Supplemental Dataset 2.** *In silico* cTP prediction of *HvAST*/*HvASL* homologous genes.

844 **Supplemental Dataset 3:** Summary of the photosynthetic and developmental related
845 traits measured using the automated phenotyping platform.

846 **Supplemental Dataset 4:** *Student's t-test* for the phenotyping experiment.

847 **DATA AVAILABILITY:**

848 The complete phenomics dataset (images, trait values and metadata) has been
849 deposited in e!DAL - The Plant Genomics & Phenomics Research Data Repository. Link
850 to the data: https://doi.ipk-gatersleben.de/DOI/a65bca88-dced-493a-bb70-
851 9952e8864672/325b7404-4ccc-40ea-a301-a9f7e4c48219/2/1847940088.

852 **ACKNOWLEDGEMENT**

AUTHOR CONTRIBUTIONS

M.L., N.S., T.B. and J.K. conceived the study. M.L., G.H., M.M., A.J. and H.T performed experiments. A.J and D.A contributed phenotyping data submission. M.L. analyzed data. M.L., N.S and T.B wrote the paper.

REFERENCES

Albrecht, V., Ingenfeld, A., and Apel, K. (2006). Characterization of the snowy cotyledon 1 mutant of Arabidopsis thaliana: the impact of chloroplast elongation factor G on chloroplast development and plant vitality. Plant Mol. Biol. **60,** 507-518.

Arend, D., Junker, A., Scholz, U., Schuler, D., Wylie, J., and Lange, M. (2016). PGP repository: a plant phenomics and genomics data publication infrastructure. Database (Oxford) **2016**.

Arend, D., Lange, M., Chen, J., Colmsee, C., Flemming, S., Hecht, D., and Scholz, U. (2014). e!DAL--a framework to store, share and publish research data. BMC Bioinformatics **15,** 214.

Armbruster, U., Labs, M., Pribil, M., Viola, S., Xu, W., Scharfenberg, M., Hertle, A.P., Rojahn, U., Jensen, P.E., Rappaport, F., Joliot, P., Dormann, P., Wanner, G., and Leister, D. (2013). Arabidopsis CURVATURE THYLAKOID1 proteins modify thylakoid architecture by inducing membrane curvature. Plant Cell **25,** 2661-2678.

Aryamanesh, N., Ruwe, H., Sanglard, L.V., Eshraghi, L., Bussell, J.D., Howell, K.A., Small, I., and des Francs-Small, C.C. (2017). The Pentatricopeptide Repeat Protein EMB2654 Is Essential for Trans-Splicing of a Chloroplast Small Ribosomal Subunit Transcript. Plant Physiol. **173,** 1164-1176.

Axelsson, E., Lundqvist, J., Sawicki, A., Nilsson, S., Schroder, I., Al-Karadaghi, S., Willows, R.D., and Hansson, M. (2006). Recessiveness and dominance in barley mutants deficient in Mg-chelatase subunit D, an AAA protein involved in chlorophyll biosynthesis. Plant Cell **18,** 3606-3616.

890 **Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li,**
891   **W.W., and Noble, W.S.** (2009). MEME SUITE: tools for motif discovery and
892   searching. Nucleic Acids Res. **37,** W202-208.
893 **Belgio, E., Ungerer, P., and Ruban, A.V.** (2015). Light-harvesting superstructures of
894   green plant chloroplasts lacking photosystems. Plant Cell Environ. **38,** 2035-
895   2047.
896 **Borner, T.** (2017). The discovery of plastid-to-nucleus retrograde signaling-a personal
897   perspective. Protoplasma **254,** 1845-1855.
898 **Börner, T., Zhelyazkova, P., Legen, J., and Schmitz-Linneweber, C.** (2014).
899   Chloroplast gene expression - RNA synthesis and processing. In Plastid Biology,
900   S.M. Theg and F.A. Wollman, eds (Dordrecht, The Netherlands: Springer), pp. 3-
901   47.
902 **Bradbeer, J.W., Atkinson, Y.E., Börner, T., and Hagemann, R.** (1979). Cytoplasmic
903   synthesis of plastid polypeptides may be controlled by plastid-synthesised RNA.
904   Nature **279,** 816-817.
905 **Budhagatapalli, N., Schedel, S., Gurushidze, M., Pencs, S., Hiekel, S., Rutten, T.,**
906   **Kusch, S., Morbitzer, R., Lahaye, T., Panstruga, R., Kumlehn, J., and**
907   **Hensel, G.** (2016). A simple test for the cleavage activity of customized
908   endonucleases in plants. Plant Methods **12,** 18.
909 **Cockram, J., Thiel, T., Steuernagel, B., Stein, N., Taudien, S., Bailey, P.C., and**
910   **O'Sullivan, D.M.** (2012). Genome dynamics explain the evolution of flowering
911   time CCT domain gene families in the Poaceae. PLoS One **7,** e45307.
912 **Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E.** (2004). WebLogo: a
913   sequence logo generator. Genome Res. **14,** 1188-1190.
914 **Cwiek-Kupczynska, H., Altmann, T., Arend, D., Arnaud, E., Chen, D., Cornut, G.,**
915   **Fiorani, F., Frohmberg, W., Junker, A., Klukas, C., Lange, M., Mazurek, C.,**
916   **Nafissi, A., Neveu, P., van Oeveren, J., Pommier, C., Poorter, H., Rocca-**
917   **Serra, P., Sansone, S.A., Scholz, U., van Schriek, M., Seren, U., Usadel, B.,**
918   **Weise, S., Kersey, P., and Krajewski, P.** (2016). Measures for interoperability
919   of phenotypic data: minimum information requirements and formatting. Plant
920   Methods **12,** 44.
921 **Daghma, D.S., Kumlehn, J., and Melzer, M.** (2011). The use of cyanobacteria as filler
922   in nitrocellulose capillaries improves ultrastructural preservation of immature
923   barley pollen upon high pressure freezing. J. Microsc. **244,** 79-84.
924 **De Bodt, S., Maere, S., and Van de Peer, Y.** (2005). Genome duplication and the
925   origin of angiosperms. Trends Ecol. Evol. **20,** 591-597.
926 **de Souza, A., Wang, J.Z., and Dehesh, K.** (2017). Retrograde Signals: Integrators of
927   Interorganellar Communication and Orchestrators of Plant Development. Annu.
928   Rev. Plant Biol. **68,** 85-108.
929 **Dekker, J.P., and Boekema, E.J.** (2005). Supramolecular organization of thylakoid
930   membrane proteins in green plants. Biochim. Biophys. Acta **1706,** 12-39.
931 **Delannoy, E., Le Ret, M., Faivre-Nitschke, E., Estavillo, G.M., Bergdoll, M., Taylor,**
932   **N.L., Pogson, B.J., Small, I., Imbault, P., and Gualberto, J.M.** (2009).
933   Arabidopsis tRNA adenosine deaminase arginine edits the wobble nucleotide of
934   chloroplast tRNAArg(ACG) and is essential for efficient chloroplast translation.
935   Plant Cell **21,** 2058-2071.

Drew, B.T., Ruhfel, B.R., Smith, S.A., Moore, M.J., Briggs, B.G., Gitzendanner, M.A., Soltis, P.S., and Soltis, D.E. (2014). Another look at the root of the angiosperms reveals a familiar tale. Syst. Biol. **63,** 368-382.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. **8,** 978-984.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151,** 1531-1545.

Fristedt, R., Willig, A., Granath, P., Crevecoeur, M., Rochaix, J.D., and Vener, A.V. (2009). Phosphorylation of photosystem II controls functional macroscopic folding of photosynthetic membranes in Arabidopsis. Plant Cell **21,** 3950-3964.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. **40,** D1178-1186.

Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U., and Stein, N. (2009). TILLING in the two-rowed barley cultivar 'Barke' reveals preferred sites of functional diversity in the gene *HvHox1*. BMC Res. Notes **2,** 258.

Gould, S.B., Waller, R.F., and McFadden, G.I. (2008). Plastid evolution. Annu. Rev. Plant Biol. **59,** 491-517.

Hagemann, R., and Scholz, F. (1962). A case of gene induced mutations of the plasmotype in barley. Theor Appl Genet **32,** 50-59.

Hall, B.G. (2013). Building phylogenetic trees from molecular data with MEGA. Mol. Biol. Evol. **30,** 1229-1235.

Hausler, R.E., Geimer, S., Kunz, H.H., Schmitz, J., Dormann, P., Bell, K., Hetfeld, S., Guballa, A., and Flugge, U.I. (2009). Chlororespiration and grana hyperstacking: how an Arabidopsis double mutant can survive despite defects in starch biosynthesis and daily carbon export from chloroplasts. Plant Physiol. **149,** 515-533.

Henningsen, K.W., Boynton, J.E., and Wettstein, D.V. (1993). Mutants at xantha and albina loci in relation to chloroplast biogenesis in barley (Hordeum vulgare L.). (Copenhagen K, Denmark: Munksgaard Export and Subscription Service).

Hensel, G., Kastner, C., Oleszczuk, S., Riechen, J., and Kumlehn, J. (2009). Agrobacterium-mediated gene transfer to cereal crop plants: current protocols for barley, wheat, triticale, and maize. Int J Plant Genomics **2009,** 835608.

Hernandez-Verdeja, T., and Strand, A. (2018). Retrograde Signals Navigate the Path to Chloroplast Development. Plant Physiol. **176,** 967-976.

Hess, W.R., Prombona, A., Fieder, B., Subramanian, A.R., and Börner, T. (1993). Chloroplast *rps15* and the *rpoB/C1/C2* gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. EMBO J. **12,** 563-571.

J. Simpson, D., Vallon, O., and Vonwettstein, D. (1989). Freeze-fracture studies on barley plastid membranesVIII. In viridis-115, a mutant completely lacking Photosystem II, oxygen evolution enhancer 1 (OEE1) and the α-subunit of cytochrome b-559 accumulate in appressed thylakoids.

J.W.A Porra, R., Thompson, W.A., and Kriedemann, P.E. (1989). Determination of Accurate Extinction Coefficients and Simultaneous Equations for Assaying Chlorophyll a and b Extracted with Four Different Solvents: Verification of the Concentration of Chlorophyll Standards by Atomic Absorption Spectroscopy.

Jang, S., Marchal, V., Panigrahi, K.C., Wenkel, S., Soppe, W., Deng, X.W., Valverde, F., and Coupland, G. (2008). Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. EMBO J. **27,** 1277-1288.

Jia, H., Liggins, J.R., and Chow, W.S. (2012). Acclimation of leaves to low light produces large grana: the origin of the predominant attractive force at work. Philos. Trans. R. Soc. Lond. B Biol. Sci. **367,** 3494-3502.

Junker, A., Muraya, M.M., Weigelt-Fischer, K., Arana-Ceballos, F., Klukas, C., Melchinger, A.E., Meyer, R.C., Riewe, D., and Altmann, T. (2014). Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. Front Plant Sci **5,** 770.

Kleine, T., and Leister, D. (2016). Retrograde signaling: Organelles go networking. Biochim. Biophys. Acta **1857,** 1313-1325.

Klukas, C., Chen, D., and Pape, J.M. (2014). Integrated Analysis Platform: An Open-Source Information System for High-Throughput Plant Phenotyping. Plant Physiol. **165,** 506-518.

Kohler, D., Helm, S., Agne, B., and Baginsky, S. (2016). Importance of Translocon Subunit Tic56 for rRNA Processing and Chloroplast Ribosome Assembly. Plant Physiol. **172,** 2429-2444.

Lafon-Placette, C., Vallejo-Marin, M., Parisod, C., Abbott, R.J., and Kohler, C. (2016). Current plant speciation research: unravelling the processes and mechanisms behind the evolution of reproductive isolation barriers. New Phytol **209,** 29-33.

Lai, J., Ma, J., Swigonova, Z., Ramakrishna, W., Linton, E., Llaca, V., Tanyolac, B., Park, Y.J., Jeong, O.Y., Bennetzen, J.L., and Messing, J. (2004). Gene loss and movement in the maize genome. Genome Res. **14,** 1924-1931.

Lee, D.W., and Hwang, I. (2018). Evolution and Design Principles of the Diverse Chloroplast Transit Peptides. Mol. Cells **41,** 161-167.

Leister, D. (2003). Chloroplast research in the genomic age. Trends Genet. **19,** 47-56.

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. **44,** W242-245.

Li, M., Hensel, G., Mascher, M., Melzer, M., Budhagatapalli, N., Rutten, T., Himmelbach, A., Beier, S., Korzun, V., Kumlehn, J., Borner, T., and Stein, N. (2019). Leaf Variegation and Impaired Chloroplast Development Caused by a Truncated CCT Domain Gene in albostrians Barley. Plant Cell **31,** 1430-1445.

Liu, J., Zhou, W., Liu, G., Yang, C., Sun, Y., Wu, W., Cao, S., Wang, C., Hai, G., Wang, Z., Bock, R., Huang, J., and Cheng, Y. (2015). The conserved endoribonuclease YbeY is required for chloroplast ribosomal RNA processing in Arabidopsis. Plant Physiol. **168,** 205-221.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science **290,** 1151-1155.

1028 **Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F.,**
1029   **Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu,**
1030   **F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang,**
1031   **D., Zheng, C., Geer, L.Y., and Bryant, S.H.** (2017). CDD/SPARCLE: functional
1032   classification of proteins via subfamily domain architectures. Nucleic Acids Res.
1033   **45,** D200-D203.
1034 **Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S.O., Wicker, T.,**
1035   **Radchuk, V., Dockter, C., Hedley, P.E., Russell, J., Bayer, M., Ramsay, L.,**
1036   **Liu, H., Haberer, G., Zhang, X.Q., Zhang, Q., Barrero, R.A., Li, L., Taudien,**
1037   **S., Groth, M., Felder, M., Hastie, A., Simkova, H., Stankova, H., Vrana, J.,**
1038   **Chan, S., Munoz-Amatriain, M., Ounit, R., Wanamaker, S., Bolser, D.,**
1039   **Colmsee, C., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J.,**
1040   **Chailyan, A., Sampath, D., Heavens, D., Clissold, L., Cao, S., Chapman, B.,**
1041   **Dai, F., Han, Y., Li, H., Li, X., Lin, C., McCooke, J.K., Tan, C., Wang, P.,**
1042   **Wang, S., Yin, S., Zhou, G., Poland, J.A., Bellgard, M.I., Borisjuk, L.,**
1043   **Houben, A., Dolezel, J., Ayling, S., Lonardi, S., Kersey, P., Langridge, P.,**
1044   **Muehlbauer, G.J., Clark, M.D., Caccamo, M., Schulman, A.H., Mayer, K.F.X.,**
1045   **Platzer, M., Close, T.J., Scholz, U., Hansson, M., Zhang, G., Braumann, I.,**
1046   **Spannagl, M., Li, C., Waugh, R., and Stein, N.** (2017). A chromosome
1047   conformation capture ordered sequence of the barley genome. Nature **544,** 427-
1048   433.
1049 **Meier, D., and K. Lichtenthaler, H.** (1981). Ultrastructural development of chloroplasts
1050   in radish seedlings grown at high- and low-light conditions and in the presence of
1051   the herbicide bentazon.
1052 **Mostowska, A.** (1986). Thylakoid and grana formation during the development of pea
1053   chloroplasts, illuminated by white, red, and blue low intensity light. Protoplasma
1054   **134,** 88-94.
1055 **Muhlhausen, S., and Kollmar, M.** (2013). Whole genome duplication events in plant
1056   evolution reconstructed and predicted using myosin motor proteins. BMC Evol.
1057   Biol. **13,** 202.
1058 **Olsson, U., Sirijovski, N., and Hansson, M.** (2004). Characterization of eight barley
1059   xantha-f mutants deficient in magnesium chelatase. Plant Physiol. Biochem. **42,**
1060   557-564.
1061 **Paterson, A.H., Bowers, J.E., and Chapman, B.A.** (2004). Ancient polyploidization
1062   predating divergence of the cereals, and its consequences for comparative
1063   genomics. Proc. Natl. Acad. Sci. U. S. A. **101,** 9903-9908.
1064 **Petsalaki, E.I., Bagos, P.G., Litou, Z.I., and Hamodrakas, S.J.** (2006). PredSL: a tool
1065   for the N-terminal sequence-based prediction of protein subcellular localization.
1066   Genomics Proteomics Bioinformatics **4,** 48-55.
1067 **Pietrzykowska, M., Suorsa, M., Semchonok, D.A., Tikkanen, M., Boekema, E.J.,**
1068   **Aro, E.M., and Jansson, S.** (2014). The light-harvesting chlorophyll a/b binding
1069   proteins Lhcb1 and Lhcb2 play complementary roles during state transitions in
1070   Arabidopsis. Plant Cell **26,** 3646-3660.
1071 **Pogson, B.J., and Albrecht, V.** (2011). Genetic dissection of chloroplast biogenesis
1072   and development: an overview. Plant Physiol. **155,** 1545-1551.

1073 **Pogson, B.J., Ganguly, D., and Albrecht-Borth, V.** (2015). Insights into chloroplast
1074   biogenesis and development. Biochim. Biophys. Acta **1847,** 1017-1024.
1075 **Pribil, M., Pesaresi, P., Hertle, A., Barbato, R., and Leister, D.** (2010). Role of plastid
1076   protein phosphatase TAP38 in LHCII dephosphorylation and thylakoid electron
1077   flow. PLoS Biol. **8,** e1000288.
1078 **Puthiyaveetil, S., van Oort, B., and Kirchhoff, H.** (2017). Surface charge dynamics in
1079   photosynthetic membranes and the structural consequences. Nat Plants **3,**
1080   17020.
1081 **Rocca-Serra, P., Brandizi, M., Maguire, E., Sklyar, N., Taylor, C., Begley, K., Field,
1082   D., Harris, S., Hide, W., Hofmann, O., Neumann, S., Sterk, P., Tong, W., and
1083   Sansone, S.A.** (2010). ISA software suite: supporting standards-compliant
1084   experimental annotation and enabling curation at the community level.
1085   Bioinformatics **26,** 2354-2356.
1086 **Rogalski, M., Schottler, M.A., Thiele, W., Schulze, W.X., and Bock, R.** (2008). Rpl33,
1087   a nonessential plastid-encoded ribosomal protein in tobacco, is required under
1088   cold stress conditions. Plant Cell **20,** 2221-2237.
1089 **Samol, I., Shapiguzov, A., Ingelsson, B., Fucile, G., Crevecoeur, M., Vener, A.V.,
1090   Rochaix, J.D., and Goldschmidt-Clermont, M.** (2012). Identification of a
1091   photosystem II phosphatase involved in light acclimation in Arabidopsis. Plant
1092   Cell **24,** 2596-2609.
1093 **Soltis, D.E., Albert, V.A., Leebens-Mack, J., Bell, C.D., Paterson, A.H., Zheng, C.,
1094   Sankoff, D., Depamphilis, C.W., Wall, P.K., and Soltis, P.S.** (2009). Polyploidy
1095   and angiosperm diversification. Am. J. Bot. **96,** 336-348.
1096 **Sun, C.W., Huang, Y.C., and Chang, H.Y.** (2009). CIA2 coordinately up-regulates
1097   protein import and synthesis in leaf chloroplasts. Plant Physiol. **150,** 879-888.
1098 **Sun, C.W., Chen, L.J., Lin, L.C., and Li, H.M.** (2001). Leaf-specific upregulation of
1099   chloroplast translocon genes by a CCT motif-containing protein, CIA2. Plant Cell
1100   **13,** 2053-2061.
1101 **Takahashi, S., and Badger, M.R.** (2011). Photoprotection in plants: a new light on
1102   photosystem II damage. Trends Plant Sci **16,** 53-60.
1103 **Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S.** (2013). MEGA6:
1104   Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. **30,** 2725-
1105   2729.
1106 **Thiel, T., Kota, R., Grosse, I., Stein, N., and Graner, A.** (2004). SNP2CAPS: a SNP
1107   and INDEL analysis tool for CAPS marker development. Nucleic Acids Res. **32,**
1108   e5.
1109 **Thiel, T., Graner, A., Waugh, R., Grosse, I., Close, T.J., and Stein, N.** (2009).
1110   Evidence and evolutionary analysis of ancient whole-genome duplication in
1111   barley predating the divergence from rice. BMC Evol. Biol. **9,** 209.
1112 **Tiller, N., and Bock, R.** (2014). The translational apparatus of plastids and its role in
1113   plant development. Mol Plant **7,** 1105-1120.
1114 **Tschiersch, H., Junker, A., Meyer, R.C., and Altmann, T.** (2017). Establishment of
1115   integrated protocols for automated high throughput kinetic chlorophyll
1116   fluorescence analyses. Plant Methods **13,** 54.

Vamosi, J.C., Magallon, S., Mayrose, I., Otto, S.P., and Sauquet, H. (2018). Macroevolutionary Patterns of Flowering Plant Speciation and Extinction. Annu. Rev. Plant Biol. **69,** 685-706.

Van de Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. Nat Rev Genet **18,** 411-424.

Walter, M., Piepenburg, K., Schottler, M.A., Petersen, K., Kahlau, S., Tiller, N., Drechsel, O., Weingartner, M., Kudla, J., and Bock, R. (2010). Knockout of the plastid RNase E leads to defective RNA processing and chloroplast ribosome deficiency. Plant J. **64,** 851-863.

Wenkel, S., Turck, F., Singer, K., Gissot, L., Le Gourrierec, J., Samach, A., and Coupland, G. (2006). CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis. Plant Cell **18,** 2971-2984.

Willi, J., Kupfer, P., Evequoz, D., Fernandez, G., Katz, A., Leumann, C., and Polacek, N. (2018). Oxidative stress damages rRNA inside the ribosome and differentially affects the catalytic center. Nucleic Acids Res. **46,** 1945-1957.

Wood, W.H., Barnett, S.F.H., Flannery, S., Hunter, C.N., and Johnson, M.P. (2019). Dynamic thylakoid stacking is regulated by LHCII phosphorylation but not its interaction with photosystem I. Plant Physiol.

Wood, W.H.J., MacGregor-Chatwin, C., Barnett, S.F.H., Mayneord, G.E., Huang, X., Hobbs, J.K., Hunter, C.N., and Johnson, M.P. (2018). Dynamic thylakoid stacking regulates the balance between linear and cyclic photosynthetic electron transfer. Nat Plants **4,** 116-127.

Zhang, Y.Y., Hao, Y.Y., Wang, Y.H., Wang, C.M., Wang, Y.L., Long, W.H., Wang, D., Liu, X., Jiang, L., and Wan, J.M. (2017). Lethal albinic seedling, encoding a threonyl-tRNA synthetase, is involved in development of plastid protein synthesis system in rice. Plant Cell Rep **36,** 1053-1064.

Zhelyazkova, P., Sharma, C.M., Forstner, K.U., Liere, K., Vogel, J., and Borner, T. (2012). The Primary Transcriptome of Barley Chloroplasts: Numerous Noncoding RNAs and the Dominating Role of the Plastid-Encoded RNA Polymerase. Plant Cell **24,** 123-136.

FIGURE LEGENDS

Figure 1. Phylogenetic analysis of *HvASL* and *HvAST* homologous genes.

The phylogenetic tree shows *Amborella trichopoda* as a sister group to all other angiosperm species. The two main branches separate the monocots and dicots, indicated by green and blue colour, respectively. Evolutionary analysis reveals a single pair of paralogs in diploids, and two pairs of paralogs in tetraploids. The paralogs of each species divide into two branches; each branch contains the corresponding orthologs for species in families *Poaceae*, *Salicaceae*, *Fabaceae*, *Crassulaceae* and

1157 *Brassicaceae*. Maintenance of these paralog pairs indicates that *HvASL* probably

1158 retained an important function in barley. The numbers above/below the branches

1159 represent bootstrap values which indicate reliability of the cluster descending from that

1160 node. The red colour node indicates where splitting of the orthologous groups occured.

1161 Positions of HvASL, HvAST, AtCIL and AtCIA2 are highlighted in bold. Family

1162 information is indicated outside the coloured stripes.

1163 **Figure 2. Functional validation of *HvCMF3* by TILLING and allelism test.**

1164 (A) TILLING screening strategy. Screening of coding regions of *HvCMF3* by three

1165 primer pairs. Red arrows indicate the relative position of the stop codons of TILLING

1166 families 4383-1 and 13082-1.

1167 (B) Summary of the identified mutations. TILLING screening revealed a total of 54 M_3

1168 mutant families with lesions in the *HvCMF3* gene, including 28 non-synonymous, 24

1169 synonymous, and 2 pre-stop mutations. Transition mutation (G to A) at position 861

1170 results in an immature stop codon in family 4383-1. Pre-stop family 13082-1 carries a

1171 transversion mutation (A to T) at position 1135. The adenine of the *HvCMF3* start codon

1172 refers as position 1.

1173 (C) Phenotype of *Hvcmf3* mutants compared with wild type cv. 'Barke' at developmental

1174 stage 3 days after germination. Leaves of *Hvcmf3-1* mutant exhibit a *xantha* phenotype.

1175 Compared to *Hvcmf3-1*, the chlorophyll-deficient phenotype of *Hvcmf3-2* mutant is less

1176 severe. The F_1 hybrid, *Hvcmf3-1/Hvcmf3-2* derived from crossing 4383-1 x 13082-1,

1177 exhibits a pale green phenotype.

1178 **Figure 3. Site-directed mutagenesis of *HvCMF3* gene by RNA-guided Cas9**

1179 **endonuclease**

1180 (A) Selection of Cas9/gRNA target sites. The two target motifs (Target Motif 1 and 2) in

1181 the anti-sense strand are underlined; the respective protospacer adjacent motif is

1182 highlighted in blue. The nucleotide in green colour indicates the position of the pre-stop

1183 mutation in the *Hvcmf3-1* mutant.

1184 (B) Alignment of *HvCMF3* sequences of wild-type and T_0 plantlets carrying mutations at

1185 target motif 1.

1186 (C) Alignment of *HvCMF3* sequences of wild-type and T_0 plantlets carrying mutations at

1187 target motif 2. The chimeric and/or heterozygous T_0 regenerant BG677E5A carries

1188 multiple mutations with each mutation shown in one single row.

1189 (D) Alignment of *HvCMF3* sequences of wild-type and T_1 homozygous mutant plants.

1190 Across panels, deletions are represented by red hyphens and insertions by red letters.

1191 The specific mutation of each plant is shown on the right of each sequence;

1192 presence/absence of wild-type allele is indicated by symbols +/-, respectively.

1193 (E) Phenotype of Cas9-induced homozygous *Hvcmf3* mutants at developmental stages

1194 3 and 10 days after germination.

1195 **Figure 4. Novel conserved functional region of HvCMF3.**

1196 (A) Alignment of 131 HvCMF3 homologous protein sequences from 66 species revealed

1197 five conserved regions which include the N-terminal chloroplast transit peptides domain,

1198 the C-terminal CCT domain and three novel conserved regions. In addition, the

1199 homologous genes contain multiple conserved peptides indicated by red dots below the

1200 position IDs. The conserved regions are marked with underline and highlighted with

1201 coloured circles. The region given in parentheses indicates the corresponding position

1202 of the conserved region in reference to HvCMF3. Alignment was manually edited by

1203 removing wrongly predicted sequence regions and by filling gaps. There were a total of

1204 675 positons left. The online tool Weblogo was adopted for graphic generation.

1205 (B) Conservation analysis of the functional region of HvCMF3 identified in this study..

1206 For each position, the overall height of the stack indicates the sequence conservation at

1207 that position, while the height of symbols within the stack indicates the relative

1208 frequency of each amino acid at that position.

1209 **Figure 5. rRNA analysis and chlorophyll content measurement**

1210 (A) Separation of cytosolic and plastid rRNAs using the Agilent high sensitivity RNA

1211 ScreenTape assay.

1212 (B) Analysis of rRNA from wild type, *Hvcmf3* mutants and *xantha* mutants using an

1213 Agilent Tapestation 4200.

1214 (C) - (D) Determination of plastid-to-cytosolic rRNA ratios. (C) 23S/25S; (D) 16S/25S.

1215   (E) Ratio of the plastid 23S rRNA to the plastid 16S rRNA.

1216   (F) – (H) Analysis of chlorophyll contents and ratio between chlorophyll *a* and

1217   chlorophyll *b.* Results are presented as means ± SE. *t-test* significant level: * $p < 0.05$,

1218   ** $p < 0.01$, *** $p < 0.001$, n.s: not significant. Three plants per genotype were analyzed.

Figure 6. Determination of photosynthetic parameters and growth dynamics of
1220   *Hvcmf3* **mutant and wild-type control plants.**

1221   (A) to (E) Measurement of photosynthetic parameters during early developmental

1222   stages. Results are presented as means ± SE. *Student's t-test* significant levels, * p

1223   <0.05, ** p <0.01. ETR, electron transport rate; Φ_{PSII}, photosystem II operating efficiency;

1224   qP, fraction of PSII centers that are 'open' based on the puddle model; F_v/F_m, maximum

1225   quantum yield of PSII photochemistry measured in the dark-adapted state; NPQ, non-

1226   photochemical quenching.

1227   (F) Plant growth dynamics. Left panel is yellow/green pixel ratio, and right panel is

1228   projected leaf area.

Figure 7. Quantification of chloroplast architecture components.

1230   (A) Diagram for demonstrating the chloroplast length, width, and surface area.

1231   (B) Illustration demonstrating the counting of thylakoid.

1232   (C) to (H) Comparison of chloroplast morphology and grana architecture between wild

1233   type and *Hvcmf3* mutants at developmental stages 3 days after germination.

1234   Chloroplast length (C), chloroplast width (D), chloroplast surface area (E), grana

1235   number (F), grana height (G), and thylakoid distance (H). Results are presented as

1236   means ± SE. *t-test* significant level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, n.s: not

1237   significant. Number of chloroplast analyzed n ≥ 24.

Figure 8. Subcellular localization of HvCMF3.

1239   (A) Schematic diagram of the constructs prepared for transient expression. *pZmUbi*,

1240   maize *UBIQUITIN1* promoter. *pCaMVd35S*, Cauliflower Mosaic Virus doubled-

1241   enhanced *35S* promoter. GFP, green fluorescent protein. mCherry, mCherry fluorescent

1242   protein; PLS, plastid localization signal, i.e. the chloroplast transit peptide (N-terminal 79
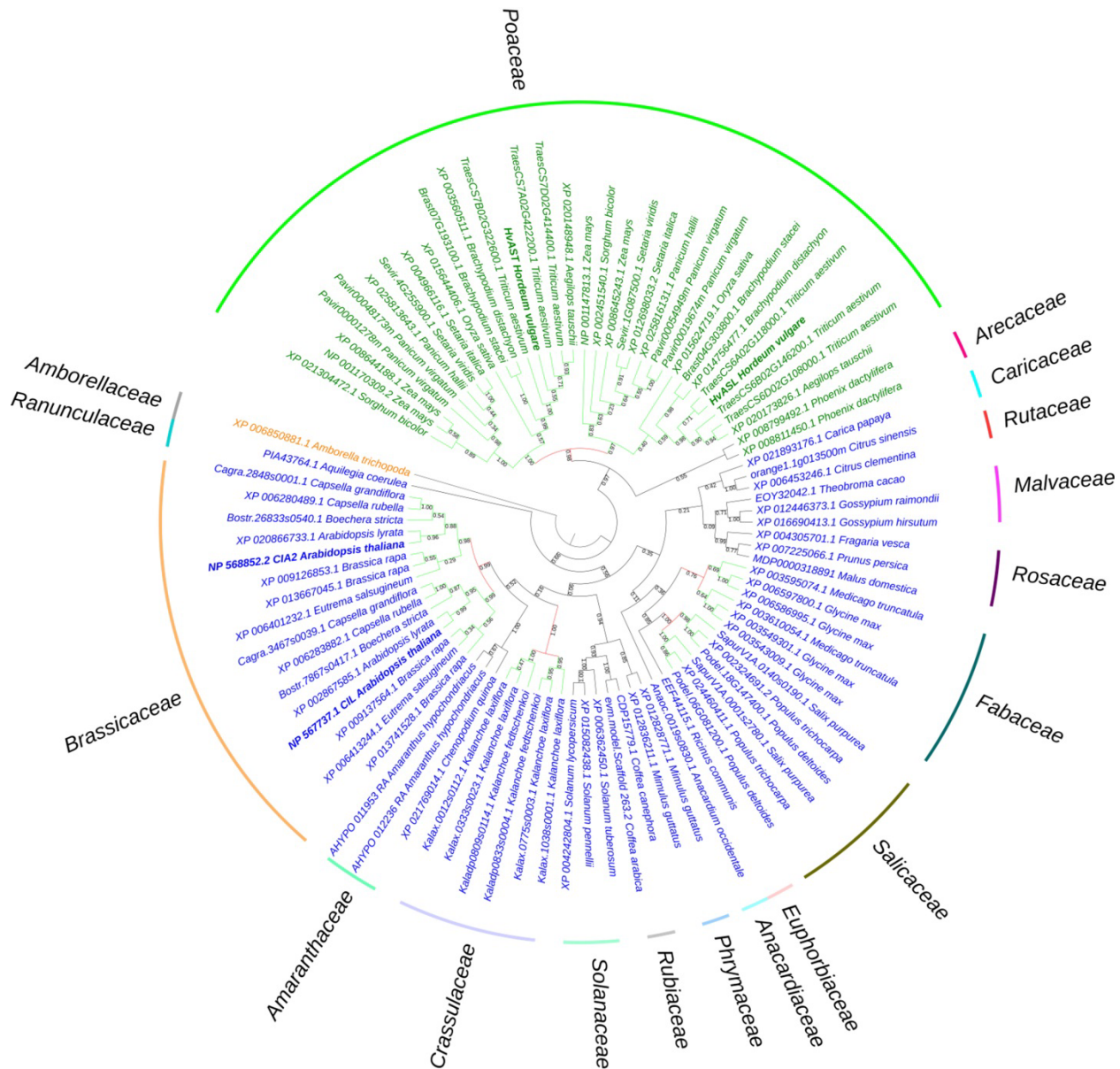
40

1243 amino acids) of the small subunit of tobacco RUBISCO. HvCMF3, coding sequence of

1244 wild-type *HvCMF3* gene. cTP_95AA_HvCMF3, N-terminal chloroplast transit peptide of

1245 HvCMF3 with a length of 95 amino acids as predicted by online tool PredSL. tNOS,

1246 *Agrobacterium nopaline synthase* terminator. The schematic drawing is not in proportion

1247 with gene length.

1248 (B) Localization of GFP control with *GFP* being driven by the maize *UBIQUITIN1*

1249 promoter.

1250 (C) Localization of the plastid marker.

1251 (D) Localization of HvCMF3:GFP. The GFP fluorescence signal is targeted both to

1252 plastid and nucleus compartments.

1253 (E) Localization of cTP_95AA_HvCMF3:GFP. The yellow arrows in the merged panels

1254 indicate the nucleus. The first leaf of 10-day-old barley seedlings was used for particle

1255 bombardment. The fluorescence was checked 24 hours after bombardment. Scale bar

1256 for all images is 20 µm.

41

1257

Figure 1. Phylogenetic analysis of *HvASL* and *HvAST* homologous genes.

The phylogenetic tree shows *Amborella trichopoda* as a sister group to all other angiosperm species. The two main branches separate the monocots and dicots, indicated by green and blue colour, respectively. Evolutionary analysis reveals a single pair of paralogs in diploids, and two pairs of paralogs in tetraploids. The paralogs of each species divide into two branches; each branch contains the corresponding orthologs for species in families *Poaceae*, *Salicaceae*, *Fabaceae*, *Crassulaceae* and *Brassicaceae*. Maintenance of these paralog pairs indicates that *HvASL* probably retained an important function in barley. The numbers above/below the branches represent bootstrap values which indicate reliability of the cluster descending from that node. The red colour node indicates where splitting of the orthologous groups occured. Positions of HvASL, HvAST, AtCIL and AtCIA2 are highlighted in bold. Family information is indicated outside the coloured stripes.
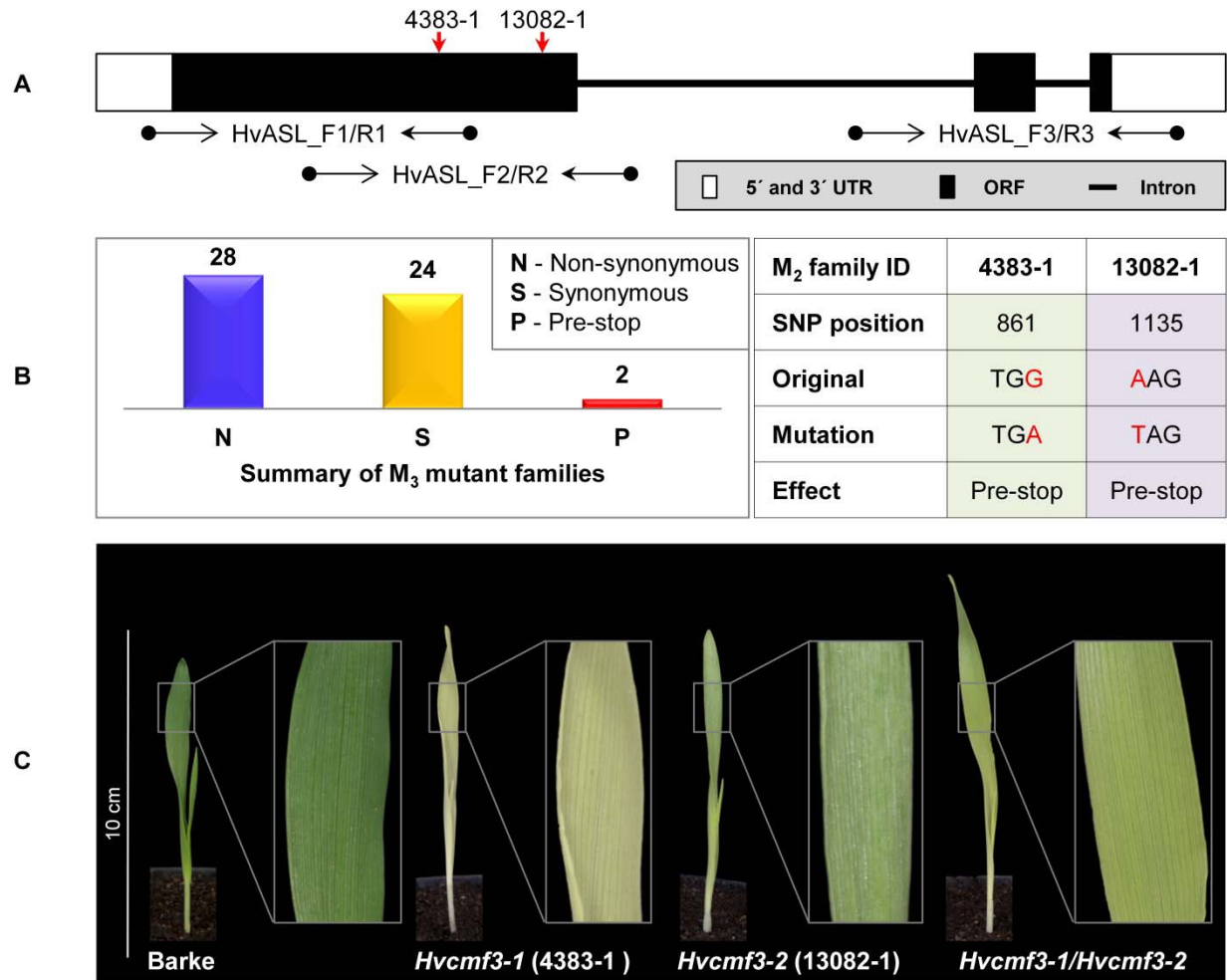
1269

Figure 2. Functional validation of *HvCMF3* by TILLING and allelism test.

(A) TILLING screening strategy. Screening of coding regions of *HvCMF3* by three primer pairs. Red arrows indicate the relative position of the stop codons of TILLING families 4383-1 and 13082-1.
(B) Summary of the identified mutations. TILLING screening revealed a total of 54 M_3 mutant families with lesions in the *HvCMF3* gene, including 28 non-synonymous, 24 synonymous, and 2 pre-stop mutations. Transition mutation (G to A) at position 861 results in an immature stop codon in family 4383-1. Pre-stop family 13082-1 carries a transversion mutation (A to T) at position 1135. The adenine of the *HvCMF3* start codon refers as position 1.
(C) Phenotype of *Hvcmf3* mutants compared with wild type cv. 'Barke' at developmental stage 3 days after germination. Leaves of *Hvcmf3-1* mutant exhibit a *xantha* phenotype. Compared to *Hvcmf3-1*, the chlorophyll-deficient phenotype of *Hvcmf3-2* mutant is less severe. The F_1 hybrid, *Hvcmf3-1/Hvcmf3-2* derived from crossing 4383-1 x 13082-1, exhibits a pale green phenotype.
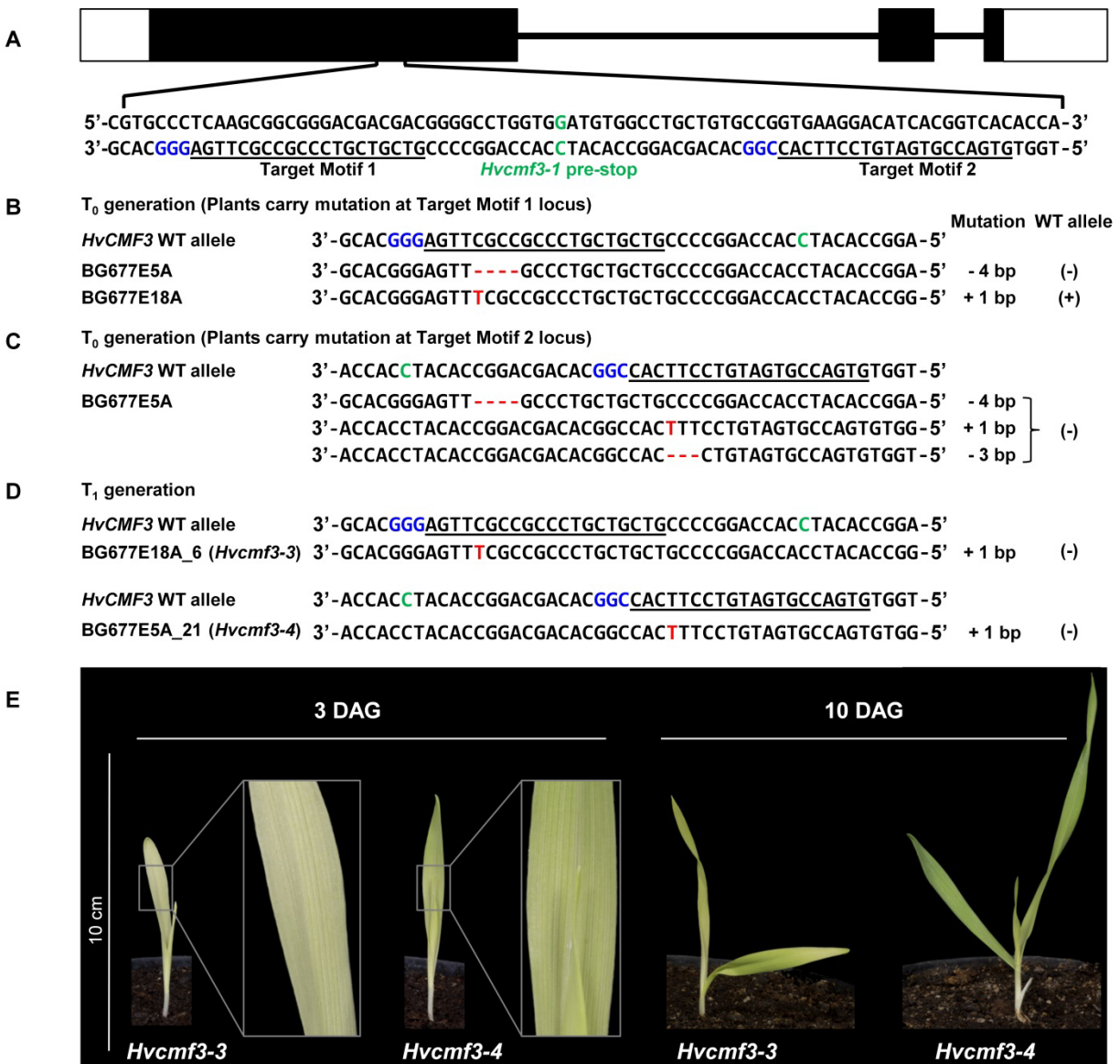
Figure 3. Site-directed mutagenesis of *HvCMF3* gene by RNA-guided Cas9 endonuclease

(A) Selection of Cas9/gRNA target sites. The two target motifs (Target Motif 1 and 2) in the anti-sense strand are underlined; the respective protospacer adjacent motif is highlighted in blue. The nucleotide in green colour indicates the position of the pre-stop mutation in the *Hvcmf3-1* mutant.

(B) Alignment of *HvCMF3* sequences of wild-type and T_0 plantlets carrying mutations at target motif 1.

(C) Alignment of *HvCMF3* sequences of wild-type and T_0 plantlets carrying mutations at target motif 2. The chimeric and/or heterozygous T_0 regenerant BG677E5A carries multiple mutations with each mutation shown in one single row.

(D) Alignment of *HvCMF3* sequences of wild-type and T_1 homozygous mutant plants. Across panels, deletions are represented by red hyphens and insertions by red letters. The specific mutation of each plant is shown on the right of each sequence; presence/absence of wild-type allele is indicated by symbols +/-, respectively.

(E) Phenotype of Cas9-induced homozygous *Hvcmf3* mutants at developmental stages 3 and 10 days after germination.

44

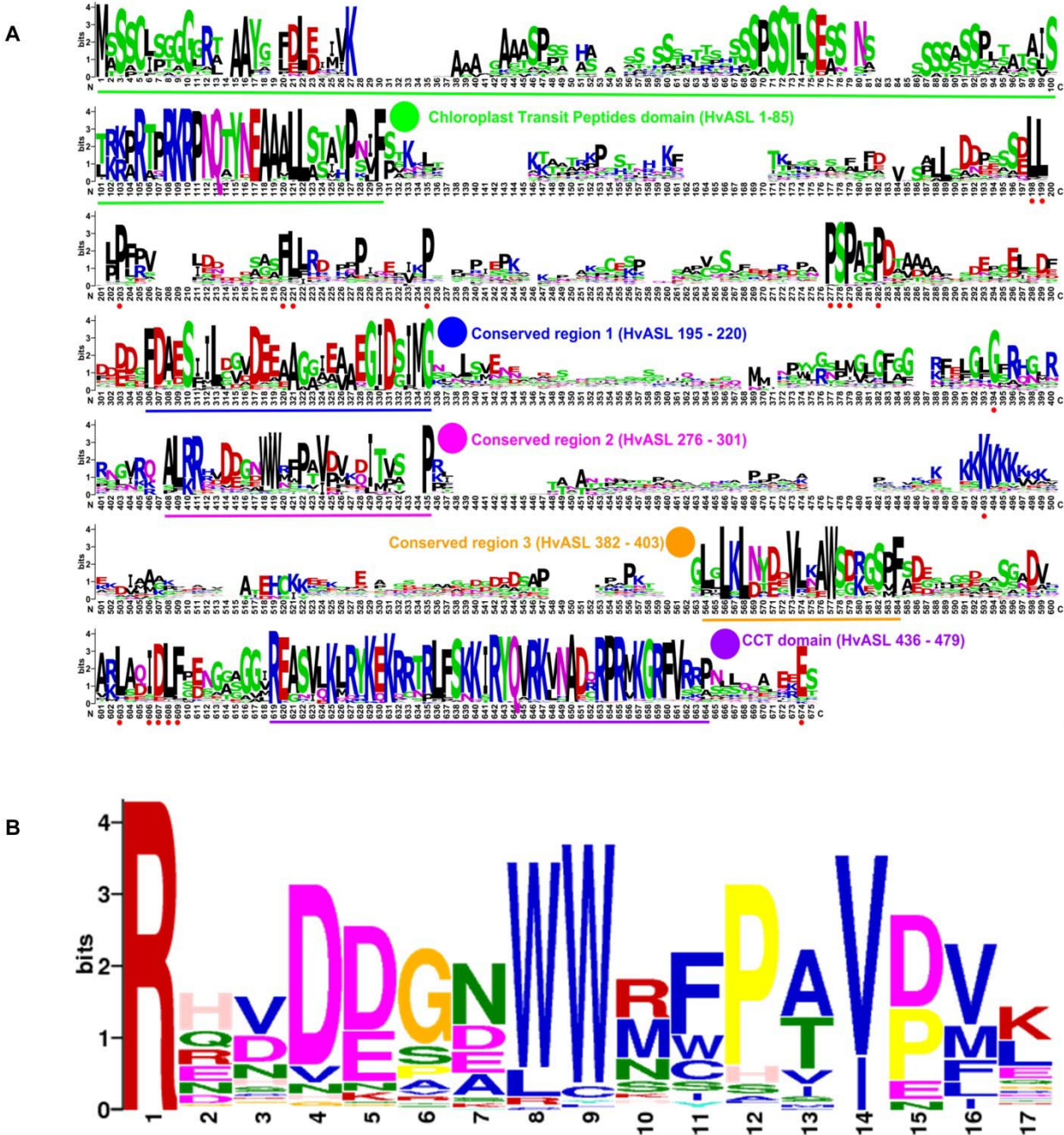**Figure 4. Novel conserved functional region of HvCMF3.**

(A) Alignment of 131 HvCMF3 homologous protein sequences from 66 species revealed five conserved regions which include the N-terminal chloroplast transit peptides domain, the C-terminal CCT domain and three novel conserved regions. In addition, the homologous genes contain multiple conserved peptides indicated by red dots below the position IDs. The conserved regions are marked with underline and highlighted with coloured circles. The region given in parentheses indicates the corresponding position of the conserved region in reference to HvCMF3. Alignment was manually edited by removing wrongly predicted sequence regions and by filling gaps. There were a total of 675 positons left. The online tool Weblogo was adopted for graphic generation.

(B) Conservation analysis of the functional region of HvCMF3 identified in this study.. For each position, the overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each amino acid at that position.

1311

Figure 5. rRNA analysis and chlorophyll content measurement

(A) Separation of cytosolic and plastid rRNAs using the Agilent high sensitivity RNA ScreenTape assay.

1314   (B) Analysis of rRNA from wild type, *Hvcmf3* mutants and *xantha* mutants using an Agilent Tapestation
1315   4200.
1316   (C) - (D) Determination of plastid-to-cytosolic rRNA ratios. (C) 23S/25S; (D) 16S/25S.
1317   (E) Ratio of the plastid 23S rRNA to the plastid 16S rRNA.
1318   (F) – (H) Analysis of chlorophyll contents and ratio between chlorophyll *a* and chlorophyll *b*. Results are
1319   presented as means ± SE. *t-test* significant level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, n.s: not significant.
1320   Three plants per genotype were analyzed.
1321

47

1322

48
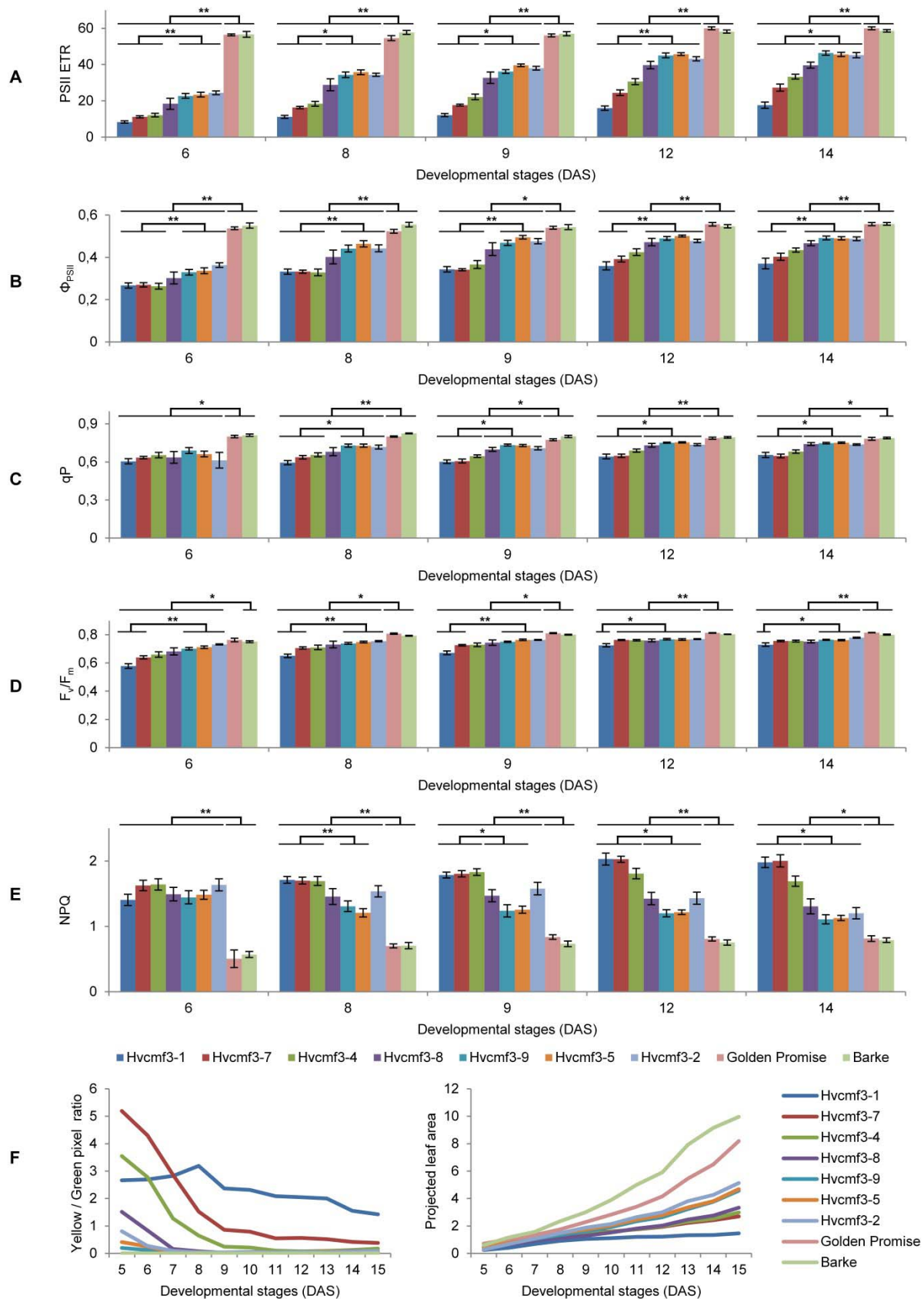
1323 **Figure 6. Determination of photosynthetic parameters and growth dynamics of *Hvcmf3* mutant**
1324 **and wild-type control plants.**

1325 (A) to (E) Measurement of photosynthetic parameters during early developmental stages. Results are
1326 presented as means ± SE. *Student's t-test* significant levels, * p <0.05, ** p <0.01. ETR, electron transport
1327 rate; Φ_{PSII}, photosystem II operating efficiency; qP, fraction of PSII centers that are 'open' based on the
1328 puddle model; F_v/F_m, maximum quantum yield of PSII photochemistry measured in the dark-adapted
1329 state; NPQ, non-photochemical quenching.
1330 (F) Plant growth dynamics. Left panel is yellow/green pixel ratio, and right panel is projected leaf area.
1331
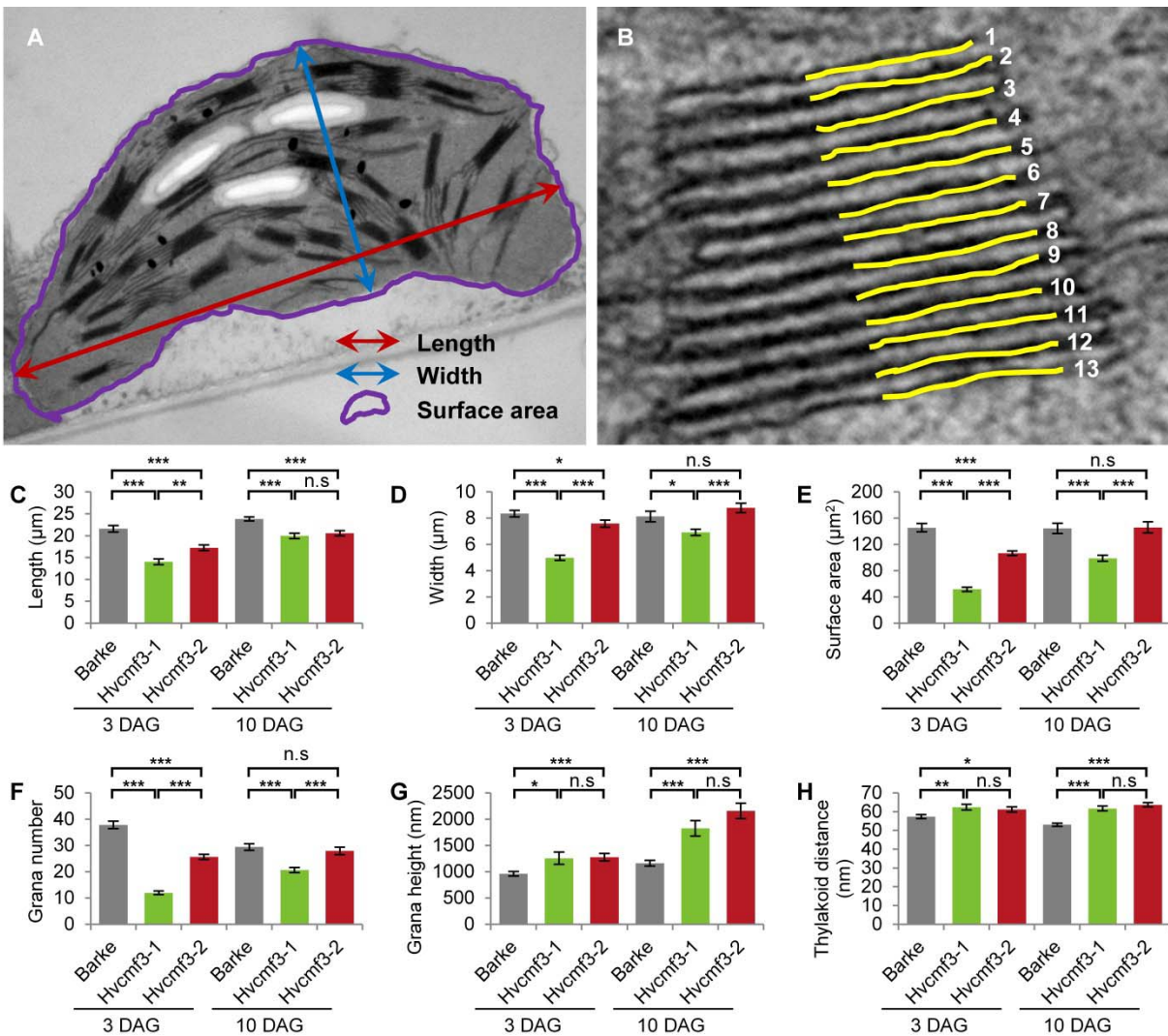
1332

Figure 7. Quantification of chloroplast architecture components.

(A) Diagram for demonstrating the chloroplast length, width, and surface area.

(B) Illustration demonstrating the counting of thylakoid.

(C) to (H) Comparison of chloroplast morphology and grana architecture between wild type and *Hvcmf3* mutants at developmental stages 3 days after germination. Chloroplast length (C), chloroplast width (D), chloroplast surface area (E), grana number (F), grana height (G), and thylakoid distance (H). Results are presented as means ± SE. *t-test* significant level: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, n.s: not significant. Number of chloroplast analyzed n ≥ 24.

1341
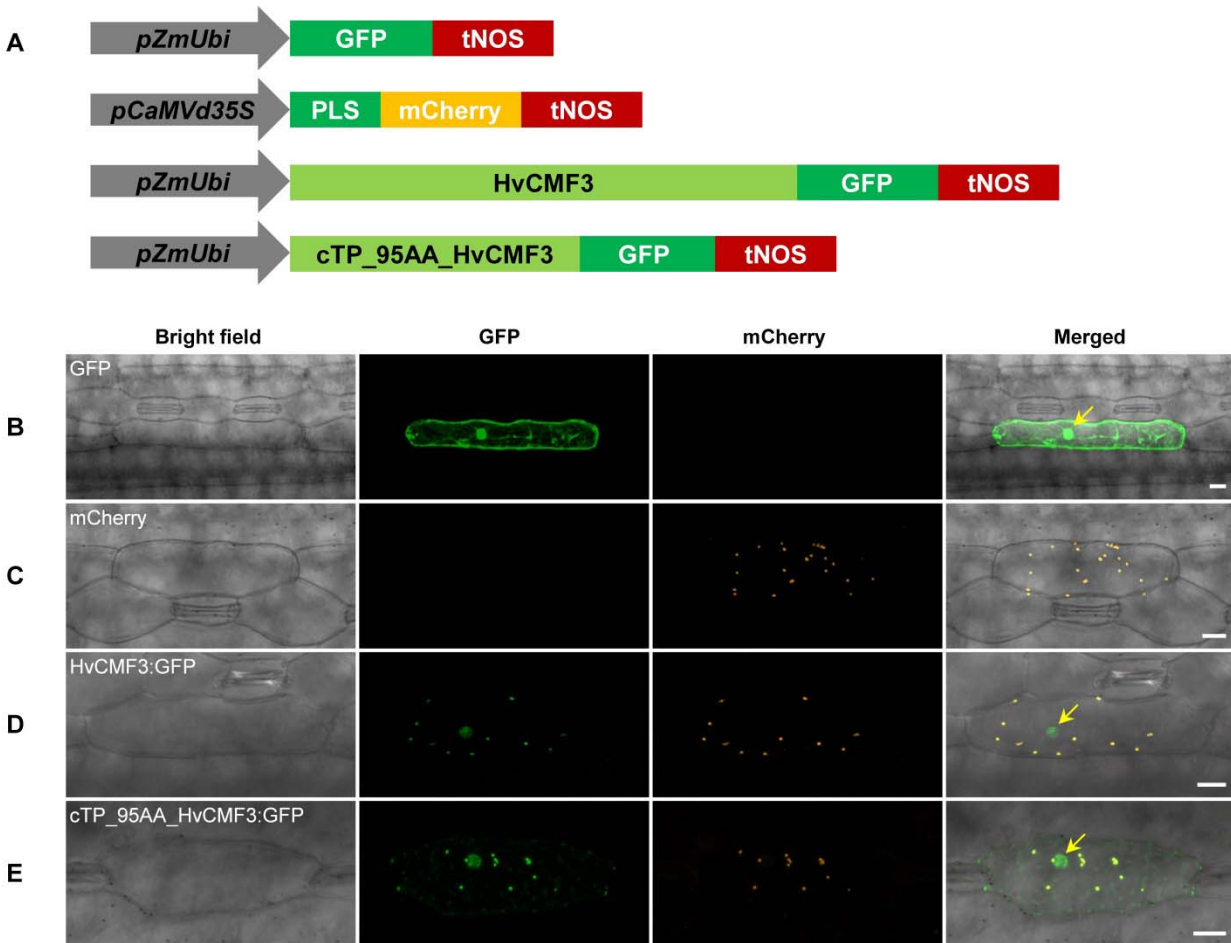
1342

Figure 8. Subcellular localization of HvCMF3.

(A) Schematic diagram of the constructs prepared for transient expression. *pZmUbi*, maize *UBIQUITIN1* promoter. *pCaMVd35S*, Cauliflower Mosaic Virus doubled-enhanced *35S* promoter. GFP, green fluorescent protein. mCherry, mCherry fluorescent protein; PLS, plastid localization signal, i.e. the chloroplast transit peptide (N-terminal 79 amino acids) of the small subunit of tobacco RUBISCO. HvCMF3, coding sequence of wild-type *HvCMF3* gene. cTP_95AA_HvCMF3, N-terminal chloroplast transit peptide of HvCMF3 with a length of 95 amino acids as predicted by online tool PredSL. tNOS, *Agrobacterium nopaline synthase* terminator. The schematic drawing is not in proportion with gene length.

(B) Localization of GFP control with *GFP* being driven by the maize *UBIQUITIN1* promoter.

(C) Localization of the plastid marker.

(D) Localization of HvCMF3:GFP. The GFP fluorescence signal is targeted both to plastid and nucleus compartments.

(E) Localization of cTP_95AA_HvCMF3:GFP. The yellow arrows in the merged panels indicate the nucleus. The first leaf of 10-day-old barley seedlings was used for particle bombardment. The fluorescence was checked 24 hours after bombardment. Scale bar for all images is 20 μm.