1 Biosynthesis and apoplast accumulation of the apocarotenoid pigment

2 azafrin in parasitizing roots of *Escobedia grandiflora*

3

4 Edison CARDONA-MEDINA[1,*], Marisa SANTOS[2], Rubens NODARI[1], Damaso

5 HORNERO-MÉNDEZ[3], Arnau PERIS[4], Darren C. J. WONG[5], José Tomás MATUS[4],

6 Manuel RODRÍGUEZ-CONCEPCIÓN[6,*]

7

8 [1] Departamento de Fitotecnia, Universidade Federal de Santa Catarina, 88034-000
9   Florianópolis, Brazil

10 [2] Departament of Botany, Universidade Federal de Santa Catarina, 88040-535
11   Florianópolis, Brazil

12 [3] Department of Food Phytochemistry, Instituto de la Grasa (IG-CSIC), 41013 Seville,
13   Spain

14 [4] Institute for Integrative Systems Biology (I2SysBio), Universitat de València-CSIC,
15   Paterna, 46908, Valencia, Spain

16 [5] Ecology and Evolution, Research School of Biology, The Australian National
17   University, 2601, Acton, Australia

18 [6] Institute for Plant Molecular and Cell Biology (IBMCP), CSIC-Universitat Politècnica
19   de València, 46022 Valencia, Spain.

20

21

22 * Corresponding authors:
23 Edison CARDONA-MEDINA, cardonam33@gmail.com
24 Manuel RODRÍGUEZ-CONCEPCIÓN, manuelrc@ibmcp.upv.es

25

26 **Word count -** Main body (total): 5354
27   Introduction: 641
28   Materials and Methods: 845
29   Results and Discussion: 3868

30

31 Number of figures: 7 (all in color but Fig. 1)

32

33 Number of tables: 0

34

35 Supporting information:
36   Methods (S1 to S4)
37   Figures (S1 to S3)
38   Dataset (S1)
39   Tables (S1 to S4)

Summary

- The herbaceous hemiparasite *Escobedia grandiflora* (Orobanchaceae) is used in traditional medicine in the Andean region. Their roots accumulate an orange pigment with a significant relevance as a cooking dye that exhibits antioxidant and cardioprotective properties.

- The present work combined metabolic and cytological analyses with *de novo* transcriptome assembly, gene expression studies, and phylogenetic analyses to confirm the chemical identity of the pigment and investigate its biosynthesis and function in *Escobedia* roots.

- The pigment was conclusively shown to be azafrin, an apocarotenoid likely derived from the cleavage of β-carotene. Candidate genes for the production of azafrin in *Escobedia* roots are proposed based on RNA-seq supported by RT-qPCR and phylogeny reconstruction analyses. In particular, our data suggest that azafrin production relies a carotenoid cleavage dioxygenase (CCD) different from CCD7 and similar to CCD4 enzymes. We also show that azafrin is delivered to the root apoplast and that it accumulates in the area where the *Escobedia* haustorium contacts the host's root, suggesting a role of azafrin in the parasitization process.

- Altogether, our work represents an unprecedented step forward in our understanding of the *Escobedia* parasitization system, but it also provides vital information towards the eventual domestication of this valuable medicinal plant.

Key Words: apocarotenoid, azafrin, *Escobedia*, haustorium, hemiparasite, *de novo* transcriptome assembly, root.

Introduction

Escobedia grandiflora (L. f.) Kuntze (Orobanchaceae) (hereafter referred to as *Escobedia*) is a perennial hemiparasitic plant native to Central and South America, where it associates with high diversity in plant communities of dry and wetland non-forested ecosystems (Burguer & Barringer, 2000; Cardona & Muriel, 2015; Cardona-Medina *et al.*, 2019; Cardona-Medina *et al.*, 2021). Wild populations of this herbaceous plant have supported several traditional uses in the Andean region. The abundant orange pigment in its root is used as a cooking dye and local medicine for hepatitis, jaundice, hyperlipidaemia, and obesity (Silva *et al.*, 2010; Muriel *et al.*, 2015). Such water-soluble pigment was tentatively proposed to be azafrin, a C_{27} apocarotenoid (Kuhn, 1935; Eschenmoser & Eugster, 1975). Azafrin, which exhibits antioxidant and cardioprotective properties (Yang *et al.*, 2018), is found at high levels in particular organs of other parasitic plants, including the roots of *Centranthera grandiflora* and the rhizomes of *Alectra parasitica*, but also in non-parasitic medicinal plants such as *Caralluma umbellata* (Agrawal *et al.*, 2014; Evanjaline & Vr, 2018; Verma *et al.*, 2019; Zhang *et al.*, 2019).

Apocarotenoids are cleavage products of carotenoids. Some of them are biologically active molecules with roles as regulators of plant development and environmental interactions, including abscisic acid (ABA), strigolactones (SL), and others yet to be fully characterized (Walter *et al.*, 2010; Hou *et al.*, 2016; Felemban *et al.*, 2019; Moreno *et al.*, 2021). Plant apocarotenoids often modulate the interaction with herbivores, arbuscular mycorrhizal (AM) fungi, and parasitic plants (Moreno *et al.*, 2021; Wang *et al.*, 2021). A well-known case is SL, which stimulate the germination of some parasitic seeds from Orobanchaceae, contribute to establishing AM symbiosis, and modulate other rhizospheric communication with symbionts and parasites (Torres-Vera *et al.*, 2016; Mutuku *et al.*, 2021). Other apocarotenoids with roles in rhizospheric interactions include blumenols, mycorradicins, zaxinone and anchorene (Moreno *et al.*, 2021; Wang *et al.*, 2021).

Biosynthesis of plant carotenoids takes place in plastids, and it begins with the formation of C_{40} 15-*cis*-phytoene through condensation of two molecules of C_{20} geranylgeranyl diphosphate (GGPP) by phytoene synthase (PSY), the first and main rate-determining enzyme of the pathway (Rodriguez-Concepcion *et al.*, 2018; Moreno *et al.*, 2021). PSY is usually encoded by small gene families in plants (Stauder *et al.*,

98   2018). Phytoene desaturation and isomerization produce lycopene, the red carotenoid

99   responsible for the colour of ripe tomatoes. The formation of β rings in the two ends of

100   the lycopene molecule generates β-carotene, the main pro-vitamin A carotenoid and the

101   proposed precursor of azafrin (Rodriguez-Concepcion *et al.*, 2018; Zhang *et al.*, 2019).

102   Cleavage of the C_{40} skeleton of carotenoids to produce apocarotenoids can either take

103   place non-enzymatically or be catalysed by carotenoid cleavage dioxygenase (CCD)

104   enzymes (Felemban *et al.*, 2019; Moreno *et al.*, 2021). In the model plant *Arabidopsis*

105   *thaliana*, CCD enzymes of the 9-*cis*-epoxycarotenoids dioxygenase (NCED) type are

106   involved in the biosynthesis of ABA whereas CCD7 and CCD8 participate in the first

107   steps of SL production. Other CCD enzymes such as plastid-localized CCD4 and

108   cytosolic CCD1 are involved in the production of other apocarotenoids, including

109   volatiles and growth regulators (Auldridge *et al.*, 2006; Hou *et al.*, 2016; Felemban *et*

110   *al.*, 2019; Moreno *et al.*, 2021). The biosynthetic pathway of azafrin was proposed to

111   begin with the isomerization of β-carotene to 9-*cis*-β-carotene by the enzyme

112   DWARF27 (D27), followed by cleavage by CCD7 to produce 10'-apo-β-carotenal, a

113   common precursor of SL (Bruno & Al-Babili, 2016; Zhang *et al.*, 2019). Then, an

114   unknown aldehyde dehydrogenase would transform aldehyde into carboxylic acid, and a

115   cytochrome P450 monooxygenase would catalyse the oxidation reactions to produce

116   azafrin (Zhang *et al.*, 2019).

117   Here we use modern mass-spectrometry technologies to ascertain the identity of the

118   *Escobedia* root pigment as azafrin, propose candidate genes involved in its production,

119   demonstrate its accumulation in the apoplastic space of root cells, and provide insights

120   on its possible function in parasitic plant-host interactions.

121

Materials and Methods

Plant material

124   Roots and dried fruits of *Escobedia grandiflora* (L. f.) Kuntze were collected from nine

125   mature individuals (i.e., in the flowering period) in a wild population located in the

126   Lagoinha do Leste Municipal Park (27°46'53.0"S, 48°29'16.1"W, 190 m.a.s.l.),

127   municipality of Florianópolis, Brazil. After collection, the roots were frozen at -80ºC,

128   lyophilized, and then pulverized with a TissueLyser (Qiagen) to obtain a fine powder

129   for further chromatography and RNA extraction. Seeds were collected from dried fruits,

130 imbibed for five days in distilled water and sown as described by Cardona-Medina *et al.*

131 (2019). Twenty imbibed seeds were sown in 2-liter pots without a host plant in

132 greenhouse conditions with a long day photoperiod. After nine months, roots were

133 collected, frozen and processed as described above for wild (host-attached) samples.

134 Furthermore, twenty imbibed seeds were sown in 5-liter pots with *Pennisetum*

135 *purpureum* Schumach (Poaceae) as a host plant (Cardona & Muriel, 2015; Cardona-

136 Medina *et al.*, 2019) and grown for nine months to observe different root and haustoria

137 development phases, as reported by Cardona-Medina *et al.* (2019).

138

139 *Azafrin identification and quantification analysis*

140 Azafrin identification was based on HPLC-DAD-MS(APCI+) analysis carried out as

141 described (Supporting Information Methods S1). For azafrin quantification, lyophilized

142 roots from *Escobedia* plants were used for pigment extraction, chromatographic

143 separation, and detection at 450 nm as described (Supporting Information Methods S1).

144 Three biological replicates were performed in every experiments.

145

146 *RNA-seq*

147 Total RNA was extracted from a pool of *Escobedia* roots grown with their hosts in a

148 wild population. Samples representing different stages of the parasitizing process (i.e.

149 root development) were extracted using the Maxwell® RSC Plant RNA kit (Promega)

150 with the Maxwell® RSC Instrument (Promega), according to the manufacturer´s

151 instructions. RNA samples were pooled in equivalent amounts and library construction

152 was performed with Illumina TruSeq Stranded mRNA kit, according to the

153 manufacturer's instructions. A single library was sequenced at a depth of 30 million

154 reads. Raw paired-end (2x150 bp) reads were processed as follows. Removal of adaptor

155 sequences, sliding-window trimming, length, and quality filtering of the raw paired-end

156 reads were performed with *fastp* v0.20.0 (Chen *et al.*, 2018) using *-w* 16 *-5 -3 -r -W* 4 *-*

157 *M* 20 *-l* 40. All other settings were left at default. *De novo* transcriptome assembly of

158 surviving reads was performed with *Trinity* v2.11.0 (Haas *et al.*, 2013) using default

159 settings. Transcriptome analysis and functional annotation was performed as described

160 (Supporting Information Methods S2).

161

162   *Phylogenetic analysis*

163   Phylogenetic reconstructions were performed with MrBayes (Huelsenbeck & Ronquist,

164   2001). The resulting topology based on a phylogram was constructed through the

165   standard stepwise pipeline. Firstly, the candidate protein sequences were aligned with

166   MUSCLE (Edgar, 2004) using default parameters and the multiple sequence alignment

167   outcome was exported in nexus format (Maddison *et al.*, 1997). MEGAX (Kumar *et al.*,

168   2018) was used to identify the best substitution protein model, i.e. with the lowest

169   Bayesian information criterion (BIC) score. Each sequence was considered as 'partially-

170   deleted' in which a candidate is discarded if it presents higher percentage of ambiguous

171   sites than the threshold specified in the Site Coverage Cutoff parameter in MEGAX

172   options (similar to the complete deletion option but with a threshold set to 100%, i.e.

173   absence of ambiguous sites). Once the multiple sequence alignment was performed, the

174   best substitution model was identified and the distribution of rates among sites was

175   selected. In our case, the best model was Jones-Taylor-Thornton (JTT) with rates

176   among sites following a gamma distribution. The phylogenetic analysis was then

177   conducted by bayesian inference, using the Markov Chain Monte Carlo algorithm

178   (MCMC) available in MrBayes. The different clusters' reliability was calculated with

179   the posterior probabilities (Gelman *et al.*, 1995), being the values higher than 0.70

180   acceptable and the values higher than 0.90 highly supported.

181

182   *RT-qPCR*

183   RNA extracted from *Escobedia* roots was used for real-time quantitative PCR (RT-

184   qPCR) as described (Supporting Information Method S3). RT-qPCR data were

185   normalized with the *Escobedia* actin gene DN8798_c0_g1_i1. Primers are listed in

186   Supporting Information Table S1. For statistical analyses of the results, a one-way

187   ANOVA was performed in which the host conditions were the explanatory variable, and

188   azafrin and relative transcript levels of candidate genes were the response variable.

189   Response variables were square-root transformed. ANOVA validation was based on the

190   Shapiro-Wilk test and the analysis of plotting residuals. Significant differences ($p \leq 0.05$)

191   were tested using Tukey´s test. The data were analyzed in R environment version 4.0.3 (

192   R Core Team, 2021).

193

194 *Structural and anatomical analyses*

195 Anatomical analyses were performed as described (Supporting Information Methods

196 S4). Haustoria were fixed in a solution with 2.5 % glutaraldehyde in a 0.1 M sodium

197 phosphate buffer and dehydrated through an ethanolic series (Ruzin, 1999). For

198 scanning electron microscopy (SEM) analyses, the fixed and dehydrated samples were

199 processed as described (Supporting Information Methods S4). The hyaline body

200 ultrastructure was visualized by transmission electron microscopy (TEM), according to

201 Pueschel (1979) after processing the samples as described (Supporting Information

202 Methods S4). Azafrin detection by confocal microscopy of fresh hand-cut root sections

203 was performed as described by D'Andrea *et al.*(2014), using the 488 nm ray line of an

204 argon laser for excitation and 500-550 nm of emission window.

205

Results and Discussion

207 *Escobedia grandiflora roots accumulate azafrin and also produce aeginetin*

208 While early studies in *Escobedia* species proposed that the orange pigment that

209 accumulates in roots is azafrin (Kuhn, 1935; Karrer & Jucker, 1948; Eschenmoser &

210 Eugster, 1975), clearcut demonstration is still missing. Consistent with those studies, the

211 HPLC analysis of a root extract from *Escobedia* plants of a wild population revealed the

212 presence of a major compound (>95%; peak 2) with an on-line UV-visible spectrum

213 with maxima at 390, 421, 436 nm (Fig. 1). This spectrum agrees with a chromophore

214 structure with seven to eight conjugated double bonds (c.d.b), in accordance with the

215 structure proposed for azafrin with a seven c.d.b. polyene chain and a conjugated

216 carbonyl group (Fig. 1a). However, the acid mobile phase likely affected absorption

217 maxima and fine structure, resulting in the UV-visible spectrum being slightly different

218 from that reported in the literature (Britton, 1991; Britton *et al.*, 2004). When peak 2

219 was collected and purified from the diode array detector outlet, the resulting UV-visible

220 spectrum in ethanol (388, 407, 431 nm) matched the one reported for azafrin in previous

221 studies (Britton, 1991; Britton *et al.*, 2004) (Fig. 1b). The chemical identity of this peak

222 was confirmed by mass spectrometry using HPLC-DAD-MS(APCI+). As shown in

223 (Fig. 1c), the mass spectrum was consistent with the formula $C_{27}H_{38}O_4$

224 (MW=426.2770), with a fragmentation pattern presenting three characteristic fragments

225 corresponding to the protonated molecule ([M+H]+, 427.27) and the loss of one and two

226   water molecules derived from the hydroxy groups ([M+H-18]+, 409.26; [M+H-18-18]+,

227   391.25). The fragment derived after the neutral loss of the carboxylic group [M-46]+ at

228   381.26 was also detected but with low relative abundance. These results conclusively

229   confirmed that peak 2 of the chromatogram, corresponding to the most abundant

230   pigment by far in *Escobedia* roots (Fig. 1a), was indeed azafrin.

231   Regarding the minor compound eluting before azafrin (peak 1; Fig. 1a), its UV-visible

232   spectrum presented maxima at 395 and 409 nm (Fig. 1d), suggesting a chromophore

233   with 6 c.d.b. (i.e., 1 c.d.b. shorter compared to azafrin). The mass spectrum of this

234   compound showed characteristic fragments: [M+H]+ at 401.26, [M+H-18]+ at 383.25,

235   [M+H-18-18]+ at 365.24 and [M-46]+ at 355.26 (Fig 1e). Remarkably, this mass

236   fragmentation profile was very similar to the one observed for azafrin but fragments

237   showed 28 u.m.a. less, which correspond to an alkene unit ($-CH_2=CH_2-$). These data

238   corroborated that the structure has a chromophore with 1 c.d.b less than azafrin. The

239   protonated molecule ([M+H]+ at 401.26) was consistent with a molecular structure with

240   a formula $C_{25}H_{36}O_4$ (MW=400.2613), which corresponds to the apocarotenoid aeginetin

241   (Fig. 1a), isolated from the roots of *Aeginetia indica* (Eschenmoser *et al.*, 1982; Britton,

242   1991), a holoparasitic herb of the same plant family as *Escobedia*, Orobanchaceae.

243   Azafrin has been identified in non-parasitic species but it is only accumulated at high

244   levels in the roots or rhizomes of root hemiparasitic plants (Agrawal *et al.*, 2014; Zhang

245   *et al.*, 2019). It is interesting to note that azafrin-overaccumulating hemiparasitic plants

246   such as *Escobedia*, *Centranthera grandiflora* (hereafter referred to as *Centranthera*) and

247   *Alectra parasitica* are classified in the same subclade (I) within the Buchnerae clade of

248   Orobanchaceae family (Nickrent, 2020). Two other species belonging to this subclade I

249   (*Melasma stricta* and *Notochilus coccineus*) were reported to have orange roots, but the

250   identity of the pigment has not been identified yet (Safford, 1999; *speciesLink network*,

251   2021). Similarly, aeginetin has only been reported in plants of the Buchnerae clade,

252   including *Aeginetia* (subclade H, highly related to subclade I), *Centranthera*

253   (Eschenmoser *et al.*, 1982; Zhang *et al.*, 2019; Nickrent, 2020), and now *Escobedia*. It

254   is possible, therefore, that the production of these apocarotenoids, and particularly the

255   accumulation of azafrin, might be play a biological function related to root parasitism in

256   this specific group of plants.

257

258 *Analysis of the Escobedia root transcriptome allows the identification of candidate*

259 *genes for carotenoid biosynthesis.*

260 To identify tentative genes involved in the synthesis of azafrin in *Escobedia*, we

261 extracted RNA from a pool of orange roots collected from different plants in a natural

262 population and used it for RNA-seq analysis. Following adapter removal, trimming, and

263 quality filtering with *fastp*, 76 million 2x150 bp paired-end reads (~ 10.8 Gb of

264 sequence data) were used for *de novo* transcriptome assembly using *Trinity*. Summary

265 statistics of the final assembly include, among others, a total of 115,882 transcripts

266 (65,701 'genes') ranging between 200 and 12,858 bp in length, a mean and N50

267 sequence length of 1,111 and 1,798 bp, a GC-content of 0.43, and a 25.5% of total

268 transcripts deemed as lowly or not expressed (i.e., FPKM < 0.5) (Supporting

269 Information Fig. S1a). Additionally, 61,325 and 48,576 transcripts showed evidence-

270 supported coding sequences predicted using *TransDecoder* and *EvidentialGene*

271 pipelines, respectively. Transcripts containing predicted CDS by *TransDecoder*

272 revealed that many were complete (i.e., full-length and containing both start and stop

273 codons) and predominantly being >1kb in length (Supporting Information Fig. S1b).

274 BUSCO assessments with the embryophyta lineage database also indicated very high

275 predicted completeness of the full transcriptome assembly and reduced CDS-only set

276 (i.e., 93.7% - 94.2% of 1,375 BUSCOs evaluated) (Supporting Information Fig. S1c).

277 These BUSCO scores rival those observed in many sequenced plant genomes with high-

278 quality gene models (Veeckman *et al.*, 2016).

279 Functional annotation according to MapMan BIN v4 categories, matching Pfam

280 domain of predicted peptides, or shared homology with plant Uniprot database

281 sequences, revealed 51,302, 51,578, and 66,335 transcripts (44 – 57% of total

282 transcripts), respectively (Supporting Information Table S2). Notably, we observed a

283 greater representation of transcripts associated to MapMan BIN15_RNA biosynthesis,

284 BIN18_Protein modification, BIN19_Protein homeostasis, BIN24_Solute transport, and

285 BIN50_Enzyme classification, among others (Supporting Information Fig. S1d). More

286 relevant for azafrin biosynthesis, BIN categories related to biosynthetic pathways of

287 carotenoids (BIN9.1.6.1) and apocarotenoids (BIN9.1.6.3), among others, were also

288 successfully assigned to transcripts (Supporting Information Fig. S1d). Additionally,

289 45,339 and 68,727 transcripts present in *Escobedia* roots shared orthology with

290 *Arabidopsis* and *Centranthera* genes, respectively (Supporting Information Table S2).

291   Together, the assembly statistics, gene completeness scores, and number of functional

292   categories and domain assignments to transcripts suggest a reasonably high assembly

293   quality of the pooled root transcriptomes.

294   According to our annotation pipelines, we identified thirty-four genes (30 predicted

295   complete and 4 partial sequences) potentially involved in carotenoid and apocarotenoid

296   pathways that were expressed in *Escobedia* roots (Fig. 2; Supporting Information Table

297   S3). The first committed step of the carotenoid pathway is the production of phytoene

298   from GGPP catalysed by PSY (Fig. 2a). PSY, the main flux-controlling enzyme of the

299   plant carotenoid pathway (Fraser *et al.*, 2002; Rodriguez-Concepcion *et al.*, 2018), is

300   usually encoded by small gene families encoding distinct isoforms associated with

301   organ- or tissue-specific production of carotenoids. For example, tomato PSY1 is

302   essential for fruit carotenoid production during ripening, while PSY2 is preferentially

303   found in photosynthetic tissues, and PSY3 functions in the root. The root-associated

304   PSY3 isoforms from dicots form a widespread phylogenetic clade found to participate

305   in arbuscular mycorrhiza (AM) interactions, whereas those from monocots form a

306   different clade and are involved in ABA formation (Stauder *et al.*, 2018). Two PSY

307   isoforms were found to be expressed in *Escobedia* roots, namely PSYa and PSYb (Fig.

308   2a). The phylogenetic comparison of the *Escobedia* PSYa and PSYb isoforms with PSY

309   sequences from *Centranthera* and two other root hemiparasitic species

310   (*Phtheirospermum japonicum* and *Striga asiatica*,) together with well-characterized

311   PSY sequences from monocots (maize, rice) and dicots (tomato, carrot, alfalfa,

312   *Arabidopsis*) led to their classification in a sub-clade of only root hemiparasitic PSY

313   sequences with a 100% of posterior probability (Fig. 3). This subclade, however, was

314   separated from the clades harbouring root-associated PSY3 sequences from dicots or

315   monocots (Fig. 3). Similar to *Escobedia*, two genes encoding PSY and belonging to the

316   same subclade were found in *Centranthera* (Fig. 3). It is interesting to note that the

317   *Centranthera* gene encoding PSYa was more actively expressed in leaves and stems

318   than in roots whereas higher levels of transcripts encoding PSYb were found in azafrin-

319   producing roots compared to leaves and stems (Fig. 2a) (Zhang *et al.*, 2019). These

320   results suggest that some non-PSY3 isoforms of PSY might have a prominent role in the

321   roots of at least some hemiparasitic plants. In the case of *Escobedia*, genes encoding

322   both PSYa and PSYb isoforms are expressed at similarly high levels in roots,

323   suggesting that azafrin production may require an active metabolic flux into the

324   carotenoid pathway.

325   In the next section of the carotenoid pathway (i.e., from phytoene to lycopene),

326   transcripts corresponding to single genes were found for phytoene desaturase (PDS), ζ-

327   carotene isomerase (Z-ISO), and carotenoid isomerase (CRTISO), whereas two

328   *Escobedia* genes were expressed encoding ζ-carotene desaturase (ZDS) (Fig. 2a).

329   Interestingly, *Arabidopsis* mutants defective in ZDS have been reported to produce an

330   apocarotenoid signal that negatively impacts plastid and leaf development (Avendaño-

331   Vázquez *et al.*, 2014; Escobar-Tovar *et al.*, 2021). Thus, the relatively high expression

332   levels of the two ZDS-encoding genes detected in *Escobedia* roots (Fig. 2a) might result

333   in high ZDS activity and hence prevent the formation of the apocarotenoid signal

334   generated in ZDS-defective mutants. After lycopene, the carotenoid pathway branches

335   out (Fig. 2a). Cyclization of the two ends of the linear lycopene molecule to produce

336   one β and one ε ring generates α-carotene. These reactions are catalysed by lycopene β

337   and ε cyclases (LCYB and LCYE, respectively). By contrast, β-carotene is synthesized

338   from lycopene when only β rings are formed by LCYB enzymes. Transcripts encoding

339   LCYE were found at much lower levels than those encoding the two LCYB isoforms

340   expressed in *Escobedia* and *Centranthera* roots, LCYBa and LCYBb (Fig. 2a),

341   suggesting a higher flux through the β, β branch compared to the β, ε branch. In

342   agreement with this conclusion, transcripts for the LUT1 enzyme, which catalyses the

343   hydroxylation of ε rings to produce lutein (the most abundant carotenoid in green

344   photosynthetic tissues), are much less abundant than those for β-ring hydroxylases

345   (BCHa and BCHb). Based on the higher abundance of transcripts for BCHa in both

346   *Escobedia* and *Centranthera*, this might be the main BCH isoform transforming β-

347   carotene into zeaxanthin (Fig. 2a). Then, two highly expressed zeaxanthin epoxidase

348   isoforms (ZEPa and ZEPb) in *Escobedia* and a single ZEP-encoding gene in

349   *Centranthera* produce violaxanthin in roots. The very low levels of transcripts found for

350   violaxanthin epoxidase (VDE) suggest that the main metabolic flux is from zeaxanthin

351   to violaxanthin in *Escobedia* and *Centranthera* roots.

352

353   *Azafrin production in Escobedia roots might rely on CCD4 rather than CCD7 enzymes.*

354   Presence of transcripts encoding NCED and other ABA biosynthetic enzymes such as

355   ABA2, ABA3, AAO3 (Fig. 2a), suggest that violaxanthin can be converted into ABA in

356   azafrin-producing *Escobedia* and *Centranthera* roots. Besides ABA, other

357   apocarotenoids are formed by the activity of CCD enzymes. Among them, SL and

358   azafrin production are proposed to share the first isomerization step from β-carotene

359   (Fig. 2a). Transcripts encoding D27 were found in azafrin-accumulating roots,

360   supporting the conclusion that β-carotene can be isomerized to 9-*cis*-β-carotene in this

361   tissue to allow the production of SL or/and azafrin (Fig. 2a). Both apocarotenoids were

362   also proposed to share the next step of the pathway, i.e. the cleavage of C_{40} 9-*cis*-β-

363   carotene by CCD7 to produce C_{27} 10'-apo-β-carotenal (Zhang *et al.*, 2019). In-depth

364   phylogenetic analysis of the CCD family from *Escobedia*, *Centranthera* and several

365   other plants, including hemiparasitic species, showed four clades designated as CCD1

366   (paraphyletic), CCD4, CCD7, and CCD8 (Fig. 4). The only gene encoding CCD1 was

367   found to be poorly expressed in *Escobedia* roots, whereas in *Centranthera* it was less

368   expressed in roots than in other tissues such as leaves and stems (Fig. 2a). The CCD4

369   clade contained seven CCD4 sequences from *Escobedia* (CCD4a to g), of which

370   CCD4b and CCD4c were most highly expressed in *Escobedia* roots (Fig. 2a). Of the

371   eight CCD4 sequences found in *Centranthera* (CCD4a to h), only CCD4d was

372   predominantly expressed in the azafrin-producing roots (Fig. 2).

373   A striking conclusion of our phylogenetic analysis is that no sequences of *Escobedia*

374   were present in the CCD7 clade, which included sequences of *Centranthera* and other

375   hemiparasitic species as well as typical CCD7 enzymes from *Arabidopsis* and tomato.

376   We hence speculated that *Escobedia* roots CCD4b or/and CCD4c might catalyse the

377   same C9-C10 cleavage reaction that CCD7 enzymes perform in other plants or tissues.

378   The CCD4 subfamily is probably the most variable group of CCD enzymes. They are

379   encoded by several genes in many species, resulting in isoforms that often differ in their

380   expression profile and substrate selectivity (Hou et al., 2016). In general, CCD4

381   enzymes have broad substrate specificity, and many of them appear to have a role in

382   carotenoid catabolism, particularly in carotenoid-sink tissues such as flowers, fruits,

383   seeds, and roots (Walter *et al.*, 2010; Rubio-Moraga *et al.*, 2014). They usually cleave

384   carotenoids (notably β-carotene) at the C9–C10/C9'-C-10' double bond, but the

385   *Arabidopsis* CCD4 enzyme also catalyses the C9–C10 cleavage of β, β xanthophylls

386   such as zeaxanthin while other CCD4 enzymes cleave asymmetrically at the C7–

387   C8/C7'-C-8' double bond (Rubio *et al.*, 2008; Huang *et al.*, 2009; Ma *et al.*, 2013;

388   Lätari *et al.*, 2015; Bruno *et al.*, 2016). Interestingly, *Arabidopsis* CCD4 has been

389   shown to catalyse the cleavage of β-carotene to all-*trans*-10'-apo-β-carotenal and, at a

390   much lower efficiency, of 9-*cis*-β-carotene to the SL precursor 9-*cis*-10'-apo-β-

391   carotenal (Bruno *et al.*, 2016). It can thus be suggested that *Escobedia* CCD4b or/and

392   CCD4c isoforms might also catalyse these reactions to deliver precursors for SL and

393   azafrin biosynthesis, hence making the participation of a CCD7 enzyme unnecessary.

394   Interestingly, in the roots of *Centranthera* transcripts encoding CCD7 were expressed at

395   much lower levels than those encoding D27 (Fig. 2a), whereas CCD4b was much more

396   highly expressed than CCD7 (Fig. 2a) (Zhang *et al.*, 2019). It is therefore possible that

397   CCD4 isoforms might produce the precursors for SL and azafrin in both *Escobedia* and

398   *Centranthera* roots. Further supporting this conclusion, our analysis indicated that

399   *Centranthera CCD4b* is the only gene that shows root-specific differential upregulation

400   compared to stem and leaf tissues among all other azafrin-related *Escobedia* orthologs

401   (Fig. 2b; Supporting Information Table S3).

402   To complement the RNA-seq analysis, transcript levels encoding CCD4b and CCD4c

403   were measured in roots of *Escobedia* plants grown either with or without hosts (Fig. 5).

404   Previous studies observed that the roots of *Escobedia* were colourless in the initial

405   developmental stages (i.e., before parasitizing a host), while the orange pigment became

406   most visible after haustoria penetration in the host root (Cardona-Medina *et al.*, 2019).

407   HPLC analyses confirmed that azafrin levels were substantially increased in *Escobedia*

408   roots when growing in the presence of host plants compared to those grown in their

409   absence (Fig. 5). The level of transcripts encoding CCD4c were similar regardless of the

410   presence of a host, whereas those encoding CCD4b were present at much higher levels

411   in azafrin-producing roots (Fig. 5). Together, these results suggest that CCD4b might be

412   the main enzyme participating in the production of azafrin.

413   C_{27} 10'-apo-β-carotenal is the substrate of CCD8 enzymes in the next step of the SL

414   biosynthesis pathway (Fig. 2a). *Escobedia* and *Centranthera* genes encoding homologs

415   for CCD8 were found to be expressed at much lower levels than the one for D27 in

416   roots (Fig. 2a). Based on these data, we hypothesise that the pathway for producing

417   strigolactones (via CCD8) is not very active in azafrin-producing *Escobedia* and

418   *Centranthera* roots. This might not be a general trend in parasitic plants. For example,

419   the tubercle of the holoparasitic plant *Phelipanche aegyptiaca* showed a high expression

420   of D27, CCD7, and CCD8 when parasitizing the host roots (Emran *et al.*, 2020).

421   However, mycorrhizal plants that produce high levels of apocarotenoids reduce their

422 production and secretion of SL, likely because abundant apocarotenoid production in

423 AM-colonized roots might generate a metabolic sink and successfully compete for SL

424 precursors (Walter *et al.*, 2010). Similarly, the presumably low flux towards SL in

425 azafrin-producing *Escobedia* and *Centranthera* roots might be related to the

426 requirement of very high levels of common precursors to support the massive

427 production of azafrin in the roots of these hemiparasitic plants. Strikingly, we were

428 unable to detect any carotenoid species in *Escobedia* roots, suggesting that the

429 carotenoid pathway in this organ is mainly directed to provide substrates for

430 apocarotenoid production, and no intermediates are accumulated. Similarly, no

431 detectable amounts of carotenoids were found in AM roots of all plants investigated

432 despite the required high flux through the carotenoid pathway (Fester *et al.*, 2002).

433

434 *Azafrin accumulates in the apoplast of the Escobedia root cortex.*

435 An anatomical analysis was next performed to investigate where azafrin accumulated in

436 the roots of *Escobedia* plants collected in the wild. The light microscopy visualization

437 of the internal structure of the root revealed that the orange pigment corresponding to

438 azafrin was not located in plastids (the site where all plant carotenoids are made) or in

439 the cell cytosol (where the synthesis of many apocarotenoids is completed) but

440 accumulated in the intercellular spaces (i.e., apoplast) of the root cortex (Fig. 6a).

441 Confocal laser scanning microscopy analyses based on the autofluorescence produced

442 by the c.b.d. system present in the polyene chain of carotenoids and apocarotenoids of

443 sufficient length such as C_{27} azafrin (D'Andrea *et al.*, 2014) confirmed that the orange

444 pigmentation detected by light microscopy was due to the presence of azafrin as both

445 autofluorescence and color signals overlapped (Fig. 6b; Supporting Information Fig.

446 S2). Light and confocal microscopy of carrot (*Daucus carota*) roots also showed an

447 overlap of color and autofluorescence signals, but in this case they were both detected

448 inside plastids (i.e., chromoplasts) as they correspond to carotenes (β-carotene and, to a

449 lower extent, α-carotene) instead of their cleavage products (Supporting Information

450 Fig. S2). By contrast, azafrin was virtually absent from the large starch-filled plastids

451 (i.e., amyloplasts) present in *Escobedia* roots (Fig. 6; Supporting Information Fig. S2).

452 Carotenoids are synthesized and accumulated in different plastid types, including

453 amyloplasts (Horner *et al.*, 2007; Sun *et al.*, 2018; Rodriguez-Concepcion *et al.*, 2018).

454 Studies of apocarotenoid formation in mycorrhizal roots have led to conclude that the

455   C_{27} products of CCD4 or/and CCD7 activities are exported from the plastid and used in

456   the cytosol as substrates of CCD1 enzymes that convert them into downstream products

457   such as C_{13} (colorless) blumenols and C_{14} (yellow) mycorrhadicins (Walter *et al.*, 2010;

458   Fiorilli *et al.*, 2019; Moreno *et al.*, 2021). Similarly, the potential C_{27} product of CCD4b

459   activity in the amyloplasts of *Escobedia* roots might be exported from the plastids to the

460   cytosol. Accumulation of C_{27} apocarotenoids is uncommon in nature, probably because

461   CCD1 enzymes normally degrade them (Floss *et al.*, 2008; Walter *et al.*, 2010). The low

462   expression level of the only gene encoding CCD1 in *Escobedia* roots (Fig. 2a) suggests

463   that most of this C_{27} intermediate might remain available for other cytosolic enzymes to

464   transform it into downstream products. The differences between azafrin and 10'-apo-β-

465   carotenal are one terminal carboxyl group and two hydroxyl groups in the cyclohexane

466   skeleton (Zhang *et al.*, 2019). The activity of cytosolic aldehyde dehydrogenase and

467   cytochrome P450 monooxygenase enzymes could transform the aldehyde group of 10'-

468   apo-β-carotenal into carboxylic acid and insert oxygen atoms, respectively, eventually

469   improving hydrophilicity (Fig. 2a). Water-soluble azafrin might then be released from

470   the root cells and accumulate in the apoplast. Based on the putative role reported for the

471   accumulation of colored apocarotenoids in AM-inoculated roots, it is possible that

472   azafrin might participate in the interaction with the rhizosphere, e.g. by providing

473   protection from oxidative damage caused by biotic or abiotic stresses (Strack & Fester,

474   2006).

475

476   *Evidence for a possible role for azafrin in the parasitization process.*

477   Following germination, the root of *Escobedia* is colourless and seedlings grow very

478   slowly until a host is parasitized, which involves both the formation of specialized

479   organs (haustoria) to penetrate the host root and the pigmentation of the root, i.e. the

480   accumulation of azafrin (Cardona & Muriel, 2015; Cardona-Medina *et al.*, 2019). The

481   haustorium is an organ characteristic of parasitic plants that have evolved in multiple

482   independent angiosperms. It contains structures for mechanical attachment to the host

483   root and vascular connections that involve the differentiation of various specialized cell

484   types (Teixeira-Costa, 2021). A detailed exploration of this structure in *Escobedia*

485   showed numerous tubular haustorial hairs associated with the orange swelling periphery

486   (Supporting Information Fig. S3). These hairs participate in securing the haustorium to

487   the epidermis of the host root and facilitate the contact and penetration into host tissues

488 (Heide-Jorgensen & Kuijt, 1995; Cui *et al.*, 2016). The haustorial opening, i.e., the area

489 that makes contact with the host root, was observed in the haustorium apex (Supporting

490 Information Fig. S3a). Longitudinal sections of the *Escobedia* haustorium attached to

491 host roots revealed a complex internal structure, presenting four recognizable regions

492 (Supporting Information Fig. S3b): haustorial base, vascular tissue, hyaline body, and

493 intrusive cells (endophyte). The haustorial base connects the parasitic root with the

494 haustorium, morphologically similar to root tissue. The vascular tissue comprised of

495 provascular cells and tracheary elements (haustorium xylem) is arranged

496 perpendicularly to the haustorial base and towards the host root xylem. Intrusive cells of

497 the haustorium penetrate and advance inside the host root, constituting the endophyte.

498 Tracheary elements of the haustorium were observed inside the host's metaxylem,

499 indicating parasitism success in the host root (Supporting Information Fig. S3b).

500 Interestingly, we noticed that every haustorium attached to host roots contained orange

501 pigment depositions corresponding to azafrin in the region directly contacting the host

502 root interface (Fig. 7; Supporting Information Fig. S3b).

503 The reason why azafrin accumulates in the haustorium-host root interface is still

504 unknown. We propose that azafrin might inhibit the host defence responses during the

505 penetration of the haustorium inside the roots. Haustorium penetration in host roots

506 involves enzymatic secretion and mechanical pressure (by haustorial hairs and cellular

507 division) that degrade and disrupt the host cells walls (Heide-Jorgensen & Kuijt, 1995;

508 Hood *et al.*, 1998; Losner-Goshen *et al.*, 1998). This process causes a wound in the host

509 root that allows the entry of haustorium intrusive cells into the host vascular system.

510 Host roots quickly respond to the wound by activating the production of reactive

511 oxygen species (ROS), which can activate programmed cell death (PCD) and trigger

512 plant defence responses (Minibayeva *et al.*, 2009; Tripathy & Oelmüller, 2012),

513 including the generation of phenolic compounds and callose deposition, induction of

514 immunity-related genes, and deposition of lignin and suberin to avoid the advance of

515 parasitization (Hiraga *et al.*, 2001; Minibayeva *et al.*, 2009; Saucet & Shirasu, 2016).

516 Evidence of necrosis involving ROS was found in resistant hosts during unsuccessful

517 penetration by the haustorium of *Orobanche cumana* (Letousey *et al.*, 2007). Thus, the

518 presence of azafrin in the haustorium might inhibit host PCD and defense responses by

519 eliminating extracellular ROS as a strategy to facilitate parasitization (Mor *et al.*, 2008).

520 Inactivation of defence responses has also been observed in biotrophic pathogens during

521 the parasitism of host plants, due to its need to proliferate in living host cells (Siddique

522 *et al.*, 2014). Likewise, endophytic fungi can produce antioxidants to circumvent

523 damage by ROS during beneficial interaction with the host plant (Hamilton *et al.*,

524 2012). Together, we speculate that azafrin accumulation in the haustorium-host root

525 interface might also play an antioxidant role to counteract the ROS-related host defence

526 responses, hence allowing the parasitism to succeed. Further work should

527 experimentally address this and other unanswered questions about the role of azafrin

528 and other apocarotenoids in the parasitization strategy of *Escobedia*. The information

529 would be vital to eventually achieve the domestication and cultivation and hence

530 exploitation of this important medicinal plant.

531

Acknowledgements

550

551

Author Contributions

ECM, RON and MRC designed the research; ECM collected plant material and performed the experiments; ECM and DHM carried out azafrin analysis; AP, JTM and DW conducted all bioinformatic analyses; ECM, MS, RON, MRC performed microscopical analyses; ECM and MRC wrote the paper. All authors contributed to interpretations and revisions of the manuscripts.

Data availability

RNA-Seq original sequence data can be found in the Sequence Read Archive (SRA) database of the NCBI under the Bioproject accession PRJNA798758.

Supporting Information

Methods S1. Azafrin detection and quantification.

Methods S2. Transcriptome analysis and functional annotation

Methods S3. RT-qPCR.

Methods S4. Structural and anatomical analyses.

Fig. S1 *De novo* transcriptome assembly from *Escobedia* roots.

Fig. S2. Carotenoid and azafrin distribution in carrot and *Escobedia* roots.

Fig S3. Overview of the *Escobedia* haustorium.

Data S1. Nucleotide sequences of *Escobedia* PSY and CCD homologs.

Table S1. Primers used for RT-qPCR.

Table S2. *De novo* transcriptome assembly using Trinity.

Table S3. Accessions and FPKM values of *Escobedia* and *Centranthera* transcripts involved in azafrin biosynthesis.

Table S4. PSY and CCD accessions used in phylogenetic analyses.

References

Agrawal P, Laddha K, Tiwari A. **2014**. Isolation and HPLC method development of azafrin from *Alectra parasitica* var. *chitrakutensis*. *Natural Product Research* **28**: 940–944.

Ahrazem O, Diretto G, Argandoña J, Rubio-Moraga Á, Julve JM, Orzáez D, Granell A, Gómez-Gómez L. **2017**. Evolutionarily distinct carotenoid cleavage dioxygenases are responsible for crocetin production in *Buddleja davidii*. *Journal of Experimental Botany* **68**: 4663–4677.

Aubry S, Kelly S, Kümpers BMC, Smith-Unna RD, Hibberd JM. **2014**. Deep evolutionary comparison of gene expression identifies parallel recruitment of trans-factors in two independent origins of C4 photosynthesis. *PLoS genetics* **10**: e1004365.

Auldridge ME, McCarty DR, Klee HJ. **2006**. Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Current Opinion in Plant Biology* **9**: 315–321.

Avendaño-Vázquez A-O, Cordoba E, Llamas E, San Román C, Nisar N, De la Torre S, Ramos-Vega M, Gutiérrez-Nava M de la L, Cazzonelli CI, Pogson BJ, *et al.* 2014. An Uncharacterized Apocarotenoid-Derived Signal Generated in ζ-Carotene Desaturase Mutants Regulates Leaf Development and the Expression of Chloroplast and Nuclear Genes in Arabidopsis. *The Plant Cell* **26**: 2524–2537.

Britton G. **1991**. Carotenoids. In: Methods in plant biochemistry. London: Academic Press, 473–518.

Britton G, Liaaen-Jensen S, Pfander H (Eds.). **2004**. *Carotenoids: Handbook*. Birkhäuser Basel.

Bruno M, Al-Babili S. **2016**. On the substrate specificity of the rice strigolactone biosynthesis enzyme DWARF27. *Planta* **243**: 1429–1440.

Bruno M, Koschmieder J, Wuest F, Schaub P, Fehling-Kaschek M, Timmer J, Beyer P, Al-Babili S. **2016**. Enzymatic study on AtCCD4 and AtCCD7 and their potential to form acyclic regulatory metabolites. *Journal of Experimental Botany* **67**: 5993–6005.

Buchfink B, Xie C, Huson DH. **2015**. Fast and sensitive protein alignment using DIAMOND. *Nature Methods* **12**: 59–60.

Burguer W, Barringer K. **2000**. Botany: New series No 4, Flora Costaricensis. *Museum of Natural History*: 39–40.

Cao FY, Yoshioka K, Desveaux D. **2011**. The roles of ABA in plant-pathogen interactions. *Journal of Plant Research* **124**: 489–499.

Cardona E, Muriel SB. **2015**. Seed germination and plant development in *Escobedia grandiflora* (Orobanchaceae): evidence of obligate hemiparasitism? *Acta Biológica Colombiana* **20**: 133–140.

Cardona-Medina EC, Rafael Shubs, Daniel de Barcellos, Fernando Joner, Rubens Nodari. 2021. Effects of root hemiparasite *Escobedia grandiflora* (Orobanchaceae) on southern Brazilian grasslands: diversity, composition, and functional groups. *Journal of Vegetation Science*.

Cardona-Medina EC, Santos M, Nodari RO. **2019**. Post-seminal structure and development of the hemiparasitic plant *Escobedia grandiflora* (Orobanchaceae). *Acta Scientiarum. Biological Sciences* **41**: 2–8.

Chen L, Guo Q, Zhu Z, Wan H, Qin Y, Zhang H. **2021**. Integrated analyses of the transcriptome and small RNA of the hemiparasitic plant *Monochasma savatieri* before and after establishment of parasite-host association. *BMC Plant Biology* **21**: 90.

Chen S, Zhou Y, Chen Y, Gu J. **2018**. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)* **34**: i884–i890.

Cui S, Wakatake T, Hashimoto K, Saucet S, Toyooka K, Yoshida S, Shirasu K. **2016**. Haustorial hairs are specialized root hairs that support parasitism in the facultative parasitic plant, *Phtheirospermum japonicum*. *Plant Physiology*: 1492–1503.

D'Andrea L, Amenós M, Rodríguez-Concepción M. **2014**. Confocal Laser Scanning Microscopy Detection of Chlorophylls and Carotenoids in Chloroplasts and Chromoplasts of Tomato Fruit. In: Rodríguez-Concepción M, ed. Methods in Molecular Biology. Plant Isoprenoids. New York, NY: Springer New York, 227–232.

Delgado-Pelayo R, Gallardo-Guerrero L, Hornero-Méndez D. **2016**. Carotenoid composition of strawberry tree (*Arbutus unedo* L.) fruits. *Food Chemistry* **199**: 165–175.

Don Gilbert G. **2013**. EvidentialGene: mRNA Transcript Assembly Software.

Edgar RC. **2004**. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.

Emran S, Nawade B, Yahyaa M, Abu Nassar J, Tholl D, Eizenberg H, Ibdah M. **2020**. Broomrape infestation in carrot (*Daucus carota*): Changes in carotenoid gene expression and carotenoid accumulation in the parasitic weed *Phelipanche aegyptiaca* and its host. *Scientific Reports* **10**: 324.

Eschenmoser W, Eugster CH. **1975**. Absolute Konfiguration von Azafrin. *Helvetica Chimica Acta* **58**: 1722–1727.

Eschenmoser W, Uebelhart P, Eugster CH. **1982**. Synthese von (+)-(5S, 6S)-Azafrin-methylester; absolute Konfiguration von Aeginetinsäure und von weiteren vicinalen Apocarotin-diolen. *Helvetica Chimica Acta* **65**: 353–364.

Escobar-Tovar L, Sierra J, Hernández-Muñoz A, McQuinn RP, Mathioni S, Cordoba E, Francs-Small CC des, Meyers BC, Pogson B, León P. **2021**. Deconvoluting apocarotenoid-mediated retrograde signaling networks regulating plastid translation and leaf development. *The Plant Journal* **105**: 1582–1599.

Evanjaline M, Vr M. **2018**. Determination of bioactive components of *Caralluma umbellata* haw. (apocynaceae) by gas chromatography and mass spectroscopy analysis. *Asian Journal of Pharmaceutical and Clinical Research* **11**: 194.

Felemban A, Braguy J, Zurbriggen MD, Al-Babili S. **2019**. Apocarotenoids involved in plant development and stress response. *Frontiers in Plant Science* **10**: 1168.

Fester T, Schmidt D, Lohse S, Walter MH, Giuliano G, Bramley PM, Fraser PD, Hause B, Strack D. **2002**. Stimulation of carotenoid metabolism in arbuscular mycorrhizal roots. *Planta* **216**: 148–154.

Fiorilli V, Wang JY, Bonfante P, Lanfranco L, Al-Babili S. **2019**. Apocarotenoids: Old and New Mediators of the Arbuscular Mycorrhizal Symbiosis. *Frontiers in Plant Science* **10**: 1186.

Floss DS, Schliemann W, Schmidt J, Strack D, Walter MH. **2008**. RNA Interference-Mediated Repression of MtCCD1 in Mycorrhizal Roots of *Medicago truncatula* Causes Accumulation of C27 Apocarotenoids, Shedding Light on the Functional Role of CCD1. *Plant Physiology* **148**: 1267–1282.

Fraser PD, Romer S, Shipton CA, Mills PB, Kiano JW, Misawa N, Drake RG, Schuch W, Bramley PM. **2002**. Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proceedings of the National Academy of Sciences* **99**: 1092–1097.

Fujioka H, Samejima H, Mizutani M, Okamoto M, Sugimoto Y. **2019**. How does *Striga hermonthica* Bewitch its hosts? *Plant Signaling & Behavior* **14**.

Gelman A, Carlin JB, Stern HS, Rubin DB. **1995**. *Bayesian Data Analysis*. New York: Chapman and Hall/CRC.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, *et al.* 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8**: 1494–1512.

Hamilton CE, Gundel PE, Helander M, Saikkonen K. **2012**. Endophytic mediation of reactive oxygen species and antioxidant activity in plants: a review. *Fungal Diversity* **54**: 1–10.

Heide-Jorgensen HS, Kuijt J. **1995**. The Haustorium of the Root Parasite *Triphysaria* (Scrophulariaceae), with special reference to xylem bridge ultrastructure. *American Journal of Botany* **82**: 782–797.

Hiraga S, Sasaki K, Ito H, Ohashi Y, Matsui H. **2001**. A Large Family of Class III Plant Peroxidases. *Plant and Cell Physiology* **42**: 462–468.

Hood ME, Condon JM, Timko MP, Riopel JL. **1998**. Primary Haustorial Development of *Striga asiatica* on host and nonhost species. *Phytopathology* **88**: 70–75.

Horner HT, Healy RA, Ren G, Fritz D, Klyne A, Seames C, Thornburg RW. **2007**. Amyloplast to chromoplast conversion in developing ornamental tobacco floral nectaries provides sugar for nectar and antioxidants for protection. *American Journal of Botany* **94**: 12–24.

Hou X, Rivers J, León P, McQuinn RP, Pogson BJ. **2016**. Synthesis and Function of Apocarotenoid Signals in Plants. *Trends in Plant Science* **21**: 792–803.

Huang F-C, Molnár P, Schwab W. **2009**. Cloning and functional characterization of carotenoid cleavage dioxygenase 4 genes. *Journal of Experimental Botany* **60**: 3011–3022.

Huelsenbeck JP, Ronquist F. **2001**. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (Oxford, England)* **17**: 754–755.

Karrer P, Jucker EM. **1948**. *Carotinoide*. Birkhäuser Basel.

Kuhn R. **1935**. Plant Pigments. *Annual review of biochemistry* **4**: 479–496.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. **2018**. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **35**: 1547–1549.

Langmead B, Salzberg SL. **2012**. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Lätari K, Wüst F, Hübner M, Schaub P, Beisel KG, Matsubara S, Beyer P, Welsch R. **2015**. Tissue-Specific Apocarotenoid Glycosylation Contributes to Carotenoid Homeostasis in Arabidopsis Leaves. *Plant Physiology* **168**: 1550–1562.

Letousey P, Zélicourt AD, Santos CVD, Thoiron S, Monteau F, Simier P, Thalouarn P, Delavault P. **2007**. Molecular analysis of resistance mechanisms to *Orobanche cumana* in sunflower. *Plant Pathology* **56**: 536–546.

Liao Y, Smyth GK, Shi W. **2014**. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics (Oxford, England)* **30**: 923–930.

Losner-Goshen D, Portnoy VH, Mayer AM, Joel DM. **1998**. Pectolytic activity by the haustorium of the parasitic plant *Orobanche* L. (Orobanchaceae) in host roots. *Annals of Botany* **81**: 319–326.

Ma G, Zhang L, Matsuta A, Matsutani K, Yamawaki K, Yahata M, Wahyudi A, Motohashi R, Kato M. **2013**. Enzymatic formation of β-citraurin from β-cryptoxanthin and Zeaxanthin by carotenoid cleavage dioxygenase4 in the flavedo of citrus fruit. *Plant Physiology* **163**: 682–695.

Maddison DR, Swofford DL, Maddison WP. **1997**. NEXUS: an extensible file format for systematic information. *Systematic Biology* **46**: 590–621.

Minibayeva F, Kolesnikov O, Chasov A, Beckett RP, Lüthje S, Vylegzhanina N, Buck F, Böttger M. **2009**. Wound-induced apoplastic peroxidase activities: their roles in the production and detoxification of reactive oxygen species. *Plant, Cell & Environment* **32**: 497–508.

Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. **2013**. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Research* **41**: e121.

Moeder W, Ung H, Mosher S, Yoshioka K. **2010**. SA-ABA antagonism in defense responses. *Plant Signaling & Behavior* **5**: 1231–1233.

Mor A, Mayer AM, Levine A. **2008**. Possible peroxidase functions in the interaction between the parasitic plant, *Orobanche aegyptiaca*, and its host, Arabidopsis thaliana. *Weed Biology and Management* **8**: 1–10.

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

Muriel SBM, Cardona EC, Arias E, Gómez AG. **2015**. Indagaciones acerca del azafrán de raiz (*Escobedia grandiflora* (l.F.) Kuntze en Antioquia - Colombia: una Especie Olvidada. *Etnobiología* **13**: 85–93.

Mutuku JM, Cui S, Yoshida S, Shirasu K. **2021**. Orobanchaceae parasite–host interactions. *New Phytologist* **230**: 46–59.

Nickrent DL. **2020**. Parasitic angiosperms: How often and how many? *TAXON* **69**: 5–27.

Pueschel CM. **1979**. Ultrastructure of tetrasporogenesis in *Plamaria palmata* (Rhodophyta). *J. Phycol* **15**: 409–424.

R: The R Project for Statistical Computing. **2021**.

Rodriguez-Concepcion M, Avalos J, Bonet ML, Boronat A, Gomez-Gomez L, Hornero-Mendez D, Limon MC, Meléndez-Martínez AJ, Olmedilla-Alonso B, Palou A, *et al.* **2018**. A global perspective on carotenoids: Metabolism, biotechnology, and benefits for nutrition and health. *Progress in Lipid Research* **70**: 62–93.

Rubio A, Rambla JL, Santaella M, Gómez MD, Orzaez D, Granell A, Gómez-Gómez L. **2008**. Cytosolic and plastoglobule-targeted carotenoid dioxygenases from *Crocus sativus* are both involved in beta-ionone release. *The Journal of Biological Chemistry* **283**: 24816–24825.

Rubio-Moraga A, Rambla JL, Fernández-de-Carmen A, Trapero-Mozos A, Ahrazem O, Orzáez D, Granell A, Gómez-Gómez L. **2014**. New target carotenoids for CCD4 enzymes are revealed with the characterization of a novel stress-induced carotenoid cleavage dioxygenase gene from *Crocus sativus*. *Plant Molecular Biology* **86**: 555–569.

Ruzin SE. **1999**. *Plant microtechnique and microscopy*. New York, NY: Oxford University press.

Safford HD. **1999**. Notas sobre a ecologia e a preservação de *Nothochilus coccineus* Radlkofer (Scrophulariaceae), espécie endêmica ao Maciço do Caparaó, ES/MG. *Acta Botanica Brasilica* **13**: 175–185.

Saucet SB, Shirasu K. **2016**. Molecular Parasitic Plant–Host Interactions. *PLoS Pathogens* **12**.

Schwacke R, Ponce-Soto GY, Krause K, Bolger AM, Arsova B, Hallab A, Gruden K, Stitt M, Bolger ME, Usadel B. **2019**. MapMan4: A Refined Protein Classification and Annotation Framework Applicable to Multi-Omics Data Analysis. *Molecular Plant* **12**: 879–892.

Siddique S, Matera C, Radakovic ZS, Hasan MS, Gutbrod P, Rozanska E, Sobczak M, Torres MA, Grundler FMW. **2014**. Parasitic worms stimulate host NADPH oxidases to produce reactive oxygen species that limit plant cell death and promote infection. *Science Signaling* **7**: ra33.

Silva MAB da, Melo LVL, Ribeiro RV, Souza JPM de, Lima JCS, Martins DT de O, Silva RM da. **2010**. Levantamento etnobotânico de plantas utilizadas como anti-hiperlipidêmicas e anorexígenas pela população de Nova Xavantina-MT, Brasil. *Revista Brasileira de Farmacognosia* **20**: 549–562.

Sivakumaran A, Akinyemi A, Mandon J, Cristescu SM, Hall MA, Harren FJM, Mur LAJ. **2016**. ABA Suppresses *Botrytis cinerea* Elicited NO Production in Tomato to Influence H2O2 Generation and Increase Host Susceptibility. *Frontiers in Plant Science* **7**: 709.

speciesLink network. **2021**.

Stauder R, Welsch R, Camagna M, Kohlen W, Balcke GU, Tissier A, Walter MH. **2018**. Strigolactone Levels in Dicot Roots Are Determined by an Ancestral Symbiosis-Regulated Clade of the PHYTOENE SYNTHASE Gene Family. *Frontiers in Plant Science* **9**: 255.

Strack D, Fester T. **2006**. Isoprenoid metabolism and plastid reorganization in arbuscular mycorrhizal roots. *The New Phytologist* **172**: 22–34.

Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, Li L. **2018**. Carotenoid Metabolism in Plants: The Role of Plastids. *Molecular Plant* **11**: 58–74.

Teixeira-Costa L. **2021**. A living bridge between two enemies: haustorium structure and evolution across parasitic flowering plants. *Brazilian Journal of Botany* **44**: 165–178.

Torres-Vera R, García JM, Pozo MJ, López-Ráez JA. **2016**. Expression of molecular markers associated to defense signaling pathways and strigolactone biosynthesis during the early interaction tomato-*Phelipanche ramosa*. *Physiological and Molecular Plant Pathology* **94**: 100–107.

de Torres-Zabala M, Truman W, Bennett MH, Lafforgue G, Mansfield JW, Rodriguez Egea P, Bögre L, Grant M. **2007**. *Pseudomonas syringae* pv. tomato hijacks the *Arabidopsis* abscisic acid signalling pathway to cause disease. *The EMBO journal* **26**: 1434–1443.

Tripathy BC, Oelmüller R. **2012**. Reactive oxygen species generation and signaling in plants. *Plant Signaling & Behavior* **7**: 1621–1633.

Veeckman E, Ruttink T, Vandepoele K. **2016**. Are We There Yet? Reliably Estimating the Completeness of Plant Genome Sequences. *The Plant Cell* **28**: 1759–1768.

Verma R, Tapwal A, Kumar D, Puri S. **2019**. Assessment of antimicrobial potential and phytochemical prolifing of ethnomedicinal plant *Bergenia ciliata* (Haw.) Sternb. in western Himalaya. *Journal of Microbiology, Biotechnology and Food Sciences* **9**: 15–20.

Walter MH, Floss DS, Strack D. **2010**. Apocarotenoids: hormones, mycorrhizal metabolites and aroma volatiles. *Planta* **232**: 1–17.

Wang JY, Lin P-Y, Al-Babili S. **2021**. On the biosynthesis and evolution of apocarotenoid plant growth regulators. *Seminars in cell & developmental biology* **109**: 3–11.

Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. **2018**. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular Biology and Evolution* **35**: 543–548.

Yang S, Chou G, Li Q. **2018**. Cardioprotective role of azafrin in against myocardial injury in rats via activation of the Nrf2-ARE pathway. *Phytomedicine: International Journal of Phytotherapy and Phytopharmacology* **47**: 12–22.

Zhang, Li, Wang, Fei, Qin. **2019**. Analysis of Centranthera grandiflora Benth Transcriptome Explores Genes of Catalpol, Acteoside and Azafrin Biosynthesis. *International Journal of Molecular Sciences* **20**: 6034.

Zheng X, Mi J, Deng X, Al-Babili S. **2021**. LC-MS-Based Profiling Provides New Insights into Apocarotenoid Biosynthesis and Modifications in Citrus Fruits. *Journal of Agricultural and Food Chemistry* **69**: 1842–1851.

Figures
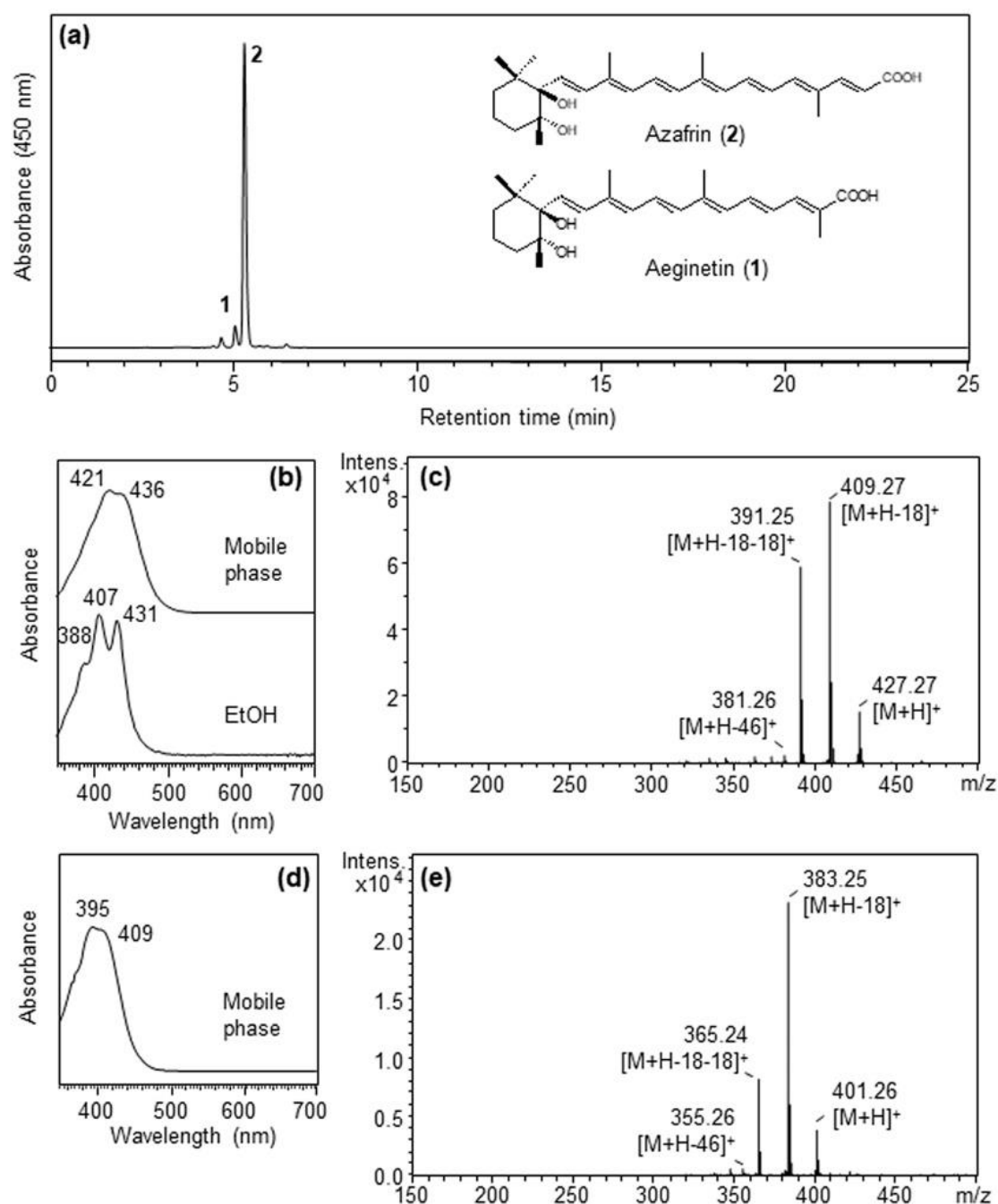
Fig. 1. *Escobedia* roots accumulate azafrin and lower levels of aeginetin. **(a)** HPLC-DAD chromatogram of *Escobedia* root pigment extract at 450 nm; **(b)** UV/Vis spectra of the major peak (**2**; azafrin) in the mobile phase and after isolation in ethanol; **(c)** mass spectrum of peak **2**; **(d)** UV/Vis spectra of the minor peak (**1**; aeginetin) in the mobile phase; **(e)** mass spectrum of peak **1**.

577
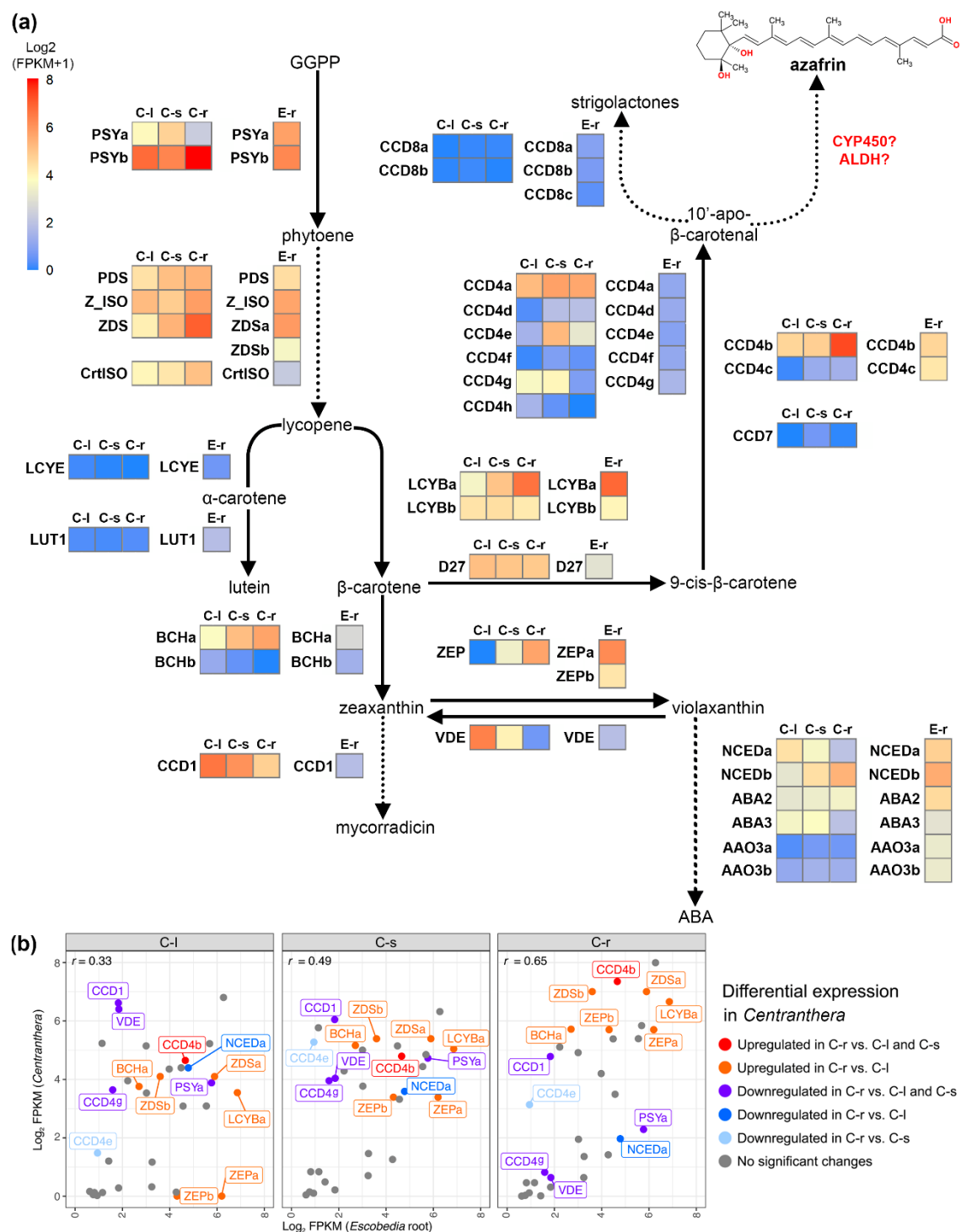
Fig. 2. Candidate genes of the azafrin biosynthetic pathway in *Escobedia* and *Centranthera*. **(a)** Proposed pathway and abundance of enzyme-encoding transcripts in *Escobedia* roots (E-r) and *Centranthera* leaf (C-l), stem (C-s) and root (C-r) tissues. Dotted lines represent multiple steps. Colours represent transcript abundance based on RNA-seq analyses (acronyms and FPKM values are listed in Supporting Information Table S3). Transcripts for the enzymes indicated in red were not identified in the root transcriptome. **(b)** Spearman correlation of gene expression values between *Escobedia* roots and the indicated *Centranthera* tissues. Scale corresponds to $\log_2$ transformed values (FPKM+1). Data for *Centranthera* were retrieved from Zhang et al. (2019) and reanalysed.
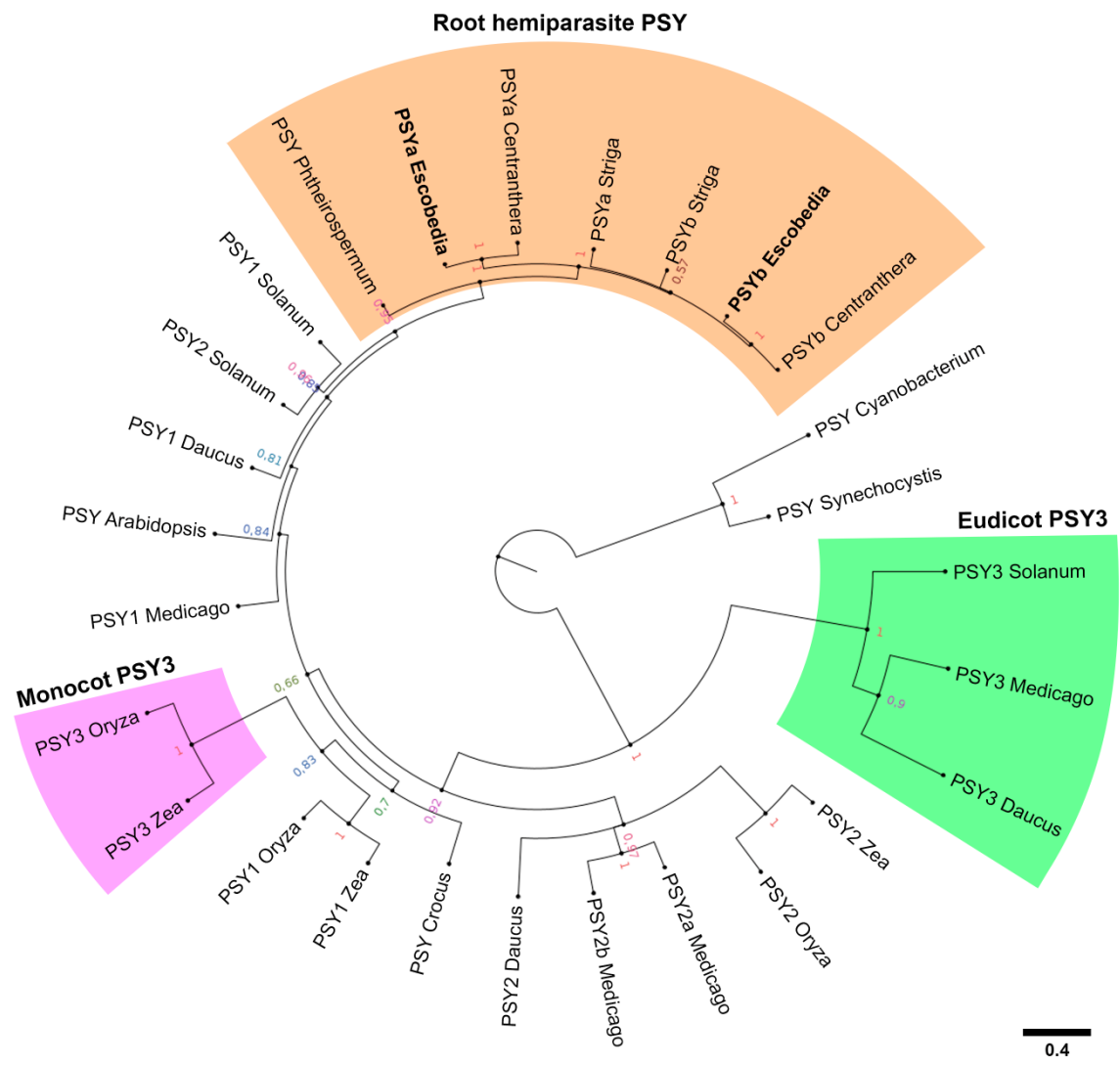
Fig. 3. Phylogenetic analysis of the phytoene synthase (PSY) family in several plants. The computed Bayesian tree includes twenty-five angiosperm species and two bacteria as outgroups. The total number of generations is 200,000. The average standard deviation of split sequences is 0.019. Distinctively separate clades are indicated with colours. *Escobedia* sequences are in bold. Gene accessions are listed in Supporting Information Table S4.

578
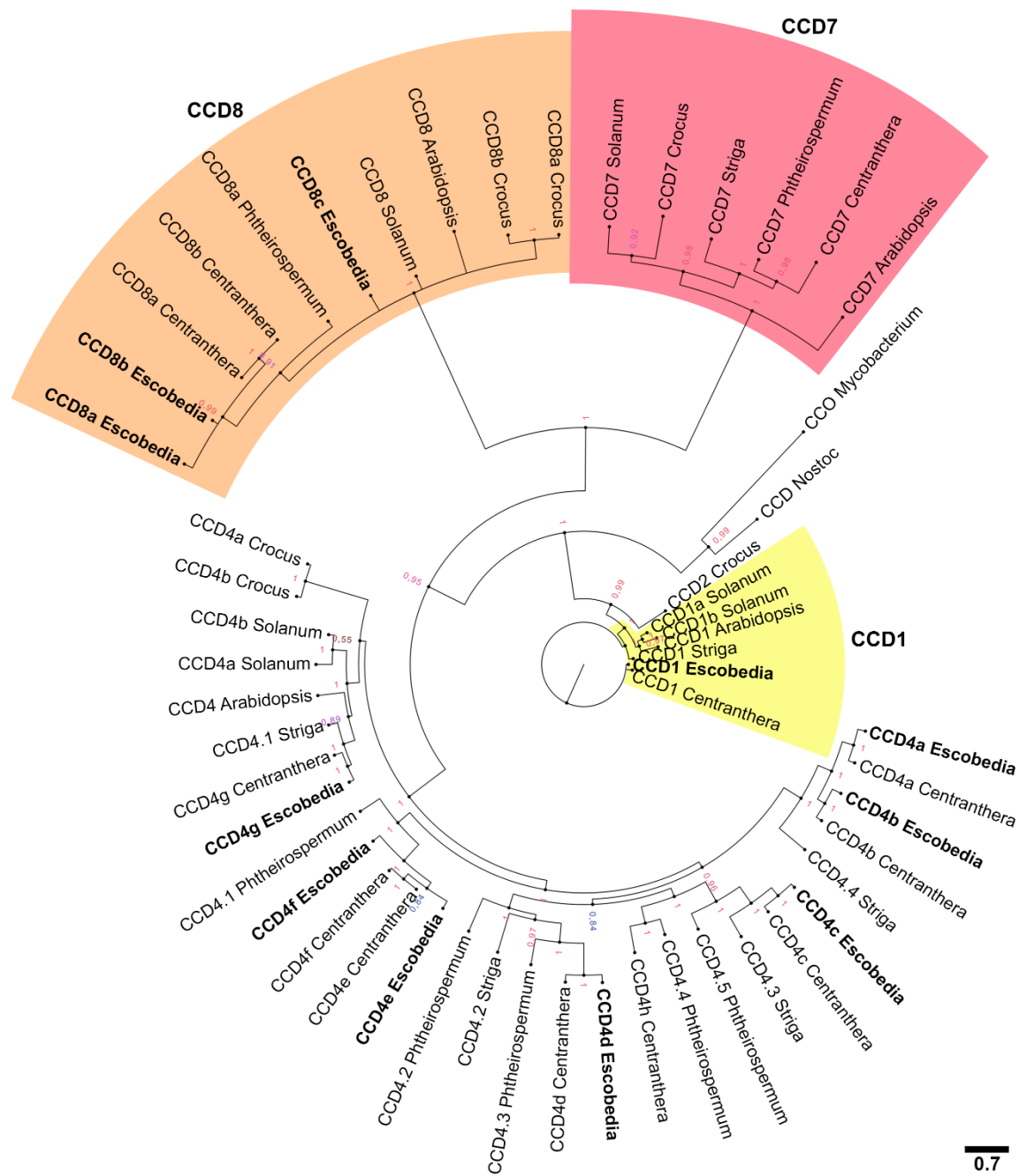
Fig. 4. Phylogenetic analysis of the carotenoid cleavage dioxygenase (CCD) family in several plants. The computed Bayesian tree includes fifty-one sequences from seven angiosperms and two bacteria as outgroups. The total number of generations is 185,000. The average standard deviation of split sequences is 0.02. Distinctively separate clades are indicated with colours. *Escobedia* are in bold. *Escobedia grandiflora* sequences are in bold text. Asterisks indicate incomplete amino acid sequences. Gene accessions are listed in Supporting Information Table S4.

579
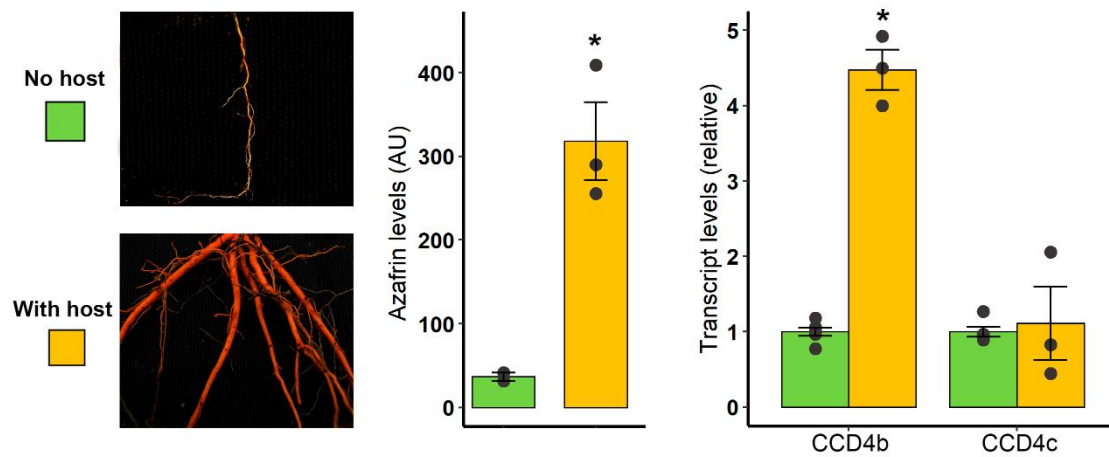
Fig. 5. Levels of azafrin and transcripts encoding CCD4 homologs in roots of *Escobedia* plants grown with (orange) and without a host (green). Asterisks represent significant differences by Tukey test ($P<0.05$). Error bars represent standard error of the means (n=3).

580
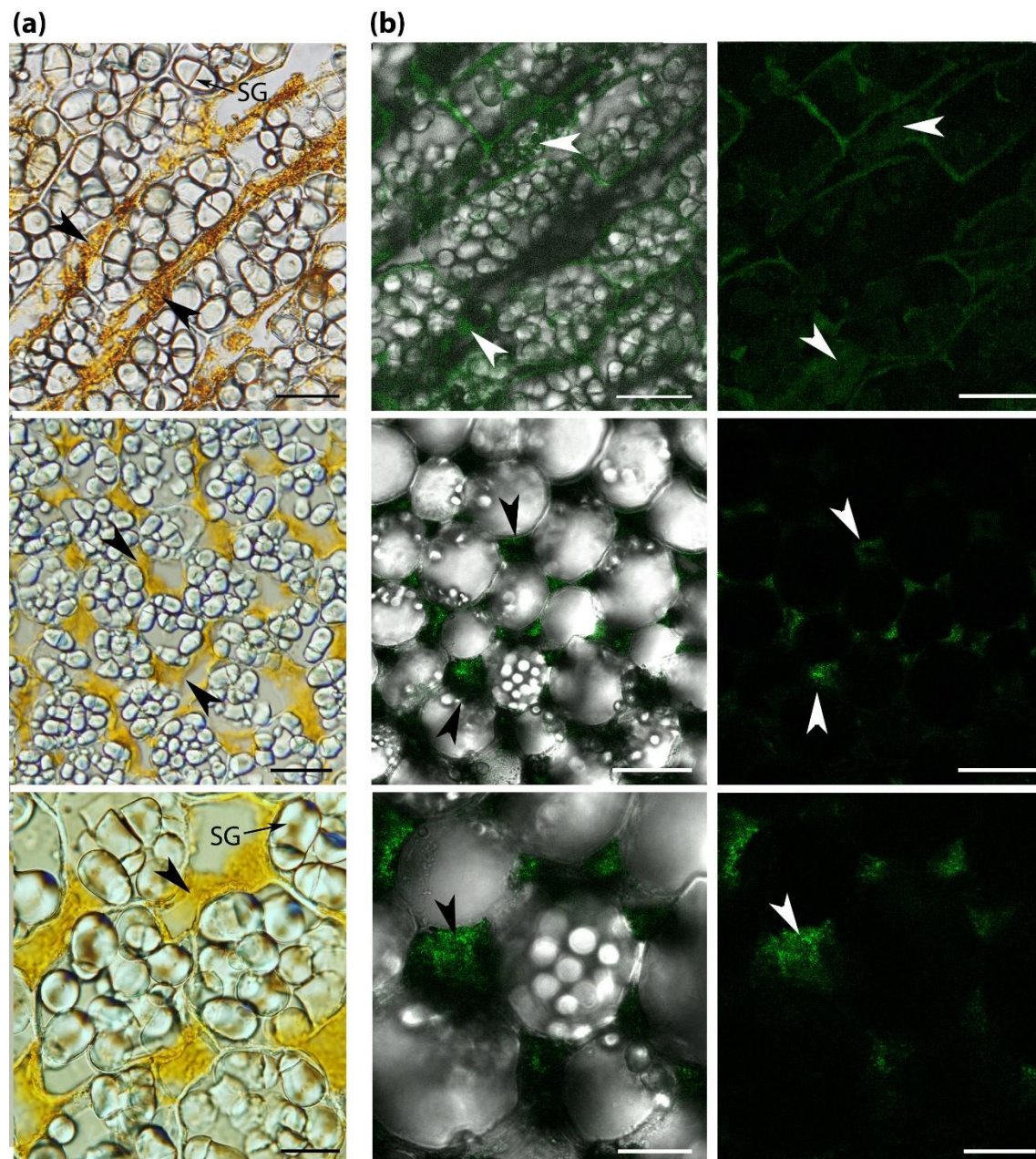
Fig. 6. Azafrin accumulates in the apoplastic space of the *Escobedia* root cortex. **(a)** Micrographs under light microscopy show azafrin as an orange pigment in the intercellular spaces (apoplast) of the cortex (arrowheads). **(b)** Micrographs under confocal microscopy show azafrin as green autofluorescence (arrowheads). Fluorescence images (right panels) are shown next to the corresponding merged micrographs of fluorescence and bright field images of the same field. Images show representative images of root longitudinal sections (upper panels) and cross-sections (central and lower panels). SG, starch grains. Scale bars: 50 µm (upper and central panels), 20 µm (lower panels).

581
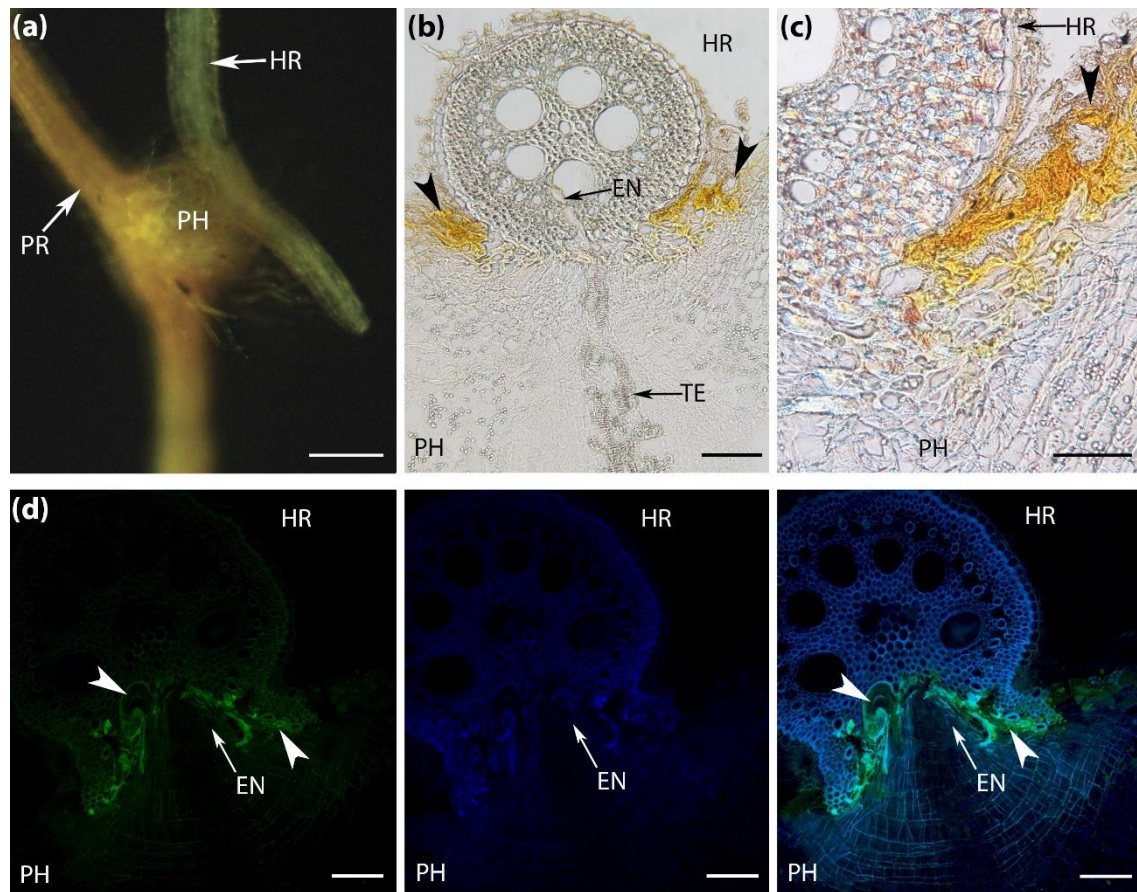
Fig. 7. Azafrin accumulates in the haustorium-host root interphase. **(a)** External view of *Escobedia* haustorium attached to a host (*Pennisetum purpureum*) root. **(b)** Representative light microscopy picture of the haustorium-host root interface showing an accumulation of the orange pigment azafrin (arrowhead). **(c)** Detail of azafrin pigment accumulation in the haustorium-host root interface (arrowhead). **(d)** Representative confocal microscopy images showing the localization of azafrin-associated fluorescence in the haustorium-host root interface. From left to right: green autofluorescence corresponding to azafrin (arrowhead); blue autofluorescence corresponding to cell walls; and merged green and blue autofluorescence. HR, host root; EN, endophyte; PH, parasite haustorium, PR, parasite root; TE, tracheary elements. Scale bars: 200 µm (a), 100 µm (b, d), 50 µm (c).