Reducing PSY activity fine tunes threshold levels of a *cis*-carotene-derived signal that regulates the PIF3/HY5 module and plastid biogenesis

Xin Hou[1], Yagiz Alagoz[2, ^], Ralf Welsch[3, $], Matthew D Mortimer[1], Barry J. Pogson[1, *], Christopher I. Cazzonelli[2, *]

[1] ARC Training Centre for Accelerated Future Crops Development, Research School of Biology, College of Sciences, The Australian National University, Canberra, ACT 2601, Australia

[2] Hawkesbury Institute for the Environment, Western Sydney University, Locked Bag 1797, Penrith, NSW, 2751, Australia

[3] Faculty of Biology II, University of Freiburg, D79104 Freiburg, Germany

Footnotes

^ Present address: King Abdullah University of Science and Technology, Division of Biological and Environmental Sciences and Engineering, Center for Desert Agriculture, The BioActives Lab, Thuwal 23955-6900, Kingdom of Saudi Arabia.

$ Present address: ScreenSYS GmbH, Engesserstrasse 4a, 79108 Freiburg, Germany.

* Author(s) responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions are: Christopher Cazzonelli (c.cazzonelli@westernsydney.edu.au) and Barry Pogson (barry.pogson@anu.edu.au).

Running Title

PSY activity regulates a *cis*-carotene signalling metabolite

Highlights

Manipulation of the PHYTOENE SYNTHASE catalytic activity in concert with its regulatory protein, ORANGE, reduces threshold levels of acyclic linear *cis*-carotenes that signal control over plastid biogenesis in dark and light grown Arabidopsis seedlings

ABSTRACT

PHYTOENE SYNTHASE (PSY) is a rate-limiting enzyme catalysing the first committed step of carotenoid biosynthesis, and changes in PSY gene expression and/or protein activity alter carotenoid composition and plastid differentiation in plants. Here we identified four genetic variants of *PSY* (*psy*$^{-4}$, *psy*$^{-90}$, *psy*$^{-130}$ and *psy*$^{-145}$) using a forward genetics approach that rescued leaf virescence phenotypes displayed by the *Arabidopsis* CAROTENOID ISOMERASE (CRTISO) mutant *ccr2* (*carotenoid and chloroplast regulation 2*) when grown under a shorter photoperiod. The four non-lethal mutations affected alternative splicing, enzyme-substrate interactions, and PSY:ORANGE multi-enzyme complex binding, constituting the dynamic posttranscriptional fine-tuning of PSY levels and activity without changing localization to the stroma and protothylakoid membranes. *psy* genetic variants did not alter overall total xanthophyll or cis-carotene accumulation in *ccr2 yet* reduced specific acyclic linear *cis*-carotenes linked to the biosynthesis of a yet-to-be-identified apocarotenoid signal. *ccr2 psy* variants modulated the ratio of PHYTOCHROME-INTERACTING FACTOR 3/ELONGATED HYPOCOTYL 5 (PIF3/HY5), displayed a normal PLB formation in etioplasts, and chlorophyll accumulation during seedling photomorphogenesis. Thus, suppressing PSY activity and impairing PSY:ORANGE protein interactions reveals how threshold specific *cis*-carotene levels can be fine-tuned through holoenzyme-metabolon interactions to control plastid development.

Keywords: PHYTOENE SYNTHASE, ORANGE, *cis*-carotene, Apocarotenoid signal, Plastid biogenesis, CAROTENOID ISOMERASE, Arabidopsis, skotomorphogenesis, photomorphogenesis

INTRODUCTION

Carotenoids are a diverse group of hydrophobic isoprenoid compounds synthesized by all photosynthetic and some non-photosynthetic organisms (Baranski and Cazzonelli, 2016). Carotenoid pigments provide colour and aroma to flowers, fruits and vegetables, attracting animals and insects for pollination and seed dispersal (Cazzonelli, 2011). Apart from very few insect species, animals are unable to biosynthesis carotenoids, and their ingestion by animals provides the substrate precursors needed to synthesise vitamin A (retinol), antioxidants, and derivative signalling metabolites that promote human health and immunity (Cazzonelli *et al.*, 2010a). Intermediate substrates can be enriched in specific mutant tissues where they are

62 reported to function as signals, yet the enzymatic steps in the pathway that can regulate their
63 threshold levels remains unclear.

64 PHYTOENE SYNTHASE (PSY) catalyses the first committed step in the carotenoid
65 biosynthetic pathway through the formation of 15 *cis*-phytoene (phytoene) from two
66 geranylgeranyl diphosphate (GGPP) molecules (Zhou *et al.*, 2022). Phytoene is then converted to
67 all-*trans* lycopene in four steps catalysed by two desaturases, PHYTOENE DESATURASE (PDS) and ζ-
68 CAROTENE DESATURASE (ZDS), and two *cis-trans* isomerases, 15-*cis*-ζ-CAROTENE ISOMERASE (ZISO)
69 and CAROTENOID ISOMERASE (CRTISO) (Alagoz *et al.*, 2018). CRTISO and light limit flux through
70 the lycopene branch point in the carotenoid pathway together with epsilon- and beta-lycopene
71 cyclase enzymes that modulate α-carotene and β-carotene biosynthesis. A series of hydroxylation
72 and/or epoxidation steps subsequently converts α-carotene to lutein and β-carotene to zeaxanthin,
73 antheraxanthin, violaxanthin and to neoxanthin, collectively comprising the most abundant
74 carotenoids found in photosynthetic leaves (Baranski and Cazzonelli, 2016) (Fig. S1). The loss-of-
75 function of most enzymatic steps in the upper *cis*-carotene pathway do not induce plant lethality
76 except for PSY, PDS, and ZDS, while the two isomersases rate-limit carotenoid accumulation in a
77 light-dependent manner (Alagoz *et al.*, 2018).

78 Carotenoid cleavage products, referred to as apocarotenoids, regulate plant development
79 and modulate responses to biotic (e.g., insect herbivory) and abiotic (e.g., light and drought)
80 stimuli (Hou *et al.*, 2016; Moreno *et al.*, 2021). Yet to be identified apocarotenoids can also be
81 derived from acyclic linear *cis*-carotenes herein referred to as a *cis*-apocarotenoid signal (*cis*-ACS)
82 (Alagoz *et al.*, 2018; Anwar *et al.*, 2021). Mutations in the Arabidopsis ZDS lead to the
83 accumulation of phytofluene and ζ-carotene isomers that, via CCD4 cleavage, trigger a *cis*-ACS
84 regulating *PHOTOSYNTHESIS ASSOCIATED NUCLEAR GENE* (*PhANG*) expression and chloroplast
85 translation that leads to impaired chloroplast biogenesis, cellular differentiation and disruptions in
86 leaf and flower meristem identity (Avendano-Vazquez *et al.*, 2014; Escobar-Tovar *et al.*, 2021;
87 McQuinn *et al.*, 2023). A *cis*-ACS likely derived from di-*cis*-ζ-carotene, neurosporene, and/or
88 prolycopene produced by the loss-of-function in *crtiso ziso* double mutants was shown to regulate
89 *PhANG* expression and plastid development during skotomorphogenesis (e.g., formation of the
90 prolamellar body; PLB) and photomorphogenesis (e.g., thylakoid structures within the chloroplast)
91 (Cazzonelli *et al.*, 2020). Neurosporene and/or prolycopene were also linked to the epistasis in
92 tomato colour mutations that involved feedback regulation of *PSY1* expression by CRTISO during
93 tomato fruit development (Kachanovsky *et al.*, 2012). Which specific *cis*-carotene (absent or

94 present) and rate-limiting enzymatic step(s) in the pathway can modulate their threshold levels
95 and flux towards generating a *cis*-ACS remains unclear.

96 Alterations in *PSY* transcription and/or translation regulate plant carotenoid content
97 (Alvarez *et al.*, 2016; Cao *et al.*, 2012; Fraser *et al.*, 2007; Maass *et al.*, 2009; Welsch *et al.*, 2010).
98 For example, the transcription of *PSY* is modulated by light signalling through multiple factors,
99 including phytochrome-interacting factors (PIFs) and gibberellin (GA)-regulated DELLA proteins
100 (Cheminant *et al.*, 2011; Toledo-Ortiz *et al.*, 2010). In Arabidopsis, the ORANGE protein (AtOR)
101 post-transcriptionally regulates PSY levels through protein-protein interactions (Zhou *et al.*, 2015).
102 The localization of PSY has also been found to dramatically affect enzymatic activity (Welsch *et al.*,
103 2000), and a single amino acid change can alter the localization of PSY and its protein activity
104 (Shumskaya *et al.*, 2012). There is a scarcity of information as to which amino acids of PSY can
105 undergo substitution and alter activity.

106 After prolycopene biosynthesis and prior to the branch of the epsilon/beta-branch in the
107 pathway, CRTISO regulates isomerisation together with light forming a unidirectional photoswitch
108 that rate-limits *cis*-carotene levels (Alagoz *et al.*, 2018; Nayak *et al.*, 2022). CRTISO appears to be
109 the key rate-limiting step inducing threshold levels of acyclic linear *cis*-carotenes that impair
110 signalling of plastid biogenesis (Dhami *et al.*, 2022). During skotomorphogenesis, the etioplasts of
111 seedling tissues develop prolamellar bodies (PLB), the characteristic paracrystalline membrane
112 structure that defines etioplasts and accelerates photomorphogenesis upon illumination (Park *et*
113 *al.*, 2002; Rodriguez-Villalon *et al.*, 2009).

114 We reported that a linear *cis*-carotene derived apocarotenoid signal (*cis*-ACS) generated by
115 the Arabidopsis *crtiso mutant* (*ccr2*; *carotenoid and chloroplast regulation 2*) that perturbed PLB
116 formation in etioplasts as well as chlorophyll accumulation and chloroplast development that lead
117 to leaf virescence in newly emerged leaf tissues (Cazzonelli *et al.*, 2020; Park *et al.*, 2002). The *cis*-
118 ACS acted in parallel to the repressor of photomorphogenesis, DEETIOLATED 1 (DET1), to
119 transcriptionally regulate the expression of *PHYTOCHROME INTERACTING FACTOR 3* (*PIF3*),
120 *ELONGATED HYPOCOTYL 5* (*HY5*) and *PROTOCHLOROPHYLLIDE OXIDOREDUCTASE* (*POR*) during
121 plastid development (Cazzonelli *et al.*, 2020). We demonstrated that the *cis*-ACS accumulated in
122 dark-grown tissues of *ccr2* or when plants were grown under a shorter photoperiod (Cazzonelli *et*
123 *al.*, 2020) (Fig. S1). A mechanism that can fine-tune *cis*-carotene levels when CRTISO activity is
124 impaired and/or reduced and modulate threshold levels remains to be discovered.

125 We hypothesised that a step in the pathway before CRTISO should induce metabolic

126 epistasis to regulate threshold levels of *cis*-carotene in a light-dependent manner. We performed a
127 forward genetics screen to restore leaf greening to virescent *ccr2* foliar tissues *ccr2* grown under a
128 short photoperiod. We mapped the casual mutations and confirmed the single nucleotide
129 polymorphisms were responsible for the phenotype revision. We functionally characterised the
130 roles of multiple PHYTOENE SYNTHASE variants controlling *cis*-carotene biosynthesis and plastid
131 biogenesis in *ccr2* etiolated and de-etiolated seedlings. We investigated how the different lesions
132 impact interactions between PSY and Arabidopsis ORANGE (AtOR), carotenoid accumulation, and
133 by inference apo-carotenoid signalling that regulates plastid development during seedling
134 transition from skotomorphogenesis to photomorphogenesis.
135

136 Material and Methods

137 *Plant Materials*

138 The *Arabidopsis thaliana* ecotype Columbia (Col-0) and ecotype Landsberg erecta (Ler-0) were
139 used as wild-type controls in this study. The Arabidopsis *ccr2* mutant, carrying a G-to-A mutation in
140 *crtiso* at the start of intron 9 that leads to miss-splicing (Park *et al.*, 2002), was in Col-0 background.
141 The *ccr2* mutant was subjected to forward genetics and a second site revertant screen through
142 mutagenizing seeds in ethyl-methane sulfonate (EMS). EMS-treated seeds were sown in the soil,
143 plants were grown and seeds were collected from pools of 5–10 M_1 plants. Approximately 40,000
144 M_2 seedlings from 30 pools of M1 seeds were screened for the emergence of green juvenile
145 rosette leaves under a 10-h photoperiod. The revertant mutants of *ccr2* showing the emergence of
146 green juvenile rosette leaves that were not virescent when grown under a 10 hr photoperiod are
147 referred to as *rccr2* (revertant of *ccr2*)

148 *Plant growth conditions and treatments*

149 For soil-grown plants, seeds were sown on seed raising mixture (Debco). Seeds were then stratified
150 for 3 d at 4°C in the dark before transferring to an environmentally controlled growth chamber set
151 to 21°C and illuminated by approximately 120 $mmol.m^{-2}.sec^{-1}$ of fluorescent lighting. Unless
152 otherwise stated, plants were grown in a 16-h photoperiod. Photoperiod shift assays were
153 performed by shifting three-week-old plants grown under a 16-h photoperiod to a 10-h
154 photoperiod for one week, and newly emerged immature leaves were scored as displaying either a
155 yellow leaf (YL) or green leaf (GL) phenotype which reflect either impaired or normal plastid

156 development, respectively.

157 For media-grown seedlings, Murashige and Skoog (MS) media (Caisson Labs; MSP01) was
158 used with half-strength of Gamborg's vitamin solution 1000X (Sigma Aldrich) and 0.5% phytagel
159 (Sigma Aldrich), and pH was adjusted to 5.8. Arabidopsis seeds were sterilized for 3 h in chlorine
160 gas in a sealed container, followed by washing seeds once with 70% ethanol and three times with
161 sterilized water. Seeds were sown onto MS media and stratified for 2 d at 4°°C in the dark to
162 synchronize germination.

163 Etiolation experiments were performed by incubating stratified seeds on MS media in the
164 dark at 21°C for 7 d, after which cotyledons were harvested under a dim green LED light. For de-
165 etiolation experiments, stratified Arabidopsis seeds were incubated in the dark at 21°C for 4 d,
166 then exposed to constant light (80 $mmol.m^{-2}.sec^{-1}$) for 72 h at 21°C. Cotyledon tissues were
167 harvested at 24-h intervals for chlorophyll quantification.

168 Seed derived calli (SDC) were generated as previously described (Mathur and Koncz, 1998).
169 Briefly, 5 mg surface-sterilized seeds were plated onto a 90 mm petri dish containing 40 mL of SDC
170 medium (4.33 $g.L^{-1}$ MS basal salts (GibcoBRL), pH 5.8, 3% w/v sucrose, 0.1% v/v Gamborg B5
171 vitamins (Sigma-Aldrich), 0.5 $mg.L^{-1}$ 2,4-D, 2 $mg.L^{-1}$ indole-3-acetic acid (Sigma-Aldrich), 0.5 $mg.L^{-1}$
172 2-isopentenyladenine (Sigma-Aldrich), 0.5% w/v phytagel). Seeds were germinated for 5 d under
173 16-h photoperiod and incubated for 16 d in the dark before harvesting tissues for HPLC analysis.
174 Callus treated with norflurazon (NFZ) were transferred onto the same medium containing 1 $\mu mol.L^{-}$
175 ^{1} NFZ prior to etiolation.

176 ***Pigment quantification***

177 Total chlorophyll was quantified as described previously (Porra *et al.*, 1989) with minor
178 modifications as previously described (Cazzonelli *et al.*, 2020). Reverse phase HPLC (Agilent 1200
179 Series) was performed using either the GraceSmart-C_{18} (4 mm, 4.6 x 250 mm column; Alltech) or
180 YMC-C_{30} (250 x 4.6 mm, S-5mm) column as described (Alagoz *et al.*, 2020). The C_{18} column was
181 used to quantify β-carotene, xanthophylls and generate *cis*-carotene chromatograms, while the C_{30}
182 column was used for an improved *cis*-carotene separation and absolute quantification. Carotenoids
183 were identified based on retention time relative to known standards and their absorbance spectra
184 at 440 nm (β-carotene, xanthophylls, pro-neurosporene, tetra-*cis*-lycopene), 400 nm (ζ-carotenes),
185 340 nm (phytofluene) and 286 nm (phytoene). Absolute quantification of xanthophyll pigments
186 and acyclic linear *cis*-carotenes in micrograms per gram fresh weight (gfw) was derived from peak

187 area using molar extinction coefficient and molecular weight as previously described (Anwar *et al.*,
188 2022). HPLC-based pigment extraction from shoot-derived callus (SDC) samples was performed as
189 previously published (Welsch *et al.*, 2008). To examine the phytoene levels in yeast alongside split
190 ubiquitin and β-galactosidase activity assays, yeast cells were lysed by sonication in 100% acetone
191 on ice. Lysates were centrifuged at 4°C, and the following extraction of phytoene and HPLC analysis
192 were performed as per described for SDC (Welsch *et al.*, 2008).

193 ***Transmission electron microscopy***

194 Cotyledons from 5-d old etiolated seedlings were harvested in dim-green safe light and fixed
195 overnight in primary fixation buffer (2.5% Glutaraldehyde, and 4% paraformaldehyde in 0.1 mol.L^{-1}
196 phosphate buffer pH 7.2) under vacuum, post-fixed in 1% osmium tetraoxide for 1 h, followed by
197 an ethanol series: 50%, 70%, 80%, 90%, 95% and 3 × 100%. After dehydration, samples were
198 incubated in epon araldite (resin):ethanol at 1:2, 1:1 and 2:1, then 3 times in 100% resin. Samples
199 were then transferred to fresh resin and hardened under nitrogen stream at 60°C for 2 d, followed
200 by sectioning of samples using Leica EM UC7 ultramicrotome (Leica Microsystems). Sections were
201 placed on copper grids, stained with 5% uranyl acetate, washed thoroughly with distilled water,
202 dried, and imaged with H7100FA transmission electron microscope (Hitachi) at 100 kV. For each of
203 the dark-grown-seedling samples, prolamellar bodies were counted from 12 fields on 3 grids.

204 ***DNA-seq library construction, Sequencing and Bioinformatics Identification of SNPs***

205 Genomic DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN). One microgram of genomic
206 DNA was sheared using the M220 Focused-Ultrasonicator (Covaris) and DNA libraries were
207 prepared using NEBNext® Ultra™ DNA Library Prep Kit (New England Biolabs) and size selected
208 (~320 bp) using AMPure XP Beads (Beckman Coulter). Libraries were then pooled and paired-end
209 sequencing (NGS) was performed using the Illumina HiSEQ1500. The total number of reads ranged
210 between 29-45 million representing approximately 22-35 times genome coverage. After
211 sequencing, the raw reads were assessed for quality using the FastQC software
212 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and subjected to trimming of
213 illumina adapters and filtering of low quality reads with Adapter Removal programme (Lindgreen,
214 2012). The reads were mapped to the Arabidopsis (TAIR9) genome with BWA mapper (Li and
215 Durbin, 2009). The resultant BWA alignment files were converted to sorted bam files using the
216 samtools v0.1.18 package (Li *et al.*, 2009) and was used as input for the subsequent SNP calling
217 analyses.

218 The SNPs were called and analysed further on both the parent and mutant lines using NGM
219 pipeline (Austin *et al.*, 2011) and SHOREmap (Schneeberger *et al.*, 2009). For the NGM pipeline,
220 the SNPs were called using samtools (v0.1.16) as instructed and processed into '.emap' files using a
221 script provided on the NGM website. The .emap files were uploaded to the NGM web-portal to
222 assess SNPs with associated discordant chastity values. To identify mutant specific SNPs, Parental
223 line SNPs-filtered for EMS changes and homozygous SNPs were defined based on the discordant
224 chastity metric. For SHOREmap, the SHORE software (Ossowski *et al.*, 2008) was used to align the
225 reads (implementing BWA) and called the SNPs (Hartwig *et al.*, 2012). SHOREmap backcross was
226 then implemented to calculate mutant allele frequencies and filter out parent SNPs and defined
227 the EMS mutational changes. Where appropriate, custom scripts were used to identify mutant
228 specific, EMS SNPs, to filter out parent SNPs and annotate the region of interest.

229 The localization of SNPs and InDels were based on the annotation of gene models provided
230 by TAIR database (http://www.arabidopsis.org/). The polymorphisms in the gene region and other
231 genome regions were annotated as genic and intergenic, respectively. The genic polymorphisms
232 were classified as CDS (coding sequences), UTR (untranslated regions), introns and splice site
233 junctions according to their localization. SNPs in the CDS were further separated into synonymous
234 and non-synonymous amino substitution. The GO/PFAM annotation data were further used to
235 functionally annotate each gene.

236 ***Plasmid construction***

237 pEARLEY::PSY-OE binary vectors were designed to overexpress wild type Arabidopsis *PSY* cDNA
238 fragment which was driven by the constitutive CaMV35S promoter. The full-length cDNA coding
239 region was chemically synthesised (Thermo Fisher Scientific) and cloned into the intermediate
240 vector pDONR221. Next, using gateway homologous recombination, the cDNA fragment was
241 cloned into pEarleyGate100 vector (Earley *et al.*, 2006). Vector construction was confirmed by
242 restriction digestion and Sanger sequencing.

243 To express recombinant PSY proteins in *E. coli*, 55 codons encoding a predicted chloroplast-
244 targeting signal were deleted from the 5'-end of the *PSY* cDNA fragments. The cDNA fragments
245 were then codon-optimised and cloned into pRSETA (ThermoFisher Scientific) through restriction
246 sites *Xho* I and *Eco*R I that were attached respectively to the 5'- and 3'-end of the chemically
247 synthesized sequences. To facilitate the purification of recombinant PSY without additional tags,
248 sequence "GAAAATCTGTATTTTCAGGGT" that encodes TEV protease (Tobacco Etch Virus nuclear-
249 inclusion-a endopeptidase) was inserted between *Xho* I site and start codon. Chemical synthesis

250 and cloning were carried out by ThermoFisher Scientific, and the plasmids were verified by
251 sequencing and restriction enzyme digestion. pRSET-PSY-WT, pRSET-PSY4, pRSET-PSY90 and pRSET-
252 PSY130 were transformed to *E. coli* strain DH5α for propagation, followed by transforming BL21
253 (DE3) strain for the expression of recombinant PSY.

254 To transiently express recombinant PSY or psy variants in *Arabidopsis* etioplasts, full length coding
255 regions of PSY, psy^{-4}, psy^{-90}, psy^{-130} and psy-$N_{181}P_{270}$ were chemically synthesized by Thermo Fisher
256 Scientific and cloned into the intermediate vector pDONR221, followed by gateway cloning to
257 pDEST15-CGFP vector (Gift from Professor James Whalen, La Trobe University) to construct pCGFP-
258 AtPSY-WT, pCGFP-AtPSY4, pCGFP-AtPSY90, pCGFP-AtPSY130 and pCGFP-AtPSY-NP.

259 ***Generation of transgenic plants***

260 The *ccr2 psy*$^{-4}$*, ccr2 psy*$^{-90}$*, ccr2 psy*$^{-130}$ *and ccr2 psy*$^{-145}$ EMS-generated double mutant lines were
261 transformed by dipping Arabidopsis flowers with Agrobacteria harbouring pEARLEY::PSY-OE binary
262 vector to generate *ccr2 psy*$^{-4}$::PSY-OE, *ccr2 psy*$^{-90}$::PSY-OE, *ccr2 psy*$^{-130}$::PSY-OE and *ccr2 psy*$^{-}$
263 ^{145}::PSY-OE transgenic lines, respectively (Clough and Bent, 1998). At least 10 independent
264 transgenic lines were generated for each double mutant by spraying seedlings grown on soil with
265 50 mg.L^{-1} of glufosinate-ammonium salt (Basta herbicide).

266 ***Expression, purification and identification of recombinant proteins***

267 *E. coli* strain BL21 (DE3) cells harbouring pRSET-PSY-WT, pRSET-PSY4, pRSET-PSY90 or pRSET-
268 PSY130 vector were inducted by adding 0.5 mmol.L^{-1} IPTG (Isopropyl β-D-1-thiogalactopyranoside,
269 final concentration) and constant shaking at 28°C. Cells were harvested after 4-h induction by
270 centrifugation. The pellets were washed with cold PBS and resuspended in buffer A (50 mmol.L^{-1}
271 NaH_2PO_4, 20 mmol.L^{-1} imidazole, 1 mol.L^{-1} NaCl, pH 7.4), and lysed by lysozyme treatment and
272 sonication on ice. After centrifugation, the supernatant containing the soluble fraction was loaded
273 onto a Pierce 5-mL Ni-NTA column (Thermo Fisher Scientific) equilibrated with buffer A. The
274 column was washed with buffer A containing low-concentration imidazole. Protein was then eluted
275 using buffer A containing 200 mmol.L^{-1} imidazole, 10 μL of which was subjected to gel
276 electrophoresis and western blot with anti-PSY polyclonal antibody to identify the recombinant
277 PSY proteins.

278 Buffer exchange to TEV digestion buffer (50 mmol.L^{-1} Tris.HCl pH 8.0, 150 mmol.L^{-1} NaCl, 20
279 mmol.L^{-1} KCl, 2 mmol.L^{-1} β-mercaptoethanol) was carried out using a Pierce desalting spin column
280 (Thermo Fisher Scientific). The column was equilibrated with TEV digestion buffer and purified

281 recombinant protein was loaded. Protein samples were then collected by centrifuging the column
282 and pooled after buffer exchange. Protein concentration was determined using Bradford reagent
283 (Bio-Rad), followed by digestion using TEV at 4°C for 14 h. A Pierce 5-mL Ni-NTA column (Thermo
284 Fisher Scientific) was equilibrated with TEV digestion buffer and TEV-cleaved recombinant PSY was
285 loaded, followed by elution of the recombinant protein with TEV digestion buffer. The
286 concentration of TEV-cleaved and purified recombinant PSY protein was determined again before
287 *in vitro* enzymatic assays.

288 ***In vitro activity assay of recombinant PSY***

289 To synthesise GGPP, *E. coli* strain TOP10 cells harbouring artificial chromosome pAC-GGPPipi were
290 cultured in chloramphenicol-containing LB medium at 28°C for 18 h. TOP 10 strain carrying pAC-
291 PHYT was used as a positive control to verify the absorbance peak of phytoene in HPLC.

292 Cells were harvested by centrifugation and 1 g of cells (wet weight) were resuspended in 2
293 mL of enzyme assay buffer (100 mmol.L^{-1} Tris.HCl pH 7.6, 10 mmol.L^{-1} $MgCl_2$, 2 mmol.L^{-1} $MnCl_2$,
294 1mmol.L^{-1} 3,3′,3′′-phosphanetriyltripropanoic acid, 20% v/v glycerol and 0.08% v/v Tween 80).
295 Cells were lysed by sonication on ice, followed by centrifugation. To the supernatant, 5 µg of
296 purified recombinant PSY protein was added, followed by incubation at 20°C in the dark for 20 min.
297 At a 4-min interval, aliquots were withdrawn, and the reaction was stopped by adding EDTA.
298 Assays were extracted by adding carotenoid extraction buffer and partitioning phytoene into the
299 ethyl acetate phase. The extractions were dried under nitrogen stream in the dark, resuspended in
300 ethyl acetate and subjected to HPLC analysis.

301 ***In vitro and in vivo activity assay of endogenous Arabidopsis PSY***

302 The immunoprecipitation of endogenous PSY was performed using P-PER® Plant Protein Extraction
303 Kit and Pierce Co-Immunoprecipitation kit (both from Thermo Fisher Scientific) following
304 manufacturer's instructions and is summarised below.

305 For total protein extraction, 160 mg leaf tissue from a 25-d old Arabidopsis plant grown
306 under 16-h photoperiod was homogenised in P-PER® Working Solution containing Halt™ Protease
307 Inhibitor Cocktail (Thermo Fisher Scientific). Following centrifugation, the lower aqueous layer
308 containing total protein was recovered for immunoprecipitation. Twenty micrograms of anti-PSY
309 antibody was coupled to AminoLink Plus coupling resin and the total protein extracted from
310 Arabidopsis leaf tissue was incubated with antibody-coupled resin overnight at 4°C. The resin was
311 then washed and proteins were eluted with elution buffer, followed by neutralisation of the eluent

312 by adding 1 mol.L^{-1} Tris (pH 9.5). Five micrograms of endogenous PSY co-immunoprecipitate (with
313 binding proteins) was subjected to enzymatic activity assay. Immunoprecipitation with normal
314 rabbit IgG (Agrisera Antibodies AS101545) was included as a negative control for the specificity of
315 the anti-PSY antibody.

316 *In vivo* PSY activity was examined by measuring phytoene levels in norflurazon-treated
317 Arabidopsis SDC. Generation of SDC and HPLC analysis was carried out as described in this study.

318 ***Split Ubiquitin Protein-Protein Interaction Assay***

319 The *PSY* cDNA sequences were truncated, and the 55 codons were deleted from the 5'-end. The
320 cDNA sequences of Arabidopsis *OR* and *GGPS11* genes were obtained as previously described
321 (Zhou *et al.*, 2015). The truncated cDNA sequences of *PSY*, *OR* and *GGPPS11* were cloned to make
322 Cub- or Nub-fusion constructs which were then transformed into yeast strain THY.AP5 (Cub) or
323 THY.AP4 (Nub). Yeast strains carrying Cub fusions were mated with strains carrying Cub fusions,
324 and the resulting diploid cells were grown on synthetic complete medium lacking leucine and
325 tryptophan (-LW). Interaction growth tests were performed with overnight cultures spotted on
326 fully selective medium (-LWAH), in a series of 1:10 dilutions with a starting OD_{600} of 2.0 after
327 growing for about 2 d at 29°C. To reduce background activation of reporter genes and visualize
328 different interaction strengths, methionine was added to the medium suppressing the expression
329 of Cub-fusion proteins (+Met; 150 µmol.L^{-1} and 1 mmol.L^{-1}). Control combinations with empty Cub-
330 or Nub-expressing vectors were included in the experiments.

331 For β-galactosidase activity assays, yeast cells were pelleted from 250 µL overnight culture
332 in complete medium and resuspended in 650 µL assay buffer (100 mmol.L^{-1} HEPES.KOH pH 7.0, 150
333 mmol.L^{-1} NaCl, 2 mmol.L^{-1} $MgCl_2$, and 1% w/v BSA), followed by adding 50 µL chloroform and 50 µL
334 0.1% (w/v) SDS and vortex. Enzymatic reactions were started by adding 125 µL of 4 mg.mL^{-1} *ortho*-
335 nitrophenyl-β-galactoside in the assay buffer. Following incubation at 30°C, until the mixture
336 turned visibly yellow, reactions were stopped by adding 1 mol.L^{-1} Na_2CO_3. Reactions were then
337 centrifuged, and OD_{420} was measured to determine the concentration of *ortho*-nitrophenol (oNP)
338 in the supernatant. The B-galactosidase activity was calculated as nmol oNP.min^{-1}.OD_{600}^{-1}, using the
339 molar extinction coefficient:oNP = 3300 g.L^{-1}.mol^{-1}. All β-galactosidase activity assays were
340 performed in triplicate.

341 ***Protoplast isolation and transient expression of recombinant PSY protein***

342 To isolate etiolated Arabidopsis protoplasts, 200 mg of etiolated cotyledons were harvested under

343 green safe light and treated in a 9-mm Petri dish with an enzyme solution (0.5% cellulase, 0.05%
344 pectinase, 600 mmol.L^{-1} mannitol, 10 mmol.L^{-1} $CaCl_2$ and 20 mmol.L^{-1} MES pH 5.6) at 21°C for 6 h
345 with gentle shaking, followed by shaking vigorously for 3 min. The protoplasts were collected
346 through a 60-µm nylon filter and washed 3 times with 20 mL cold washing buffer (600 mmol.L^{-1}
347 mannitol, 10 mmol.L^{-1} $CaCl_2$ and 20 mmol.L^{-1} MES pH 5.6), by centrifugation at 120× g for 5 min
348 and resuspending the pellet. The number of protoplasts per mL was determined using a
349 haemocytometer. In 150 µL of 600 mmol.L^{-1} mannitol containing 10 mmol.L^{-1} $CaCl_2$, 20 µg of PSY-
350 CGFP plasmid was added to 10^6 protoplasts and mixed gently for 5 s. Five hundred microliters of 40%
351 polyethylene glycol 6000 solution (in 500 mmol.L^{-1} mannitol and 100 mmol.L^{-1} $Ca(NO_3)_2$) was
352 added and the contents were mixed gently for 15 s, followed by dilution with 4.5 mL mannitol/MES
353 solution (500 mmol.L^{-1} mannitol, 15 mmol.L^{-1} $MgCl_2$ and 0.1% MES pH 5.6) and incubation at 21°C
354 for 20 min. The protoplasts were pelleted by centrifugation at 120× g and washed with 600
355 mmol.L^{-1} mannitol and 10 mmol.L^{-1} $CaCl_2$, followed by incubation for 16 h at 25°C in the dark. The
356 protoplasts were subjected to western blot using isolated stromal and membrane fractions.

357 ***Isolation of stromal and membrane fractions from etioplasts***

358 About 5× 10^6 etiolated protoplasts transiently expressing recombinant PSY were centrifuged at
359 room temperature at 250× g, and then at 4°C at 2,000× g. The pellet was resuspended in 10 mL
360 isolation buffer and layered on 40% (w/v) Percoll® (Sigma-Aldrich) followed by centrifugation at
361 4°C. The Percoll® layer was removed, and the etioplast pellet was washed twice with isolation
362 buffer. Etioplasts were resuspended in 600 µL hypotonic buffer (10 mmol.L^{-1}Tris.HCl pH 7.0 and 4
363 mmol.L^{-1} $MgCl_2$) on ice for lysis. PLBs were then pelleted through centrifuging the lysate at 3,000×
364 g at 4°C, and the supernatant was collected as the stromal fraction 10 µL of which was used for
365 western blot. The pellet was resuspended in 5 mL hypotonic buffer and subjected to an ultrasonic
366 bath treatment on ice. The suspension was layered on a sucrose step gradient (15 mL of 1 mol.L^{-1}
367 sucrose under 10 mL of 600 mmol.L^{-1} sucrose in hypotonic buffer) and centrifuged at 4°C for 3 h at
368 75,000× g. Protothylakoids accumulating at the interface of the gradient were collected, diluted
369 with an equal volume of hypotonic buffer and concentrated by ultracentrifugation at 125,000× g
370 for 1 h at 4°C; the protothylakoids which pelleted below the gradient were harvested and
371 combined with the fraction collected from the interface. The combined protothylakoid fractions
372 were resuspended in 600 µL hypotonic buffer and 10 µL of each sample was used for western blot.

373 ***Protein extraction and quantification***

374 Fifty to one hundred milligrams of etiolated Arabidopsis seedlings (7-d old) were harvested under
375 dim-green safe light ground to fine powder, or around 100 mg green leaf tissues of 25-d old plants
376 were harvested under normal light. Total protein was extracted using a TCA-acetone method as
377 previously described (Mechin *et al.,* 2007). The concentration of protein was measured using
378 Bradford reagent (Bio-Rad) and adjusted to 2 $\mu g.\mu L^{-1}$. A serial dilution was used to determine
379 western blot sensitivity for each antibody and determine the optimal concentration for
380 quantification. To examine recombinant PSY protein from *E. coli* cell lysates, endogenous PSY from
381 stromal or membrane fractions of Arabidopsis etioplasts, 10 µL of sample was used for gel
382 electrophoresis; to examine endogenous PSY or OR protein levels in Arabidopsis leaf tissues, 10 µg
383 total protein was used; to detect OR protein from co-immunoprecipitation, 10 µL of PSY co-
384 immunoprecipitate was used; to examine PIF3 or HY5 protein levels, 5 µg total protein was used.
385 Proteins run on a gel were transferred to a PVDF membrane (Bio-Rad). Membranes were blocked
386 and then incubated with anti-PSY antibody (1:1000, Agrisera Antibodies AS163991), anti-OR
387 antibody (1:1000, gift from Prof. Li Li, College of Agriculture and Life Sciences, Cornell University),
388 anti-HY5 antibody (Agrisera Antibodies AS121867, 1:1000) or anti-PIF3 antibody (Agrisera
389 Antibodies AS163954, 1:2000) for 2 h at room temperature. Membranes were then washed and
390 incubated with HRP-conjugated Goat anti-Rabbit IgG (1:5000, Agrisera Antibodies AS09602) at
391 room temperature for 90 min, or for PIF3 with HRP-conjugated Rabbit anti-Goat IgG (Agrisera
392 Antibodies AS09605, 1:5000) for 90 min. Membranes were re-probed using anti-Actin poly- clonal
393 antibody (Agrisera Antibodies AS132640, 1:3000) and HRP-conjugated Goat anti-Rabbit IgG
394 (Agrisera Antibodies AS09602, 1:5000) for internal protein normalisation.

395 ***Enzyme-linked immunosorbent assay (ELISA)***

396 Clear Flat-Bottom Immuno 96-well plates (Thermo Fisher Scientific) were coated with Arabidopsis
397 total protein (20 $\mu g.mL^{-1}$) diluted in carbonate buffer (pH 9.6) at 4°C overnight, followed by
398 washing the plates with PBST (pH 7.4). Plates were then blocked for 1 h at room temperature and
399 washed with PBST. Plates were incubated with anti-PSY (1 $\mu g.mL^{-1}$, Agrisera Antibodies AS163991)
400 or anti-OR (1 $\mu g.mL^{-1}$, gift from Prof. Li Li), washed and incubated with HRP-conjugated Goat anti-
401 Rabbit IgG (0.4 $\mu g.mL^{-1}$, Agrisera Antibodies AS09602). All antibodies were diluted in PBS-TY buffer
402 PBS-TY buffer (PBS with 0.05% v/v Triton X-100 and 1% w/v yeast extract, pH7.4) and all
403 incubations were performed at 37°C for 1 h. After antibody incubations and washing the plates,

404 substrate solution containing 0.5 $mg.mL^{-1}$ O-phenylenediamine dihydrochloride (Sigma-Aldrich) in
405 phosphate-citrate buffer (pH 5.0) and 0.02% (v/v) H_2O_2 was added and incubated for 5 min at room
406 temperature. The reactions were stopped with $2mol.L^{-1}$ H_2SO_4, and the absorbance at 492 nm was
407 measured on a TECAN M1000PRO plate reader (Tecan Group). Solubilisation buffer diluted in
408 carbonate buffer was used as blank control, and a serial dilution of purified recombinant PSY
409 protein in carbonate buffer was used as a positive control. All ELISAs were done in triplicates.

410 ***Quantitative RT-PCR (qRT-PCR)***

411 The total RNA was extracted using Spectrum™ Plant Total RNA kit (Sigma-Aldrich) as per the
412 manufacturer's protocol. First strand cDNA synthesis was performed using 1 µg total RNA, Oligo
413 dT18 primer and Transcriptor First Strand cDNA synthesis kit (Roche) as per the manufacturer's
414 instructions. The qRT-PCR was performed using 2 µL of primer mix (2 µM for each primer), 1 µL
415 1/15 diluted cDNA template, 5 µL LightCycler 480 SYBR Green I Master mix and 2 µL sterile milli-Q
416 water. For each sample, three technical replicates for each of the three biological replicates were
417 tested. The relative gene expression levels were calculated by using relative quantification (Target
418 Eff Ct(Wt-target)/Reference Eff Ct(Wt-target)) and fit point analysis (Pfaffl, 2001). Protein
419 Phosphatase 2A (PP2A; At1g13320) gene was used as reference gene for normalisation
420 (Czechowski *et al.*, 2005). *PP2A* has been validated against Cyclophilin (At2g29960) and TIP41
421 (At4g34270) as secondary reference genes for different Arabidopsis tissues (Cazzonelli *et al.*, 2014;
422 Cazzonelli *et al.*, 2010b). DNA plasmids harboring wild type transcripts or splice variant cDNA was
423 used as a reference to generate a standard curve and optimise the qRT-PCR conditions. All primer
424 sequences are listed in Table S1.

425 ***Protein Modelling***

426 The Arabidopsis PSY structural models were built using the Colabfold (Mirdita *et al.*, 2022)
427 implementation of Alphafold2 (Jumper *et al.*, 2021) and ESMFold (Lin *et al.*, 2023). The ColabFold
428 structure was inferred from the Arabidopsis PSY amino acid sequence with the signal peptide
429 removed, using the 'no template' method, and the predicted structure was relaxed using amber
430 force fields. Other settings were as the default. Five structures were predicted and ranked
431 according to the predicted local distance difference test (pLDDT) (Mariani *et al.*, 2011) confidence
432 measure with the highest ranked model selected with a score of 88. The ESMFold model was
433 made using one copy and three recycles with model confidence pLDDT score of 87. While the
434 ESMFold model had higher pLLDT scores for the N- and C-termini and the N-terminal variable loop

435 region, the pLDDT scores for the conserved core of the enzyme were marginally lower (confidence
436 data is included with both structural models in the supplementary materials). As such the
437 ColabFold model was chosen, and the fold was validated by comparing the model to the crystal
438 structure of Enteroxoxxus hirae dehydrosqualene synthase complexed with three Mg2+ ions and
439 the Farnesyl thiopyrophosphate – FPS ligand (PDB accession: 5IYS). Structural alignments were
440 made using PyMOL (Schrödinger, 2015) and alignments of the ColabFold model with the ESMFold
441 model and 5IYS crystal structures are included in the supplementary materials. Figures were made
442 using PyMOL.

443

444 RESULTS

445 *Mutations in PSY reduce cis-carotene levels and restore plastid development in ccr2*

446 Our forward genetics approach uncovered uncharacterised second-site mutations (referred to as
447 revertant *ccr2*; *rccr2*$^{-4}$, *rccr2*$^{-90}$, *rccr2*$^{-130}$ and *rccr2*$^{-145}$) capable of restoring leaf greening to *ccr2*
448 plants that would otherwise show virescence reflected by a yellow leaf (YL) phenotype when
449 grown under a short (10 h) photoperiod (Fig. 1A, Fig. S2A). All *rccr2* lines showed enhanced
450 chlorophyll levels in young emerging leaves (~2-fold compared to *ccr2*) and in cotyledons following
451 de-etiolation matching WT levels (Fig. 1B-C) and yet displayed reduced foliar lutein levels like *ccr2*
452 (see Figure 3B in Cazzonelli et al., 2020). In dark-grown etiolated seedlings, the four *rccr2* lines
453 displayed a PLB structure similar to, or partially restored, when compared to WT (Fig. 1D-E).
454 Therefore, restoration of PLB formation in etiolated *rccr2* tissues leads to normal chlorophyll
455 accumulation in de-etiolated seedlings and foliar tissues from plants grown under a shorter
456 photoperiod.

457 Next-generation sequencing of *rccr2*$^{-4}$, *rccr2*$^{-90}$, *rccr2*$^{-130}$ and *rccr2*$^{-145}$ revealed each
458 harboured a single mutation in the *PSY* (At5g17230) gene (Fig. 2A and Table S2). The single
459 nucleotide polymorphisms (SNPs) were confirmed by Sanger sequencing (Fig. 2B). *rccr2*$^{-4}$ and
460 *rccr2*$^{-90}$ had G→A mutations at exons 4 and 5 leading to M266I and A352T amino acid changes in
461 PSY respectively referred to as *psy*-4 and psy-*90*. In *rccr2*$^{-130}$ a C→T mutation at exon 3 resulted in
462 a substitution of P178 to S, referred to as *psy-130*. In *rccr2*$^{-145,}$ the G→A mutation at the exon
463 2/intron 3 border generated an alternative splice site that retained intron 3 and potentially
464 introduced a premature stop codon and truncate protein referred to as *psy-145* (Table S2). Albino
465 seedling phenotypes of *rccr2*$^{-145}$ growing in soil were observed (20% of a homozygous population),

466 revealing an atypical segregation of *psy*$^{-145}$ splice variant dominance (Fig. S3A). Both the full-length
467 and truncated versions of the *PSY* transcripts were evident in *rccr2*$^{-145}$, although the full-length
468 transcript was 3-fold higher than the spliced variant in green rosette leaves (Fig. S3B-C). Therefore,
469 the four *psy* gene mutations retain functionality of PSY and only *rccr2*$^{-145}$ displays partial lethality.

470 However, PSY protein levels were reduced in *ccr2 psy* variants relative to WT and *ccr2*
471 (Figure 3A). To verify the PSY casual mutations were responsible for restoring plastid development
472 in *ccr2*, a functional copy of PSY regulated by the CaMV35s promoter was overexpressed (PSY-OE)
473 in the four *psy ccr2* double mutants, *ccr2*, and WT. Western blot analysis confirmed that PSY
474 protein levels in 4-week old foliar tissues were overexpressed at similar levels in WT, *ccr2* and *ccr2*
475 *psy* double mutant lines harbouring PSY-OE (Fig. 3A). The overexpression of PSY restored the
476 yellow leaf virescence in all *ccr2 psy*::*PSY*-OE lines grown under 10-h photoperiod (Fig. S4A).
477 Consistent with the yellow leaf (YL) phenotype, a two-fold reduction of total chlorophyll was
478 observed in the YL of *PSY*-OE lines compared to older green leaves from WT or WT::*PSY*-OE lines
479 (Fig. S4B). Chloroplast development appears to have been complemented by genetic perturbations
480 that lower PSY levels, raising questions as to what regulates virescence in *ccr2*.

481 We quantified linear *cis*-carotene content in etiolated tissues from *ccr2 psy* double mutants
482 (Fig. 3). Total *cis*-carotene content in the *ccr2 psy* was reduced by 40% – 70% relative to *ccr2*.
483 Overexpression of PSY in *ccr2 psy* enhanced total *cis*-carotene content back to *ccr2* levels. *cis*-
484 carotene content in *ccr2* was not affected by *PSY* overexpression, and *cis*-carotenes were not
485 detected in WT (Fig. 3B). Intriguingly, the total *cis*-carotene pool was not significantly different
486 between *ccr2 psy-145* and *ccr2 psy-145* harboring PSY-OE indicating the phenotypes are not
487 proportional to the total *cis*-carotene pool.

488 We investigated if there is a threshold to the complementation such as an individual *cis*-
489 carotene being present or absent. In comparison to *ccr2*, *ccr2 psy* double mutants (*ccr2 psy-4*, *ccr2*
490 *psy-130,* and *ccr2 psy-145*) showed a significant reduction in the levels of phytoene (4 – 6 fold), di-
491 *cis*-phytofluene (4 – 6 fold), tri-*cis*-ζ-carotene (3 – 5 fold), di-*cis*-ζ-carotene (3 – 5 fold) and tri-*cis*-
492 neurosporene (2 fold) (Fig. 3C). The levels of tetra-*cis*-lycopene (prolycopene; P-Lyc) in *ccr2 psy*$^{-130}$
493 and *ccr2 psy*$^{-145}$ were comparable to *ccr2*, and marginally lower in *ccr2 psy*$^{-4}$ (Fig. 3C). Therefore,
494 the restored plastid biogenesis in *ccr2* by mutations that lower PSY protein levels correlate with a
495 threshold reduction in *cis*-carotenes other than prolycopene.

496

497 ***psy variant alleles restore PIF3 and HY5 transcript and protein levels in ccr2***

498 We questioned if the *cis*-carotenes had a structural or signalling function by investigating
499 how the *psy* variants regulated *PIF3* and *HY5* in etiolated *ccr2* seedlings. Transcript levels of *PIF3*
500 and *HY5* in *ccr2* were 6-fold higher, and 2-fold lower in *ccr2*, respectively (Cazzonelli *et al.*, 2020). In
501 the four *ccr2 psy* mutants, the transcript levels of both *PIF3* and *HY5* were like WT (Fig. 4A). In
502 comparison to WT, PIF3 protein levels were dramatically up-regulated in *ccr2*, yet in the four
503 double mutants, levels were comparable to WT. HY5 protein levels were significantly lower in *ccr2*
504 at trace levels, while all four *ccr2 psy* double mutants displayed considerably higher HY5 protein
505 levels like WT (Fig. 4B). The higher PIF3 to HY5 ratio of transcripts in etiolated *ccr2* seedlings,
506 corroborates with protein levels and the four PSY mutant variants restore WT levels of PIF3 and
507 HY5 levels. Therefore, a threshold level of specific *cis*-carotenes can be linked to the generation of
508 a signalling metabolite.

509

510 ***psy variant alleles show reduced PSY enzyme activities***

511 We examined phytoene synthase activity of the four *psy* variant alleles. First, we used an *in*
512 *vivo* shoot-derived calli (SDC) assay where seeds are germinated in the light to develop seed-
513 derived callus and thereafter transferred to darkness for 14 days (Schaub *et al.*, 2018). During
514 darkness ongoing carotenoid biosynthesis is compensated by degradation and norflurazon which
515 inhibits PDS activity enriches for phytoene and reflects the rate of phytoene synthase activity as a
516 proxy for PSY protein levels. SDC of WT and *ccr2* accumulated similar levels of phytoene, which
517 were 1.5- to 3-fold higher than in the four *ccr2 psy* double mutants (Fig. 5A). Endogenous phytoene
518 synthase was purified from WT, *ccr2* and *ccr2 psy* foliar tissues by immunoprecipitation and *in vitro*
519 activity assays showed that PSY variants synthesised less phytoene indicating impaired PSY activity
520 (Fig. 5B). Using an *E. coli* carotenoid expression system generating GGPP (Cunningham and Gantt,
521 2007), we tested the *in vitro* activity of purified recombinant versions of PSY and its mutant
522 variants (*psy-4*, *psy-90*, *psy-130*) to convert GGPP to phytoene. In agreeance with pervious assays,
523 *psy-4* and *psy-90* displayed significantly reduced phytoene levels, while *psy-130* was marginally
524 lower. Collectively, the *psy* mutant variants showed impaired PSY activity.

525 We used computational modelling to predict how the amino acid substitutions of psy-4,
526 psy-90 and psy-130 (Fig. 2A) alter enzyme activity compared to a structural model of PSY (Fig. 6). In
527 PSY-130, a proline to serine substitution (P178S) could impact metal binding at aspartate residues

528 172 and 176, providing higher flexibility to form a hairpin. The P178S mutation may impact the
529 metal binding by aspartate 302, which is proximal to proline 178 of the predicted 3D structure.
530 These aspartate residues reside within the $D_{172}ELVD_{176}$ and $D_{298}VGED_{302}$ regions conserved
531 isoprenoid synthase proteins shown to form an active site and bind substrates (Pandit *et al.*, 2000).
532 In *psy-4*, the methionine substitution to isoleucine (P266I) is situated close to the lower substrate
533 pocket. Alanine 352, which was mutated to threonine in psy-90 (A353T) models steric interactions
534 with the adjacent α-helix. This could impact substrate binding because of the proximity to the 3D
535 structured substrate pocket (Fig. 6). Therefore, the psy-4 and psy-90 variants likely alter the
536 structure of the active site and/or substrate binding, corroborating with reduced phytoene
537 synthase catalytic activity demonstrated using *in vivo* and *in vitro* assays.

538

539 ***psy mutant alleles affect protein-protein interactions between PSY and OR***

540 A split ubiquitin system (SUS) based on yeast-two-hybrid (Y2H) was used to evaluate
541 whether the PSY amino acid substitutions alters protein-protein interaction between PSY and
542 Arabidopsis ORANGE (AtOR) that can post-transcriptionally regulate PSY levels (Zhou *et al.*, 2015).
543 A clear interaction was observed between wild-type PSY and AtOR by growth in selective media,
544 and β-galactosidase activity of yeast strains co-expressing different PSY versions fused to C-
545 terminal ubiquitin moiety (Cub) and AtOR fused to N-terminal ubiquitin moiety (Nub) (Fig. 7A). A
546 suppressed interaction was detected when P266I (PSY-4) or A353T (PSY-90), but not P178S (PSY-
547 130), was fused to Cub, while psy-90 generally failed to interact with AtOR (Fig. 7A and 7B).

548 We tested the interaction between PSY versions and Arabidopsis GGPP SYNTHASE 11
549 (AtGGPPS11), which was suggested to be an essential protein interacting with PSY to produce
550 phytoene (Ruiz-Sola *et al.*, 2016). The protein-protein interaction between A353T (PSY-90) and
551 AtGGPPS11 was negatively affected, although P266I (PSY-4) and P178S (PSY-130) variants showed
552 WT interactions with AtGGPPS11 (Fig. 7A and 7B). Consistent with PSY activity assays, the
553 production of phytoene in yeast cells expressing A353T (PSY-90) plus AtGGPPS11 was blocked,
554 while P266I (PSY-4) or P178S (PSY-130) plus AtGGPPS11 showed a -2 to 4-fold reduction in
555 phytoene levels compared to PSY (Fig. S5).

556 To confirm the reduced interaction between AtOR and PSY variants, we quantified AtOR in
557 the PSY co-immunoprecipitate mixture using western blot analysis. Considerably lower levels of
558 AtOR were observed in the A353T (PSY-90) co-immunoprecipitate, and a clear reduction was

559 evident in the P266I (PSY-4) co-immunoprecipitate evidencing that these two variants, but not
560 P178S (PSY-130) or PSY-140 can impair the physical interaction with AtOR (Fig. 7C).

561 We next quantified PSY and AtOR protein levels in total protein extracts of WT, *ccr2* and
562 *ccr2 psy* variant leaf tissues. PSY protein levels in *ccr2 psy-4* and *ccr2 psy-90* were lower (Fig. 8),
563 and enzyme-linked immunosorbent assays (ELISAs) confirmed 2- and 5-fold reductions,
564 respectively (Fig. S6). The levels of AtOR protein were also lower in *ccr2 psy-4* and *ccr2 psy-90* by
565 3- to 4-fold (Fig. 8 and Fig. S6). Western blot and ELISA assays showed that both AtPSY and AtOR
566 protein levels in *ccr2 psy-130* and *ccr2 psy*-145 were like WT (Fig. S6). We concluded that the PSY-
567 AtOR interaction was impaired in *ccr2 psy-4* and *ccr2 psy-90,* leading to a reduction in PSY and
568 AtOR protein levels.

569

570 ***PSY localization was not affected in psy mutant variant alleles***

571 To examine the sub-organellular localization of the PSY versions, we transiently expressed
572 recombinant protein of the PSY versions in protoplasts isolated from cotyledons of Arabidopsis
573 etiolated seedlings and detected PSY or psy variants in PLB as well as stroma fractions of etioplasts.
574 As a control, we included psy-$N_{181}P_{270}$ (psy-NP), the Arabidopsis equivalent variant of ZmPSY-
575 $N_{168}P_{257}$ that was reported to display an altered localization in etioplasts and chloroplasts. PSY
576 takes two topological forms: membrane-bound and stromal, and we detected endogenous PSY in
577 both PLB and stroma fractions (Fig. 9). There was no significant difference among the pattern of
578 distribution of the PSY and PSY variants in the PLB and stroma fractions of etioplasts (Fig. 9). The
579 transiently expressed recombinant PSY or PSY variants appeared to affect PSY levels in the stroma
580 only (Fig. 9).

581

582 **DISCUSSION**

583 Here we define new mutations that disrupt PSY enzyme activity at the entry point to
584 carotenoid biosynthesis that act epistatically to *crtiso* mutations with respect to carotenoid profiles
585 and visual phenotypes, namely impaired leaf greening. A forward genetics screen identified four
586 non-lethal *PSY* mutant alleles harbouring a single point mutation that reduced PSY activity, protein
587 levels, and AtOR interactions, which underpin substrate supply into the carotenoid biosynthetic
588 pathway (Fig. 10). The psy variants significantly reduced threshold levels of specific *cis*-carotenes in

589 *ccr2* and presumably the production of a *cis*-ACS regulating plastid biogenesis and the PIF3/HY5
590 module (Fig. 10). Single amino acid sequence changes can profoundly alter PSY activity and
591 identification, and optimization of which key amino acid residues reduce activity will help to not
592 only unravel the intrinsic features of PSY activity but facilitate modelling of highly efficient PSY for
593 development of carotenoid enriched crops.

594

595 ***Mutations in PSY alter RNA splicing, enzyme activity, protein levels, and multi-***
596 ***enzyme interactions***

597 Mechanisms comprised of alternative splicing, enzyme-substrate interactions, multi-
598 enzyme complex binding, constitute dynamic regulation to fine-tune PSY activity and reduce linear
599 *cis*-carotene biosynthesis without impacting overall carotenoid end-product accumulation. Single
600 amino acid changes in PSY have been shown to increase enzymatic activity and up-regulate
601 carotenoid biosynthesis (Shumskaya *et al.*, 2012; Welsch *et al.*, 2010). Here we defined four unique
602 mutations that impaired PSY operations in *ccr2,* leading to the restoration of plastid biogenesis (Fig.
603 1). Alternative splicing changes in *PSY* transcript sequence length have been shown to produce
604 variants with different translation efficiency and/or distinct enzyme activity to control the
605 functional PSY in Arabidopsis leaves (Alvarez *et al.*, 2016; Zhou *et al.*, 2022). The mutation in *psy-*
606 145 altered intron-3 splicing leading to a truncated protein evident by 20% of *ccr2 psy-145*
607 homozygous seedlings displaying an albino phenotype (Fig. 2, S2, Table S2). The remaining 80% of
608 *ccr2 psy-145* seedlings remained green, harbouring both the spliced and non-spliced variants of
609 PSY, retained interactions between PSY and AtOR without affecting their protein levels (Fig. 3, 6, 7
610 and 8), yet reduced phytoene levels *in vivo* and *in vitro* activity assays (Fig 5). The reduced
611 phytoene levels in *ccr2 psy-145* compared to *ccr2* could be due to the inefficiency in intron-3
612 splicing (Fig. S3), keeping consistent with the reduction in *cis*-carotene levels in *ccr2 psy-145* (Fig. 3).

613 Three non-lethal mutations impaired PSY operations by reducing PSY activity for *psy*-4, *psy-*
614 90 and *psy*-130 when tested *in vivo* (SDC) and *in vitro* (*E.coli* assays) (Fig 5). The amino acid
615 substitutions M266I (psy-4), A352T (psy-90), P178S (psy-130) modified PSY operations to lower
616 accumulation and *cis*-carotene in *ccr2* (Fig. 3). The AtPSY $_{172}$DELVD$_{176}$ and $_{298}$DVGED$_{302}$ sequences
617 were predicted to form an active site and bind phosphate groups of GGPP, and are conserved
618 among isoprenoid synthase superfamily (Pandit *et al.*, 2000; Shumskaya *et al.*, 2012; Zhou *et al.*,
619 2022). Our modelling of the PSY enzyme showed that these two sequences were positioned at the

620 entry to the substrate pocket (Fig. 6). The P178S mutation of psy-130 (Fig. 2, Table S2) is between
621 the $_{172}$DELVD$_{176}$ sequence and $S_{181,}$ which is an essential amino acid for the activity and localization
622 of PSY (Shumskaya *et al.*, 2012). P178S could alter the structure of the substrate pocket and/or
623 impact putative metal binding at D172, D176 and D302 within the two conserved PSY protein
624 domains (Fig. 6). The M266I mutation in psy-4 is close to the bottom segment of the substrate
625 pocket and likely effects substrate fitting and/or reaction catalysis. The A352T mutation of psy-90 is
626 also close to the substrate pocket (Fig. 6) and may interfere with the formation of a preferable
627 structure of the pocket. The transient expression of psy-4, psy-90, psy-130, and Atpsy-$N_{181}P_{270}$ (psy-
628 NP) variants in Arabidopsis protoplasts showed similar localisation in PLB and stroma fractions
629 when compared to the WT PSY (Fig. 9). While Atpsy-$N_{181}P_{270}$ can be localised to plastoglobuli in
630 rice and maize protoplasts (Shumskaya *et al.*, 2012) this was not the case in Arabidopsis. The P178S
631 (psy-130), M266I (psy-4), and A352T (psy-90) mutations do not affect plastid localisation yet impair
632 PSY enzyme-substrate interactions that reduces *cis*-carotene biosynthesis.

633 The PSY:AtOR interaction was reduced in *psy*-4 and *psy*-90 (Fig 7, 8, 9), thereby reducing the
634 levels of both proteins and phytoene levels in their respective *ccr2 psy* variant backgrounds (Fig. 5A,
635 5B, 8, S6). The reduced PSY:AtOR interaction and lower protein levels are consistent with a
636 previous report (Zhou *et al.*, 2015), where a decrease in AtOR protein levels negatively affected the
637 recruitment of inactive PSY populations from the stroma to the membrane and consequently
638 lowered PSY activity. Furthermore, reduced interaction between PSY and OR increases the
639 proportion of PSY which is subjected to Clp-mediated degradation, thus reducing the overall
640 amount of enzymatically active PSY (Welsch *et al.*, 2018). A nonsense mutation in *Cucumis melo*
641 OR protein (*Cmor*-lowβ) was also able to lower tetra-*cis*-lycopene levels in *crtiso* melon fruit tissues
642 (Chayut *et al.*, 2017). Therefore, the impairment in PSY-AtOR interaction in *ccr2 psy-4* and *ccr2 psy-*
643 *90* leads to a reduction in PSY and AtOR protein levels reveals a new rate-limiting step to modulate
644 the biosynthesis of *cis*-carotenes and an associated signalling metabolite.

645 For the biosynthesis of carotenoids, a large multi-enzyme complex appears to involve PSY
646 and GGPPS (Camara, 1993; Cunningham and Gantt, 1998; Fraser *et al.*, 2000; Shumskaya and
647 Wurtzel, 2013). Within the metabolon, GGPPS may channel its product GGPP to PSY as substrate;
648 hence, interaction with GGPPS is required for PSY activity (Ruiz-Sola *et al.*, 2016). The A352T
649 mutation in psy-90 perturbed the PSY-AtGGPPS11 interaction such that in yeast cells phytoene
650 synthesis was almost abolished evidencing that full PSY activity requires interaction with GGPPS11
651 (Fig. 7 and S5). Our study herein demonstrates that PSY, AtOR and GGPPSS11 constitute a multi-

652 enzyme complex that acts to bottleneck the supply of linear *cis*-carotene substrates required to
653 generate a *cis*-ACS that regulates plastid development (Fig. 10).
654

655 ***psy variants reduce di-cis-ζ-carotene and tri-cis-neurosporene substrates***

656 Acyclic linear *cis*-carotene levels in *ccr2 psy* variants were greatly reduced below the threshold in
657 *ccr2* that signals control over plastid biogenesis. The psy variants reduced total linear *cis*-carotene
658 levels in etiolated seedlings by 2- to 3-fold, yet there was only a significant reduction in phytoene,
659 di-*cis*-phytofluene, tri-*cis*-ζ-carotene, di-*cis*-ζ-carotenes and tri-*cis*-neurosporene, but not tetra-*cis*-
660 lycopene, ruling the later out as a substrate for ACS (Fig 3, 4). We rule out phytoene, phytofluene,
661 and tri-*cis*-ζ-carotene as potential substrates for generating a *cis*-ACS since the *ccr2 ziso* double
662 mutants continue to accumulate these substrates (Cazzonelli *et al.*, 2020) and just like *ccr2 psy*
663 variants show restored plastid biogenesis. Evidence using chemical inhibition of CCD activity
664 supported that tri-*cis*-neurosporene and di-*cis*-ζ-carotene were potential substrate(s) for *in vivo*
665 cleavage into a signalling metabolite (Cazzonelli *et al.*, 2020). Collectively both these studies reveal
666 that di-*cis*-ζ-carotene and/or tri-*cis*-neurosporene are the preferred substrate(s) of a *cis*-ACS (Fig.
667 10).
668

669 ***psy variants modulate the PIF3/HY5 regulatory hub***

670 PIF3 and HY5 are key regulatory transcription factors controlling the dark-to-light transition
671 of plants in concert with carotenogenesis. A high PIF3/HY5 ratio maintains skotomorphogenesis in
672 etiolated seedlings, whereas a low PIF3/HY5 ratio promotes photomorphogenesis in de-etiolated
673 seedlings (Dong *et al.*, 2014; Job and Datta, 2020; Osterlund *et al.*, 2000). PIF3 and HY5 negatively
674 and positively regulate *PSY* expression, respectively, affecting carotenoid biosynthesis (Fig. 10)
675 (Toledo-Ortiz *et al.*, 2010; Toledo-Ortiz *et al.*, 2014). DET1 and constitutive photomorphogenic 1
676 (COP1) modulate the PIF3/HY5 regulatory hub by post-transcriptionally controlling the levels of
677 PIF3 and HY5 (Lau and Deng, 2012; Llorente *et al.*, 2017; Stephenson *et al.*, 2009; Xu *et al.*, 2016).
678 Using a chemical inhibitor of CCD cleavage, we previously showed that a *cis*-ACS can act
679 independent of DET1 to transcriptionally regulate *PIF3* and *HY5* mRNA expression and their
680 corresponding protein levels (Cazzonelli *et al.*, 2020). Here we evidenced that limiting a threshold
681 level of specific *cis*-carotenes in *ccr2* can restore the ratio of PIF3 and HY5 in dark grown seedlings
682 such that it mimics a WT plant. Interestingly, the modulation of a higher PIF3/HY5 protein ratio (ie.,

683 transcript and protein levels of *PIF3* were up-regulated, while *HY5* was down-regulated) does not
684 affect *PSY* protein levels in *ccr2* etiolated tissues. Rather the modification of PSY protein
685 levels/activity in the four *ccr2 psy* mutants can lower the PIF3/HY5 ratio by decreasing *PIF3* and
686 increasing *HY5* transcript and protein levels (Fig. 4). The reduction of PSY protein/activity levels in
687 *psy* variants likely fine-tunes threshold levels of a *cis*-carotene derived signal that regulates the
688 PIF3/HY5 module without affecting end-product carotenoid accumulation.

689 In summary, a lack of carotenoid isomerisation and shorter photoperiod can trigger
690 threshold levels of *cis*-carotenes that signal transcriptional regulation of the PIF3/HY5 module to
691 control plastid biogenesis during both skotomorphogenesis and photomorphogenesis. Non-lethal
692 point mutations in four *psy* genetic variants manipulated PSY alternative splicing, activity, PSY-AtOR
693 interactions, and PSY protein levels reduced levels of di-*cis*-ζ-carotene and/or tri-*cis*-neurosporene,
694 presumably below a threshold required for ACS production. Given these are the most likely
695 substrates for a yet-to-be-identified *cis*-ACS it remains of interest as to the physiological and
696 biochemical basis of this threshold, which are more than quantities normally required for the
697 production of a hormone. The epistasis between CRTISO and PSY is paramount to controlling *cis*-
698 carotene signalling in plants when light levels are limiting (Fig. 10).

699

700 **Acknowledgements**

701 The Authors thank Julian Koschmieder, Dennis Schlossarek and Carmen Schubert from the Faculty
702 of Biology II, University of Freiburg, D79104 Freiburg, Germany for their technical support to
703 quantify pigments in shoot derived calli. We thank Jiwon Lee (Centre for Advanced Microscopy,
704 The Australian National University, Canberra, ACT 2601, Australia) for technical support in
705 Transmission electron microscopy. We thank John Rivers, Kai Xun Chan (The Australian National
706 University), and Dr Ryan McQuinn (Western Sydney University) for insightful discussions.

707 **Author contributions**

708 Xin Hou; Data curation, Software, Formal analysis, Validation, Investigation, Methodology, Writing
709 (original draft), review, and editing, Prepared figures and tables, Performed most experiments.

710 Christopher I Cazzonelli; Conceptualization, Resources, Formal analysis, Supervision, Funding
711 acquisition, Validation, Investigation, Visualization, Methodology, Project administration,
712 Performed preliminary experiments, Figure preparation, Writing, review, and editing, Supervised
713 YA and XH.

714 Ralf Welsch and Julian Koschmieder; Data curation, Formal analysis, Investigation, Methodology,
715 review, and editing, Contributed to Fig. 5.

716 Yagiz Alagoz; Data curation, Formal analysis, Investigation, Methodology, review, and editing,
717 Contributed Fig. 3 and 10.

718 Matthew D Mortimer; Investigation, Methodology, Writing – method for protein modelling, review,
719 and editing, Contributed to Fig. 6.

720 Barry J Pogson; Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition,
721 Validation, Investigation, Visualization, Methodology, Project administration, Writing, review, and
722 editing, Supervised XH.

723 **Conflict of interest**

724 The authors have no conflict of interest to declare.

725 **Funding**

726 B.J.P. and M.D.M were supported by an Australian Research Council Australian Laureate Fellowship
727 (FL190100056) and the Australian Research Council Training Centre for Accelerated Future Crops
728 Development (IC210100047).

729 **Data availability**

730 All data supporting the findings of this study are available within the paper and within its
731 supplementary materials published online. The Arabidopsis Genome Initiative locus number for
732 the major gene discussed in this article is as following: *PSY* (At5g17230). The model of *PSY* gene is
733 drawn from At5g17230.1.

734

REFERENCES

Alagoz Y, Dhami N, Mitchell C, Cazzonelli CI. 2020. cis/trans Carotenoid Extraction, Purification, Detection, Quantification, and Profiling in Plant Tissues. Methods Mol Biol **2083**, 145-163.

Alagoz Y, Nayak P, Dhami N, Cazzonelli CI. 2018. cis-carotene biosynthesis, evolution and regulation in plants: The emergence of novel signaling metabolites. Arch Biochem Biophys **654**, 172-184.

Alvarez D, Voss B, Maass D, Wuest F, Schaub P, Beyer P, Welsch R. 2016. Carotenogenesis Is Regulated by 5 ' UTR-Mediated Translation of Phytoene Synthase Splice Variants. Plant Physiology **172**, 2314-2326.

Anwar S, Brenya E, Alagoz Y, Cazzonelli CI. 2021. Epigenetic Control of Carotenogenesis During

Plant Development. Critical Reviews in Plant Sciences **40**, 23-48.
Anwar S, Nayak J, Alagoz Y, Wojtalewicz D, Cazzonelli CI. 2022. Purification and use of carotenoid standards to quantify cis-trans geometrical carotenoid isomers in plant tissues. In: Wurtzel ET, ed. *Methods in Enzymology Carotenoids: Carotenoid and apocarotenoid analysis*: Elsevier.
Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW, McCourt P, Guttman DS. 2011. Next-generation mapping of Arabidopsis genes. Plant J **67**, 715-725.
Avendano-Vazquez AO, Cordoba E, Llamas E, San Roman C, Nisar N, De la Torre S, Ramos-Vega M, Gutierrez-Nava MD, Cazzonelli CI, Pogson BJ, Leon P. 2014. An Uncharacterized Apocarotenoid-Derived Signal Generated in zeta-Carotene Desaturase Mutants Regulates Leaf Development and the Expression of Chloroplast and Nuclear Genes in Arabidopsis. Plant Cell **26**, 2524-2537.
Baranski R, Cazzonelli CI. 2016. Carotenoid biosynthesis and regulation in plants. In: Kaczor A, Baranska M, eds. *Carotenoids: Nutrition, Analysis and Technology*. Hoboken: Wiley-Blackwell, 161-189.
Camara B. 1993. Plant Phytoene Synthase Complex - Component Enzymes, Immunology, and Biogenesis. Methods in Enzymology **214**, 352-365.
Cao H, Zhang J, Xu J, Ye J, Yun Z, Xu Q, Deng X. 2012. Comprehending crystalline beta-carotene accumulation by comparing engineered cell models and the natural carotenoid-rich system of citrus. J Exp Bot **63**, 4403-4417.
Cazzonelli CI. 2011. Carotenoids in nature: insights from plants and beyond. Functional Plant Biology **38**, 833-847.
Cazzonelli CI, Hou X, Alagoz Y, Rivers J, Dhami N, Lee J, Marri S, Pogson BJ. 2020. A cis-carotene derived apocarotenoid regulates etioplast and chloroplast development. Elife **9**, e45310.
Cazzonelli CI, Nisar N, Hussain D, Carmody ME, Pogson BJ. 2010a. Biosynthesis and Regulation of Carotenoids in Plants-Micronutrients, Vitamins and Health Benefits. Plant Developmental Biology: Biotechnological Perspectives Vol 2, 117-137.
Cazzonelli CI, Nisar N, Roberts AC, Murray KD, Borevitz JO, Pogson BJ. 2014. A chromatin modifying enzyme, SDG8, is involved in morphological, gene expression, and epigenetic responses to mechanical stimulation. Front Plant Sci **5**, 533.
Cazzonelli CI, Roberts AC, Carmody ME, Pogson BJ. 2010b. Transcriptional control of SET DOMAIN GROUP 8 and CAROTENOID ISOMERASE during Arabidopsis development. Mol Plant **3**, 174-191.
Chayut N, Yuan H, Ohali S, Meir A, Sa'ar U, Tzuri G, Zheng Y, Mazourek M, Gepstein S, Zhou X, Portnoy V, Lewinsohn E, Schaffer AA, Katzir N, Fei Z, Welsch R, Li L, Burger J, Tadmor Y. 2017. Distinct Mechanisms of the ORANGE Protein in Controlling Carotenoid Flux. Plant Physiol **173**, 376-389.
Cheminant S, Wild M, Bouvier F, Pelletier S, Renou JP, Erhardt M, Hayes S, Terry MJ, Genschik P, Achard P. 2011. DELLAs regulate chlorophyll and carotenoid biosynthesis to prevent photooxidative damage during seedling deetiolation in Arabidopsis. Plant Cell **23**, 1849-1860.
Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16**, 735-743.
Cunningham FX, Gantt E. 1998. Genes and Enzymes of Carotenoid Biosynthesis in Plants. Annu Rev Plant Physiol Plant Mol Biol **49**, 557-583.
Cunningham FX, Jr., Gantt E. 2007. A portfolio of plasmids for identification and analysis of carotenoid pathway enzymes: Adonis aestivalis as a case study. Photosynth Res **92**, 245-259.
Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol **139**, 5-17.
Dhami N, Pogson BJ, Tissue DT, Cazzonelli CI. 2022. A foliar pigment-based bioassay for interrogating chloroplast signalling revealed that carotenoid isomerisation regulates chlorophyll

abundance. Plant Methods **18**, 18.
Dong J, Tang D, Gao Z, Yu R, Li K, He H, Terzaghi W, Deng XW, Chen H. 2014. Arabidopsis DE-ETIOLATED1 represses photomorphogenesis by positively regulating phytochrome-interacting factors in the dark. Plant Cell **26**, 3630-3645.
Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS. 2006. Gateway-compatible vectors for plant functional genomics and proteomics. Plant J **45**, 616-629.
Escobar-Tovar L, Sierra J, Hernández-Muñoz A, McQuinn RP, Mathioni S, Cordoba E, Colas des Francs-Small C, Meyers BC, Pogson B, León P. 2021. Deconvoluting apocarotenoid-mediated retrograde signaling networks regulating plastid translation and leaf development. Plant J **105**, 1582-1599.
Fraser PD, Enfissi EM, Halket JM, Truesdale MR, Yu D, Gerrish C, Bramley PM. 2007. Manipulation of phytoene levels in tomato fruit: effects on isoprenoids, plastids, and intermediary metabolism. Plant Cell **19**, 3194-3211.
Fraser PD, Schuch W, Bramley PM. 2000. Phytoene synthase from tomato (Lycopersicon esculentum) chloroplasts - partial purification and biochemical properties. Planta **211**, 361-369.
Hartwig B, James GV, Konrad K, Schneeberger K, Turck F. 2012. Fast isogenic mapping-by-sequencing of ethyl methanesulfonate-induced mutant bulks. Plant Physiol **160**, 591-600.
Hou X, Rivers J, Leon P, McQuinn RP, Pogson BJ. 2016. Synthesis and Function of Apocarotenoid Signals in Plants. Trends Plant Sci **21**, 792-803.
Job N, Datta S. 2020. PIF3/HY5 module regulates BBX11 to suppress protochlorophyllide levels in dark and promote photomorphogenesis in light. New Phytol.
Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature **596**, 583-589.
Kachanovsky DE, Filler S, Isaacson T, Hirschberg J. 2012. Epistasis in tomato color mutations involves regulation of phytoene synthase 1 expression by cis-carotenoids. Proceedings of the National Academy of Sciences of the United States of America **109**, 19021-19026.
Lau OS, Deng XW. 2012. The photomorphogenic repressors COP1 and DET1: 20 years later. Trends Plant Sci **17**, 584-593.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25**, 1754-1760.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**, 2078-2079.
Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, dos Santos Costa A, Fazel-Zarandi M, Sercu T, Candido S, Rives A. 2023. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science **379**, 1123-1130.
Lindgreen S. 2012. AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res Notes **5**, 337.
Llorente B, Martinez-Garcia JF, Stange C, Rodriguez-Concepcion M. 2017. Illuminating colors: regulation of carotenoid biosynthesis and accumulation by light. Curr Opin Plant Biol **37**, 49-55.
Maass D, Arango J, Wust F, Beyer P, Welsch R. 2009. Carotenoid crystal formation in Arabidopsis and carrot roots caused by increased phytoene synthase protein levels. PLOS ONE **4**, e6373.
Mariani V, Kiefer F, Schmidt T, Haas J, Schwede T. 2011. Assessment of template based protein structure predictions in CASP9. Proteins **79 Suppl 10**, 37-58.
Mathur J, Koncz C. 1998. Callus culture and regeneration. In: Martinez-Zapater J, Salinas J, eds. *Arabidopsis protocols: Methods in molecular biology*, Vol. 82. Totowa: Humana Press, Inc, 31-34.

McQuinn RP, Leroux J, Sierra J, Escobar-Tovar L, Frusciante S, Finnegan EJ, Diretto G, Giuliano G, Giovannoni JJ, Leon P, Pogson BJ. 2023. Deregulation of zeta-carotene desaturase in Arabidopsis and tomato exposes a unique carotenoid-derived redundant regulation of floral meristem identity and function. Plant J **114**, 783-804.

Mechin V, Damerval C, Zivy M. 2007. Total protein extraction with TCA-acetone. Methods Mol Biol **355**, 1-8.

Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. Nat Methods **19**, 679-682.

Moreno JC, Mi J, Alagoz Y, Al-Babili S. 2021. Plant apocarotenoids: from retrograde signaling to interspecific communication. Plant J **105**, 351-375.

Nayak JJ, Anwar S, Krishna P, Chen ZH, Plett JM, Foo E, Cazzonelli CI. 2022. Tangerine tomato roots show increased accumulation of acyclic carotenoids, less abscisic acid, drought sensitivity, and impaired endomycorrhizal colonization. Plant Sci **321**, 111308.

Ossowski S, Schneeberger K, Clark RM, Lanz C, Warthmann N, Weigel D. 2008. Sequencing of natural strains of Arabidopsis thaliana with short reads. Genome Research **18**, 2024-2033.

Osterlund MT, Hardtke CS, Wei N, Deng XW. 2000. Targeted destabilization of HY5 during light-regulated development of Arabidopsis. Nature **405**, 462-466.

Pandit J, Danley DE, Schulte GK, Mazzalupo S, Pauly TA, Hayward CM, Hamanaka ES, Thompson JF, Harwood HJ, Jr. 2000. Crystal structure of human squalene synthase. A key enzyme in cholesterol biosynthesis. J Biol Chem **275**, 30610-30617.

Park H, Kreunen SS, Cuttriss AJ, DellaPenna D, Pogson BJ. 2002. Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis. Plant Cell **14**, 321-332.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res **29**, e45.

Porra RJ, Thompson WA, Kriedemann PE. 1989. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim. Biophys. Acta **975**, 384-394.

Rodriguez-Villalon A, Gas E, Rodriguez-Concepcion M. 2009. Colors in the dark: a model for the regulation of carotenoid biosynthesis in etioplasts. Plant Signal Behav **4**, 965-967.

Ruiz-Sola MA, Coman D, Beck G, Barja MV, Colinas M, Graf A, Welsch R, Rutimann P, Buhlmann P, Bigler L, Gruissem W, Rodriguez-Concepcion M, Vranova E. 2016. Arabidopsis GERANYLGERANYL DIPHOSPHATE SYNTHASE 11 is a hub isozyme required for the production of most photosynthesis-related isoprenoids. New Phytol **209**, 252-264.

Schaub P, Rodriguez-Franco M, Cazzonelli CI, Álvarez D, Wüst F, Welsch R. 2018. Establishment of an Arabidopsis callus system to study the interrelations of biosynthesis, degradation and accumulation of carotenoids. PLOS ONE **13**, e0192158.

Schneeberger K, Ossowski S, Lanz C, Juul T, Petersen AH, Nielsen KL, Jorgensen JE, Weigel D, Andersen SU. 2009. SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nat Methods **6**, 550-551.

Schrödinger LLC. 2015. The PyMOL Molecular Graphics System.

Shumskaya M, Bradbury LM, Monaco RR, Wurtzel ET. 2012. Plastid localization of the key carotenoid enzyme phytoene synthase is altered by isozyme, allelic variation, and activity. Plant Cell **24**, 3725-3741.

Shumskaya M, Wurtzel ET. 2013. The carotenoid biosynthetic pathway: thinking in all dimensions. Plant Sci **208**, 58-63.

Stephenson PG, Fankhauser C, Terry MJ. 2009. PIF3 is a repressor of chloroplast development. Proc Natl Acad Sci U S A **106**, 7654-7659.

Toledo-Ortiz G, Huq E, Rodriguez-Concepcion M. 2010. Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. Proc Natl Acad Sci U S A **107**, 11626-11631.

Toledo-Ortiz G, Johansson H, Lee KP, Bou-Torrent J, Stewart K, Steel G, Rodríguez-Concepción M, Halliday KJ. 2014. The HY5-PIF Regulatory Module Coordinates Light and Temperature Control of Photosynthetic Gene Transcription. PLOS Genetics **10**, e1004416.

Welsch R, Arango J, Bar C, Salazar B, Al-Babili S, Beltran J, Chavarriaga P, Ceballos H, Tohme J, Beyer P. 2010. Provitamin A accumulation in cassava (Manihot esculenta) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. Plant Cell **22**, 3348-3356.

Welsch R, Beyer P, Hugueney P, Kleinig H, von Lintig J. 2000. Regulation and activation of phytoene synthase, a key enzyme in carotenoid biosynthesis, during photomorphogenesis. Planta **211**, 846-854.

Welsch R, Wust F, Bar C, Al-Babili S, Beyer P. 2008. A third phytoene synthase is devoted to abiotic stress-induced abscisic acid formation in rice and defines functional diversification of phytoene synthase genes. Plant Physiology **147**, 367-380.

Welsch R, Zhou X, Yuan H, Álvarez D, Sun T, Schlossarek D, Yang Y, Shen G, Zhang H, Rodriguez-Concepcion M, Thannhauser TW, Li L. 2018. Clp Protease and OR Directly Control the Proteostasis of Phytoene Synthase, the Crucial Enzyme for Carotenoid Biosynthesis in Arabidopsis. Molecular Plant **11**, 149-162.

Xu X, Chi W, Sun X, Feng P, Guo H, Li J, Lin R, Lu C, Wang H, Leister D, Zhang L. 2016. Convergence of light and chloroplast signals for de-etiolation through ABI4-HY5 and COP1. Nat Plants **2**, 16066.

Zhou X, Rao S, Wrightstone E, Sun T, Lui ACW, Welsch R, Li L. 2022. Phytoene Synthase: The Key Rate-Limiting Enzyme of Carotenoid Biosynthesis in Plants. Front Plant Sci **13**, 884720.

Zhou X, Welsch R, Yang Y, Alvarez D, Riediger M, Yuan H, Fish T, Liu J, Thannhauser TW, Li L. 2015. Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. Proc Natl Acad Sci U S A **112**, 3558-3563.

FIGURE LEGENDS

Fig. 1: PSY variants restore greening and chlorophyll levels in *ccr2* leaves and cotyledons as well as PLB formation in etiolated seedlings. A) Three-week-old WT, *ccr2*, and r*ccr2* mutant variant (r*ccr2*$^{-4}$, r*ccr2*$^{-90}$, r*ccr2*$^{-130}$, r*ccr2*$^{-145}$) plants were grown in multiple independent experiments under a 10-h photoperiod and a representative image from each genotype is displayed. **B**) Total chlorophyll in leaf tissues from WT, *ccr2* and *rccr2* plants growing under 8-h photoperiod. Error bars denote the standard error of the mean (n=5). **C**) Total chlorophyll levels in cotyledons from de-etiolated seedlings (grown in darkness for 4 d) exposed to continuous white light for 3 days and quantified at 0, 24, 48 and 72 h post illumination Error bars denote standard error of the mean (n=20 seedlings). *R*; *rccr2*. Stars denote significance by comparing *rccr2* lines to WT using a one-way ANOVA ($p < 0.05$). **D**) Transmission electron microscopy (TEM) images of representative etioplasts from 5-d old etiolated WT, *ccr2*, and r*ccr2* mutant variant (r*ccr2*$^{-4}$, r*ccr2*$^{-90}$, r*ccr2*$^{-130}$, r*ccr2*$^{-145}$) seedlings. Images are representatives of > 15 plastids from > 5 TEM sections. **E**) Ratios of etioplasts containing PLBs. Letters denote post-hoc Tukey groups following a one-way ANOVA test.

Fig. 2: Sequencing of *rccr2* lines identifies mutations within the *PSY* gene that retains protein function. A) Schematic structure of wild-type *PSY* DNA, exon structure of the protein and position of the amino acid variant in PSY for r*ccr2*$^{-4}$, r*ccr2*$^{-90}$, and r*ccr2*$^{-130}$. UTR; untranslated leader region, aa; amino acid, P178S; proline at position 178 to serine (r*ccr2*$^{-130}$), M266I; methionine 266 to isoleucine (r*ccr2*$^{-4}$), A352T; alanine 352 to threonine (r*ccr2*$^{-90}$). **B**) Sequencing of *PSY* genomic DNA from leaves of multiple independent plants highlights the mutation position in the four *ccr2 psy*

lines.

Fig. 3: *cis*-carotene and PSY protein levels in WT, *ccr2*, *ccr2 psy*, and *PSY* overexpression lines. A) Protein western blot confirming the overexpression of PSY protein in the four *ccr2 psy* double mutants (-) and harboring a transgene overexpressing *PSY* (+). For each sample, 10 µg of total protein extracted from leaf tissues of 4-week-old Arabidopsis plants and the western blots were probed using anti-PSY polyclonal antibody, stripped and re-probed with the ACTIN housekeeper protein. **B**) Total linear *cis*-carotene levels in etiolated (7-d) WT, *ccr2*, and *ccr2 psy* double mutants (-) and harboring a transgene overexpressing *PSY* (+). Error bars denote the standard error of means (n=3), and stars denote significant differences in comparison to *ccr2* ($P < 0.05$; one-way ANOVA). **C**) Linear *cis*-carotene levels in 7-d old etiolated tissues from WT, *ccr2* and *ccr2 psy* double mutants. Error bars denote the standard error of means (n=3). tc-ζ-c: tri-*cis*-ζ-carotene; dc-ζ-c: di-*cis*-ζ-carotene; P-Neu: tri-*cis*-neurosporene; P-Lyc: tetra-*cis*-lycopene; 15c-Phy: 15 *cis*-phytoene; 9,15dc-Phf: 9, 15 di-*cis*-phytofluene.

Fig. 4: Transcript and protein levels of HY5 and PIF3 in WT, *ccr2* and *ccr2 psy*. A) Transcript levels of *PIF3* and *HY5* in etiolated 7-d seedling tissues from WT, *ccr2*, *ccr2 psy-4*, *ccr2 psy-90*, *ccr2 psy-130* and *ccr2 psy-145*. Error bars denote the standard error of the mean (n=3). Stars denote significance by comparing mutant lines to WT using a one-way ANOVA ($p < 0.05$). **B**) Representative western blot images showing protein levels of PIF3 and HY5 in WT, *ccr2* and the *ccr2 psy* etiolated seedling tissues. For each sample, five micrograms of total protein was loaded to separate gels, blotted with PIF or HY5 antibodies and membrane re-probed using anti-Actin antibody as an internal loading control.

Fig. 5: In vivo/vitro enzymatic assays of PSY activity and carotenoid levels in WT, *ccr2* and *ccr2 psy*. A) *In vivo* PSY activity assay using seed-derived callus (SDC) treated with 1 $\mu mol.L^{-1}$ norflurazon. Norflurazon inhibits PDS and leads to the accumulation of phytoene, providing a measure of PSY activity. **B**) *In vitro* activity of endogenous PSY (or mutant) from WT, *ccr2* and *ccr2 psy* double mutants. PSY was purified from mature leaves by co-immunoprecipitation, and 5 µg was used for enzymatic activity assays measuring phytoene levels. **C**) *In vitro* enzyme activity assay showing phytoene and recombinant PSY protein levels purified from *E. coli* cells expressing WT, *psy-4*, *psy-90*, *psy-130* and *psy-145* variants. Untransformed *E. coli* BL21 (DE3) cells were used as a negative control. Western blots showing anti-PSY and anti-6× Histidine (His) levels to confirm expression and purification of recombinant PSY. 5 µg of protein was loaded to the gel. Data is shown as mean ± standard errors of three biological replicates in all measurements. *Difference is significant compared to WT ($P < 0.05$ in one-way ANOVA).

Fig. 6: Structural mapping of amino acid substitutions in a 3D PSY structural model. Cartoon representation of a 3D structural model representation built using Colabfold implementation of Alphafold2 and ESMFold. Sub-panels show mutations with different angles and features for clarity. PyMOL cavity settings are Cavities and Pockets (Culled), Cavity Detection Radius at seven angstroms, and Cavity Detection Cutoff at three solvent radii for clarity. **A**) Structural model showing the mutations in ball-representation and putative metal-binding aspartate residues in stick representation. **B**) P178S mutation (shown in ball-representation) in psy-130 and putative metal-binding aspartate residues (shown in stick-representation). This mutation may impact the metal binding by D172 and D176 located in the same hairpin, and that by D302 which is close to P178 in the predicted 3D structure. **C**) A352T (upper left) in psy-90 and M266I (lower right) in psy-4 are shown in ball-representation. Substrate pocket (centre) and voids are shown in tan colour. The A352T substitution may affect substrate binding and the residue's steric interaction with its

adjacent α-helix. The M266I mutation is close to the bottom segment of the substrate pocket.

Fig. 7: PSY amino acid substitutions alter AtPSY protein-protein interaction with AtOR and AtGGPPS11. A) Protein-protein interaction assays using split ubiquitin system. *AtPSY* gene variants were fused to the C-terminal ubiquitin moiety (Cub) and co-expressed with AtOR (OR) or AtGGPPS11 (G11) fused to the N-terminal ubiquitin moiety (Nub). Yeasts were spotted onto either nonselective (-LW) or fully selective medium (-LWAH) in 10-fold dilution series. In order to reduce background activation of reporter genes and visualize different interaction strengths, methionine (150 µmol.L^{-1} and 1 mmol.L^{-1}) was added to the media reducing expression of Cub fusion proteins (+150 µM M and +1mM M, respectively). Control combinations with empty Cub expressing vectors are included below. **B**) β-Galactosidase activity of yeast strains co-expressing different PSY versions fused to Cub and AtOR or ArGGPPS11 fused with Nub or Nub only (Control), respectively. Enzyme activity was determined by oNPG assay and is given in nmol oNP.min^{-1}.OD_{600}^{-1}. Error bars indicate standard error of means (n=3). Stars denote significant differences ($p < 0.05$) in comparison to WT (ANOVA). **C**) Co-immunoprecipitation analysis of AtPSY and AtOR proteins. Equal amounts of total protein extracted from each leaf sample were immunoprecipitated using an anti-PSY antibody or normal rabbit IgG as a negative control. Ten microliters of immunoprecipitate was subjected to western blot using anti-OR antibody. Half of the protein before immunoprecipitation was kept for input and subjected to western blotting with anti-PSY and anti-OR antibodies. Experiments were performed in triplicate and representative results were displayed.

Fig. 8: Western blot analysis of PSY and OR protein levels in *ccr2 psy* variants. Ten micrograms of total protein extracted from leaf tissue of WT, *ccr2* or *ccr2 psy* variants was used for western blot analysis to quantify endogenous AtPSY and AtOR protein levels using anti-PSY and anti-OR antibodies, respectively. Actin was included as an internal reference and detected using anti-Actin antibody. The blot shown is representative of three western blot replicates.

Fig. 9: PSY levels in PLB and stroma fractions of Arabidopsis etioplasts before and after the expression of recombinant proteins. Protoplasts were isolated from the cotyledons of etiolated 7-d old seedlings and incubated in the dark with a PSY-CGFP plasmid for 16 h. Ten microliters of PLB or stroma fraction of etioplasts was subjected to western blot using anti-PSY antibody. Ct: control with no expression of recombinant PSY.

Fig. 10. Manipulation of PSY protein activity provides a central target to modify a linear *cis*-carotene-derived apocarotenoid signal (*cis*-ACS) that modulates plastid biogenesis. Variants in PSY activity affect splicing (alternative leading to truncated protein), enzyme activity (enzyme-substrate binding), and protein-protein interactions (multi-enzyme complex binding between PSY, OR, and/or GGPS11) impacting upon protein levels that constitute dynamic regulations to fine-tune the level of a *cis*-ACS without impacting overall carotenoid end-product accumulation. *cis*-ACS up-regulates PIF3 and suppresses HY5 protein levels that fine-tune PSY control over *cis*-ACS biosynthesis. CRTISO activity and longer photoperiods (i.e., efficient photoisomerization) reduce *cis*-carotene biosynthesis and the Arabidopsis *crtiso* loss-in-function *ccr2* mutant causes *cis*-carotene accumulation that triggers *cis*-ACS to block prolamellar body (PLB) formation in etioplasts and causes psudeochloroplasts to form when mutants are grown under shorter photoperiods (Cazzonelli *et al.*, 2020).

SUPPLEMENTAL FIGURES

Supplemental Fig. 1: Simplified diagram of the carotenoid biosynthesis pathway. Genes encoding the catalysing enzymes of each step before all-*trans*-lycopene are labelled to the left of the pathway and mutants to the right (red line indicates blocked step). GGPP: geranylgeranyl diphosphate; PSY: phytoene synthase; phytoene: *cis*-phytoene; phytofluene: di-*cis*-phytofluene; PDS: phytoene desaturase; ZISO: 15-*cis*-ζ-carotene isomerase; ZDS: ζ-carotene desaturase; CRTISO: carotenoid isomerase. Norflurazon is a herbicide that inhibits PDS activity. *cis*-ACS: linear *cis*-carotene derived apocarotenoid signal (ACS).

Supplemental Fig. 2: Percentage of leaf virescence leaves in *rccr2* lines reverted the leaf-yellowing phenotype in *ccr2* and led to albino phenotype in one line. A) Percentage of yellow leaf area of WT, *ccr2* and *rccr2* (r) lines. Plants were grown under 8-h photoperiod. Star denotes a significant difference compared to WT ($p < 0.05$ in one-way ANOVA).

Supplemental Fig. 3: Characterisation of *ccr2 psy-145* and alternative splicing of *PSY*. A) Albino seedlings displayed by *rccr2*$^{-145}$ *(ccr2 psy-145* - red circle) seedlings were grown under a 16-h photoperiod. **B**) Reverse transcription PCR (RT-PCR) showing alternative splicing of intron 3 in *ccr2 psy*$^{-145}$ leaf tissues. Exons 2 and 3 were partially amplified, and products were visualised on a 1% agarose gel. Spliced shows 221 bp (WT and *ccr2*) and unspliced 762 bp (*ccr2 psy*$^{-145}$) amplicons. 100 bp ladder is displayed. **C**) Quantitative RT-PCR (qRT-PCR) of *PSY* mRNA amplifying part of intron 3 (162 bp) in *ccr2 psy*$^{-145}$ (+) or a region (155bp) spanning exon 2 and 3 when intron 3 is spliced out in WT and *ccr2* leaf tissues. Standard error bars are shown (n=10). RNA levels were normalized to Protein Phosphatase 2A (AT1G13320) reference gene validated for mRNA normalisation in leaves using Cyclophilin (At2g29960) and TIP41 (At4g34270) secondary reference genes (Cazzonelli *et al.*, 2014; Cazzonelli *et al.*, 2010b).

Supplemental Fig. 4: Overexpression of functional *PSY* restores *ccr2* mutant phenotypes in *ccr2 psy* double mutant lines. A) Representative images of rosettes show yellow virescence in newly emerged leaves from *ccr2* and *ccr2 psy* lines compared to the WT control and WT::*PSY*-OE. T4 generation transgenic plants were grown under a 10-h photoperiod for 3 weeks, and images from 50-100 plants were analysed for five independent lines. *PSY*-OE; *PSY* overexpression. **B**) Total chlorophyll in leaves from T4 generation transgenic plants. Plants were grown under an 8-h photoperiod. Error bars denote the standard error of means (n=5). Star denotes a significant difference compared to WT-OE ($p < 0.05$ in one-way ANOVA). YL; yellow leaf.

Supplemental Fig. 5: Split ubiquitin assays showing phytoene levels generated in yeast cells expressing PSY+AtGGPPS11 combinations. Phytoene levels were measured using HPLC. Error bars indicate standard error of means (n=3). Star denotes a significant difference compared to WT ($p < 0.05$; one-way ANOVA).

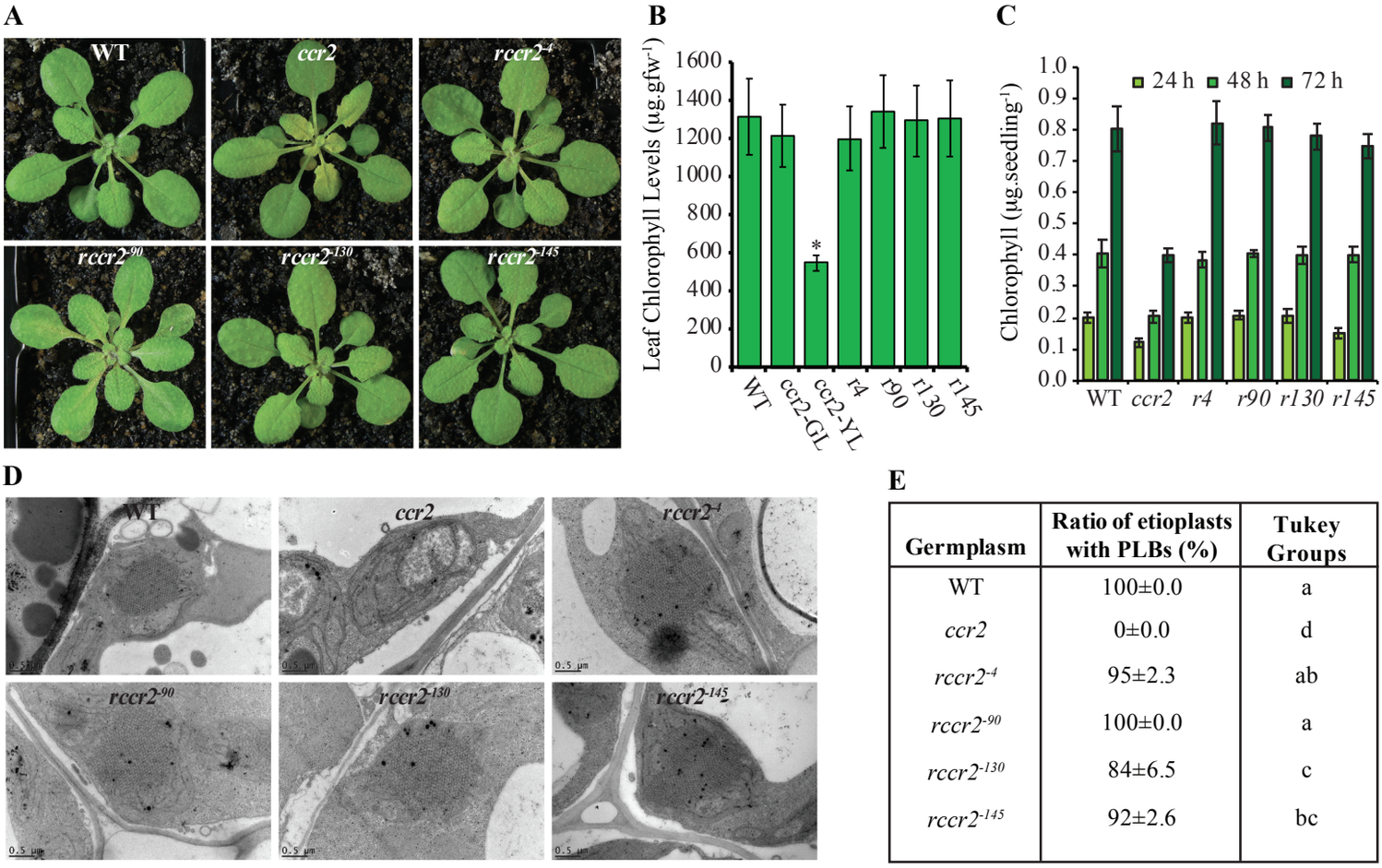
Supplemental Fig. 6: ELISA of PSY and AtOR protein levels. Each well was coated with total protein from Arabidopsis leaf tissue and incubated with anti-PSY or anti-OR antibodies. Following the addition of substrate solution containing O-phenylenediamine dihydrochloride and H_2O_2 absorbances were measured at 492 nm. Values were averaged from five biological replicates, and error bars denote the standard error of means (n=5). Stars denote the significant difference compared to WT ($p < 0.05$; one-way ANOVA).

SUPPLEMENTAL TABLES

Supplemental Table 1: Primers used in this study

Supplemental Table 2: Genomic information of the mutations identified in *PSY* from *rccr2* lines.

SNP; single nucleotide polymorphism.



E

Germplasm	Ratio of etioplasts with PLBs (%)	Tukey Groups
WT	100±0.0	a
ccr2	0±0.0	d
rccr2-4	95±2.3	ab
rccr2-90	100±0.0	a
rccr2-130	84±6.5	c
rccr2-145	92±2.6	bc

Fig. 1: PSY variants restore greening and chlorophyll levels in ccr2 leaves and cotyledons as well as PLB formation in etiolated seedlings. A) Three-week-old WT, ccr2, and rccr2 mutant variant (rccr2-4, rccr2-90, rccr2-130, rccr2-145) plants were grown in multiple independent experiments under a 10-h photoperiod and a representative image from each genotype is displayed. B) Total chlorophyll in leaf tissues from WT, ccr2 and rccr2 plants growing under 8-h photoperiod. Error bars denote the standard error of the mean (n=5). C) Total chlorophyll levels in cotyledons from de-etiolated seedlings (grown in darkness for 4 d) exposed to continuous white light for 3 days and quantified at 0, 24, 48 and 72 h post illumination Error bars denote standard error of the mean (n=20 seedlings). R; rccr2. Stars denote significance by comparing rccr2 lines to WT using a one-way ANOVA ($p < 0.05$). D) Transmission electron microscopy (TEM) images of representative etioplasts from 5-d old etiolated WT, ccr2, and rccr2 mutant variant (rccr2-4, rccr2-90, rccr2-130, rccr2-145) seedlings. Images are representatives of > 15 plastids from > 5 TEM sections. E) Ratios of etioplasts containing PLBs. Letters denote post-hoc Tukey groups following a one-way ANOVA test.

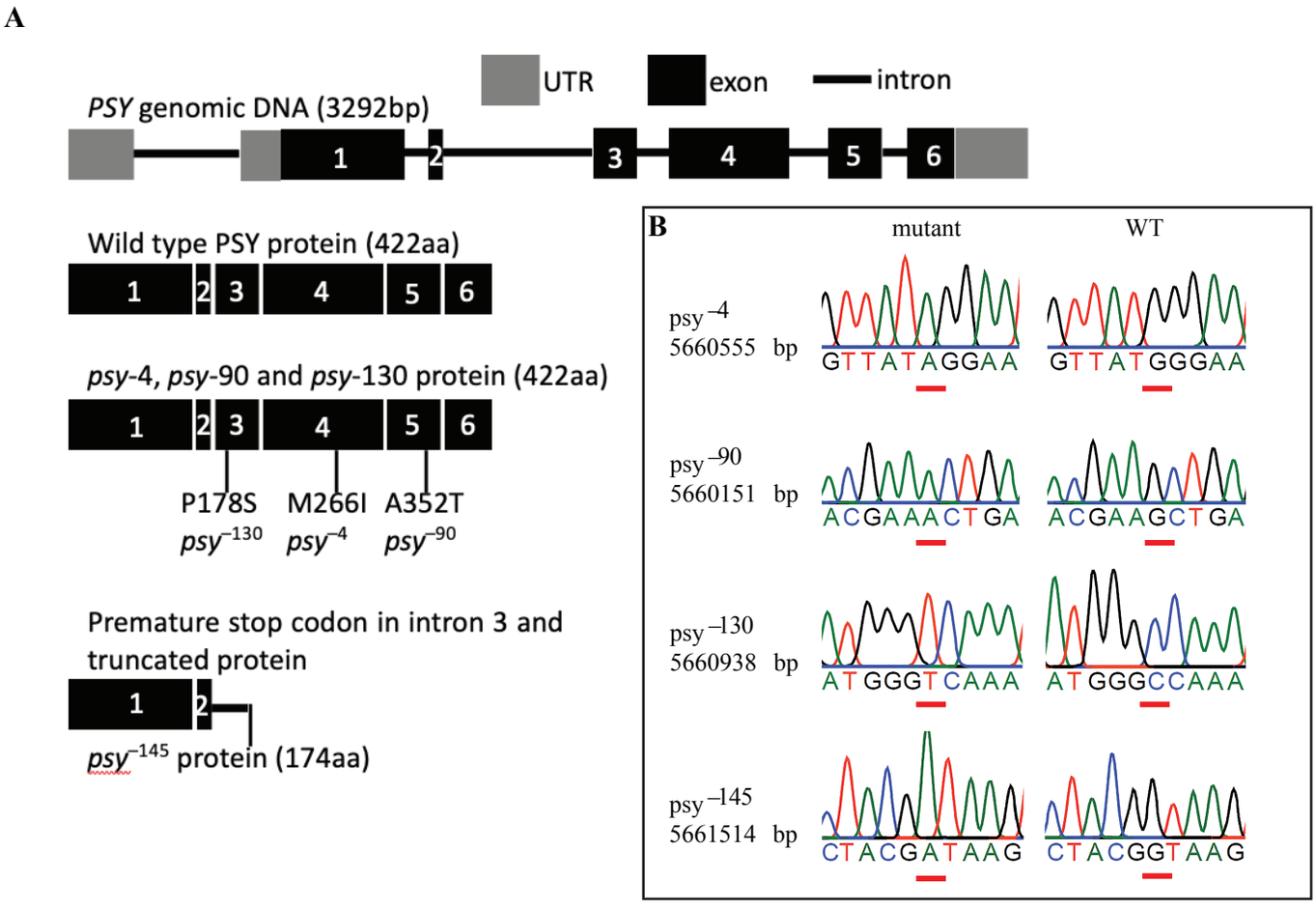


Fig. 2: Sequencing of rccr2 lines identifies mutations within the PSY gene the retains protein function. A) Schematic structure of wild-type PSY DNA, exon structure of the protein and position of the amino acid variant in PSY for rccr2-4, rccr2-90, and rccr2-130. UTR; untranslated leader region, aa; amino acid, P178S; proline at position 178 to serine (rccr2-130), M266I; methionine 266 to isoleucine (rccr2-4), A352T; alanine 352 to threonine (rccr2-90). B) Sequencing of PSY genomic DNA from leaves of multiple independent plants highlights the mutation position in the four ccr2 psy lines.

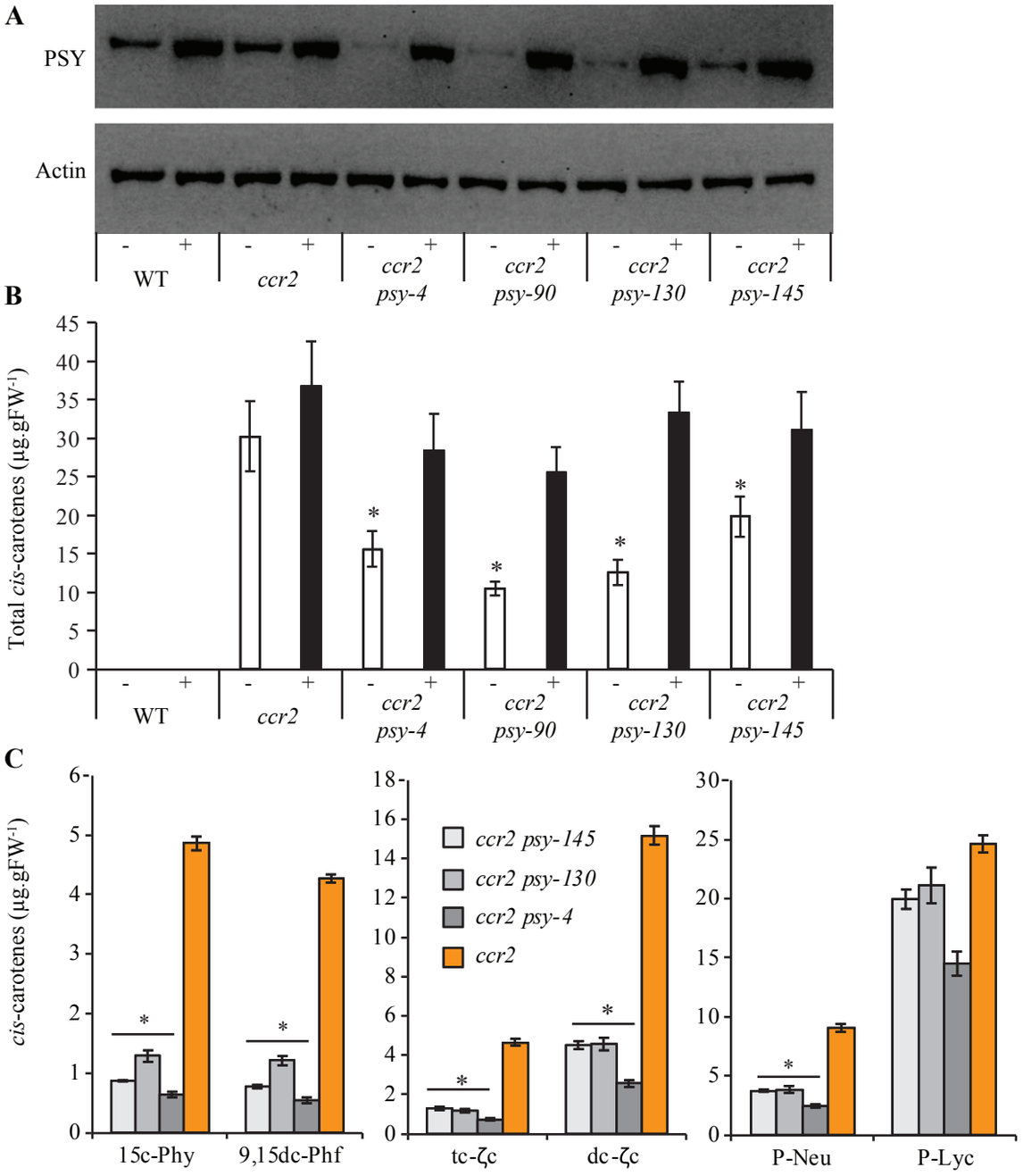


Fig. 3: cis-carotene and PSY protein levels in WT, ccr2, ccr2 psy, and PSY overexpression lines. A) Protein western blot confirming the overexpression of PSY protein in the four ccr2 psy double mutants (-) and harboring a transgene overexpressing PSY (+). For each sample, 10 µg of total protein extracted from leaf tissues of 4-week-old Arabidopsis plants and the western blots were probed using anti-PSY polyclonal antibody, stripped and re-probed with the ACTIN housekeeper protein. B) Total linear cis-carotene levels in etiolated (7-d) WT, ccr2, and ccr2 psy double mutants (-) and harboring a transgene overexpressing PSY (+). Error bars denote the standard error of means (n=3), and stars denote significant differences in comparison to ccr2 ($P < 0.05$; one-way ANOVA). C) Linear cis-carotene levels in 7-d old etiolated tissues from WT, ccr2 and ccr2 psy double mutants. Error bars denote the standard error of means (n=3). tc-ζ-c: tri-cis-ζ-carotene; dc-ζ-c: di-cis-ζ-carotene; P-Neu: tri-cis-neurosporene; P-Lyc: tetra-cis-lycopene; 15c-Phy: 15 cis-phytoene; 9,15dc-Phf: 9, 15 di-cis-phytofluene.

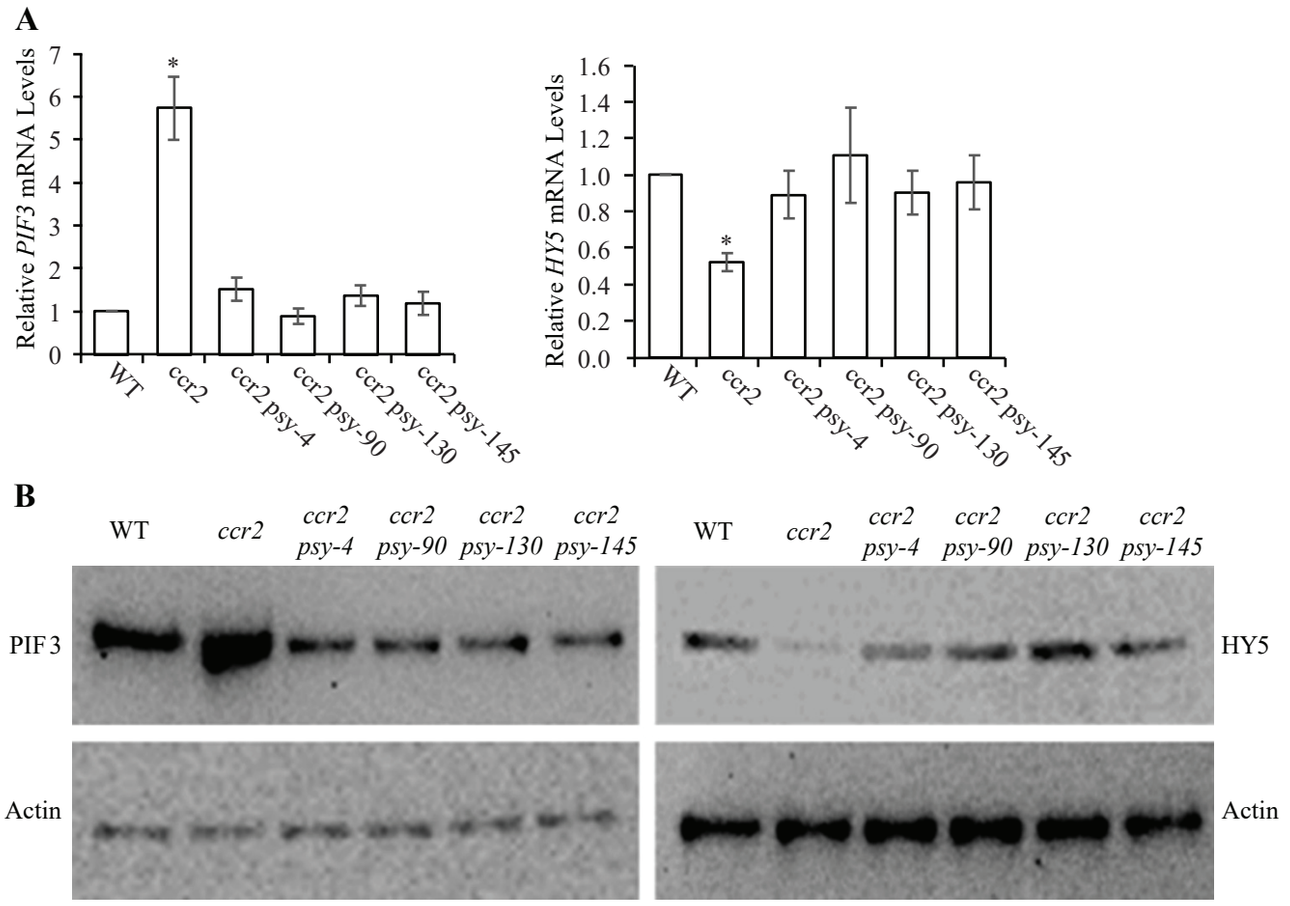


Fig. 4: Transcript and protein levels of HY5 and PIF3 in WT, ccr2 and ccr2 psy. A) Transcript levels of PIF3 and HY5 in etiolated 7-d seedling tissues from WT, ccr2, ccr2 psy-4, ccr2 psy-90, ccr2 psy-130 and ccr2 psy-145. Error bars denote standard error of the mean (n=3). Stars denote significance by comparing mutant lines to WT using a one-way ANOVA ($p < 0.05$). B) Representative western blot images showing protein levels of PIF3 and HY5 in WT, ccr2 and the ccr2 psy etiolated seedling tissues. For each sample, five micrograms of total protein was loaded to separate gels, blotted with PIF or HY5 antibodies and membrane re-probed using anti-Actin antibody as an internal loading control.

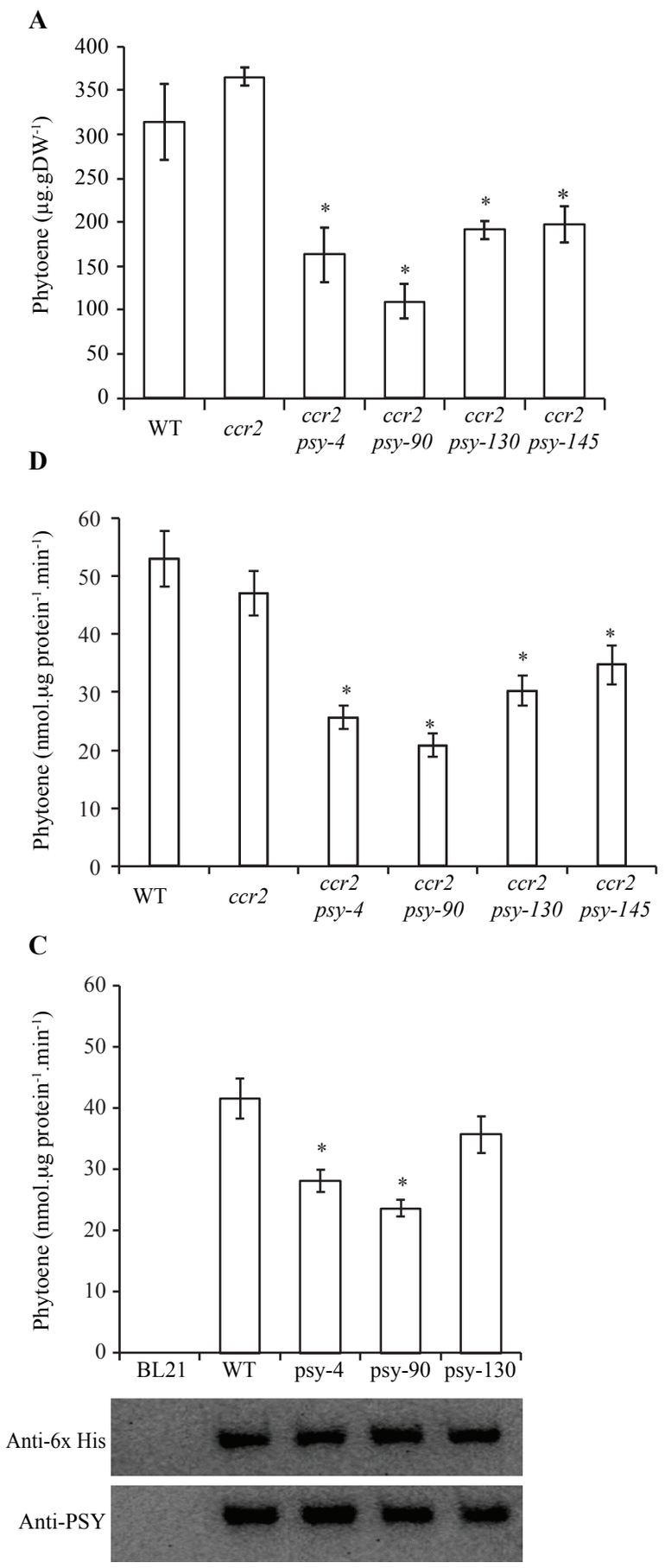


Fig. 5: In vivo/vitro enzymatic assays of PSY activity and carotenoid levels in WT, ccr2 and ccr2 psy. A) In vivo PSY activity assay using seed-derived callus (SDC) treated with 1 µmol.L-1 norflurazon. Norflurazon inhibits PDS and leads to the accumulation of phytoene, providing a measure of PSY activity. B) In vitro activity of endogenous PSY (or mutant) from WT, ccr2 and ccr2 psy double mutants. PSY was purified from mature leaves by co-immunoprecipitation, and 5 µg was used for enzymatic activity assays measuring phytoene levels. C) In vitro enzyme activity assay showing phytoene and recombinant PSY protein levels purified from E. coli cells expressing WT, psy-4, psy-90, psy-130 and psy-145 variants. Untransformed E. coli BL21 (DE3) cells were used as a negative control. Western blots showing anti-PSY and anti-6× Histidine (His) levels to confirm expression and purification of recombinant PSY. 5 µg of protein was loaded to the gel. Data is shown as mean ± standard errors of three biological replicates in all measurements. *Difference is significant compared to WT ($P < 0.05$ in one-way ANOVA).

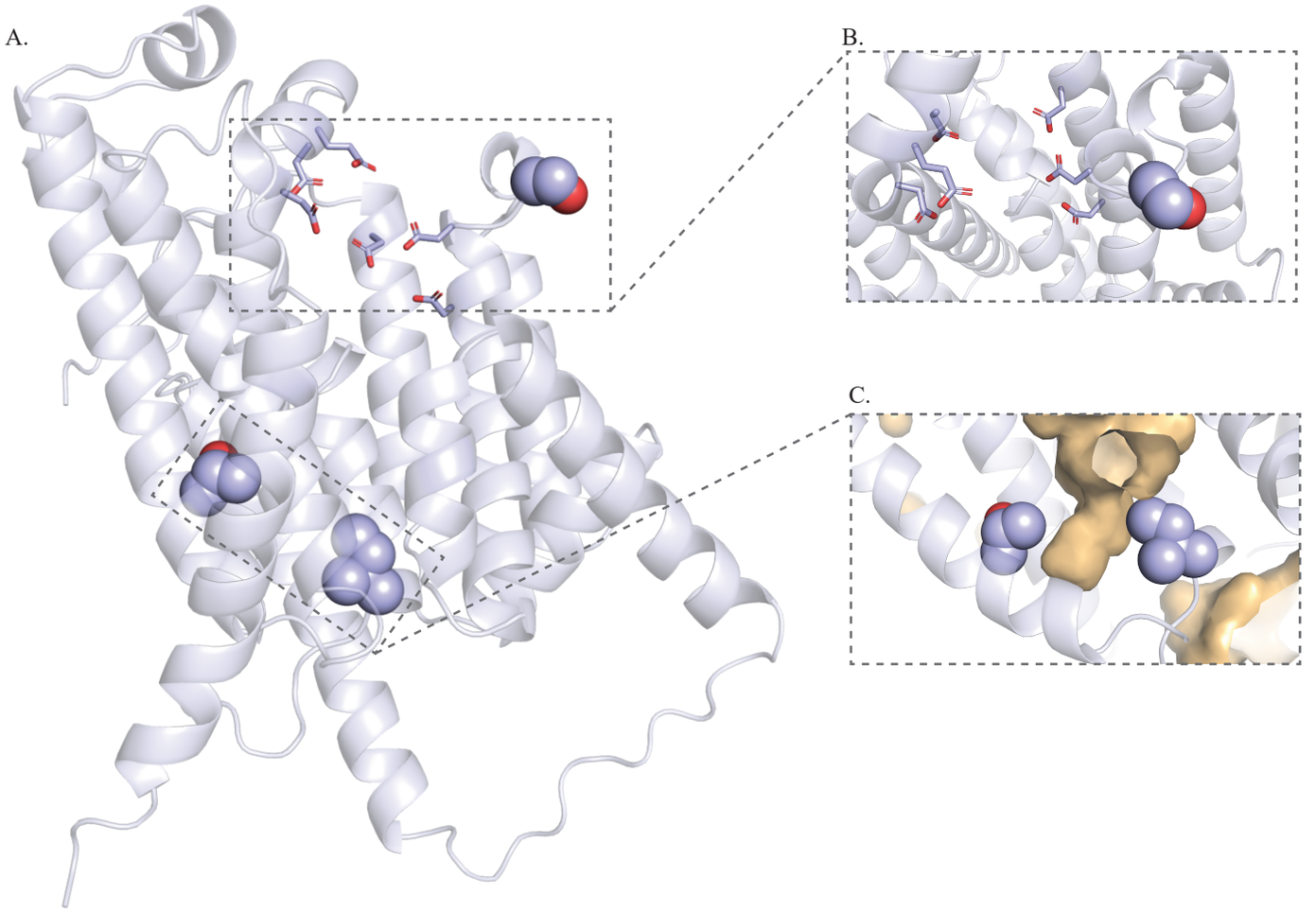


Fig. 6: Structural mapping of amino acid substitutions in a 3D PSY structural model. Cartoon representation of a 3D structural model representation built using Colabfold implementation of Alphafold2 and ESMFold. Sub-panels show mutations with different angles and features for clarity. PyMOL cavity settings are Cavities and Pockets (Culled), Cavity Detection Radius at seven angstroms, and Cavity Detection Cutoff at three solvent radii, for clarity. A) Structural model showing the mutations in ball-representation, and putative metal-binding aspartate residues in stick representation. B) P178S mutation (shown in ball-representation) in psy-130 and putative metal-binding aspartate residues (shown in stick-representation). This mutation may impact the metal binding by D172 and D176 located in the same hairpin, and that by D302 which is close to P178 in the predicted 3D structure. C) A352T (upper left) in psy-90 and M266I (lower right) in psy-4 are shown in ball-representation. Substrate pocket (centre) and voids are shown in tan colour. The A352T substitution may affect substrate binding and the residue's steric interaction with its adjacent α-helix. The M266I mutation is close to the bottom segment of the substrate pocket.

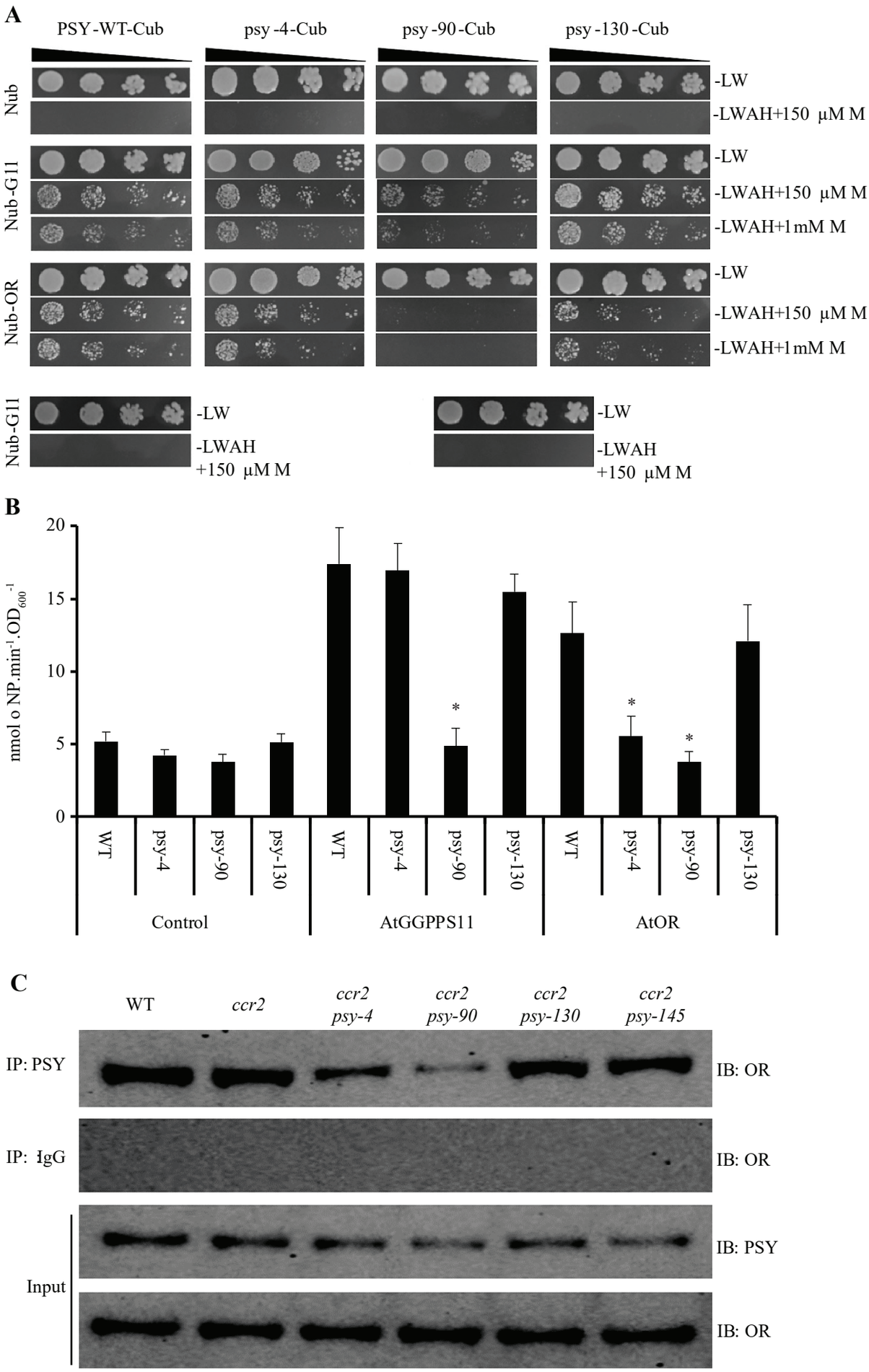


Fig. 7: PSY amino acid substitutions alter AtPSY protein-protein interaction with AtOR and AtGGPPS11. A) Protein-protein interaction assays using split ubiquitin system. AtPSY gene variants were fused to the C-terminal ubiquitin moiety (Cub) and co-expressed with AtOR (OR) or AtGGPPS11 (G11) fused to the N-terminal ubiquitin moiety (Nub). Yeasts were spotted onto either nonselective (-LW) or fully selective medium (-LWAH) in 10-fold dilution series. In order to reduce background activation of reporter genes and visualize different interaction strengths, methionine (150 µmol.L-1 and 1 mmol.L-1) was added to the media reducing expression of Cub fusion proteins (+150 µM M and +1mM M, respectively). Control combinations with empty Cub expressing vector are included below. B) β-Galactosidase activity of yeast strains co-expressing different PSY versions fused to Cub and AtOR or ArGGPPS11 fused with Nub or Nub only (Control), respectively. Enzyme activity was determined by oNPG assay and is given in nmol oNP.min-1.OD600-1. Error bars indicate standard error of means (n=3). Stars denote significant differences ($p < 0.05$) comparison to WT (ANOVA). C) Co-immunoprecipitation analysis of AtPSY and AtOR proteins. Equal amounts of total protein extracted from each leaf sample were immunoprecipitated using anti-PSY antibody or using normal rabbit IgG as a negative control. Ten microliters of immunoprecipitate was subjected to western blot using anti-OR antibody. Half of the protein before immunoprecipitation was kept for input, and subjected to western blotting with anti-PSY and anti-OR antibodies. Experiments were performed in triplicate and representative results displayed.

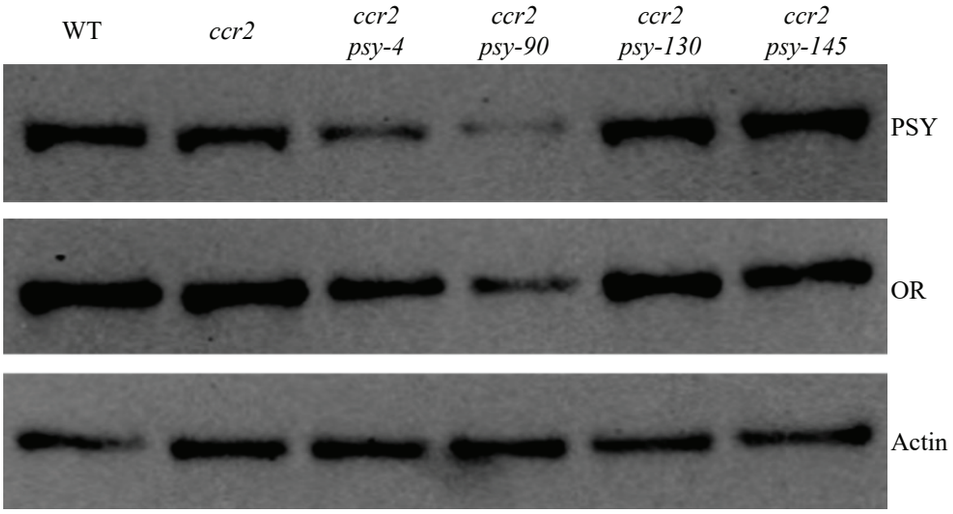


Fig. 8: Western blot analysis of PSY and OR protein levels in ccr2 psy variants. Ten micrograms of total protein extracted from leaf tissue of WT, ccr2 or ccr2 psy variants was used for western blot analysis to quantify endogenous AtPSY and AtOR protein levels using anti-PSY and anti-OR antibodies, respectively. Actin was included as an internal reference and detected using anti-Actin antibody. The blot shown is representative of three western blot replicates

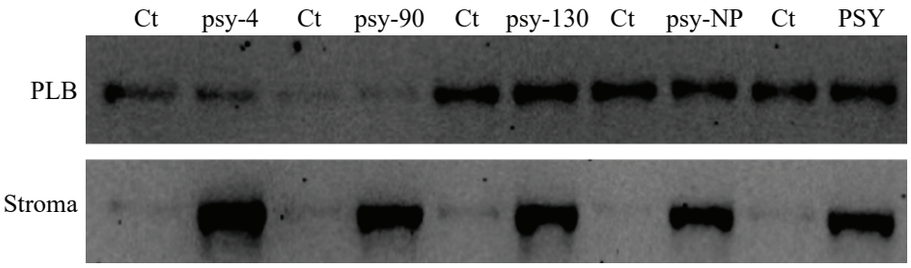


Fig. 9: PSY levels in PLB and stroma fractions of Arabidopsis etioplasts before and after the expression of recombinant proteins. Protoplasts were isolated from the cotyledons of etiolated 7-d old seedlings and incubated in the dark with a PSY-CGFP plasmid for 16 h. Ten microliters of PLB or stroma fraction of etioplasts was subjected to western blot using anti-PSY antibody. Ct: control with no expression of recombinant PSY.

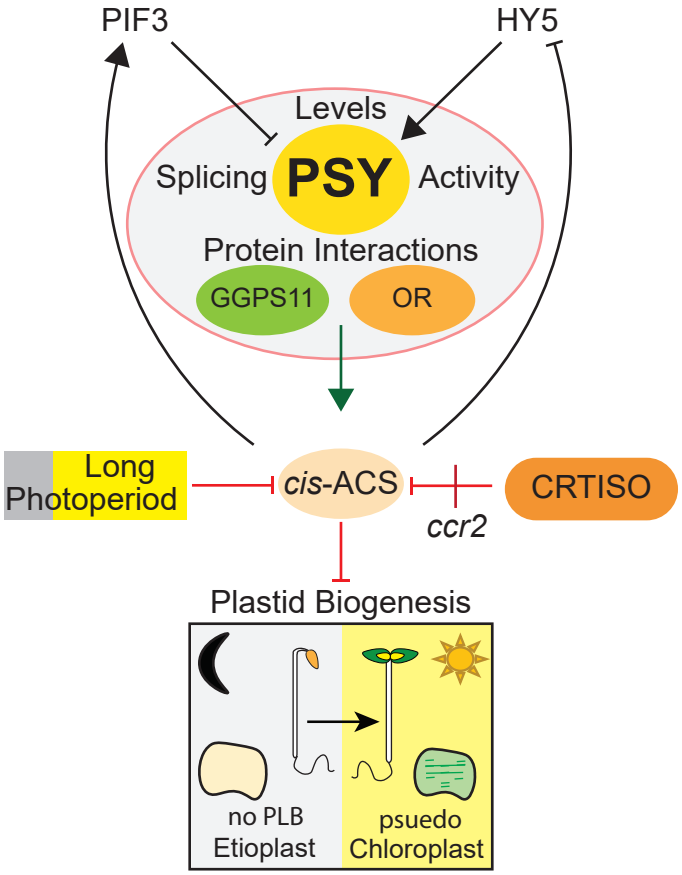


Fig. 10. Manipulation of PSY protein activity provides a central target to modify a cis-carotene derived apocarotenoid signal that modulates plastid biogenesis. Variants in PSY activity effect splicing (alternative leading to truncated protein), activity (enzyme-substrate binding), and protein interactions (multi-enzyme complex binding between PSY, OR, and/or GGPS11) impacting upon protein levels that constitute dynamic regulations to fine-tune the level of a cis-carotene derived apocarotenoid signal (cis-ACS) without impacting overall carotenoid end-product accumulation. cis-ACS upregulates PIF3 and suppresses HY5 protein levels, that fine tune PSY control over cis-ACS biosynthesis. CRTISO activity and longer photoperiods (ie., efficient photoisomerization) reduce cis-carotene biosynthesis and the Arabidopsis crtiso loss-in-function ccr2 mutant causes cis-carotene accumulation that triggers cis-ACS to block prolamellar body (PLB) formation in etioplasts and causes psudeochloroplasts to form when mutants are grown under shorter photoperiods {Cazzonelli, 2020 #28}.