DR HIROSHI MAEDA (Orcid ID : 0000-0003-0246-694X)

Article type : MS - Regular Manuscript

Relaxation of tyrosine pathway regulation underlies the evolution of betalain pigmentation in Caryophyllales

Samuel Lopez-Nieves[1], Ya Yang[2], Alfonso Timoneda[3], Minmin Wang[1], Tao Feng[3], Stephen A. Smith[2], Samuel F. Brockington[3] and Hiroshi A. Maeda[1]

[1]Department of Botany, University of Wisconsin–Madison, Madison, WI 53706, USA; [2]Department of Ecology & Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA; [3]Department of Plant Sciences, University of Cambridge, Cambridge, CB2 3EA, UK

Author for correspondence:

Hiroshi A. Maeda

Tel: +1 608 262 5833

Email: maeda2@wisc.edu

Received: *21 June 2017*

Accepted: *25 August 2017*

Summary

- Diverse natural products are synthesized in plants by specialized metabolic enzymes, which are often lineage-specific and derived from gene duplication followed by functional divergence. However, little is known about the contribution of primary metabolism to the evolution of specialized metabolic pathways.
- Betalain pigments, uniquely found in the plant order Caryophyllales, are synthesized from the aromatic amino acid *L*-tyrosine (Tyr) and replaced the otherwise ubiquitous phenylalanine-derived anthocyanins. This study combined biochemical, molecular and phylogenetic analyses and uncovered coordinated evolution of Tyr and betalain biosynthetic pathways in Caryophyllales.
- We found that *Beta vulgaris*, which produces high levels of betalains, synthesizes Tyr via plastidic arogenate dehydrogenases ($TyrA_a$/ADH) encoded by two *ADH* genes (*BvADHα* and *BvADHβ*). Unlike BvADHβ and other plant ADHs that are strongly inhibited by Tyr, BvADHα exhibited relaxed sensitivity to Tyr. Also, Tyr-insensitive BvADHα orthologs arose during the evolution of betalain pigmentation in the core Caryophyllales and later experienced relaxed selection and gene loss in lineages that reverted from betalain to anthocyanin pigmentation, such as Caryophyllaceae.
- These results suggest that relaxation of Tyr pathway regulation increased Tyr production and contributed to the evolution of betalain pigmentation, highlighting the significance of upstream primary metabolic regulation for the diversification of specialized plant

metabolism.

Key words: anthocyanins, arogenate dehydrogenase (ADH/TyrA_a), betalains, Caryophyllales, metabolic pathway evolution, tyrosine biosynthesis.

Introduction

Plants synthesize numerous specialized metabolites (also known as secondary metabolites), which play crucial roles in plant adaptation. In contrast to well-documented diversification of plant enzymes directly involved in specialized metabolism (Chen *et al.*, 2011; Mizutani & Ohta, 2010; Pichersky & Lewinsohn, 2011; Weng, 2014; Moghe & Last, 2015), relatively little is known about the evolution of primary metabolic enzymes that provide precursors to the production of various specialized metabolites.

L-Tyrosine (Tyr) is an essential aromatic amino acid required for protein biosynthesis in all organisms; however, it is synthesized *de novo* only in bacteria, fungi, and plants, but not in animals. Consequently, animals have to consume Tyr or *L*-phenylalanine (Phe) that can be hydroxylated to Tyr (Pencharz *et al.*, 2007). Besides protein biosynthesis, plants also use Tyr to produce a diverse array of specialized metabolites that are important for defense (e.g. dhurrin, Gleadow & Møller, 2014), stress tolerance (e.g. tocopherols, Mene-Saffrane *et al.*, 2010), and pollinator attraction (e.g., betalains, Tanaka *et al.*, 2008). Notably, humans have a long history of utilizing Tyr-derived specialized metabolites, such as the psychedelic alkaloid mescaline derived from the cactus *Lophophora williamsii* (Ibarra-Laclette *et al.*, 2015) and the analgesic morphine derived from *Papaver somniferum* (opium poppy, Beaudoin & Facchini, 2014; Millgate *et al.*, 2004).

Tyr is synthesized from prephenate, which is converted from the final product of the

shikimate pathway, chorismate (Maeda & Dudareva, 2012; Siehl, 1999; Tzin, V. & Galili, 2010). In most bacteria and fungi, prephenate is oxidatively decarboxylated by prephenate dehydrogenase ($TyrA_p$/PDH, hereafter referred only as PDH; EC 1.3.1.12) to 4-hydroxyphenylpyruvate (HPP), which is transaminated to Tyr (Bentley, 1990, **Fig. 1a**). On the other hand, most plants first transaminate prephenate into arogenate and subsequently decarboxylate into Tyr by arogenate dehydrogenase ($TyrA_a$/ADH, hereafter referred only as ADH; EC 1.3.1.78, Rippert & Matringe, 2002a,b), both steps occurring in the plastids (Dal Cin *et al.*, 2011; Rippert *et al.*, 2009; **Fig. 1a**). The Tyr pathway is usually highly regulated at PDH and ADH. These homologous enzymes are strongly feedback inhibited by Tyr and control carbon flow between the two competing Tyr and Phe pathways (Gaines *et al.*, 1982; Bentley, 1990; Rippert & Matringe, 2002a,b; **Fig. 1b**). A recent report showed that, in addition to plastidic ADH enzymes, some plants possess a PDH enzyme(s) that is not inhibited by Tyr and is localized to the cytosol (Rubin & Jensen, 1979; Schenck *et al.*, 2015; 2017; Siehl, 1999). Clearly, there is evolutionary variation in the Tyr pathway(s) in different plant lineages that warrants investigation.

Betalains are a class of Tyr-derived pigments that, within the flowering plants, occur exclusively in the order Caryophyllales where they replace the otherwise ubiquitous anthocyanins (Mabry, 1964; Tanaka *et al.*, 2008). Within Caryophyllales, the majority of families are betalain pigmented. In two families, Molluginaceae and Caryophyllaceae, however, evolutionary reversions from betalain to anthocyanin pigmentation have occurred (Brockington *et al.*, 2015), highlighting the fact that these two classes of water-soluble pigments have never been found in the same organism (Bate-Smith, 1962; Mabry, 1964; Clement & Mabry, 1996; Brockington *et al.*, 2011). Betalains and anthocyanins are synthesized from Tyr and Phe, respectively, but have similar physiological functions in pollinator attraction and stress tolerance

(Tanaka *et al.*, 2008). Betalains are also used as a natural food dye (E162) and have anticancer and antidiabetic properties (Neelwarne & Halagur, 2012; Lee *et al.*, 2014; Khan, 2015). Furthermore, intermediates in the betalain pathway are important pharmaceuticals (e.g. *L*-dihydroxyphenylalanine (*L*-DOPA) for the treatment of Parkinson's disease) or are substrates for other pharmaceutical agents (e.g. the production of dopamine and isoquinoline alkaloids such as morphine). Consequently, understanding the coordinated regulation of Tyr and betalain biosynthesis has the potential to enhance the production of Tyr, and the yield of Tyr-derived plant natural products important for human health and nutrition.

Betalain biosynthesis starts with hydroxylation of Tyr to *L*-DOPA by at least three closely related cytochrome P450 enzymes (CYP76AD1, CYP76AD5, and CYP76AD6; **Fig. 1a**) (Polturak *et al.*, 2016; Sunnadeniya *et al.*, 2016). *L*-DOPA is further converted into betalamic acid or cyclo-DOPA by *L*-DOPA dioxygenases (DODA, Christinet *et al.*, 2004; Gandía-Herrero & García-Carmona, 2012) or CYP76AD1 (Hatlestad *et al.*, 2012), respectively (**Fig. 1a**). Betalamic acid then spontaneously reacts with cyclo-DOPA or amines to produce various forms of betacyanins or betaxanthins, respectively, which are usually further glycosylated. Recent studies found that the two key enzymes within the betalain pathway, DODA, and CYP76AD1, duplicated just before the emergence of betalain pigmentation (Brockington *et al.*, 2015). Subsequently, one of the duplicated copies (DODAα and CYP76AD1α) in both genes became specialized for betalain biosynthesis and were lost or downregulated in the anthocyanin-producing families such as Molluginaceae and Caryophyllaceae (Brockington *et al.*, 2015). Despite recent and rapid progress in understanding the betalain pathway enzymes and their evolution, little is known about the regulation of primary Tyr metabolism in relation to the evolution of this novel Tyr-dependent betalain pathway.

Here we first investigated the Tyr biosynthetic pathway and its regulation in table beet

(*Beta vulgaris*), which produces high levels of betalains (Goldman, 1996). Using comparative genomics, biochemical, and cellular analyses, we found plastidic ADH enzymes from *B. vulgaris* that exhibit relaxed sensitivity to Tyr inhibition *in vitro* and *in vivo*. Phylogenetic analysis combined with recombinant enzyme characterization further demonstrated that de-regulated ADH enzymes emerged during the evolution of betalain pigmentations in the core Caryophyllales, and were lost or downregulated following disappearance of betalains. Furthermore, transient expression of the de-regulated ADH in *Nicotiana benthamiana* led to high accumulation of Tyr *in planta*. The results revealed the important contribution of primary Tyr pathway regulation to the unique evolution of a plant specialized metabolic pathway, betalain biosynthesis.

Materials and Methods

Plant source and growth conditions

B. vulgaris L. varieties, red beet (W357B), yellow beet (Touch Stone), and white beet (Blankoma), were provided by Dr Irwin Goldman from the University of Wisconsin-Madison, Department of Horticulture (Goldman, 1996), whereas sugar beet (Big Buck) and sea beet (PI 562585) were commercial sugar beets obtained from the Heirloom Seeds (West Finley, PA, USA) and the National Plant Germplasm System (NPGS), respectively. Spinach (*Spinacia oleraceae*), Pigeonberry (*Rivina humilis*), four o'clock (*Mirabilis jalapa*), and common purslane (*Portulaca oleracea*) were grown from seed with a growing mix soil (Fafard®, Agawam, MA, USA) in a growth chamber under 12 h light (100 μmol m^{-2} s^{-1}) **[Author, μE has been altered to μmol m^{-2} s^{-1} to fit journal style. Please check.]**, 22ºC and 60% humidity. After 1 month of growth, their leaves were harvested for RNA extraction.

Identification and cloning of ADH homologs from Caryophyllales

BLASTP searches were performed using the protein sequences of ADH and PDH enzymes from *A. thaliana* (AtADH1/At5g34930, NP_173023; AtADH2/At1g15710, NP_198343), *Glycine max* (GmPDH, KM507071), *Synechocystis* sp. PCC6803 (SyADH, WP_010872597), *Escherichia coli* (EcPDH, WP_052912694), *Aquifex aeolicus* (AaPDH, WP_010881139) as queries against the sugar beet genome (*Beta vulgaris* http://bvseq.molgen.mpg.de/) (Supporting Information **Fig. S1b**). Potential ADH candidates were identified based on a broad phylogenetic analysis that included various plant ADH and PDH sequences (**Fig. S1c**).

Genomic DNA was extracted using Tris-sodium chloride-EDTA/sodium dodecyl sulfate buffer and precipitated with isopropanol and 200 mM ammonium acetate. For RNA isolation, the method described by Wang *et al.* (2011) was used with some modifications. The tissues were ground in a mortar with liquid nitrogen and powder polyvinylpyrrolidone (PVP). After addition of 700 µl fresh pre-warmed lysis buffer (2% CTAB, 2 M NaCl, 100 mM Tris-HCl pH 8, 25 mM EDTA and 5% β-mercaptoethanol), the samples were shaken vigorously for 2 min and incubated in a water bath at 65°C for 5 min. The RNA was converted into complementary DNA (cDNA) using the High-Capacity cDNA Reverse Transcription Kit (Applied Biotechnology, USA) and SuperScript IV Reverse Transcriptase with oligo dT_{20} primer or random primers (Invitrogen, USA).

Cloning primers were designed with the Invitrogen primer design (http://tools.lifetechnologies.com/content.cfm?pageid=9716) and the PCR In-Fusion® primers designing program (http://bioinfo.clontech.com/infusion/convertPcrPrimersInit.do, Clontech, Mount View, CA). All ADH candidate genes, except for PoADHα (see below), were PCR amplified from cDNA using gene-specific primers (**Table S1**) and Phusion DNA polymerase (Thermo, Waltham, MA, USA) with the following conditions: initial denaturation at 95°C for 5

min, 35 cycles of amplification at 95°C for 30 s, 58°C for 30 s, 72°C for 30 s, with a final extension at 72°C for 10 min. The PCR fragments were purified using QIAquick gel extraction kit (Qiagen, Valencia, CA) and were inserted into the pGEX-2T vector (GE Healthcare) at EcoRI and BamHI sites using the In-Fusion cloning method (Clontech). PoADHα was gene synthesized (Biomatik, Cambridge, Ontario, Canada) and directly cloned into the same pGEX-2T vector. For generation of His-tagged proteins, the cloned PCR fragments were inserted into the pET28a vector (Novagen, Madison WI, USA) at NdeI and EcoRI site.

Recombinant enzyme expression and purification

The His-tagged recombinant protein expression was carried out as we described previously (Dornfeld *et al.*, 2014). For GST-tagged recombinant protein expression, the cloned pGEX-2T vectors were introduced into Rosetta-2 *E. coli* competent cells (Novagen, Madison, WI, USA) and cultured overnight at 37°C, 200 rpm in 10 ml LB medium containing Ampicillin (100 μg ml^{-1}). The 10 ml of the overnight culture were transferred to 1 l LB medium with Ampicillin (100 μg ml^{-1} and further incubated at 37°C and 200 rpm until the OD_{600} reached 0.3. The temperature was then changed to 18°C and, after 1 h, isopropyl β-D-1-thiogalactopyranoside (IPTG, 400 mM final concentration) was added to induce recombinant protein expression. After overnight incubation at 18°C under constant shaking at 200 rpm, cultures were harvested by centrifugation at 2,000 ***g*** for 10 min at 4°C, and the pellet was washed with 0.9% NaCl solution. The samples were harvested and resuspended in 25 ml of lysis buffer (phosphate-buffered saline (PBS) pH 7.4, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM dithiothreitol (DTT) and plant proteases inhibitor cocktail (Amresco, Solon, OH, USA)). The resuspended cells were sonicated for periods of 20 s for 5 min. The cell lysate was centrifuged at 10,000 ***g*** for 30 min at 4°C, and the supernatant was applied to Fast Protein Liquid Chromatography (FPLC, AKTApure25 FPLC

system, GE Healthcare) equipped with GSTtrap^{TM}FF (GE Healthcare, USA). Prior and after injection, the column was washed with five times bed volume wash buffer A (PBS, pH 7.6) followed by five times bed volume of wash buffer B (10 mM glutathione, 1.54 g of reduced glutathione dissolved in 500 ml of 50 mM Tris-HCl, pH 8). The recombinant enzymes containing GST-tag were eluted with ten-bed volumes of the elution buffer B and collected into Eppendorf tubes containing 500 µl. Recombinant enzymes eluted in the fraction five and six, which were combined and desalted using a gel filtration column (Sephadex G50-80 resin, Sigma-Aldrich, St Louis, MO, USA) in the reaction buffer (200 mM HEPES (pH 7.6), 50 mM KCl, 10% ethylene glycol). Enzyme concentrations were measured using Bradford assay (Bio-Rad, Des Plaines, IL, USA) and the enzyme purity was estimated by running on SDS-PAGE gel and analyzing with ImageJ (http://imagej.nih.gov/ij/).

ADH and PDH activity assays *in vitro*

ADH and PDH activity from beet tissues (**Fig. S2a,b**) were analyzed by using the leaves and stem/root crude protein extract of red beet (*W357B*). The beets were grown in a glasshouse for 12 wk with a temperature of 22–25ºC and 16 h of ambient and supplemented lights. Protein extraction was performed by grinding 1 g of tissues in liquid nitrogen and resuspending the powder in the extraction buffer (200 mM HEPES (pH 7.6), 50 mM KCl, 10% ethylene glycol, 1 mM PMSF, 1 mM DTT and plant proteases inhibitor cocktail (Ameresco)). The extracts were desalted using the gel filtration column (Sephadex G50-80 resin, Sigma-Aldrich, St Louis, MO, USA) into the reaction buffer. The ADH or PDH assays were performed by mixing the desalted protein extract with 1 mM $NADP^+$ and 1 mM *L*-arogenate or prephenate in a total volume of 10 µl or 25 µl, respectively. *L*-Arogenate was prepared by enzymatic conversion from prephenate (Sigma-Aldrich, St Louis, MO, USA), as previously described (Schenck *et al.*, 2015). The

reactions were started by adding the enzyme (crude extract or recombinant enzyme) and incubated at 37°C for 45 min. The reaction was stopped with two times volume of methanol. The same ADH and PDH assay protocols were used for initial characterization of purified recombinant BvADH enzymes

For detection of Tyr product from the ADH assays, 10 µl of the reaction mixture was first derivatized with the equal volume of the 40.26 mM OPA solution (5.4 mg OPA (Sigma-Aldrich, St Louis, MO, USA) mixed in 100 µl methanol, 5 µl 2-mercaptoethanol and 900 µl 0.4M boric acid) for 3 min, injected to high pressure liquid chromatography (HPLC, Agilent 1260) equipped with the Eclipse XDH-C18 column (5 µm, 3.0 x 150 mm, Agilent, USA), and separated by a 30 min linear gradient from 20–45% methanol in 0.1% ammonium acetate at a flow rate of 0.8 ml min^{-1}. The substrate and product of ADH assays (Tyr and arogenate, respectively) were detected by a fluorescence detector (Agilent, USA) with excitation at 360 nm and emission of 455 nm. For PDH assays, the reactions were stopped by addition of $NaBH_4$, which converts the reaction product HPP into hydroxyphenyllactic acid (HPLA), followed by neutralization with 100 µl of 6 N HCl as described by Schenck et al. 2015. The HPLC was equipped with ZORBAX SB-C18 column (Agilent, USA) using a 6 min isocratic elution at 25% methanol in 0.1% phosphoric acid, followed by a 20 min linear gradient of 25–60% methanol at a flow rate of 1.0 ml min^{-1}. The HPLA were monitored by absorption at 270 nm.

To test the electron donor and substrate preferences of purified recombinant enzymes, the ADH and PDH reactions were performed as described above, except for 12 min with 400 µM *L*-arogenate and 1 mM cofactor (NAD^+ or $NADP^+$). The reaction was stopped by placing the tubes on ice and immediately measured for the production of the reduced cofactor, NAD(P)H, at 340 nm by spectrophotometer (NanoDrop 2000, Thermo Scientific, USA). The quantification was based on the standard curve of authentic NADPH.

To examine Tyr sensitivity of the purified recombinant enzymes, ADH assay was performed as described previously (Schenck *et al.*, 2015) but in the presence or absence of different concentrations of *L*-Tyr. Tyr was first dissolved in 0.025 N NaOH at 100 mM (as the water solubility of Tyr is very low, < 2 mM), which was diluted to 4 mM to 10 μM final concentration in 0.0025 N NaOH. The reactions contained 500 mM HEPES (pH 7.6) to maintain the final pH at 7.6. The production of reduced cofactor (NADPH) was monitored at 340 nm using a spectrophotometer every 2 min for 10 min. In addition, other effectors (*L*-Phe, *L*-Trp, and betanin) were used to test possible inhibition of the enzyme ADH activity at a final concentration of 1 mM. All of the reactions were performed under non-saturated condition, where activity increased linearly depending on reaction times and enzyme concentrations.

Transient expression of *BvADHα* and *BvADHβ* in *Nicotiana benthamiana*

ADHα and *ADHβ* sequences used for *N. benthamiana* agroinfiltration were amplified from *Beta vulgaris* var. *vulgaris* variety '*Boltardy*' (Chiltern Seeds, UK) swollen hypocotyl and leaf tissue cDNA libraries respectively, which were prepared using BioScript Reverse Transcriptase (Bioline Reagents, London, UK). Transcripts were amplified by PCR using gene specific primers (**Table S1**) and Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific, Waltham, MA, USA). Vectors for transient transformation were constructed with Golden Gate cloning using the MoClo Tool Kit (Weber *et al.*, 2011; Addgene, Cambridge, MA, USA), with the BpiI and BsaI restriction sites eliminated after cloning. The turboGFP sequence used in this assay was a variant codon-optimized for plants contained in the MoClo Plant Parts Kit (Engler *et al.*, 2014; Addgene). *BvADHα*, *BvADHβ*, and *turboGFP* sequences were ultimately cloned into the pICH86988 binary vector under control of the Cauliflower Mosaic Virus 35S promoter and the *Agrobacterium tumefaciens* octopine synthase (OCS) terminator.

Transient gene expression assays in *N. benthamiana* were performed according to the previously described agroinfiltration method with some modifications (Sparkes *et al*., 2006). All constructs were transformed into the *Agrobacterium tumefaciens* GV3101 strain, and grown in LB media supplemented with kanamycin (50 mg l^{-1}), gentamycin (25 mg l^{-1}) and rifampicin (50 mg l^{-1}) until reaching an OD_{600} of 1.5. Cultures were then brought to a final OD_{600} of 0.5 in infiltration media (10 mM $MgCl_2$, 0.1 mM acetosyringone, 10 mM MES at pH 5.6) for 3 h before infiltration. Infiltration spots corresponding to 35S::*BvADHα*, 35S::*BvADHβ*, and 35S::t*urboGFP* were performed in the same leaves of 6-wk-old *N. benthamiana* plants alternating the position of the spots between plants in a clockwise manner to account for intra-leaf variation (Barshandy *et al*., 2015). Infiltrated tissue was sampled 3 d post-infiltration from five biological replicates for tyrosine quantification and qRT-PCR analysis.

For quantification of tyrosine and other amino acids, *c.* 40 mg fresh weight tissues were harvested, lyophilized, sent from the University of Cambridge (UK) to the University of Wisconsin-Madison (USA), and analyzed exactly as described. Tyrosine and other amino acids were extracted and measured as described previously (Wang *et al.*, 2017). Amino acid standards (Sigma-Aldrich, St Louis, MO, USA) of 4–1000 μM were prepared the same way to make standard curves.

Phylogenetic analysis

Amino acid sequences from genomes (full open reading frame) and transcriptomes (full or partial open reading frame) of Brockington *et al*. (2015) were used for phylogenetic analysis following methods described in Brockington *et al*. (2015) with minor modifications. In addition, we carried out analysis of dN/dS ratio in *ADHα* to test for relaxed selection in anthocyanic lineages (**Table S2**;

see **Methods S1** for a full description**)**.

Subcellular localization of GFP-fused ADH enzymes

The subcellular localization experiments of GFP-fused ADH enzymes were conducted as we described previously (Schenck *et al.*, 2015).

Accession numbers

The Genbank accession numbers for the sequences mentioned in this article are: BvADHβ *W357B* red beet variety (KY207366), BvADHβ *Boltardy* red beet variety (MF346292), BvADHβ *Big Buck* sugar beet variety (KY207367), BvADHβ *Touch Stone* yellow beet variety (KY207368), BvADHβ *Blankoma* white beet variety (KY207369), BvADHβ *Sea beet* PI562585 variety (KY207370), BvADHα *Big Buck* sugar beet variety (KY207371), BvADHα *W357B* red beet variety (KY207372), BvADHα *Boltardy* red beet variety (MF346291), BvADHα *Blankoma* white beet variety (KY207373), BvADHα *Touch Stone* yellow beet variety (KY207374), BvADHα *Sea beet* PI562585 variety (KY207375), SoADHβ (KY207376), SoADHα (KY207378), NaADHβ (KY207377), MjADHα (KU881770), RhADHα (KY207379), PoADHα (KY207380), SmADHα (KY274179), PpADHα (KY274180), and HlADHα (KY274181).

Results

B. vulgaris has two ADH enzymes.

To first investigate how *B. vulgaris* synthesizes Tyr, protein crude extracts of red beet leaf and root/stem tissues were analyzed for ADH and PDH activity, the production of Tyr or HPP from arogenate or prephenate, respectively. Tyr was produced from arogenate in the red beet extracts

of both leaves and roots/stems (**Fig. S2a**) similar to soybean leaf extract, which was previously shown to have both ADH and PDH activity (Schenck *et al.*, 2015). On the other hand, unlike the soybean leaf extract, HPP production was not detected in the leaf and root/stem extracts of red beet (**Fig. S2b**). These results showed that red beet has ADH but not PDH activity.

To identify the gene(s) responsible for the ADH activity in *B. vulgaris*, previously reported plant and microbial *ADH* and *PDH* genes (Bonvin *et al.*, 2006; Hudson *et al*., 1984; Legrand *et al.*, 2006; Rippert & Matringe, 2002a,b; Schenck *et al.*, 2015; **Fig. S1b**) were used to BLAST against the genome of sugar beet, another cultivar of *B. vulgaris* (Dohm *et al.*, 2014) (assembly v.1.2 http://bvseq.molgen.mpg.de). Two *B. vulgaris* sequences homologous to these *ADH* and *PDH* genes were found on chromosome 8 of the *B. vulgaris* genome 25.3 kbp apart (**Fig. S1a**). They were more similar to plant *ADHs* and *PDHs* (59–61% similarity at amino acid levels) than bacterial ones (24–40% similarity, **Fig. S1b**). Within plants, the two *ADH* candidate genes from *B. vulgaris* both belong to the canonical ADH clade containing Arabidopsis *ADHs* (Rippert & Matringe, 2002a,b), rather than the non-canonical clade containing legume PDHs (Schenck *et al.*, 2015; 2017), and appear to be derived from a recent duplication within the order Caryophyllales (**Fig. S1c,d**).

For biochemical characterization, these two putative BvADHs were expressed in *E. coli* as recombinant enzymes, which were further purified using affinity chromatography and subjected to ADH and PDH assays. Both of the beet recombinant enzymes showed ADH activity (i.e. the production of Tyr from arogenate; **Fig. 1b**) and strongly preferred $NADP^+$ over NAD^+ (**Fig. S3**) similar to other plant ADH enzymes and activities (Gaines *et al*., 1986; Rippert & Matringe, 2002a,b). On the other hand, neither of the beet enzymes exhibited detectable PDH activity (**Fig. S2c**), which is consistent with the lack of PDH activity in beet tissues (**Fig. S2b**) and also confirmed the absence of *E. coli* PDH contamination (Hudson *et al.*, 1984). Therefore,

these two genes were designated as *B. vulgaris arogenate dehydrogenases* (*BvADHα* and *BvADHβ*).

Both BvADHs are plastid localized but only *BvADHα* expression is correlated with betalain pathway genes

Most plant enzymes involved in the aromatic amino acid pathways are localized within the plastids (Dal Cin *et al.*, 2011; Maeda & Dudareva, 2012; Rippert *et al.*, 2009), and both BvADH proteins also have a predicted N-terminal plastid transit peptide (**Fig. S4**). To experimentally determine the subcellular localization of BvADHs, a green fluorescent protein (GFP) was fused to the C-terminal of BvADHs, expressed in Arabidopsis protoplasts, and analyzed for their localization using confocal microscopy. The fluorescence signal of GFP fused with BvADHα or BvADHβ overlapped with chlorophyll autofluorescence, which was different from the free GFP control and similar to GFP fused with plastidic Arabidopsis ADH (Rippert *et al.*, 2009) (AtADH2, **Fig. 1c**). These results suggest that both BvADHs are targeted to the plastids and that Tyr is mainly produced by the plastidic arogenate pathway in *B. vulgaris*.

To examine expression patterns of *BvADHs*, especially in comparison to the betalain pathway genes, expression levels of *BvADHα* and *BvADHβ* were analyzed and compared with those of *DODAα*, *CYP76AD1α*, and *BvMYB1* in cotyledon and hypocotyl tissues of sugar and red beets (**Fig. 1d**). Consistent with previous studies (Hatlestad *et al.*, 2012; 2015), *DODAα* and *CYP76AD1α*, as well as *BvMYB1* transcription factor, were much more highly expressed in red than sugar beet. Interestingly, *BvADHα* expression showed similar trends and was significantly higher in red than sugar beet in both cotyledon and hypocotyl tissues. On the other hand, *BvADHβ* expression levels were very similar between genotypes in both tissue types (**Fig. 1d**). These results showed that expression of *BvADHα*, but not *BvADHβ*, is correlated with those of

betalain pathway genes in *B. vulgaris*.

BvADHα but not BvADHβ exhibits relaxed sensitivity to Tyr

Both ADH and PDH enzymes are usually inhibited by Tyr in most organisms (Bentley, 1990; Connelly & Conn, 1986; Gaines *et al.*, 1982; Rippert & Matringe, 2002a,b; Sun, 2009). To determine if the BvADHs are also feedback regulated by Tyr, ADH activity of the recombinant BvADH enzymes were analyzed in the presence and absence of Tyr as an effector molecule. The ADH activity of glutathione S-transferase (GST)-tagged BvADHβ was inhibited by 80% and 100% in the presence of 100 μM and 1 mM Tyr, respectively (**Fig. 2**), similar to the Tyr-sensitive Arabidopsis AtADH2 (Rippert & Matringe, 2002a,b). By contrast, ADH activity of BvADHα was reduced only by half at 1 mM Tyr (**Fig. 2**). Similar results were obtained for histidine (His)-tagged ADH enzymes, where BvADHα showed much less sensitivity to Tyr than AtADH2 (**Fig. S5**), though the expression of His-tagged BvADHβ was not successful. Other aromatic amino acids (Phe and tryptophan) as well as betanin, the major betacyanin accumulated in red beet, did not significantly reduce the ADH activity of BvADHα, BvADHβ, or AtADH2 at 1 mM (**Fig. S6)**. These results revealed that BvADHα, but not BvADHβ, has relaxed sensitivity to Tyr inhibition.

Heterologous expression of *BvADHα* but not *BvADHβ* increase Tyr accumulation in plants

To test if BvADHα having relaxed sensitivity to Tyr can enhance the production of Tyr *in planta*, *BvADHα* and *BvADHβ* were transiently expressed in *N. benthamiana* through Agrobacteria infiltration (**Fig. S7a**; Sparkes *et al*., 2006) and their impacts on Tyr production were analyzed. A control vector expressing GFP was also infiltrated as a negative control (**Fig. S7a**). *BvADHα* expression resulted in >10-fold increase in Tyr levels relative to the GFP control, while the

increase of Tyr due to *BvADHβ* expression was not significantly different (**Figs 3a; Table S3)**. Interestingly, phenylalanine (Phe) levels were decreased significantly under *BvADHα*, but not *BvADHβ* expression (**Fig. 3b**). Other amino acid levels were largely unaffected by *BvADHα* or *BvADHβ* expression (**Table S3**). These results demonstrate that *BvADHα* expression leads to elevated accumulation of Tyr *in planta*.

BvADHα orthologs emerged during the evolution of betalain pigmentation in Caryophyllales

Domestication has modified metabolic traits in various crops (Hanson *et al.*, 1996; Rapp *et al.*, 2010; Rong *et al.*, 2014). Thus, we hypothesized that the BvADHα enzyme with relaxed Tyr regulation was selected during domestication and intensification of color in table beets, that have been used at least since the Roman times (Biancardi *et al.*, 2012; Dohm *et al.*, 2014). To test this hypothesis, the nucleotide and protein sequences of BvADHα (and BvADHβ) were compared among different domesticated beets, red beet (*W357B*), sugar beet (*Big Buck*), yellow beet (*Touch Stone*), and white beet (*Blankoma*), as well as their wild relative, sea beet (Biancardi *et al.*, 2012) (*Beta vulgaris* subsp. *maritima*). Several single nucleotide polymorphisms (SNPs) were detected among different lines in both *BvADHα* and *BvADHβ* (**Fig. S4a,b**). However, only a few of them affected the amino acid sequences and were within and near the N-terminal signal peptide of BvADHα and BvADHβ, respectively (**Fig. S4c,d**). Thus, the mature enzyme regions of BvADHα were unaltered during domestication.

To further test if the ADHα enzymes with reduced Tyr sensitivity are restricted to the species *B. vulgaris*, the corresponding genes for BvADHα and BvADHβ were cloned from a closely related species within the same Amaranthaceae family, spinach (*Spinacia oleracea*), whose draft genome is available (http://bvseq.molgen.mpg.de). Spinach ADHα and ADHβ orthologs (SoADHα and SoADHβ) had 77 and 83% identity at amino acid levels to the

corresponding BvADHs in the mature enzymatic regions. The recombinant enzymes of spinach ADHs showed similar Tyr sensitivity to beet ADHs: SoADHα, but not SoADHβ, exhibited reduced Tyr sensitivity (**Fig. 2)**. These results suggest that the reduced Tyr sensitivity of BvADHα at least at the enzyme level was not the result of selection during domestication of beet cultivars, but was already present in the common ancestor of the beet and spinach ADHα enzymes.

To determine the origin and molecular evolution of *BvADHα*, we mined genome and transcriptomic data across the Caryophyllales for *ADH* orthologs and performed a phylogenetic analysis (**Figs 4a, S1e**). The results indicate that a gene duplication event on the branch leading to stem Caryophyllales produced *ADHα* and *ADHβ* lineages. While *ADHβ* orthologs were expressed across the entire Caryophyllales, expression of *ADHα* closely parallels betalain production in Caryophyllales. *ADHα* expression is undetectable from the anthocyanic clade that diverged before the earliest inferred origin of betalain synthesis (hereafter referred to as non-core Caryophyllales; Brockington *et al*., 2009). Two families in the Caryophyllales, Molluginaceae and Caryophyllaceae have reverted from betalain to anthocyanin pigmentation (Brockington *et al.*, 2011, 2015). Presence of the *ADHα* orthologs in the transcriptomes of Molluginaceae and Caryophyllaceae was much less common than the presence of *BvADHβ* **(Fig. 4a,b)**. Thus the presence of *ADHα*, but not *ADHβ*, closely mirrors the distribution of betalain pigmentation across Caryophyllales, similar to the pattern in two other genes of the betalain pathway, *CYP76AD1α* and *DODAα* (Brockington *et al*., 2015).

Betalain-producing species have deregulated BvADHα enzyme and elevated Tyr levels

To further test experimentally if *ADHα* orthologs across Caryophyllales share the unique property of reduced Tyr inhibition, *ADH* genes from representative members of Caryophyllales

(Brockington *et al.*, 2011) were cloned and the Tyr sensitivity of encoded enzymes was evaluated. An ADHβ enzyme from the anthocyanin-producing non-core Caryophyllales, *Nepenthes ventricosa* x *alata* (NaADHβ, Nepenthaceae; **Fig. 4b**), was strongly inhibited by Tyr (**Fig. 5**) similar to beet and spinach ADHβ (**Fig. 2**). On the other hand, ADHα orthologs from betalain-producing families, *Rivina humilis* (RhADHα, Rivinaceae), *Mirabilis jalapa* (MjADHα, Nyctaginaceae), and *Portulaca oleracea* (PoADHα, Portulacaceae), all shared relaxed Tyr inhibition and retained 42% to 68% of ADH activity even at 1 mM Tyr (**Fig. 5**).

To test if Tyr-insensitivity of the recombinant ADHα enzyme is also detectable *in vivo*, Tyr sensitivity of leaf ADH activity was analyzed from species containing *ADHα* (i.e. spinach) and ones lacking *ADHα* (i.e. *Arabidopsis thaliana*; *Dianthus barbatus*, Caryophyllaceae). Spinach rather than beet was used due to its cleaner background during HPLC-based enzyme assay. As shown in **Table 1 and Fig. S8**, ADH activity of *Arabidopsis* and *Dianthus barbatus* tissues was strongly inhibited (92–95%) by 0.5 mM of Tyr effector, whereas that of spinach was much more resistant to Tyr inhibition (only *c*. 21% inhibited), consistent with the presence of SoADHα with relaxed sensitivity to Tyr (**Fig. 2**).

To further test if the presence of deregulated ADHα leads to increased Tyr accumulation in betalain-producing species, Tyr levels were quantified in young leaves of a variety of Caryophyllales species with or without ADHα and also in *A. thaliana* as a comparison. Anthocyanin-producing species from non-core Caryophyllales (e.g. *Nepenthes ventricosa* x *alata*) and Caryophyllaceae (e.g. *Dianthus barbatus*) had Tyr levels (2.1 to 8.8 nmol/gFW) comparable to that of Arabidopsis (5.3 nmol g^{-1} FW). On the other hand, while large variations were observed, betalain-producing ADHα-containing species all had significantly higher Tyr levels (from 12 to 180 nmol g^{-1} FW) than Arabidopsis (**Fig. 4c**). These results demonstrate that betalain-producing species have ADHα with relaxed sensitivity to Tyr inhibition and accumulate

elevated levels of Tyr.

ADHα orthologs underwent relaxed selection and gene loss in lineages that have reverted from betalain to anthocyanin pigmentation

Interestingly, when *ADHα* orthologs were recovered from Caryophyllaceae or Molluginaceae transcriptomic data, they were often recovered in partial sequences, indicating general low abundance. Within the Caryophyllaceae, *ADHα* orthologs was only detected in the subfamily Paronychioideae (Greenberg & Donoghue, 2011), which forms a grade paraphyletic to the rest of the family. To test for relaxed selection in anthocyanic lineages we further examined a subset of *ADHα* orthologs with sequences either verified by Sanger sequencing or by transcriptome read mapping and manual inspection of read coverage (Methods S1**)**. Although no obvious acceleration of substitution was observed in Caryophyllaceae from nucleotide coding sequences (CDS; **Fig. S9a**), there was apparent acceleration in their amino acid sequences (**Fig. S9b**). Furthermore, the dN/dS ratio in Caryophyllaceae *ADHα* (0.166) was elevated compared to the rate among betalain-producing *ADHα* (0.0743) under the Partitioned MG94xREV Model, assuming homogenous synonymous and nonsynonymous rates across sites. In addition, we found evidence of relaxed selection (as opposed to intensification of positive selection) that contributes to the increase in nonsynonymous rate in Caryophyllaceae under the RELAX framework (P = 5.6E-8, **Table S4**) (Wertheim *et al.*, 2014). Moreover, the genome assembly of the anthocyanic carnation (*Dianthus caryophyllus*, Caryophyllaceae subfamily Caryophylloideae that nested within subfaily Paronychioideae, Greenberg & Donoghue, 2011; Yagi *et al.*, 2014) lacked *ADHα* ortholog and only contained *ADHβ* ortholog, suggesting complete gene loss of *ADHα* in the subfamily Caryophylloideae (Greenberg & Donoghue, 2011). Species within the anthocyanic Caryophyllaceae, therefore, exhibit the transition from relaxed selection to gene loss of *ADHα*

orthologs, which associates with the loss of betalain pigmentation in Caryophyllaceae.

Discussion

This study found that *B. vulgaris* has ADH but no PDH enzymes or activity (**Figs 1b, S2, S3**). This is similar to most plants (Connelly & Conn, 1986; Gaines *et al.*, 1982; Rippert & Matringe, 2002a,b) but different from legumes that have both ADH and PDH (Rubin & Jensen, 1979; Schenck *et al.*, 2015; 2017; Siehl, 1999). Thus, *B. vulgaris* synthesizes Tyr via the ADH pathway that occurs within the plastids (Rippert *et al.*, 2009) (**Fig. 1c**). We also found that *B. vulgaris* possesses two paralogous genes encoding the ADH enzymes, namely ADHα and ADHβ (**Fig. S1c**,**d**). Interestingly, ADHα but not ADHβ exhibited relaxed sensitivity to Tyr inhibition (**Fig. 2**). Although recent studies reported that the legume PDH enzymes are also Tyr insensitive (Schenck *et al.*, 2015; 2017), BvADHα and legume PDHs have two major differences. First, legume PDHs are localized in the cytosol (Schenck *et al.*, 2015), whereas BvADHα (and BvADHβ) was targeted to the plastids (**Fig. 1c**). Second, legume PDHs completely lost Tyr sensitivity (Schenck *et al.*, 2015) but BvADHα was still inhibited by Tyr at higher concentrations (**Figs 2, S5**). The maintenance of inhibition at higher concentration is likely necessary because Phe biosynthesis is also localized within the plastids, and thus BvADHα is directly competing for the arogenate substrate with Phe biosynthesis (**Fig. 1a**). Complete loss of ADH regulation by Tyr would, therefore, deplete Phe and essential Phe-derived compounds (e.g., proteins, lignin).

Other insensitive ADH/PDH enzymes have been previously found in microorganisms (Legrand *et al.*, 2006) and the structural analyses of Tyr sensitive and insensitive enzymes identified histidine 217 as a possible residue responsible for its Tyr sensitivity (Legrand *et al.*, 2006; Sun *et al.*, 2009). Also, phylogeny-guided structure-function analysis revealed that converting a single active site aspartate 222 residue into a non-acidic residue played a key role in

the evolution of the legume PDH enzymes and simultaneously introduced prephenate substrate specificity and Tyr insensitivity (Schenck *et al.*, 2017). However, the corresponding histidine and aspartate residues are still present in BvADHα (**Fig. S10**), suggesting that different mechanisms, and as yet unidentified residues are involved in the relaxed Tyr sensitivity of BvADHα.

Previous analyses of molecular evolution of DODAα and CYP76AD1α, two enzymes which convert Tyr into betalains (Christinet *et al.*, 2004; Gandía-Herrero & García-Carmona, 2012; Hatlestad *et al.*, 2012), revealed that both of these genes arose through gene duplication, just before the origin of betalain pigmentation in Caryophyllales (Brockington *et al.*, 2015). Similarly, this study found that *ADHα* orthologs arose by gene duplication, before the emergence of *DODAα* and *CYP76AD1α* (**Fig. 4a,b**), intimately associated with the origin of betalain pigmentation. One of the duplicated copies, *ADHα*, underwent neofunctionalization and became much less sensitive to Tyr inhibition, which is the key regulatory mechanism of Tyr biosynthesis (Maeda & Dudareva, 2012; Rippert & Matringe, 2002a,b). ADHα enzymes with relaxed Tyr sensitivity are maintained in all betalain-producing species of Caryophyllales, at least the ones that we analyzed (**Figs 2, 5**). Furthermore, the expression pattern of *BvADHα* is distinct from that of *BvADHβ* and similar to those of the betalain biosynthetic genes (*DODAα* and *CYP76AD1α*) and *MYB1* transcription factor (**Fig. 1d**), suggesting that the alteration of ADHα enzyme property was accompanied by changes in its expression profile. Although similar examples of biochemical and transcriptional changes during the evolution of plant specialized metabolic enzymes/genes have been reported (Kajikawa *et al*., 2017; Moghe & Last, 2015; Panchy *et al*., 2016; Weng *et al*., 2012; Xu *et al*., 2017), here we revealed a unique example of coordinated evolution of primary amino acid pathway (i.e. Tyr biosynthesis) and its downstream specialized metabolism (i.e. betalain biosynthesis).

In the anthocyanic Caryophyllaceae, the transition of betalain pigmentation to

anthocyanin pigmentation was associated with down-regulation, relaxed natural selection, and deletion of *ADHα* (**Figs 4, S1c,d, S9; Table S4**). Similar down-regulation and deletion of genes were also observed during the loss of flower petals (Zhang *et al.*, 2013) and arbuscular mycorrhizal symbiosis (Delaux *et al.*, 2014) in various plant lineages. Together these lines of evidence suggest that maintenance of the ADHα is superfluous, following loss of betalain pigmentation. The ultimate cause of reversion of betalain to anthocyanin pigmentation in multiple lineages within the core Caryophyllales is currently unknown. It may be due to a number of factors, including: i) metabolic cost of nitrogen-containing alkaloid betalain pigments, ii) shift in pollinator populations that are attracted by unique spectra (e.g. blue) of some anthocyanins, iii) increased demand for other Phe-derived compounds (e.g. tannins, flavonoids), or iv) simple genetic drift enabled by the presence of still intact Phe, phenylpropanoid, core flavonoid pathways in betalain-producing plants (Brockington *et al*, 2011;, Shimada *et al*., 2005; Xu et al., 2016).

A mechanism underlying the mutually exclusive distribution of betalain and anthocyanin pigments has long fascinated evolutionary biologists (Brockington *et al*, 2011; Des Marais, 2015). Our analyses now provide one possible explanation. The relaxation of the Tyr-mediated feedback inhibition may direct more carbon flow towards Tyr, and away from Phe biosynthesis (**Fig. 1a**), as demonstrated by increased Tyr and decreased Phe levels upon transient expression of *ADHα* (**Fig. 3**). This may create a surplus of Tyr at the expense of Phe-derived products such as anthocyanins. Furthermore, betalain-producing, *ADHα*-containing core Caryophyllales species accumulated more Tyr than plants not possessing *ADHα* (**Fig. 4c**). The involvement of other factors such as transcriptional regulation of betalain, anthocyanin, and Tyr/Phe pathway genes remain to be examined (Hatlestad *et al*., 2015; Ambawat *et al*, 2013), however our data provide a fascinating insight into the contribution of Tyr biosynthesis regulation to the evolution of a novel

betalain pigment biosynthesis.

Prior heterologous reconstructions of specialized metabolic pathways resulted in significant accumulations of Tyr-derived plant natural products, such as a cyanogenic glycoside, dhurrin, in Arabidopsis (*c*. 4% per dry weight, Tattersall *et al*., 2001; Kristensen *et al*., 2005) and betalains in tobacco (330 mg kg^{-1} approaching red beet extract of 760 mg kg^{-1}, Polturak *et al.*, 2016). In other cases, however, *DODA* and *CYP76AD1* expression in Arabidopsis still required feeding of Tyr for betalain production (Harris *et al*., 2012; Sunnadeniya et al., 2016). Therefore, "pulling" a precursor (e.g. Tyr) may not be always enough to efficiently produce its downstream product, and "pushing" the precursor supply may be also important. Indeed, in red beets, increased Tyr levels have a strong positive correlation with enhanced accumulation of betalains (Wang *et al*., 2017), suggesting that elevated production of Tyr plays important role in overall production of betalains. Over 100-fold increase in Tyr accumulation observed in *N. benthamiana* leaves expressing *ADHα* (**Fig. 3**) further demonstrates an exciting opportunity to introduce Caryophyllales ADHα enzymes into other plants and microbes, deregulate Tyr biosynthesis, and boost the availability of Tyr and the production of Tyr-derived products (e.g., vitamin E, isoquinoline alkaloids including morphine).

Acknowledgements

We thank S. Swanson for the help with confocal microscopy, S. R. Franco for writing assistance, A. Dewanjee for SoADH proteins expression, I. Goldman for the *B. vulgaris* seeds, G. J. Hatlestad and A. M. Lloyd for sharing the RNAseq data of table beet W357B, and to the Cambridge University Botanic Garden for access to living collections. This work is supported by the National Science Foundation Graduate Research Fellowship (grant no. DGE-1256259) to S.L.N. and the Agriculture and Food Research Initiative competitive grant (2015-67013-22955)

from the USDA National Institute of Food and Agriculture to H.A.M.

Author contributions

S.L.N. and H.A.M. designed research; Y.Y. and S.A.S. performed phylogenetic analysis; T.F. and S.F.B. characterized the *P. polygonifolia*, *S. marina*, and *H. latifolia ADHα*, A.T. and S.F.B. performed transient and RT-PCR of *BvADHα* and *BvADHβ* expression in *N. benthamiana*, M.W. quantified amino acid levels of the *N. benthamiana* tissues and analyzed Tyr sensitivity of ADH activity from plant tissues, and S.L.N. performed the rest of the experiments. S.L.N., Y.Y., S.F.B., M.W. and H.A.M. analyzed data, S.L.N., Y.Y., S.F.B. and H.A.M. wrote the paper.

References

Ambawat S, Sharma P, Yadav NR, Yadav RC. **2013**. MYB transcription factor genes as regulators for plant responses: an overview. *Phyisyiology and Molecular Biology of Plants* **19**: 307–321.

Bate-Smith EC. **1962**. The phenolic constituents of plants and their taxonomic significance. *Botanical Journal of the Linnean Society* **58**: 95–173.

Barshandy H, Jalkanen S, Teeri TH. 2015. Within leaf variation is the largest source of variation in agroinfiltration of *Nicotiana benthamiana*. *Plant Methods* **11**: 47.

Beaudoin GAW, Facchini PJ. **2014**. Benzylisoquinoline alkaloid biosynthesis in opium poppy. *Planta* **240**: 19–32.

Bentley R. **1990**. The ahikimate pathway – a metabolic tree with many branche. *Critical Reviews in Biochemistry and Molecular Biology* **25**: 307–84.

Biancardi E, Panella LW, and Lewellen R. 2012. *Beta maritima: the origin of beets*. New York, USA: Springer.

Bonvin J, Aponte RA, Marcantonio M, Singh S, Christendat D, Turnbull JL. **2006**.

Biochemical characterization of prephenate dehydrogenase from the hyperthermophilic bacterium *Aquifex aeolicus*. *Protein Science* **15**: 1417–32.

Brockington SF, Walker RH, Glover BJ, Soltis PS, Soltis DE. **2011**. Complex pigment evolution in the Caryophyllales. *New Phytologist* **190**: 854-864.

Brockington SF, Yang Y, Gandia-Herrero F, Covshoff S, Hibberd JM, Sage RF, Wong GKS, Moore MJ, Smith SA. **2015**. Lineage-specific gene radiations underlie the evolution of novel betalain pigmentation in Caryophyllales. *New Phytologist* **207**: 1170–1180.

Byng G, Whitaker R, Elick C, Jensen ROYA. **1981**. Enzymology of *L*-tyrosine biosynthesis in corn (*Zea mays*). *Phytochemistry* **20**: 1289–1292.

Chen F, Tholl D, Bohlmann J, Pichersky E. **2011**. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant Journal* **66**: 212–229.

Christinet L, Burdet F, Zaiko M, Hinz U, Zrÿd JP. **2004**. Characterization and functional identification of a novel plant 4,5-extradiol dioxygenase involved in betalain pigment biosynthesis in *Portulaca grandiflora*. *Plant Physiology* **134**: 265-274.

Clement JS, Mabry TJ. **1996**. Pigment evolution in the caryophyllales: a systematic overview. *Botanica Acta* **109**: 360–367.

Connelly JA, Conn EE. **1986**. Tyrosine biosynthesis in *Sorghum bicolor*: isolation and regulatory properties of arogenate dehydrogenase. *Zeitschrift für Naturforschung C* **41**: 69–78.

Dal Cin V, Tieman DM, Tohge T, McQuinn R, de Vos RCH, Osorio S, Schmelz E, Taylor MG, Smits-Kroon MT, Schuurink RC *et al.* 2011. Identification of genes in the phenylalanine metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit. *The Plant Cell* **23**: 2738–2753.

Delaux PM, Varala K, Edger PP, Coruzzi GM, Pires JC, Ané JM. **2014**. Comparative

phylogenomics uncovers the impact of symbiotic associations on host genome evolution. *PLoS Genetics* **10**: e1004487.

Brockington SF, Alexandre R, Ramdial J, Moore MJ, Crawley S, Dhingra A, Hilu K, Soltis PS. 2009. Phylogeny of the Caryophyllales sensu lato: Revisiting hypotheses on pollination biology and perianth differentiation in the core Caryophyllales. *International Journal of Plant Sciences* **170**: 627-643.

Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, Tafer H, Rupp O, Sörensen TR, Stracke R, Reinhardt R *et al.* 2014. The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). *Nature* **505**: 546–549.

Dornfeld C, Weisberg AJ, C RK, Dudareva N, Jelesko JG, Maeda HA. **2014**. Phylobiochemical characterization of class-Ib aspartate/prephenate aminotransferases reveals evolution of the plant arogenate phenylalanine pathway. *The Plant Cell* **26**: 3101–3114.

Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JDG, Patron NJ, and Marillonnet S. 2014. A golden gate modular cloning toolbox for plants. *ACS Synthetic Biol*ogy **3**: 839–843.

Gaines, C.G., Byng GS, Whitaker RJ and, Jensen RA. **1982**. L-Tyrosine regulation and biosynthesis via arogenate dehydrogenase in suspension-cultured cells of *Nicotiana silvestris* Speg. et Comes. *Planta* **156**: 233–240.

Gandía-Herrero F, García-Carmona F. **2012**. Characterization of recombinant *Beta vulgaris* 4,5-DOPA-extradiol-dioxygenase active in the biosynthesis of betalains. *Planta* **236**: 91–100.

Gleadow RM, Møller BL. **2014**. Cyanogenic glycosides: synthesis, physiology, and phenotypic plasticity. *Annual Review of Plant Biology* **65**: 155–185.

Greenberg AK and Donoghue MJ. 2011. Molecular systematics and character evolution in Caryophyllaceae. *TAXON* **60**: 1637–1652

Goldman IL. **1996**. A list of germplasm releases from the University of Wisconsin table beet breeding program. *HortScience* **31**: 880–881.

Hanson M, Gaut BS, Stec AO, Fuerstenberg SI, Goodman MM, Coe EH, Doebley JF. **1996**. Evolution of anthocyanin biosynthesis in maize kernels: the role of regulatory and enzymatic loci. *Genetics* **143**: 1395–1407.

Hatlestad GJ, Sunnadeniya RM, Akhavan N, Gonzalez A, Goldman IL, McGrath JM, Lloyd AM. **2012**. The beet R locus encodes a new cytochrome P450 required for red betalain production. *Nature Genetics* **44**: 816–820.

Hudson GS, Wong V, Davidson B. **1984**. Chorismate mutase/prephenate dehydrogenase from *Escherichia coli* K12: purification, characterization, and identification of a reactive cysteine. *Biochemistry* **23**: 6240–6249.

Ibarra-Laclette E, Zamudio-Hernández F, Pérez-Torres CA, Albert VA, Ramírez-Chávez E, Molina-Torres J, Fernández-Cortes A, Calderón-Vázquez C, Olivares-Romero JL, Herrera-Estrella A *et al.* 2015. *De novo* sequencing and analysis of *Lophophora williamsii* transcriptome, and searching for putative genes involved in mescaline biosynthesis. *BMC Genomics* **16**: 657.

Khan MI. **2015**. Plant betalains: safety, antioxidant activity, clinical efficacy, and bioavailability. *Comprehensive Reviews in Food Science and Food Safety* **15**: 316–330.

Kajikawa M, Sierro N, Kawaguchi H, BakaherN, Ivanov NV, Hashimotp T, Shoji T. **2017**. Genomic insights into the evolution of the nicotine biosynthesis pathway in tobacco. *Plant Physiology* **4**: 999–1011.

Kristensen C, Morant M, Olsen CE, Ekstrom CT, Galbraith DW, Moller BL, Bak S. **2005**. Metabolic engineering of dhurrin in transgenic *Arabidopsis* plants with marginal inadvertent effects on the metabolome and transcriptome. *Proceedings of the National Academy of Sciences,*

USA **102**: 1779–1784.

Lee EJ, An D, Nguyen CTT, Patill BS, Kim J, Yoo KS. 2014. Betalain and betaine composition of greenhouse- or field-produced beetroot (*Beta vulgar*is L.) and inhibition of HepG2 cell proliferation. *Journal of Agriculture and Food Chemistry* **62**: 1324–1331.

Legrand P, Dumas R, Seux M, Rippert P, Ravelli R, Ferrer JL, Matringe M. **2006**. Biochemical characterization and crystal structure of *Synechocystis* arogenate dehydrogenase provide insights into catalytic reaction. *Structure* **14**: 767–776.

Mabry TJ. **1964**. The betacyanins, a new class of red violet pigments, and their phylogenetic significance. In: Leone CA, ed. *Taxonomic biochemistry, physiology, and serology*. New York, NY, USA: Ronald Press: 239-254.

Maeda H, Dudareva N. **2012**. The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annual Review of Plant Biology* **63**: 73–105.

Mene-Saffrane L, Jones AD, DellaPenna D. **2010**. Plastochromanol-8 and tocopherols are essential lipid-soluble antioxidants during seed desiccation and quiescence in Arabidopsis. *Proceedings of the National Academy of Sciences, USA* **107**: 17815–17820.

Millgate AG, Pogson BJ, Wilson IW, Kutchan TM, Zenk MH, Gerlach WL, Fist AJ, Larkin PJ. **2004**. Analgesia: Morphine-pathway block in top1 poppies. *Nature* **431**: 413–414.

Mizutani M, Ohta D. **2010**. Diversification of P450 genes during land plant evolution. *Annual Review of Plant Biology* **61**: 291–315.

Moghe GD, Last RL. **2015**. Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiology* **169**: 1512–23.

Neelwarne B, Halagur SB. 2012. Red beet: an overview. In B. Neelwarne, ed. *Red beet biotechnology: food and pharmaceutical applications*. New York, New York USA: Springer, 1-43.

Panchy N, Lehti-Shiu M, Shiu S-H. **2016**. Evolution of gene duplication in plants. *Plant Physiology* **171**: 2294-2316.

Pencharz PB, Hsu JW-C, Ball RO. **2007**. Aromatic amino acid requirements in healthy human subjects. *The Journal of Nutrition* **137**: 1576S–1578S.

Pichersky E, Lewinsohn E. **2011**. Convergent evolution in plant specialized metabolism. *Annual Review of Plant Biology* **62**: 549–566.

Polturak G, Breitel D, Grossman N, Sarrion-Perdigones A, Weithorn E, Pliner M, Orzaez D, Granell A, Rogachev I, Aharoni A. **2016**. Elucidation of the first committed step in betalain biosynthesis enables the heterologous engineering of betalain pigments in plants. *New Phytologist* **210**: 269–283.

Rapp RA, Haigler CH, Flagel L, Hovav RH, Udall JA, Wendel JF. **2010**. Gene expression in developing fibres of upland cotton (*Gossypium hirsutum* L.) was massively altered by domestication. *BMC Biology* **8**: 1–15.

Rippert P, Matringe M. **2002a**. Purification and kinetic analysis of the two recombinant arogenate dehydrogenase isoforms of *Arabidopsis thaliana*. *European Journal of Biochemistry* **269**: 4753–4761.

Rippert P, Matringe M. **2002b**. Molecular and biochemical characterization of an *Arabidopsis thaliana* arogenate dehydrogenase with two. *Plant Molecular Biology* **48**: 361–368.

Rippert P, Puyaubert J, Grisollet D, Derrier L, Matringe M. **2009**. Tyrosine and phenylalanine are synthesized within the plastids in Arabidopsis. *Plant Physiology* **149**: 1251–1260.

Rong J, Lammers Y, Strasburg JL, Schidlo NS, Ariyurek Y, de Jong TJ, Klinkhamer PGL, Smulders MJM, Vrieling K. **2014**. New insights into domestication of carrot from root transcriptome analyses. *BMC Genomics* **15**: 895.

Rubin JL, Jensen RA. **1979**. Enzymology of *L*-tyrosine biosynthesis in mung bean (*Vigna radiata* [L.] Wilczek). *Plant Physiology* **64**: 727–734.

Schenck CA, Holland CK, Schneider MR, Men Y, Lee SG, Jez JM, Maeda HA. **2017**. Molecular Basis of the evolution of alternative tyrosine biosynthetic routes in plants. *Nature Chemocal Biology* **13**: 1029–1035.

Schenck CA, Chen S, Siehl DL, Maeda HA. **2015**. Non-plastidic, tyrosine-insensitive prephenate dehydrogenases from legumes. *Nature Chemical Biology* **11**: 52–57.

Shimada S, Inoue YT, Sakuta M. **2005**. Anthocyanidin synthase in non-anthocyanin-producing Caryophyllales species. *The Plant Journal* **44**: 950–959.

Siehl DL. **1999**. The biosynthesis of tryptophan, tyrosine, and phenylalanine from chorismate. In: Singh B, ed. *Plant amino acids: biochemistry and biotechnology*. New York, USA: CRC Press, , 171–204.

Sparkes IA, Runions J, Kearns A, Hawes C. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols* **1**: 2019-2025.

Sun W, Shahinas D, Bonvin J, Hou W, Kimber MS, Turnbull J, Christendat D. 2009. The crystal structure of *Aquifex aeolicus* prephenate dehydrogenase reveals the mode of tyrosine inhibition. *Journal of Biological Chemistry* **284**: 13223–13232.

Sunnadeniya R, Bean A, Brown M, Akhavan N, Hatlestad G, Gonzalez A, Symonds VV, Lloyd A. **2016**. Tyrosine hydroxylation in betalain pigment biosynthesis is performed by cytochrome P450 enzymes in beets (*Beta vulgaris*). *PLoS ONE* **11**: 1–16.

Tanaka Y, Sasaki N, Ohmiya A. **2008**. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant Journal* **54:** 733-749.

Tattersall DB, Bak S, Jones PR, Olsen CE, Nielsen JK, Hansen ML, Høj PB, Møller BL.

2001. Resistance to an herbivore through engineered cyanogenic glucoside synthesis. *Science* **293:** 1826-1828.

Tzin V, Galili G. 2010. The biosynthetic pathways for shikimate and aromatic amino acids in *Arabidopsis thaliana*. *The Arabidopsis Book/American Society of Plant Biologists* **8**: e0132.

Wang M, Lopez-Nieves S, Goldman IL, Maeda HA. 2017. Limited tyrosine utilization explains lower betalain contents in yellow than red table beet genotypes. *Journal of Agricultural and Food Chemistry* **65**: 4305-4313.

Wang X, Xiao H, Chen G, Zhao X, Huang C, Chen C, Wang F. 2011. Isolation of high-quality RNA from *Reaumuria soongorica*, a desert plant rich in secondary metabolites. *Molecular Biotechnology* **48**: 165-172.

Weber E, Engler C, Gruetzner R, Werner S, and Marillonnet S. 2011. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* **18**: 6:e16765.

Weng JK. 2014. The evolutionary paths towards complexity: A metabolic perspective. *New Phytologist* **201**: 1141–1149.

Weng JK, Philippe RN, Noel JP. 2012. The rise of chemodiversity in plants. *Science* **336**; 1667-1670.

Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2014. RELAX: detecting relaxed selection in a phylogenetic framework. *Molecular Biology and Evolution* **32**: 1–13.

Xu S, Brockmöller T, Navarro-Quezada A, Kuhl H, Gase K, Ling Z, Zhou W, Kreitzer C, Stanke M, Tang H *et al*. 2017. Wild tobacco genomes reveal the evolution of nicotine biosynthesis. *Proceedings of the National Academy of Sciences, USA* **114**: 6133–6138.

Xu S, Huang Q, Lin C, Lin L, Zhou Q, Lin F, He E. 2016. Transcriptome comparison reveals candidate genes responsible for the betalain-/anthocyanidin-production in bougainvilleas.

Functional Plant Biology **43**: 278–286.

Yagi M, Kosugi S, Hirakawa H, Ohmiya A, Tanase K, Harada T, Kishimoto K, Nakayama M, Ichimura K, Onozaki T *et al.* 2014. Sequence analysis of the genome of carnation (*Dianthus caryophyllus* L.). *DNA Research* **21**: 231–241.

Zhang R, Guo C, Zhang W, Wang P, Li L, Duan X, Du Q, Zhao L, Shan H, Hodges SA *et al.* 2013. Disruption of the petal identity gene APETALA3-3 is highly correlated with loss of petals within the buttercup family (Ranunculaceae). *Proceedings of the National Academy of Sciences, USA* **110**: 5074–5079.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Physical location, homology, and phylogeny of BvADHα and BvADHβ.

Fig. S2 ADH but not PDH activity was detected from *B. vulgaris* tissues or recombinant enzyme**.**

Fig. S3 BvADHs prefer $NADP^+$ over NAD^+ as cofactor.

Fig. S4 No amino acid changes were found in the mature protein coding region of BvADHα among different *B. vulgaris* varieties.

Fig. S5 Recombinant His-tagged BvADHα also exhibits reduced sensitivity to Tyr relative to AtADH2.

Fig. S6 BvADHs are not inhibited by phenylalanine, tryptophan, and betanin.

Fig. S7 Transgene expression and tyrosine levels of individual leaf samples of infiltrated *Nicotiana benthamiana*.

Fig. S8 Tyr sensitivity of ADH activity from plant tissues.

Fig. S9 ADHα sequences used for texting relax selection.

Fig. S10 Histidine 217 residue responsible for Tyr sensitivity of *Aquifex aeolicus* PDH (AaPDH) is still present in BvADHα.

Table S1 Primers used as indicated in the text and methods

Table S2 Sequences of Caryophyllales (ingroups) and non-Caryophyllales (outgroups) used in this study

Table S3 Amino acid levels of *Nicotiana benthamiana* leaves expressing *GFP*, *BvADHα,* or *BvADHβ*

Table S4 RELAX analysis support the acceleration in amino acid substitution in Caryophyllales is due to relaxed purifying selection, instead of intensified positive selection

Notes S1 Alignment of ADH amino acid sequences inferred using MAFFT. **[Author, please insert a mention to Notes S1 to S4 in the main text, to fit journal style.]**

Notes S2 ADHα codon alignment used in RELAX analysis.

Notes S3 Trimmed ADHα codon alignment used in RELAX analysis.

Notes S4 Trimmed ADHα amino acid alignment used in RELAX analysis.

Fig. 1 *Beta vulgaris* have two ADH enzymes localized in the plastids. (a) Tyrosine and betalain biosynthetic pathways in plants. *L*-Tyrosine (Tyr) can be synthesized from prephenate via

arogenate dehydrogenase (ADH/TyrA$_a$) or prephenate dehydrogenase (PDH/TyrA$_p$). Tyr is exported from the plastid to cytosol and then converted to *L*-dihydroxyphenylalanine (*L*-DOPA) by CYP76AD1α, CYP76AD5, and CYP76AD6 (CYP76AD1α/5/6). *L*-DOPA is then eventually converted to betalains, red betacyanins and yellow betaxanthins. Biosynthesis of Tyr competes for arogenate or prephenate substrate with that of *L*-phenylalanine (Phe), the precursor of anthocyanins. Blue lines denote feedback regulation by Tyr. DODA, *L*-DOPA dioxygenase. (b) Arogenate substrate was incubated with the purified recombinant enzymes of BvADHα or BvADHβ together with $NADP^+$ cofactor and the production of Tyr was analyzed. The HPLC traces were offset for presentation. *Arabidopsis thaliana* ADH2 (AtADH2) was used as a control for the ADH assay. (c) Green fluorescence protein (GFP) was fused at the C-terminal of BvADHα and BvADHβ and transiently expressed in Arabidopsis protoplasts. Free GFP and GFP-fused Arabidopsis ADH2 (AtADH2) were used as controls for cytosolic and plastidic localization, respectively. Representative images show GFP fluorescence and chlorophyll autofluorescence in green and magenta, respectively. Bars, 10 μm. (d) Expression levels of *BvADHα* and *BvADHβ* were compared with those of betalain pathway genes in the cotyledon and hypocotyl of 7-d-old sugar beet and red beet (W357B). Significant differences between the two genotypes: *, $P < 0.05$, (Student's *t*-test). Bars represent percent expression relative to the sample with the highest expression. Data are means of three biological replicates ± SE. nd, not detectable.

Fig. 2 Beet and spinach ADHα but not ADHβ have reduced sensitivity to Tyr. ADH activity was measured at different Tyr concentrations using $NADP^+$ cofactor and purified recombinant ADH enzymes of beet (BvADHα, BvADHβ), spinach (SoADHα, SoADHβ), and Arabidopsis (AtADH2). Data are expressed as the percentage of respective control activity without Tyr (0

μM) and means of three independent experiments ± SE. nd, not detectable; nt, not tested.

Fig. 3 Heterologous expression of *BvADHα* but not *BvADHβ* increases tyrosine levels in *Nicotiana benthamiana*. *Agrobacterium tumefaciens* carrying the construct of *35S::GFP*, *35S::BvADHα*, or *35S::BvADHβ* was infiltrated to *N. benthamiana* leaves, which were analyzed for amino acid contents using GC-MS. Levels of tyrosine (a) and phenylalanine (b). Significant differences from the *35S::GFP* control: *, $P < 0.05$ (Student's *t*-test). Data are means ± SE (n = 5).

Fig. 4 Phylogenetic distribution of ADHα in Caryophyllales. The blue and pink branches represent anthocyanin and betalain-producing families, respectively, while families with unclear/unidentified pigmentation are shown in gray. (a) Maximum-likelihood phylogeny of *ADH* genes in Caryophyllales. Scale bar indicates inferred number of amino acid substitution per site. ADH enzymes characterized in this study are indicated at the end of each branch. (b) Presence and absence of *BvADHα* and *BvADHβ* orthologs detected from genome or transcriptome data was mapped to the family-level phylogenetic tree of the Caryophyllales order. Filled circles denote that corresponding orthologs were detected in all species within the family, whereas partially filled circles indicate that the filled portion of the species within each family had corresponding orthologs. Open circles denote no corresponding orthologs were detected. Red lines indicate estimating timings of duplication events of *ADH* and betalain pathway genes (*CYP76AD1* and *DODA*). Dash lines (-) represent families with no available transcriptomic or genomic data. **(c)** Tyr contents were analyzed in various Caryophyllales species. *Arabidopsis thaliana* was used as outgroup. Orange bars indicate species having *ADHα* orthologs. Young leaf tissues were used for all samples except a Cactaceae species, in which flowers were used to avoid

succulent tissues. Significant difference from Arabidopsis (*, $P<0.05$) based on fixed effect model (see the Materials and Methods section). Also, a statistical analysis based on the mixed effect model showed significant differences between two groups, plants with and without ADHα ($P<0.0001$). Bars represent means ± SE (n = four biological replicates).

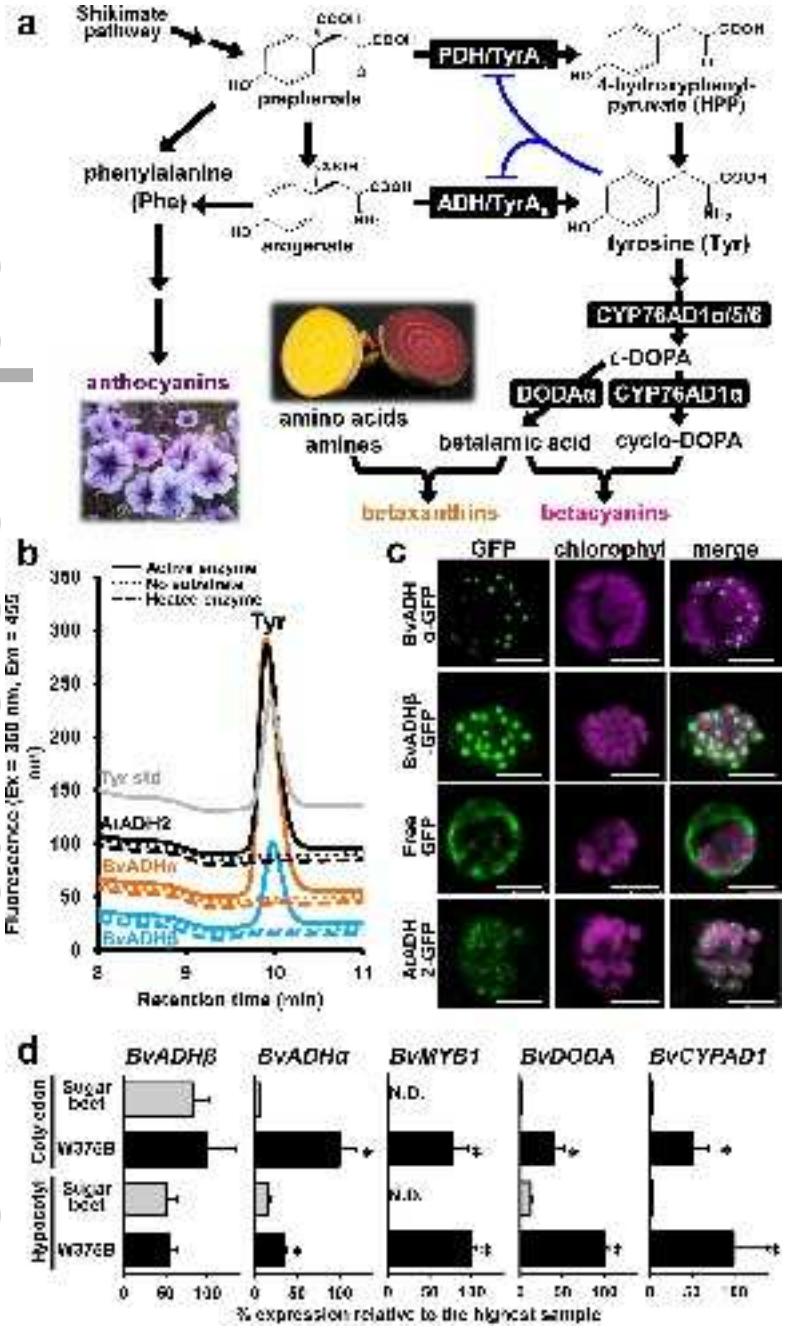
Fig. 5 ADHα from various species of core Caryophyllales also exhibit relaxed sensitivity. ADH activity was measured under different Tyr concentrations using purified recombinant ADH enzymes of *Nepenthes ventricosa* × *alata* (NaADHβ), *Rivina humilis* (RhADHα), *Mirabilis jalapa* (MjADHα), and *Portulaca oleracea* (PoADHα) ADH. Data are expressed as the percentage of respective control activity without Tyr (0 μM) and the mean of three independent experiments ± SE. nd, not detectable; nt, not tested.

Table 1 Tyr sensitivity of ADH activity from plant tissue extracts

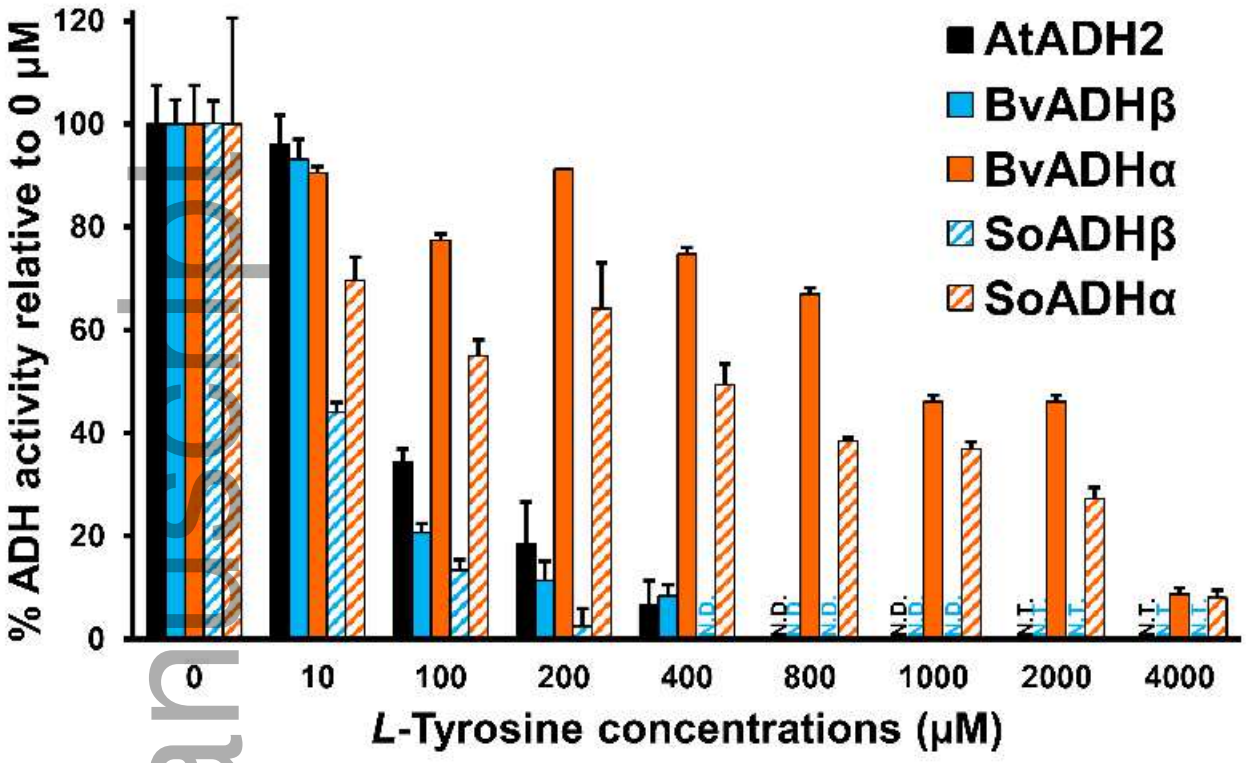
Species	**ADH activity (nmol mg^{-1} protein)**		**Inhibition (%)**
	0 mM 3-fluoro-Tyr	0.5 mM 3-fluoro-Tyr	
Spinach oleracea	66.4 ± 5.0	52.7 ± 1.9	20.7%
Dianthus barbatus	18.1 ± 0.3	0.9 ± 0.2	95.0%
Arabidopsis thaliana	93.5 ± 5.2	7.8 ± 0.5	91.6%

Total protein extracts of spinach, *Dianthus barbatus*, and Arabidopsis leaf tissues were used to analyze ADH activity in the presence and absence of 0.5 mM Tyr analog (3-fluoro-Tyr), which were used to calculate percent inhibition. ADH activity was measured with 1 mM arogenate

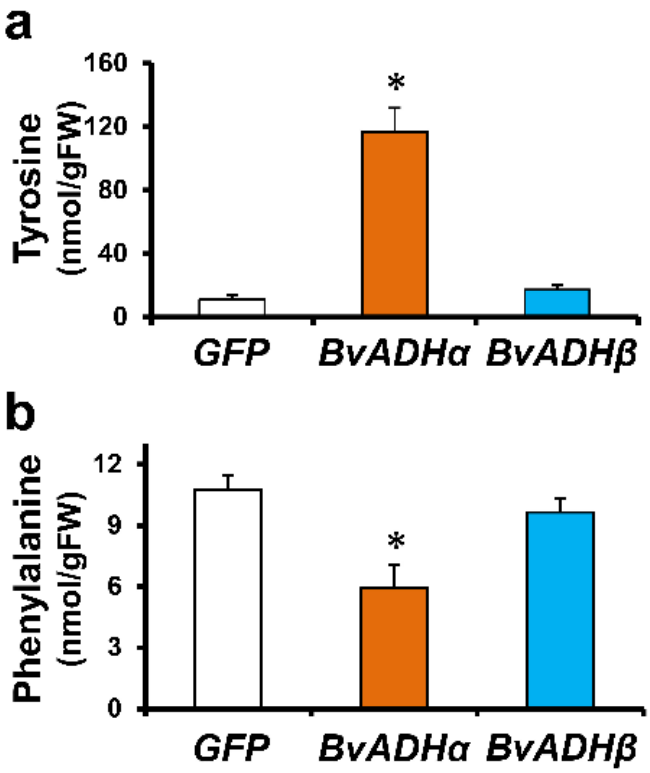
substrate and 1 mM $NADP^+$ cofactor during 2 h incubation (see Supporting Information **Fig. S8**). Data are means ± SE (n = 4).



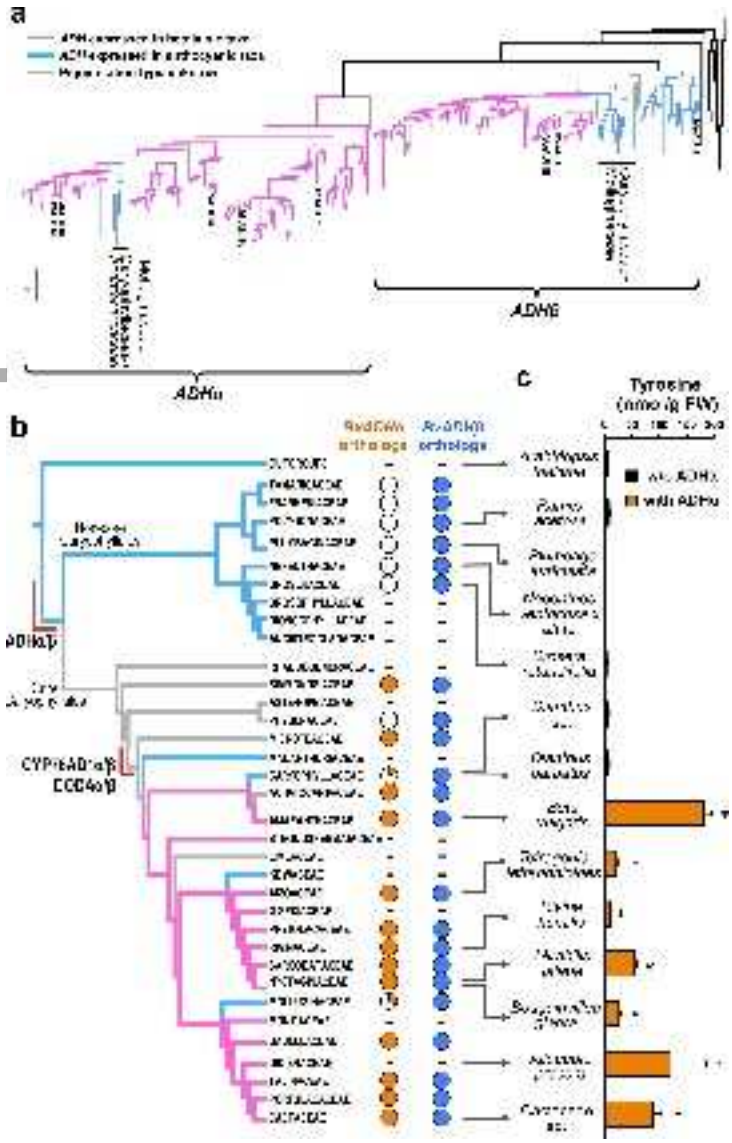
nph_14822_f1.jpg



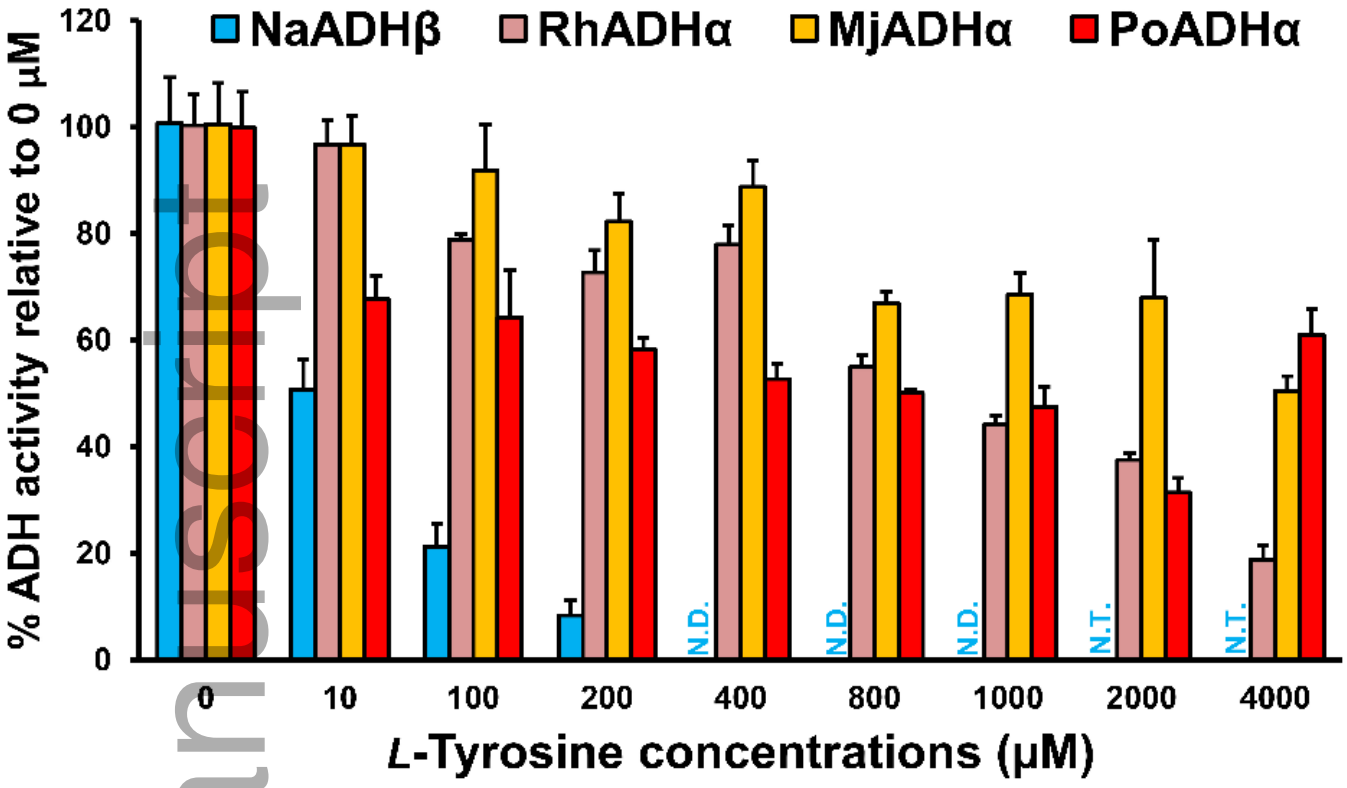
nph_14822_f2.tif



nph_14822_f3.tif



nph_14822_f4.jpg



nph_14822_f5.tif