BpCYP76AD15 is involved in betaxanthin biosynthesis in bougainvillea callus

Authors:

Sho Ohno*, Rika Kokado, Rikako Makishima, Motoaki Doi

Affiliations:

Graduate School of Agriculture, Kyoto University, Kyoto, Kyoto, 606-8502 Japan

***Corresponding author: Sho Ohno

Email: ohno.sho.3c@kyoto-u.ac.jp

ORCID: orcid.org/0000-0002-0810-0327

Main conclusion: *BpCYP76AD15* is involved in betaxanthin biosynthesis in callus, but not in bracts, in bougainvillea.

Author Contribution statement

SO conceived the study. SO, MD designed the experiments. SO, RK and RM conducted the experiments. SO wrote the manuscript. All the authors read and approved the manuscript.

Abstract

Bougainvillea (*Bougainvillea peruviana*) is a climbing tropical ornamental tree belonging to Nyctaginaceae. Pigments that conferring colorful bracts in bougainvillea are betalains, and that conferring yellow color are betaxanthins. In general, for red to purple betacyanin biosynthesis, α clade CYP76AD that has tyrosine hydroxylase and DOPA oxygenase

activity is required, while for betaxanthin biosynthesis, β clade CYP76AD that has only tyrosine hydroxylase is required. To date, betaxanthin biosynthesis pathway genes have not been identified yet in bougainvillea. Since bougainvillea is phylogenetically close to four-O-clock (*Mirabilis jalapa*), and it was reported that β clade CYP76AD, *MjCYP76AD15*, is involved in floral betaxanthin biosynthesis in four-O-clock. Thus, we hypothesized that orthologous gene of *MjCYP76AD15* in bougainvillea might be involved in bract betaxanthin biosynthesis. To test the hypothesis, we attempted to identify β clade CYP76AD gene from yellow bracts by RNA-seq, however we could not. Instead, we found that callus accumulated betaxanthin and that β clade CYP76AD gene, *BpCYP76AD15*, was expressed in callus. We validated *BpCYP76AD15* function by transgenic approach (agro-infiltration and overexpression in transgenic tobacco), and it was suggested that *BpCYP76AD15* is involved in betaxanthin biosynthesis in callus, but not in bracts in bougainvillea. Interestingly, our data also indicate the existence of two pathways for betaxanthin biosynthesis (β clade CYP76AD-dependent and -independent), and the latter pathway is important for betaxanthin biosynthesis in bougainvillea bracts.

Key words:

Betalain, bract, DOPA-betaxanthin, Nyctaginaceae, Yellow pigment

Abbreviations

ADH: arogenate dehydrogenase

B6GT: betanidin-6-*O*-glucosyltransferase

cDOPA5GT: cyclo-DOPA 5-*O*-glucosyltransferase

CYP76AD: cytochrome p450 76AD

DODA: DOPA 4,5-dioxygenase

HPLC: high-performance liquid chromatography

L-DOPA: 3,4-dihydroxy-L-phenyalanine

RACE: rapid amplification of cDNA ends

Introduction

Bougainvillea (*Bougainvillea peruviana*) is a climbing tropical ornamental tree belonging to Nyctaginaceae. It is popular with colorful flowers such as red, pink, white, purple, and yellow pigmented bracts. Pigments that conferring colorful bracts in bougainvillea are betalains. Previous studies identified the structure of betalains in bougainvillea bracts such as gomphrenin I (betanidin 6-*O*-*ß*-glucoside), bougainvillein-v (betanidin 6-*O*-*ß*-sophoroside) and dopaxanthin (Heuer et al. 1994; Kugler et al. 2007).

Betalain is observed in the core Caryophyllales plants, except for Caryophyllaceae and Molluginaceae (Clement and Mabry 1996; Tanaka et al. 2008; Gandia-Herrero and Garcia-Carmona 2013). Betalains never co-exist with another water-soluble reddish pigment anthocyanins in nature. Betalain pigments are divided into two classes, betacyanin that confers red to purple color, and betaxanthin that confers yellow color.

For betacyanin biosynthesis, arogenate is converted to tyrosine by arogenate dehydrogenase (ADH) in the first step, and tyrosine is hydroxylated to form 3,4-dihydroxy-L-phenyalanine (L-DOPA) by α clade or β clade cytochrome p450 76AD (CYP76AD). Then, L-DOPA is subsequently converted to betalamic acid by DOPA 4,5-dioxygenase (DODA). Alternatively, L-DOPA is oxidized and cyclized to cyclo-DOPA

by α clade CYP76AD. cyclo-DOPA is then glycosylated by cyclo-DOPA 5-*O*-glucosyltransferase (cDOPA5GT) to form cDOPA 5-*O*-glucodside and which spontaneously condenses with betalamic acid forming the betanidin-5-glucoside (betanin). Or, cyclo-DOPA spontaneously condenses with betalamic acid to form the betanidin and then glycosylated by betanidin-5-*O*-glucosyltransferase or betanidin-6-*O*-glucosyltransferase (B6GT) at the 5′O or 6′O position to form betacyanin such as betanidin-5-*O*-glucoside (betanin) or betanidin 6-*O*-glucoside (gomphrenin I), respectively (Fig. 1) (Polturak and Aharoni 2018).

For betaxanthin biosynthesis, L-DOPA is converted to betalamic acid by DODA, and the betalamic acid is spontaneously condensed with amine or amino acid to form betaxanthins. In this procedure, β clade CYP76AD is involved, which functions as a tyrosine hydroxylase but lacks L-DOPA oxidase activity. This L-DOPA oxidase activity is a key difference between the α clade CYP76AD and the β clade CYP76AD. Identification of β clade CYP76AD is limited to three genes, BvCYP76AD5 and BvCYP76AD6 has been identified in beetroot (*Beta vulgaris*) (Polturak et al. 2016; Sunnadeniya et al. 2016), and MjCYP76AD15 has been identified for floral betaxanthin biosynthesis in *Mirabilis jalapa* (Polturak et al. 2018).

In the previous study, *BpCYP76AD1* and *BpDODA1* are involved in betacyanin biosynthesis in bougainvillea (Ohno et al. 2021). However, genes of betaxanthin biosynthesis pathway have not been identified yet. Since bougainvillea is phylogenetically close to *Mirabilis* that both belong to Nyctaginaceae, we hypothesized that *MjCYP76AD15* orthologous gene in bougainvillea might be involved in floral betaxanthin biosynthesis. To test the hypothesis, we isolated *BpCYP76AD15* and conducted functional validation using transgenic approach. Interestingly, *BpCYP76AD15*

expression was detected from callus but not from yellow bracts, indicating other gene(s) are involved in floral betaxanthin biosynthesis in bougainvillea.

Materials and methods

Plant materials

B. peruviana 'California Gold' (yellow), 'Golden Gold' (yellow), 'Thimma' (pink-white bicolor) and 'San Diego Red' (red) purchased from Yubujima (Yaeyama, Japan) or Nangoku-no-mori (Kagoshima, Japan) were used for the experiment. These plants were grown in the greenhouse in the experimental field of Kyoto University (Kyoto, Japan). Bracts at different developmental stages were collected for the study. Stage 0: around 0.5 cm length; stage 1: around 1.0 cm length; stage 2: around 1.5 cm length; and stage 3: around 3.0 cm length.

Callus induction

The procedure for callus induction was according to Anand et al. (2017). In summary, petioles and pedicels were sterilized with 5% sodium hypochlorite and placed onto modified MS media containing 6.0 ppm 2,4-D, 3.0% (w/v) sucrose and 0.55% (w/v) agar. Induced callus was subcultured to the same media by every 4-8 weeks.

High-performance liquid chromatography (HPLC) analysis

Bracts at stage 3 and callus were homogenized with a mortar and a pestle under liquid nitrogen, then 1 mL of 100 % methanol was added. Then, the extract was dried and

dissolved in water. Benthamiana tobacco leaves 5-7 days after infiltration were sampled and homogenized with a mortar and a pestle under liquid nitrogen, then 1 mL extraction solution of 0.05 % citric acid was added. These extracts were centrifuged at 4°C for 15 min at 15,000 rpm, and the supernatant was collected and 20 µL of the solution was injected in the HPLC apparatus. HPLC analysis was performed on HPLC Shimazu series, SCL-10AVP , SPD-M10AVP, CTO-10AVP, SIL-10ADVP, LC-10ADVP, FCV-10ALVP, and DGU-14A (LC solutions software; Shimazu Corp., Kyoto, Japan). A C18 column (4.6 mm x 250 mm) (Nihon Waters K.K., Tokyo, Japan) maintained at 30°C was used. The detection wavelengths were 470 nm for betaxanthin and 535 nm for betacyanin. Eluent A was 1.0 % formic acid dissolved in water and eluent B was 80 % acetonitrile dissolved in water. The analysis period for each sample was 60 min and comprised 0 min with 2.0 %, 60 min with 33 %with eluent B at a flow rate of 1 mL min^{-1}.

Isolation of BpCYP76AD15 sequence

To isolate *CYP76AD15* homologous gene in bougainvillea, one primer set (F: AAGGTTAAACTACCCCCGGGTCCGA, R: CCCGAAAGGCAATAGCTCAAAATCA) was designed based on *MjCYP76AD15* sequence (KM516798). We obtained 1,221bp of *BpCYP76AD15* partial fragment from 'Thimma' callus cDNA. Then, 5' part sequence of the *BpCYP76AD15* was determined by inverse PCR digesting with *Eco*RV, and 3' part sequence of the *BpCYP76AD15* was determined by 3' rapid amplification of cDNA ends (RACE). Primers used for identification of *BpCYP76AD15* gene were shown in Table S1.

Phylogenetic analysis

A phylogenetic tree was constructed with the open reading frames (ORFs) or putative amino acid sequences of different *CYP76AD* genes using neighbor-joining method, and bootstrap consensus inferred with 100 replicates in MEGA 11 (https://www.megasoftware.net) (Tamura et al. 2021; Saitou and Nei 1987). The accession numbers for amino acid sequences used in the phylogenesis were as follows: *Bougainvillea peruviana* BpCYP76AD1-1 (BCD59210), BpCYP76AD2 (BCD59213), BpCYP76AD15A (BDZ29443), BpCYP76AD15B (BDZ29444); *Arabidopsis thaliana* AtCYP76C2 (NP_182081), AtCYP94B3 (NP_190421); *Basella alba* BaCYP76AD13 (AJD87469); *Beta vulgaris* BvCYP76AD1 (HQ656023), BvCYP76AD5 (KM592961), BvCYP76AD6 (KM592962), BvCYP71A1-like (XP_010695999), BvG8H (XP_010676619); *Brassica oleracea* var. *italica* BoCYP79F1 (ALR85710); *Catharanthus roseus* CrCYP707A (AYM55788), CrCYP76B6 (geraniol 10-hydroxylase: CAC80883); *Cleretum bellidiforme* CbCYP76AD12 (AJD87468), CbCYP76AD16 (AJD87472); *Dianthus caryophyllus* DcCYP76AD19 (AMA07825); *Eschscholzia californica* EcCYP80B1 ((S)-N-methylcoclaurine 3'-hydroxylase: AAC39453); *Glycine max* GmCYP71D10p (AAB94588); *Glycine soja* GsCYP76C4 (KHN04086); *Glycyrrhiza echinate* GeCYP93B1 (BAA22423); *Mirabilis jalapa* MjCYP76 (ALT04745), MjCYP76AD3 (HQ656026), MjCYP76AD7 (AJD87463), MjCYP76AD15 (KM516798); *Mollugo verticillata* MvCYP76AD17 (AMA07822), MvCYP76AD18 (AMA07824), MvCYP76F84 (AMA07823); *Morus notabilis* MnCYP71A1 (XP_024025581); *Nicotiana tomentosiformis* NtCYP82E4 (ABM46920); *Opuntia ficus-indica* OfCYP76AD9 (AJD87465); *Phytolacca americana* PaCYP76AD11 (AJD87467); *Salvia miltiorrhiza* SmCYP71AP14 (AJD25152); *Solanum lycopersicum* SlCYP85A3 (NP_001234520); *Vitis vinifera* VvCYP76T21 (AOE22895), VvCYP76F14 (AOE22894); *Zea mays* ZmCYP90D2 (ACG30621).

174

175 *RNA-seq analysis*

176 Total RNA was extracted from ‘California Gold’ and ‘Golden Gold’ bracts at stage 1
177 using Sepasol RNA I Super G (Nacalai Tesque, Kyoto, Japan), and then purified with a
178 high-salt solution for precipitation (Takara Bio Inc., Otsu, Japan). Three RNA samples
179 for each cultivar were mixed equally. The mixed RNA samples were sequenced using
180 Illumina Novaseq6000 with 101-bp paired end. A total of 58,658,866 and 60,587,152 raw
181 reads were obtained for ‘California Gold’ and ‘Golden Gold’, respectively. Data from
182 each library was *de novo* assembled by Trinity (Grabherr et al. 2011) with default settings,
183 and genes homologous to *MjCYP76AD15* were searched by local blast program using
184 *MjCYP76AD15* (KM516798) as a query sequence. Also, genes which have high
185 homology with CYP76AD1, DODA, cDOPA5GT and B6GT were searched by local blast
186 program using *BpCYP76AD1-1* (LC542869), *BpCYP76AD1-like*, *BpCYP76AD2*
187 (LC542872), *BpDODA1* (LC542873), *BpDODA3* (LC542877), *MjcDOPA5GT*
188 (LC542880), *Dorotheanthus bellidiformis betanidin 6-O-glucosyltransferase*
189 (AF374004) as query sequencies.

190

191 *Gene expression analysis*

192 Total RNA was extracted from bracts and callus using Sepasol RNA I Super G (Nacalai
193 Tesque), purified with a high-salt solution for precipitation (Takara Bio Inc.) and reverse
194 transcribed with ReverTra Ace (Toyobo, Osaka, Japan), following which 1 μL of 10-fold
195 diluted RT product was used as a template for qRT-PCR. qRT-PCR was performed with
196 THUNDERBIRD SYBR qPCR Mix (Toyobo) according to the manufacturer’s
197 instructions using the LightCycler 480 system (Roche Diagnostics K.K., Tokyo, Japan).
198 The qRT-PCR was performed as follows: 95 °C for 2 min, followed by 40 cycles at 95 °C

for 10 s, 55 °C for 5 s and 72 °C for 20 s. Single-target product amplification was checked using a melting curve. Quantification was carried out with standard curve method and *BpActin* was used for a reference gene. Primers used for qRT-PCR are shown in Table S2. To analyze *BpCYP76AD15* gene expression in various tissues, RT-PCR was conducted. For RT-PCR analysis, total RNA was extracted from callus, bracts, leaves and stems using Sepasol RNA I Super G (Nacalai Tesque). After reverse transcription with ReverTra Ace (Toyobo), PCR was performed with KOD-FX Neo polymerase (Toyobo). The PCR was set as follows: 94 °C for 2 min, followed by 30–35 cycles at 98 °C for 10 s, 55 °C for 30 s, and 68 °C for 2 min. Primers used for RT-PCR are shown in Table S3.

Agroinfiltration and *production of transgenic tobacco plants*

cDNA of *BpCYP76AD15A*, *BpCYP76AD15B* and *BpDODA1* was subcloned to a pDONR221 (Invitrogen, Carlsbad, CA, USA) vector and then recombinated to a pGWB2 binary vector (Nakagawa et al. 2007) using Gateway system. All constructs were transformed into the *A. tumefaciens* EHA105 strain. The *BETA-GLUCURONIDASE* (*GUS*) coding sequence was used for control assays. Primers used for gateway cloning are shown in Table S4.

Agroinfiltration in tobacco (*Nicotiana benthamiana*) plants were performed according to Polturak et al. (2016). Agrobacterium harbouring transgene were infiltrated solely or co-infiltrated to 3-4 weeks seedling. For co-infiltration, each agrobacteria suspension which has around 1.0 at OD_{600} were mixed in a 1:1 ratio before infiltration. The tomato bushy stunt virus p19 silencing suppressor expressed in the pDGB3alpha2 35S:P19:Tnos (GB1203) vector (addgene) was also mixed before infiltration. Leaves used for subsequent pigment extraction and RT-PCR were sampled from 5-7 days post infiltration. Three biological replicates for each experiment were sampled, each

consisting of three different leaves. The primers that were used for RT-PCR are shown in Table S5.

Transgenic tobacco plants were obtained by standard *Agrobacterium tumefaciens* EHA105 strain leaf disk transformation method as described by Horsch et al. (1989) using *Nicotiana tabacum*. The generated transgenic T_0 plants harboring a copy of the transgenes were selected by genomic PCR. For genotyping, genomic DNA was extracted using SDS method. Transgenic plants overexpressing both *BpCYP76AD15* and *BpDODA1* were obtained by crossing T_0 plants. The primers that were used for genomic PCR and RT-PCR for transgene validation are shown in Table S6.

Results

Pigment analysis

For betaxanthin in bracts, a peak at 17.0 min was detected from 'California Gold', 'Golden Gold', 'San Diego Red', but almost none was detected from 'Thimma' by HPLC (Fig. 2b). This peak has maximum absorption at 473 nm and m/z was 391 by LC-MS (data not shown), indicating that this betaxanthin is DOPA-betaxanthin as reported in Polturak et al. (2018). For betacyanin, a large peak at 44.0 min was detected from 'San Diego Red' and 'Thimma', but not from 'California Gold', 'Golden Gold' on 535 nm (Fig. 2c).

When we cultured petioles and pedicels on modified MS plate supplemented with 6.0 ppm 2,4-D according to Anand et al. (2017), all cultivars produced yellow callus regardless of bract color (Fig. 2a). A peak at 15.7 min was detected from all the tested

cultivars by HPLC analysis (Fig. 2b). This peak has maximum absorption at 480 nm analyzed by a spectrophotometer. Unfortunately, we could not obtain good MS spectrum for this peak, but these results indicated different types of betaxanthin accumulated in bracts and callus.

Isolation of BpCYP76AD15 gene

In the previous study, Polturak et al. (2018) reported that *MjCYP76AD15* is involved in floral betaxanthin biosynthesis in *M. jalapa*. Therefore, we attempted to clone *MjCYP76AD15* orthologous genes from bougainvillea bracts using primers designed based on *MjCYP76AD15*, however it was unsuccessful. We also performed RNA-seq analysis on stage 1 bracts of 'California Gold' and 'Golden Gold', and *de novo* assembled the data, however genes homologous to *MjCYP76AD15* were not detected by local blast analysis (Table S7). Eventually, *BpCYP76AD15* partial transcript sequence was isolated from 'Thimma' callus using primers designed based on *MjCYP76AD15*, and the 5' and 3' sequences were determined by inverse PCR and 3' RACE. Two cDNA sequences probably allelic variants with some SNPs and different 3' UTR sequence were identified. Thus, these two variants were named *BpCYP76AD15A* and *BpCYP76AD15B*. *BpCYP76AD15A* encodes 502 putative amino acids, whereas *BpCYP76AD15B* encodes 503 putative amino acids, and shared 95% identity. Putative amino acid sequence of *BpCYP76AD15A* and *BpCYP76AD15B* share 83% and 84% homology to MjCYP76AD15, 74% and 75% to BvCYP76AD5 and 74% and 75% to BvCYP76AD6, respectively (Fig. 3a). The phylogenetic tree of CYP family proteins indicated both *BpCYP76AD15A* and *BpCYP76AD15B* were classified into CYP76AD β clade (Fig. 3b).

Expression analysis of betalain biosynthetic genes in bracts and callus

Genes with high homology to betalain biosynthetic genes were identified by *de novo* assembly of RNA-seq analysis. In addition to already identified ten genes (*BpADH*, *BpCYP76AD1*, *BpCYP76AD1-like*, *BpCYP76AD15A*, *BpCYP76AD15B*, *BpDODA1*, *BpDODA2*, *BpDODA3*, *BpDODA4* and *BpcDOPA5GT*), six genes (*BpCYP76AD2*, *BpDODA2-like*, *BpcDOPA5GT-like1*, *BpcDOPA5GT-like2*, *BpcDOPA5GT-like3* and *BpB6GT*) were newly identified. We analyzed expression patterns of these 16 genes in bracts and callus among four cultivars (Fig. 4). Though, some genes showed significantly different expression, but the difference among four cultivars were not drastic. The most remarkable differences were found in expression difference between bracts and callus. In callus, expressions of *BpCYP76AD1* and *BpcDOPA5GT-like2* were not detected, whereas in bracts, *BpCYP76AD15A* and *BpCYP76AD15B* were not detected even in yellow bract cultivars (Fig. 4). Further, to analyze *BpCYP76AD15* gene expression in leaves and stems, RT-PCR was conducted. Slight expression was detected in some leaves and stems, but expression level was lower than in callus (Fig. S1).

Functional validation of BpCYP76AD15 by transgenic approach

To validate *BpCYP76AD15* function, agro-infiltration was demonstrated (Fig. 5a). Expression of introduced gene was confirmed by RT-PCR (Fig. 5b). The leaf color of *BpCYP76AD15A*, *BpCYP76AD15B* or *BpDODA1* infiltrated with p19 was not changed dramatically, however the leaf color of co-infiltration of *BpCYP76AD15A*, *BpDODA1* and p19, or *BpCYP76AD15B*, *BpDODA1* and p19 turned yellow (Fig. 5a). Color of leaf extract was yellow (Fig. 5c) and peaks at 15.7 min and 19.0 min at 470 nm were detected by HPLC (Fig. 5d). The peak at 15.7 min has maximum absorption at 480 nm, indicating this betaxanthin is identical to that of callus.

Next, we produced transgenic tobacco plants (*N. tabacum*) expressing both *BpCYP76AD15B* and *BpDODA1* by crossing single overexpression lines. Expression of introduced genes was confirmed by RT-PCR (Fig. 6a). Flower color of *BpCYP76AD15B* and *BpDODA1* overexpressing plants was yellowish compared to *BpCYP76AD15B* or *BpDODA1* single overexpressing plants (Fig. 6b). Flower extracts were also yellow in both *BpCYP76AD15B* and *BpDODA1* overexpressing flowers, while flower extracts of single overexpression of *BpCYP76AD15B* or *BpDODA1* did not turn yellow (Fig. 6c). Several peaks at 470 nm were detected from *BpCYP76AD15B* and *BpDODA1* overexpressing flowers by HPLC (Fig. 6d). From these results, it was demonstrated that *BpCYP76AD15* is involved in betaxanthin biosynthesis in callus, but not in bracts.

Discussion

In this study, we demonstrated the involvement of *BpCYP76AD15*, β clade *CYP76AD*, in betaxanthin biosynthesis in bougainvillea callus. Recently, in addition to BvCYP76AD1 (Hatlestad et al. 2012) and MjCYP76AD3 (Sunnadeniya et al. 2016), several α clade *CYP76AD* genes were functionally validated such as *AmCYP76AD1* in *Amaranthus tricolor* (Chang et al. 2021) and *CqCYP76AD1-1* in *Chenopodium quinoa* (Imamura et al. 2018). However, compared to α clade *CYP76AD* genes, less is known about β clade *CYP76AD* genes. To best of our knowledge, functional characterization of β clade *CYP76AD* has been limited to BvCYP76AD5, BvCYP76AD6 and MjCYP76AD15 (Polturak et al. 2016, 2018; Sunnadeniya et al. 2016). Here, β clade *CYP76AD* gene in bougainvillea, *BpCYP76AD15* was isolated from callus cDNA that shares high homology

with *MjCYP76AD15* (Fig. 3). Ectopic overexpression of *BpCYP76AD15* in combination with *BpDODA1* resulted in betaxanthin accumulation by agro-infiltration and transgenic plants (Fig. 5 and 6), indicating *BpCYP76AD15* is involved in betaxanthin biosynthesis. Interestingly, *BpCYP76AD15* gene expression was detected only in callus but not in bracts even in yellow bract cultivars by qRT-PCR and RT-PCR (Fig. 4, Fig S1). In addition, not only *BpCYP76AD15* but also other β clade *CYP76AD*-like gene was not detected from RNA-seq analysis of stage 1 bracts of 'California Gold' and 'Golden Gold' (Table S7). Therefore, these results indicated that *BpCYP76AD15* is involved in betaxanthin biosynthesis in callus, but not in bracts.

In betalain biosynthesizing plants such as portulaca (*Portulaca* sp.), beets, celosia (*Celosia argentea*) and quinoa, betalains were detected from callus or cultured cells (Kishima et al. 1991; Nazmul et al. 2003; Leathers et al. 1992; Akita et al. 2000; Guadarrama-Flores et al. 2015, Henarejos-Escudero et al. 2018). In these species, both betacyanin and betaxanthin are detected from callus or cultured cells, and environmental stimuli such as light, auxins, nitrogen sources and Fe^{2+} concentration effected on betalain amounts (Kishima et al. 1991, 1995; Leathers et al. 1992). However, in the case of bougainvillea, all cultivars produced yellow callus derived from betaxanthin regardless of bract color (Fig. 2a). This indicates that betaxanthin may have some physiological function in callus, and it will be elucidated in the future.

Our data also indicates differential regulation of betalain biosynthesis between bracts and callus. In beets, *BvCYP76AD5* and *BvCYP76AD6* expression was detected from red hypocotyl (Polturak et al. 2016; Sunnadeiya et al. 2016), and *MjCYP76AD15* expression was detected from petals, stamen, anthers, stigma and leaf in *M. jalapa* (Polturak et al. 2018). Thus, in these species, α clade and β clade *CYP76AD* genes expressed in the same organ, however, in bougainvillea, clear organ specificity was

detected that α clade *CYP76AD* expression was detected in bracts while β clade *CYP76AD* expression was detected in callus (Fig. 4, Fig S1).

Interestingly, our data also indicate the existence of two pathways for betaxanthin biosynthesis (β clade CYP76AD-dependent and -independent), and the latter pathway is important for betaxanthin biosynthesis in bougainvillea bracts. DeLoache et al. (2015) reported that F309L mutation in BvCYP76AD1 abolished L-DOPA oxidase activity and enhance tyrosine hydroxylase activity to produce high amount of betaxanthin. We analyzed *BpCYP76AD1* cDNA sequence expressing in 'California Gold' and 'Golden Gold' bracts, and identified two cDNA sequences, *BpCYP76AD1-2* and *BpCYP76AD1-3*, from both 'California Gold' and 'Golden Gold', that was corresponded to genomic PCR result in Ohno et al. (2021). *BpCYP76AD1-2* in yellow cultivars has three SNPs with 'Thimma' *BpCYP76AD1-2*, while *BpCYP76AD1-3* in yellow cultivars was identical to 'Thimma' *BpCYP76AD1-3*. However, F309 position was conserved in the both *BpCYP76AD1* cDNA sequences in yellow cultivars (data not shown). Therefore, further analysis is required to identify the genes involved in betaxanthin biosynthesis in bougainvillea bracts.

Supplementary data

Fig. S1. Expression of *BpCYP76AD15* genes in various tissues

Table S1 Primers used for the isolation of *BpCYP76AD15* gene

Table S2 Primers used for qRT-PCR

Table S3 Primers used for RT-PCR

Table S4 Primers used for gateway cloning

Table S5 Primers used for validation of transgene expression

Table S6 Primers used for genotyping of transgenic tobacco plants

Table S7 TPM values of betalain biosynthesis related genes in the bracts of 'California Gold' and 'Golden Gold'

Acknowledgements

Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics (Japan).

Conflict of interests

The authors declare no conflict of interests.

Data availability statement

The data that support the findings of this study are openly available in the Genbank and Sequence Read Archive under the accession number PRJDB15488 (https://www.ncbi.nlm.nih.gov/bioproject/PRJDB15488). Accession numbers: *BpCYP76AD15A* cDNA (LC760651), BpCYP76AD15B cDNA (LC760652). RNA-seq: 'California Gold' bract at stage 1 (DRR452711); 'Golden Gold' bract at stage 1

(DRR452712).

References

Akita T, Hina Y, Nishi T (2000) Production of betacyanins by a cell suspension culture of table beet (*Beta vulgaris* L.). Biosci Biotechnol Biochem 64:1807–1812. https://doi.org/10.1271/bbb.64.1807

Anand P, Singh KP, Prasad KV, Kaur C, Verma AK (2017) Betalain estimation and callus induction in different explants of *Bougainvillea* spp. Indian J Agric Sci 87:191–196. https://epubs.icar.org.in/index.php/IJAgS/article/view/67564

Chang YC, Chiu YC, Tsao NW, Chou YL, Tan CM, Chiang YH, Liao PC, Lee YC, Hsieh LC, Wang SY, Yang JY (2021) Elucidation of the core betalain biosynthesis pathway in *Amaranthus tricolor*. Sci Rep 11:6086. https://doi.org/10.1038/s41598-021-85486-x

Clement J, Mabry T (1996) Pigment evolution in the *Caryophyllales*: a systematic overview. Botanica Acta 109:360–367. https://doi.org/10.1111/j.1438-8677.1996.tb00584.x

DeLoache W, Russ Z, Narcross, L, Gonzales AM, Martin VJJ, Dueber JE (2015) An enzyme-coupled biosensor enables (*S*)-reticuline production in yeast from glucose. Nat Chem Biol 11:465–471. https://doi.org/10.1038/nchembio.1816

Gandía-Herrero F and García-Carmona F (2013) Biosynthesis of betalains: yellow and violet plant pigments. Trends Plant Sci 18:334–343. https://doi.org/10.1016/j.tplants.2013.01.003

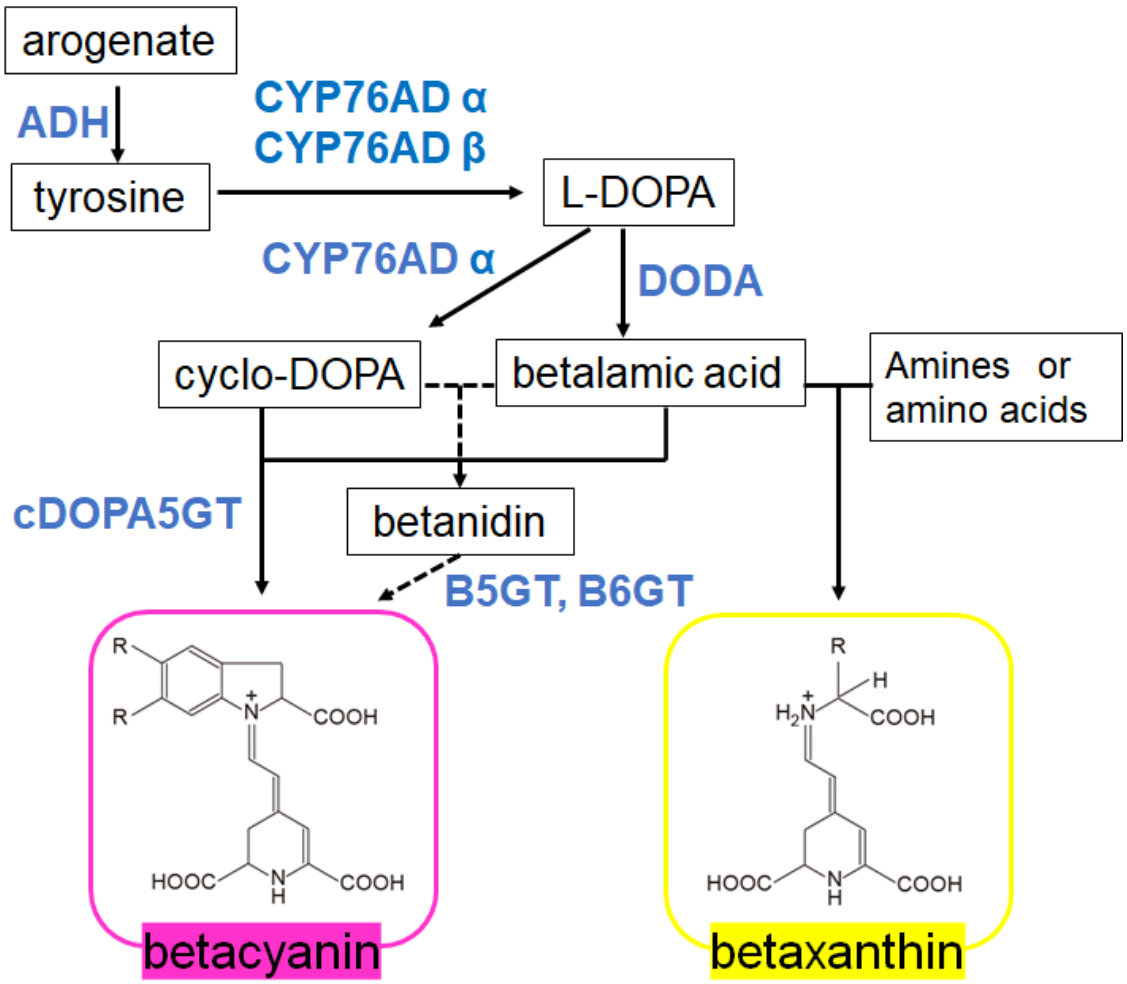
423 Grabherr MG, Haas BJ, Yassour M et al. (2011) Full-length transcriptome assembly from
424 RNA-Seq data without a reference genome. Nature Biotechnol 29:644–652.
425 https://doi.org/10.1038/nbt.1883
426 Guadarrama-Flores B, Rodríguez-Monroy M, Cruz-Sosa F, García-Carmona F, Gandía-
427 Herrero F (2015) Production of dihydroxylated betalains and dopamine in cell
428 suspension cultures of *Celosia argentea* var. plumosa. J Agric Food Chem 63:2741–
429 2749. https://doi.org/10.1021/acs.jafc.5b00065
430 Hatlestad GJ, Sunnadeniya RM, Akhavan NA, Gonzalez A, Goldman IL, McGrath JM,
431 Lloyd AM (2012) The beet *R* locus encodes a new cytochrome P450 required for red
432 betalain production. Nat Genet 44:816–820. https://doi.org/10.1038/ng.2297
433 Henarejos-Escudero P, Guadarrama-Flores B, Guerrero-Rubio MA, Gómez-Pando LR,
434 García-Carmona F, Gandía-Herrero F (2018) Development of betalain producing
435 callus lines from colored quinoa varieties (*Chenopodium quinoa* Willd). J Agric
436 Food Chem 66:467–474. https://doi.org/10.1021/acs.jafc.7b04642
437 Heuer S, Richter S, Metzger JW, Wray V, Nimtzt M, Strack D (1994) Betacyanins from
438 bracts of *Bougainvillea glabra*. Phytochemistry 37:761–767.
439 https://doi.org/10.1016/S0031-9422(00)90354-6
440 Horsch RB, Fry J, Hoffmann N, Neidermeyer J, Rogers SG, Fraley RT (1989) Leaf disc
441 transformation. In: Gelvin SB, Schilperoort RA, Verma DPS, eds. Plant Molecular
442 Biology Manual, pp 63–71. https://doi.org/10.1007/978-94-009-0951-9_5
443 Imamura T, Takagi H, Miyazato A, Ohki S, Mizukoshi H, Mori M (2018) Isolation and
444 characterization of the betalain biosynthesis gene involved in hypocotyl
445 pigmentation of the allotetraploid *Chenopodium quinoa*. Biochem Biophys Res
446 Commun 496:280–286. https://doi.org/10.1016/j.bbrc.2018.01.041
447 Kishima Y, Nozaki K, Akashi R, Adachi T (1991) Light-inducible pigmentation

448 in *Portulaca* callus; selection of a high betalain producing cell line. Plant Cell
449 Rep 10:304–307. https://doi.org/10.1007/BF00193147
450 Kishima Y, Shimaya A, Adachi T (1995) Evidence that blue light induces betalain
451 pigmentation in *Portulaca* callus. Plant Cell Tiss Organ Cult 43:67–70.
452 https://doi.org/10.1007/BF00042673
453 Kugler F, Florian CS, Reinhold C (2007) Characterisation of betalain patterns of
454 differently coloured inflorescences from *Gomphrena globose* L. and *Bougainvillea*
455 sp. by HPLC-DAD-ESI-MS^n. Anal Bioanal Chem 387:637–648.
456 https://doi.org/10.1007/s00216-006-0897-0
457 Leathers RR, Davin C, Zrÿd JP (1992) Betalain producing cell cultures of *Beta vulgaris*
458 L. var. bikores monogerm (red beet). In Vitro Cell Dev Biol –Plant 28:39–45.
459 https://doi.org/10.1007/BF02823016
460 Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamurai M, Niwa Y, Toyoda K, Matsuoka
461 K, Jimbo T, Kumura T (2007) Development of series of gateway binary vectors,
462 pGWBs, for realizing efficient construction of fusion genes for plant transformation.
463 J Biosci Bioeng 103:34–41. https://doi.org/10.1263/jbb.104.34
464 Nazmul M, Bhuiyan H, Adachi T (2003) Stimulation of betacyanin synthesis through
465 exogenous methyl jasmonate and other elicitors in suspension-cultured cells of
466 Portulaca. J Plant Physiol 160:1117–1124. https://doi.org/10.1078/0176-1617-01044
467 Ohno S, Makishima R, Doi M (2021) Post-transcriptional gene silencing
468 of *CYP76AD* controls betalain biosynthesis in bracts of bougainvillea. J Exp Bot
469 72:6949–6962. https://doi.org/10.1093/jxb/erab340
470 Polturak G, Aharoni A (2018) “La Vie en Rose”: biosynthesis, sources, and applications
471 of betalain pigments. Mol Plant 11:7–22.
472 https://doi.org/10.1016/j.molp.2017.10.008

473 Polturak G, Heinig U, Grossman N, Battat M, Leshkowitz D, Malitsky S, Rogachev I,
474 Aharoni A (2018) Transcriptome and metabolic profiling provides insights into
475 betalain biosynthesis and evolution in *Mirabilis jalapa*. Mol Plant 11:189–204.
476 https://doi.org/10.1016/j.molp.2017.12.002
477 Polturak G, Breitel D, Grossman N, Sarrion-Perdigones A, Weithorn E, Pliner M, Orzaez
478 D, Granell A, Rogachev I, Aharoni A (2016) Elucidation of the first committed step
479 in betalain biosynthesis enables the heterologous engineering of betalain pigments
480 in plants. New Phytol 210:269–283. https://doi.org/10.1111/nph.13796
481 Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing
482 phylogenetic trees. Mol Biol Evol 4:406–425.
483 https://doi.org/10.1093/oxfordjournals.molbev.a040454
484 Sunnadeniya R, Bean A, Brown M, Akhavan N, Hatlestad G, Gonzalez A, Symonds VV
485 Lloyd A (2016) Tyrosine hydroxylation in betalain pigment biosynthesis is
486 performed by cytochrome P450 enzymes in beets (*Beta vulgaris*). PLoS ONE 11:1–
487 16. https://doi.org/10.1371/journal.pone.0149417
488 Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics
489 Analysis Version 11. Mol Biol Evol 38:3022–3027.
490 https://doi.org/10.1093/molbev/msab120
491 Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins,
492 betalains and carotenoids. Plant J 54:733–749. https://doi.org/10.1111/j.1365-
493 313X.2008.03447.x
494

495 **Fig. 1**

496

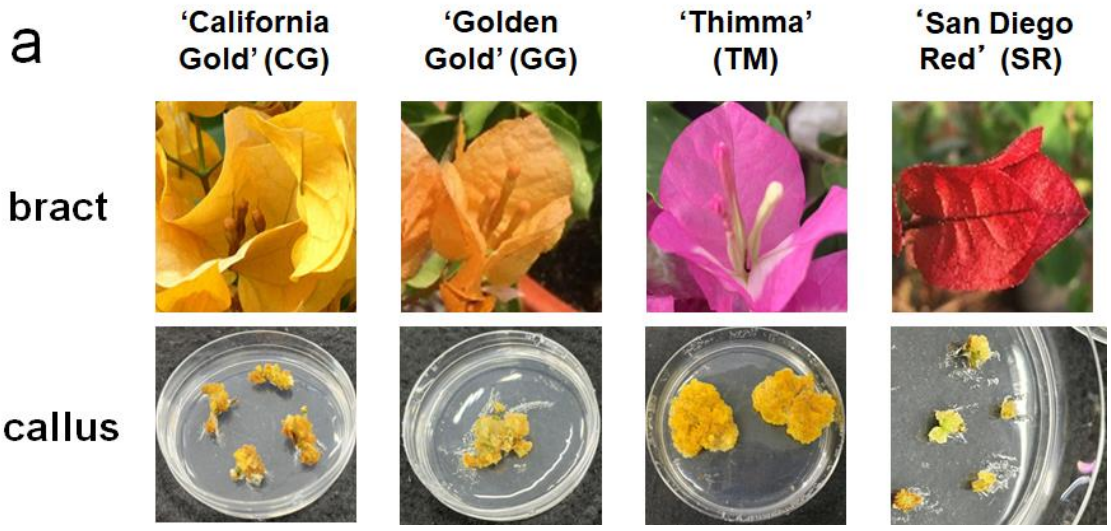


497

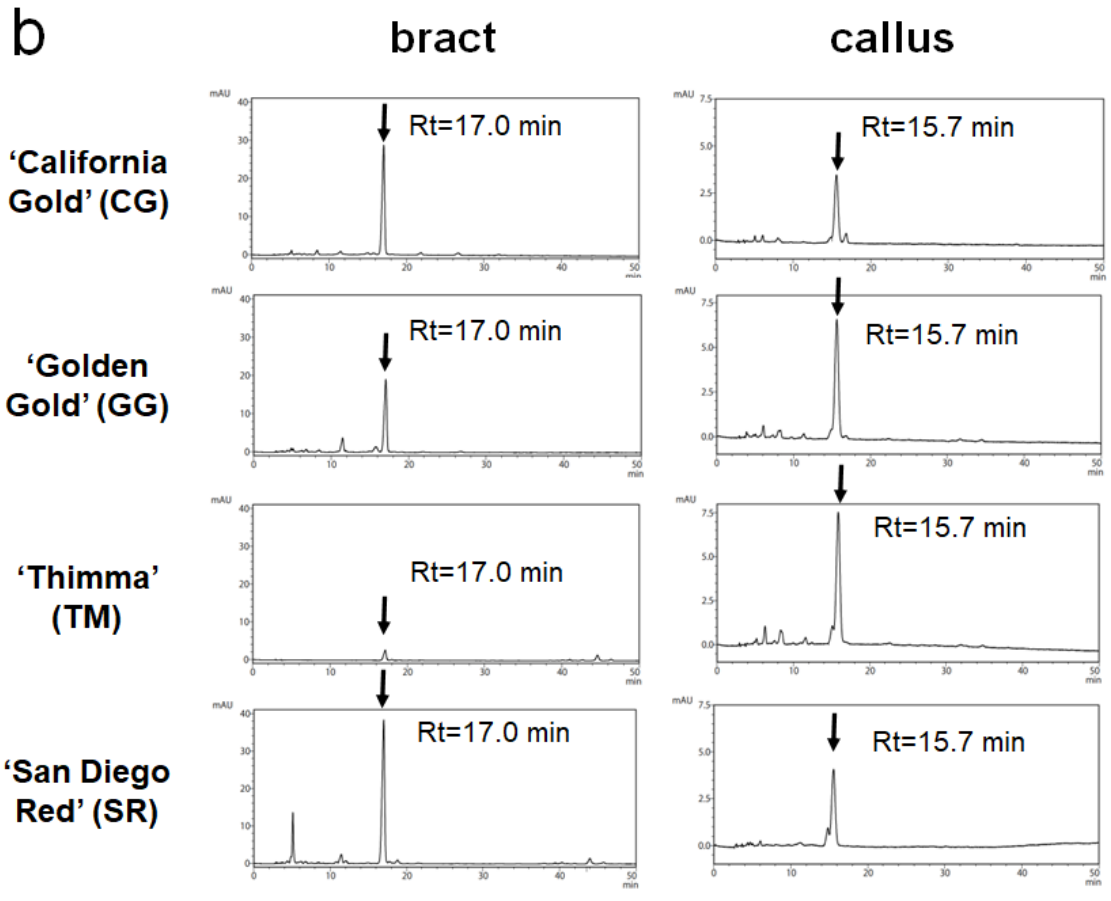
498 Fig.1 Betalain biosynthetic pathway. ADH: arogenate dehydrogenase, CYP76AD:
499 cytochrome p450 76AD, DODA: L-DOPA 4,5-dioxygenase, cDOPA5GT: cyclo-DOPA
500 5-*O*-glucosyltransferase, B5GT: betanidin-5-*O*-glucosyltransferase, B6GT: betanidin-6-
501 *O*-glucosyltransferase

502

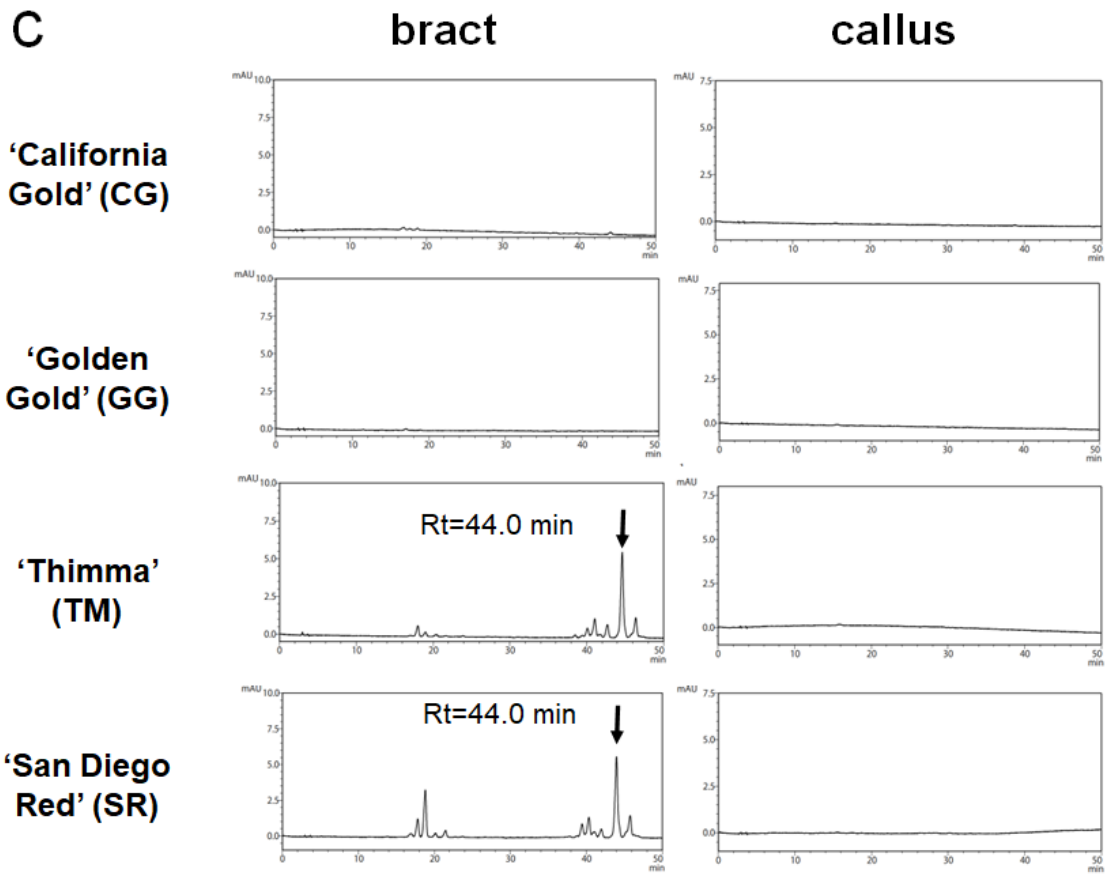
503 **Fig. 2**



504



505

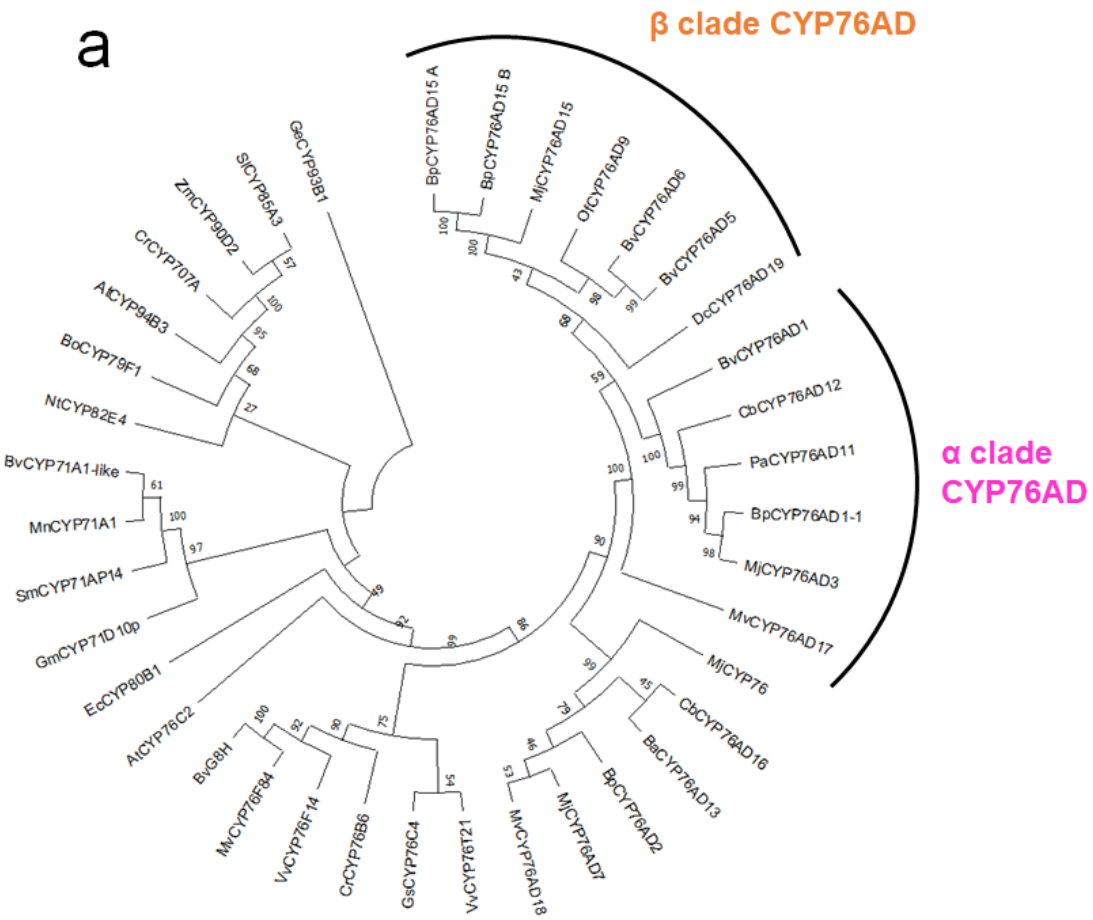


506

507 Fig. 2 Pigment analysis in bracts and callus. **a** Cultivars used in this study. **b** HPLC
508 analysis on 470 nm for betaxanthin in bracts and callus. **c** HPLC analysis on 535 nm for
509 betacyanin in bracts and callus

510

Fig. 3


a
β clade CYP76AD
α clade CYP76AD
BpCYP76AD15 A
BpCYP76AD15 B
MjCYP76AD15
OfCYP76AD9
BvCYP76AD6
BvCYP76AD5
DcCYP76AD19
BvCYP76AD1
CbCYP76AD12
PaCYP76AD11
BpCYP76AD1-1
MjCYP76AD3
MvCYP76AD17
MjCYP76
CbCYP76AD16
BaCYP76AD13
BpCYP76AD2
MjCYP76AD7
MvCYP76AD18
VvCYP76T21
GsCYP76C4
CrCYP76B6
VvCYP76F14
MvCYP76F84
BvG8H
AtCYP76C2
EcCYP80B1
GmCYP71D10p
SmCYP71AP14
MnCYP71A1
BvCYP71A1-like
NtCYP82E4
BoCYP79F1
AtCYP94B3
CrCYP707A
ZmCYP90D2
StCYP85A3
GsCYP93B1

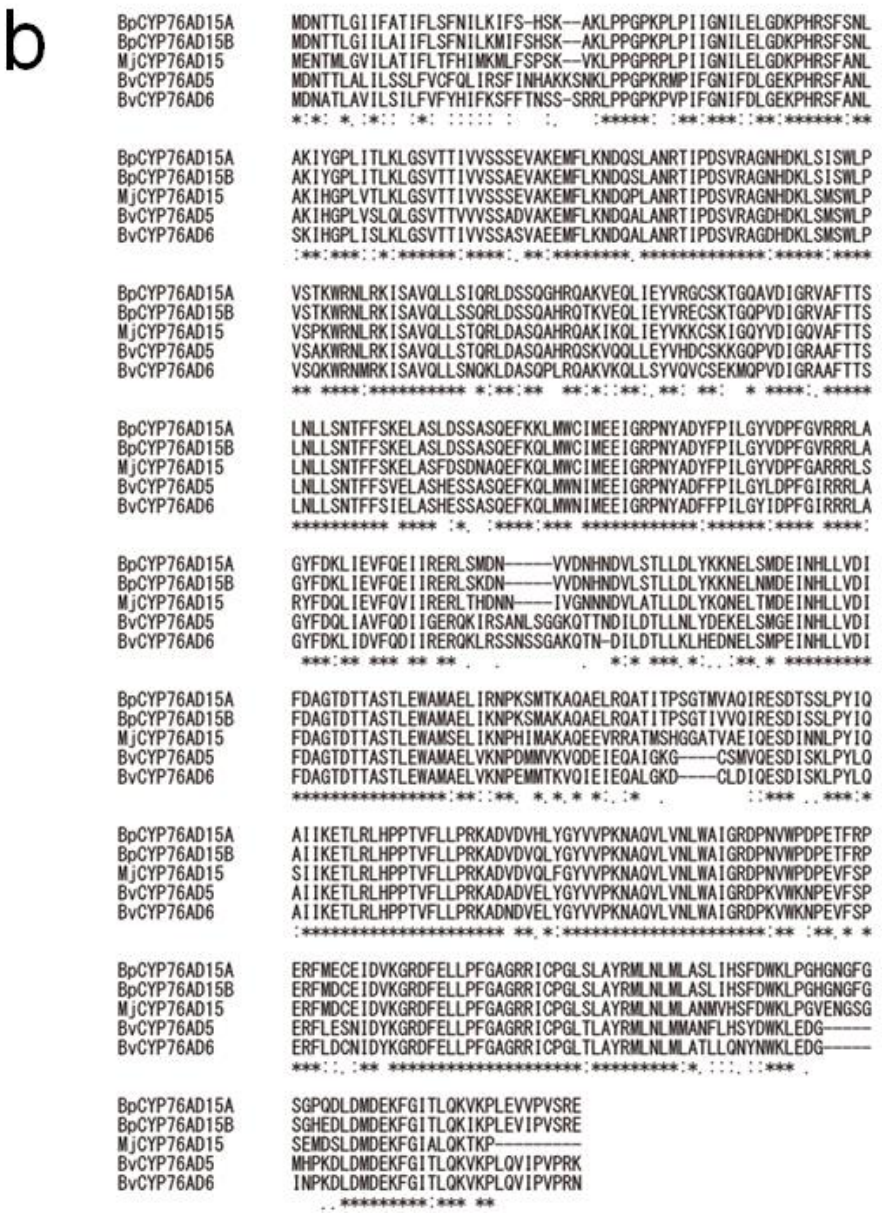


Fig. 3 Phylogenetic analysis of BpCYP76AD15. **a** Phylogenetic tree for plant CYP76 family protein. Full amino acid sequences were aligned using ClustalW, and the tree was constructed by the neighbor-joining method. Bootstrap values obtained with 100 repetitions are indicated on each branch. **b** Alignment of putative amino acid sequence of BpCYP76AD15A, BpCYP76AD15B, MjCYP76AD15, BvCYP76AD5 and BvCYP76AD6

Fig. 4

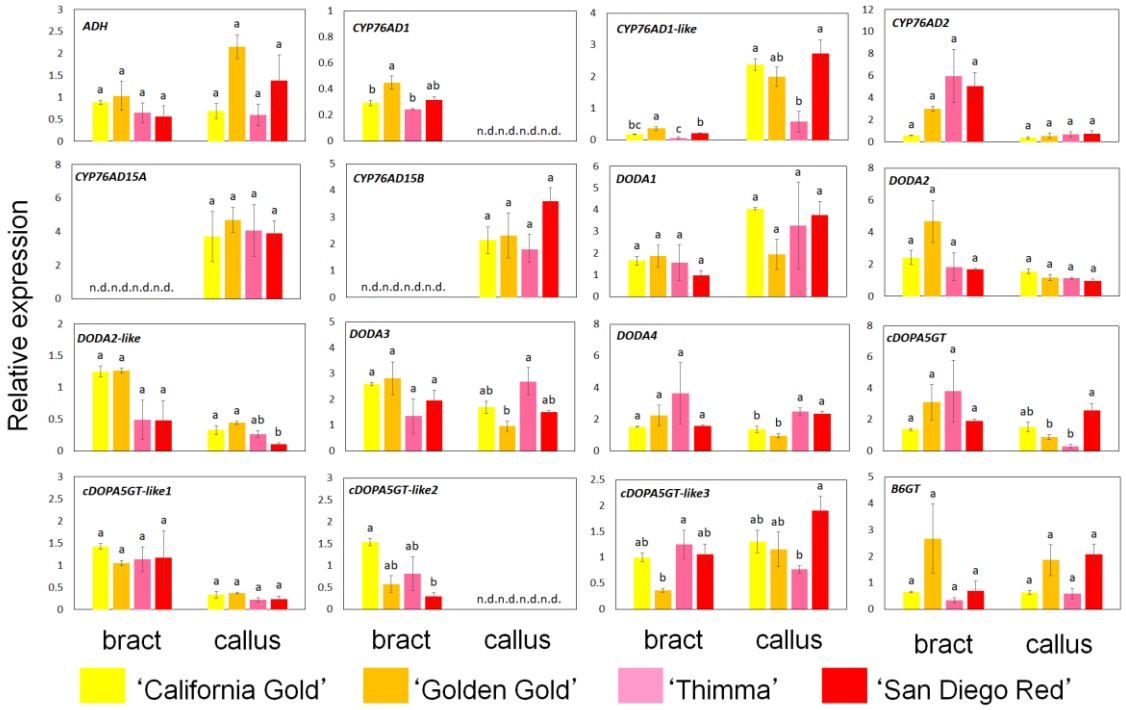
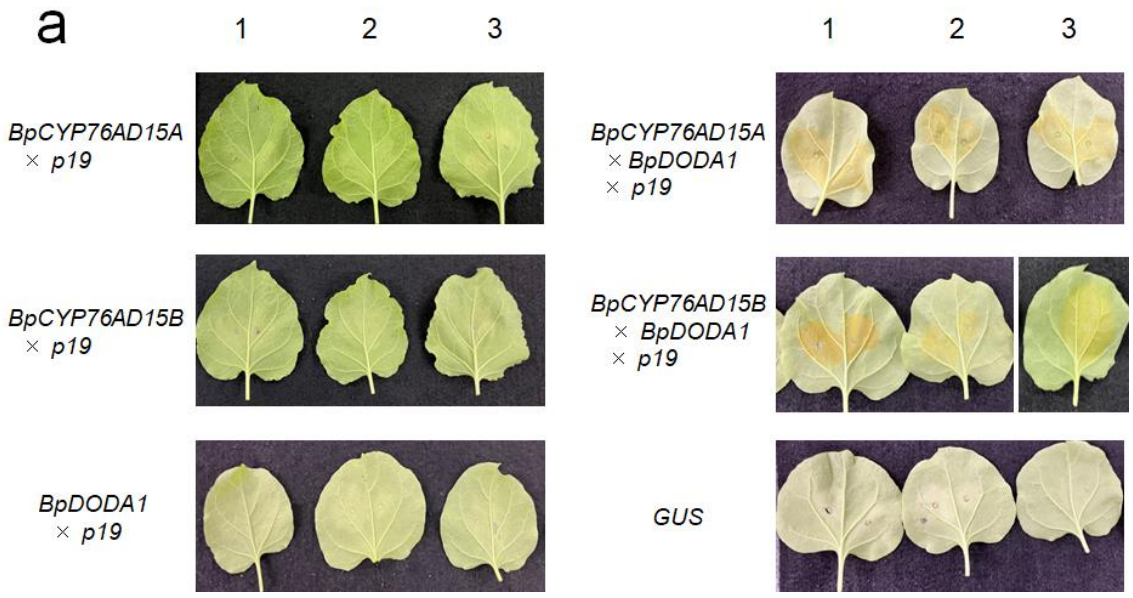


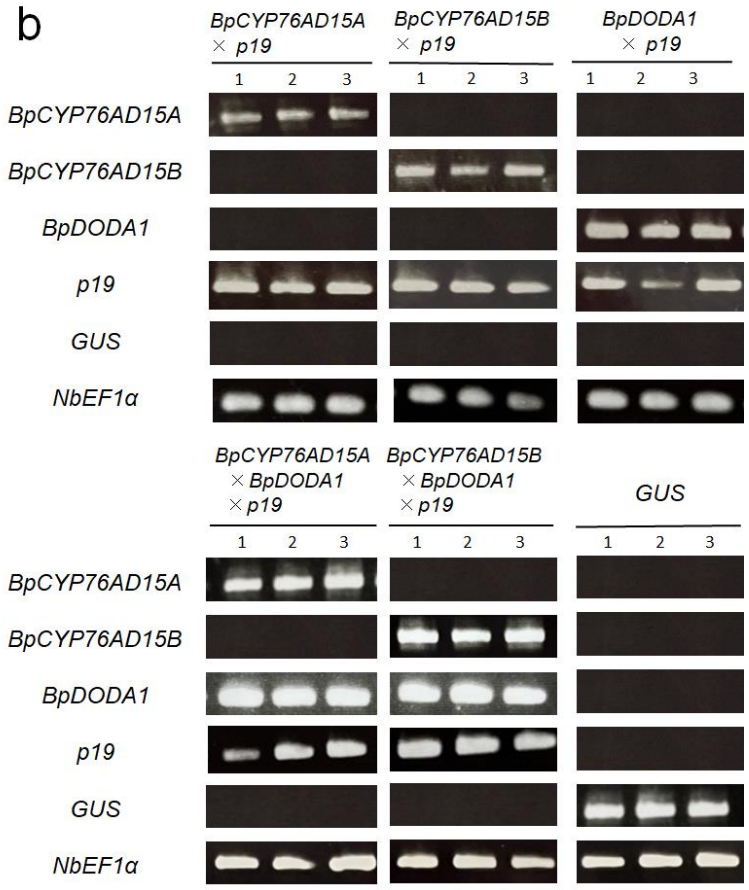
Fig. 4 Relative expression of betalain biosynthetic genes in 1-2 cm bracts and in callus among four cultivars. *BpActin* was used as an internal control. Bars represent standard errors (n = 3). n.d. indicates not detected. Different letters above the bars indicate significant differences by Tukey's HSD test ($P < 0.05$)

539 **Fig. 5**

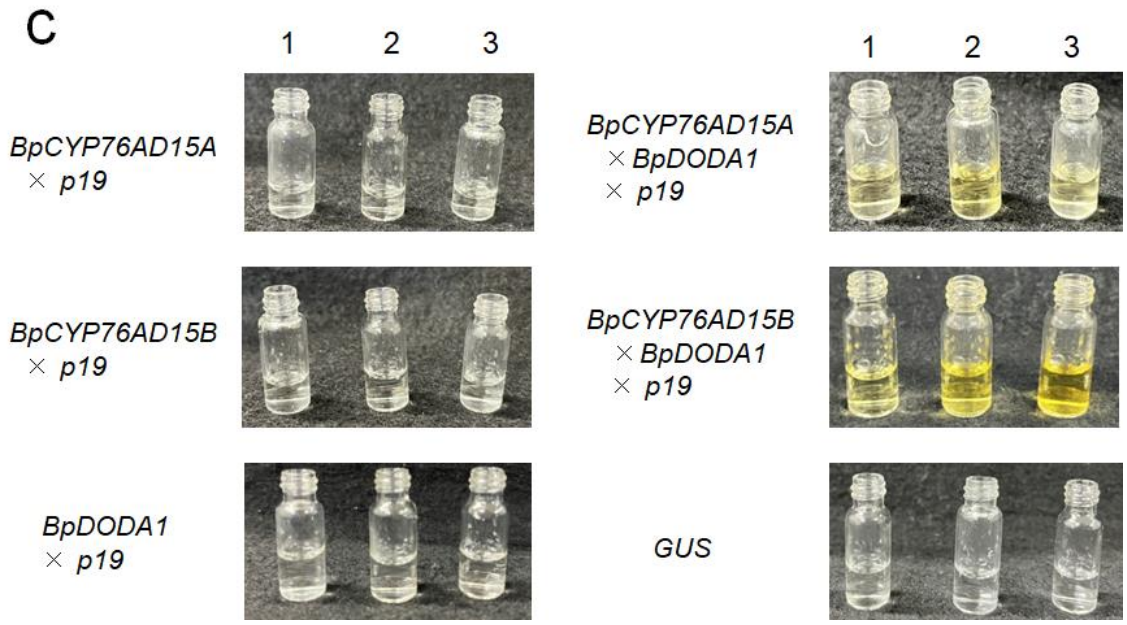


540

541

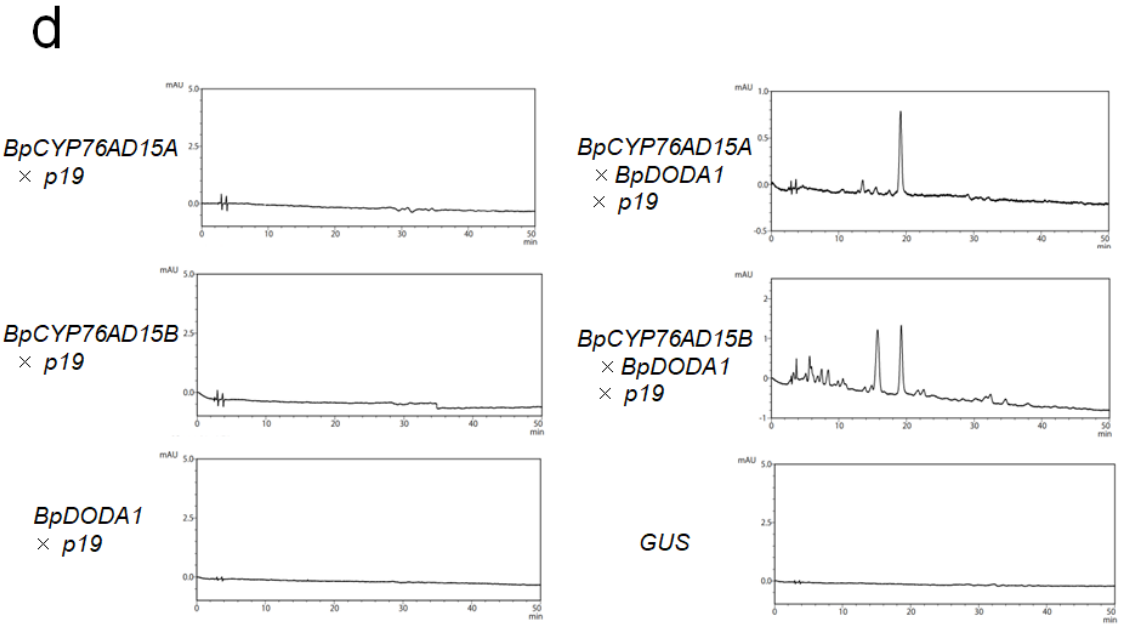


542



543

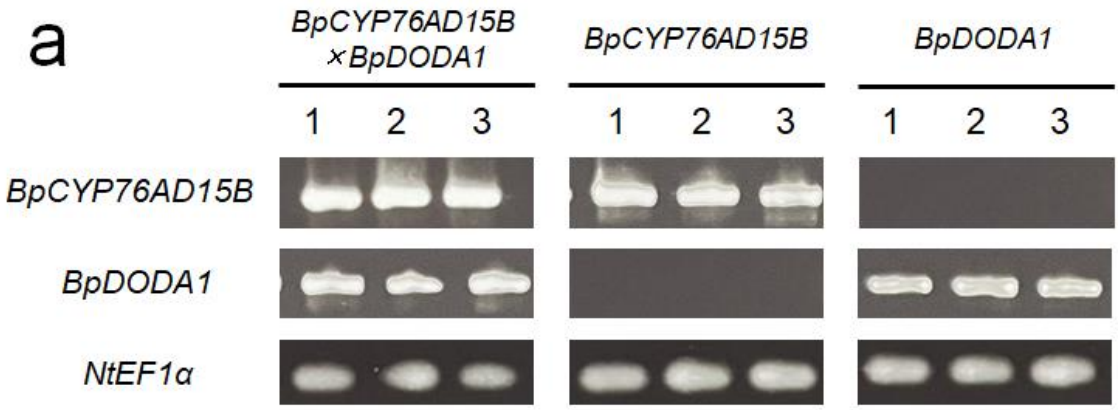
544

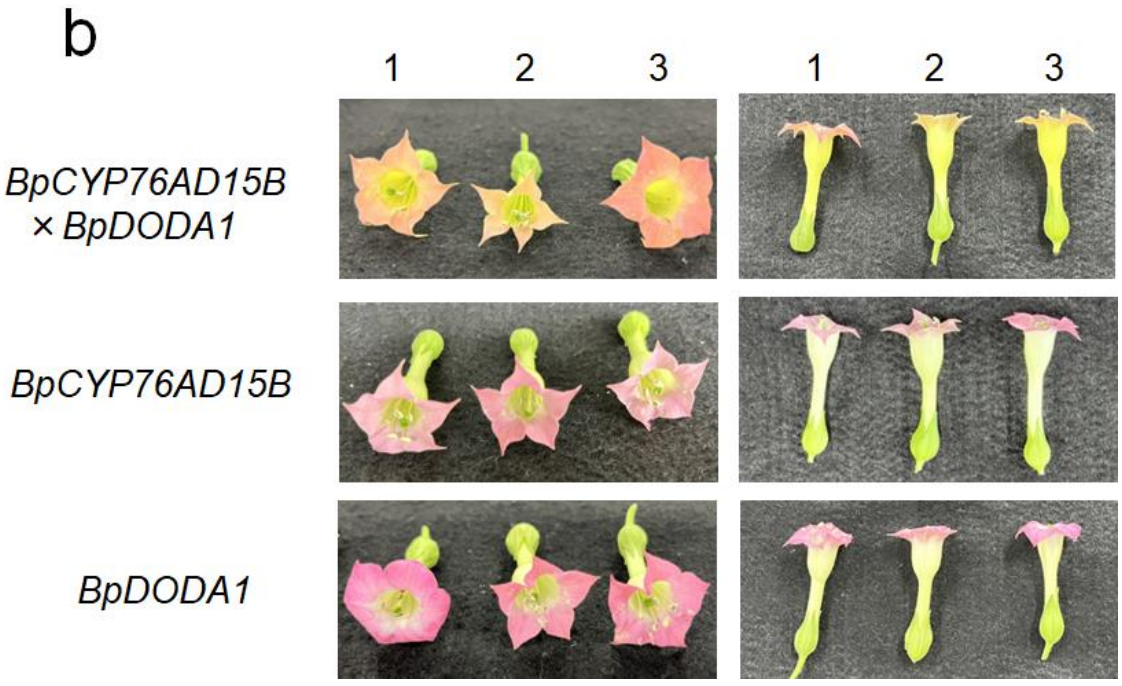


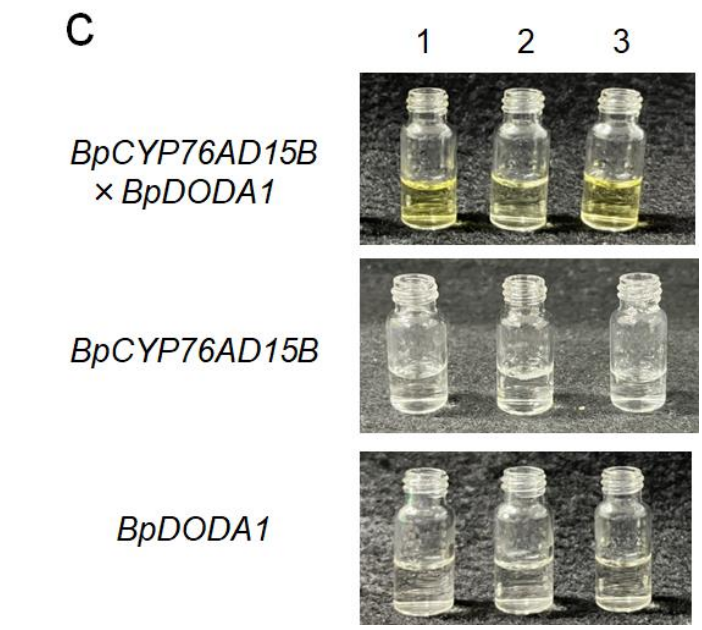
545

546 Fig. 5 Transient over-expression analysis by agroinfiltration in leaves. **a** Photos of
547 infiltrated leaves. **b** RT-PCR validation of transgenic over-expression in infiltrated leaves.
548 **c** Color of leaf extracts. **d** HPLC chromatograms of over-expression leaves analyzed at
549 470 nm

Fig. 6


a
BpCYP76AD15B ×BpDODA1
BpCYP76AD15B
BpDODA1
1 2 3 1 2 3 1 2 3
BpCYP76AD15B
BpDODA1
NtEF1α


b
1 2 3 1 2 3
BpCYP76AD15B × BpDODA1
BpCYP76AD15B
BpDODA1



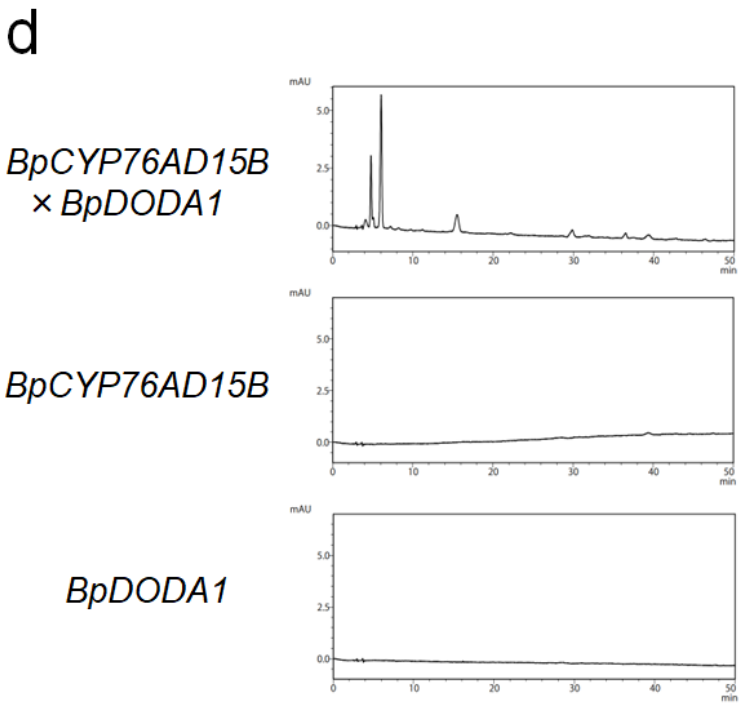


Fig. 6 Over-expression analysis in transgenic *N. tabacum* plants. **a** RT-PCR validation of transgene expression in flowers. **b** Flower color of transgenic over-expression flowers. **c** Color of flower extract from over-expression transgenic plants. **d** HPLC chromatograms of over-expression flowers analyzed at 470 nm