

# Two independently evolved natural mutations additively deregulate TyrA enzymes and boost tyrosine production *in planta*

Samuel Lopez-Nieves<sup>1,2</sup> , Jorge El-Azaz<sup>1</sup>, Yusen Men<sup>1</sup>, Cynthia K. Holland<sup>3</sup> , Tao Feng<sup>2</sup>, Samuel F. Brockington<sup>2</sup> , Joseph M. Jez<sup>4</sup>  and Hiroshi A. Maeda<sup>1,\*</sup> 

<sup>1</sup>Department of Botany, University of Wisconsin–Madison, Madison, WI 53706, USA,

<sup>2</sup>Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK,

<sup>3</sup>Department of Biology, Williams College, Williamstown, MA 01267, USA, and

<sup>4</sup>Department of Biology, Washington University in St Louis, St Louis, MO 63130, USA

Received 17 September 2021; revised 29 October 2021; accepted 15 November 2021.

\*For correspondence (e-mail maeda2@wisc.edu).

## SUMMARY

L-Tyrosine is an essential amino acid for protein synthesis and is also used in plants to synthesize diverse natural products. Plants primarily synthesize tyrosine via TyrA arogenate dehydrogenase (TyrA<sub>a</sub> or ADH), which are typically strongly feedback inhibited by tyrosine. However, two plant lineages, Fabaceae (legumes) and Caryophyllales, have TyrA enzymes that exhibit relaxed sensitivity to tyrosine inhibition and are associated with elevated production of tyrosine-derived compounds, such as betalain pigments uniquely produced in core Caryophyllales. Although we previously showed that a single D222N substitution is primarily responsible for the deregulation of legume TyrAs, it is unknown when and how the deregulated Caryophyllales TyrA emerged. Here, through phylogeny-guided TyrA structure–function analysis, we found that functionally deregulated TyrAs evolved early in the core Caryophyllales before the origin of betalains, where the E208D amino acid substitution in the active site, which is at a different and opposite location from D222N found in legume TyrAs, played a key role in the TyrA functionalization. Unlike legumes, however, additional substitutions on non-active site residues further contributed to the deregulation of TyrAs in Caryophyllales. The introduction of a mutation analogous to E208D partially deregulated tyrosine-sensitive TyrAs, such as Arabidopsis TyrA2 (AtTyrA2). Moreover, the combined introduction of D222N and E208D additively deregulated AtTyrA2, for which the expression in *Nicotiana benthamiana* led to highly elevated accumulation of tyrosine *in planta*. The present study demonstrates that phylogeny-guided characterization of key residues underlying primary metabolic innovations can provide powerful tools to boost the production of essential plant natural products.

**Keywords:** TyrA arogenate dehydrogenase (TyrA<sub>a</sub>/ADH), anthocyanins, betalains, Caryophyllales, tyrosine biosynthesis, site-directed mutagenesis.

## INTRODUCTION

Recent developments in synthetic biology and gene editing technologies provide exciting opportunities to improve chemical production in various organisms (Cravens et al., 2019; Nishida and Kondo, 2021; Owen et al., 2017; Pyne et al., 2019; Xu and Qi, 2019), including in plants that can naturally produce abundant and diverse phytochemicals directly from atmospheric CO<sub>2</sub> and soil nutrients using sunlight as energy (Maeda, 2019b; Shih, 2018). Although specialized metabolic genes/enzymes have been rapidly identified from a variety of plants (Nett et al., 2020; Irmisch, et al., 2020; Jacobowitz and Weng 2020; Jozwiak et al.,

2020), limited tools are available to modulate plant primary metabolism because of the assumption that primary metabolic enzymes are highly conserved among different plant species (Moghe and Last 2015; Moore et al., 2019). There are, however, some examples of lineage-specific diversification of primary metabolism (Maeda, 2019a) that can be harnessed to modulate ‘hard-to-engineer’ plant primary metabolism (Maeda, 2019b) and provide a unique opportunity to understand how the overall metabolic networks have evolved in plants (Maeda and Fernie, 2021).

Besides serving as an essential building block for proteins, the aromatic amino acid L-tyrosine is a key precursor

for synthesis of numerous plant natural products (Maeda and Dudareva, 2012; Schenck and Maeda, 2018; Tzin and Galili, 2010), such as isoquinoline alkaloids, catecholamines, and betalain pigments (Figure 1a). Plants usually synthesize tyrosine via TyrA aroenate dehydrogenase (TyrA<sub>a</sub> or ADH), such as in *Arabidopsis* (de Oliveira et al., 2019; Rippert and Matringe, 2002a, b), although some plants in the Fabaceae (legume) family, such as *Glycine max* (soybean) and *Medicago truncatula* (barrelclover), use an alternative pathway requiring TyrA prephenate dehydrogenase (TyrA<sub>p</sub> or PDH) to synthesize tyrosine (Schenck et al., 2015; 2017) (Figure 1a). TyrA enzymes are typically feedback inhibited by tyrosine with a half-maximal inhibitory concentration ( $IC_{50}$ )  $\leq 100 \mu\text{M}$  (Connelly and Conn, 1986; Rippert and Matringe, 2002a; Schenck and Maeda, 2018), whereas legume TyrA<sub>p</sub> enzymes are insensitive to tyrosine inhibition (Schenck et al., 2015), supporting the increased production of tyrosine in some legumes (Coley et al., 2019). Phylogeny-guided structure–function analyses further identified that the substitution of an active site aspartate to asparagine (D222N) confers the prephenate substrate specificity and tyrosine insensitivity of legume TyrA<sub>p</sub> (Schenck et al., 2017a). Thus, a single amino acid change underlies the functionalization of the TyrA enzymes and hence the elevated production of tyrosine-derived compounds in legume species.

Many plant species in the order Caryophyllales, such as table beets (*Beta vulgaris*), uniquely produce tyrosine-derived red to yellow pigments, betalains, in place of ubiquitous red to purple anthocyanin pigments derived from a different aromatic amino acid, phenylalanine (Figure 1a). Besides the evolution of specialized metabolic enzymes in betalain-producing core Caryophyllales (Brockington et al., 2015; Hatlestad et al., 2012; Sheehan et al., 2020), TyrA<sub>a</sub> genes were found to be duplicated into TyrA $\alpha$  and TyrA $\beta$  (or ADH $\alpha$  and ADH $\beta$ ) early in the core Caryophyllales before the origin of betalains (Figure 1b). TyrA $\beta$  is expressed constitutively, whereas TyrA $\alpha$  is co-expressed with betalain biosynthetic genes in betalain producing tissues (Lopez-Nieves et al., 2018). Biochemical characterization revealed that the TyrA $\alpha$  of betalain-producing species, such as table beets (*B. vulgaris*, BvTyrA $\alpha$ ) and spinach (*Spinacia oleracea*, SoTyrA $\alpha$ ), exhibit relaxed sensitivity to tyrosine with  $IC_{50} > 500 \mu\text{M}$ , whereas TyrA $\beta$  enzymes of core or non-core Caryophyllales are still inhibited by tyrosine with  $IC_{50} < 100 \mu\text{M}$  (Lopez-Nieves et al., 2018). Thus, the relaxed regulation of TyrA $\alpha$  likely increased the availability of tyrosine, which is further converted to the betalain pigments uniquely produced in the core Caryophyllales (Lopez-Nieves et al., 2018) (Figure 1a). When and how functionally deregulated TyrA $\alpha$  emerged during the evolution of the order Caryophyllales remains to be determined.

To address these questions, the present study carried out biochemical characterization of TyrA $\alpha$  enzymes, with a

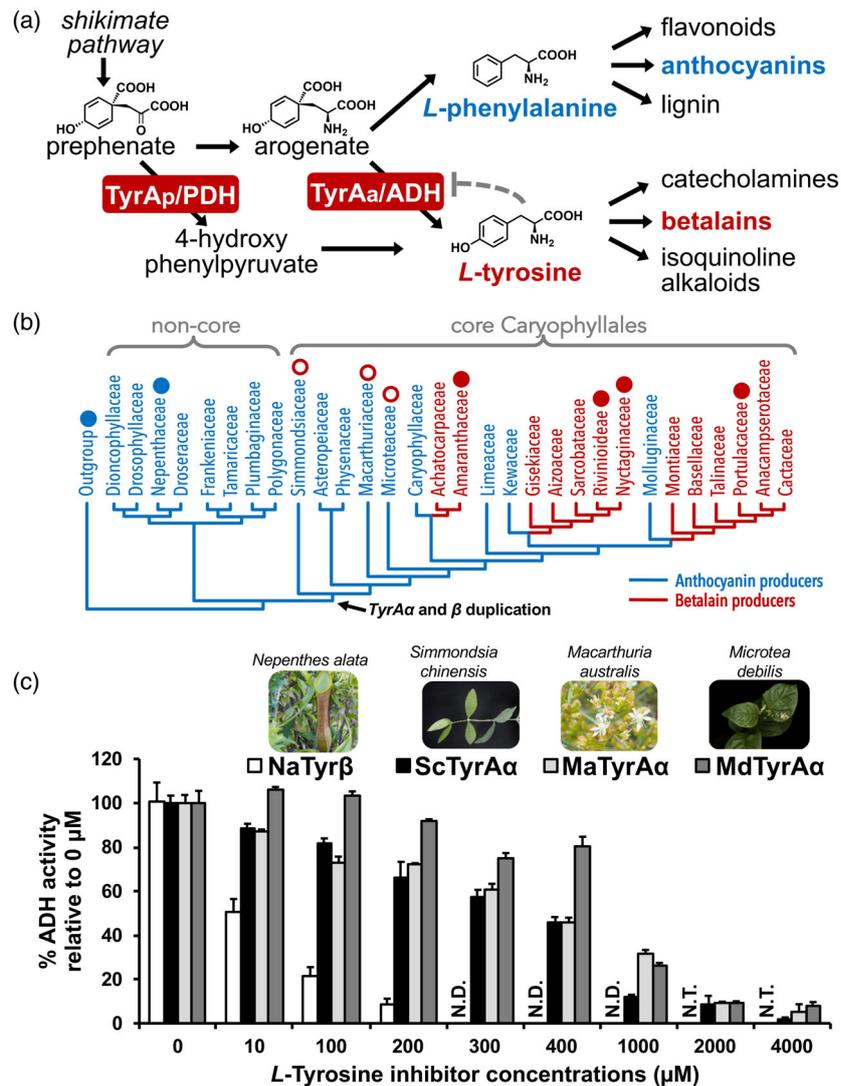
particular focus on non betalain-producing core Caryophyllales species, which are sister to betalain-producing lineages (Figure 1b). These TyrA $\alpha$  enzymes showed relaxed sensitivity to tyrosine, suggesting that the deregulation of tyrosine biosynthesis indeed occurred before the evolution of the betalain biosynthetic pathway. Through sequence comparison of closely-related TyrAs but having distinct properties (i.e. tyrosine sensitivity), we determined that the substitution of one active site residue, E208D, is critical for TyrA $\alpha$  deregulation. We also found that substitutions of multiple non-active site residues further enhance the relaxation of TyrA $\alpha$  feedback inhibition in core Caryophyllales. The introduction of E208D together with D222N (the key substitutions of deregulated TyrA of Caryophyllales and legumes, respectively) abolished feedback inhibition in distantly related tyrosine-sensitive TyrA2 from *Arabidopsis thaliana* (AtTyrA2). Moreover, expression of these tyrosine-insensitive AtTyrA2 variants in *Nicotiana benthamiana* resulted in hyper accumulation of tyrosine *in planta*.

## RESULTS

### Deregulated TyrA $\alpha$ enzymes evolved early in core Caryophyllales before the origin of betalain pigments

Mining of transcriptome data from over 100 plant species previously found that the presence of the TyrA $\alpha$  genes positively associated with the presence of betalain pigmentation and that these TyrA $\alpha$  genes from betalain-producing families (e.g. Amaranthaceae, Nyctaginaceae, Rivinaceae, Portulacaceae) (Figure 1b) indeed encode TyrA $\alpha$  with relaxed sensitivity to tyrosine (Lopez-Nieves et al., 2018). Non betalain-producing families, such as Simmondsiaceae, Macarthuraceae, and Microteaceae, which split early from the rest of the core Caryophyllales (Timoneda et al. 2020; Thulin et al., 2016; Walker et al., 2018), also have the TyrA $\alpha$  genes (Lopez-Nieves et al., 2018) (Figure 1b). These observations raise questions about the function of the TyrA $\alpha$  genes in the early-diverging group in core Caryophyllales prior to the origin of betalains.

To test whether these TyrA $\alpha$  genes indeed encode TyrA $\alpha$  enzymes that have relaxed sensitivity to tyrosine inhibition (e.g. *B. vulgaris* TyrA $\alpha$ ), the TyrA $\alpha$  genes were cloned from available non betalain-producing species of Simmondsiaceae, Macarthuraceae, and Microteaceae, which include *Simmondsia chinensis* (commonly known as Jojoba), *Macarthuria australis*, and *Microtea debilis* (ScTyrA $\alpha$ , MaTyrA $\alpha$ , and MdTyrA $\alpha$ , respectively). These TyrA $\alpha$  proteins were expressed in *Escherichia coli*, and their recombinant enzymes were purified and characterized for their tyrosine sensitivity. When their aroenate dehydrogenase activity was measured in the absence and presence of varied concentration of tyrosine, all three enzymes still showed relaxed sensitivity to tyrosine feedback inhibition. ScTyrA $\alpha$ , MaTyrA $\alpha$ , and MdTyrA $\alpha$  enzymes retained 12, 31,



**Figure 1.** Non betalain producers of core Caryophyllales also have a functionally deregulated TyrA $\alpha$  enzyme.

(a) Plants can mainly synthesize L-tyrosine from aroenate by aroenate dehydrogenases (TyrA $\alpha$ /ADH, red box), which are typically feedback inhibited by tyrosine (gray dashed line) to balance the use of aroenate for production of tyrosine and phenylalanine, key precursors of various specialized metabolites including betalain and anthocyanin pigments, respectively.

(b) Many families of core Caryophyllales (red) produce betalain pigments (red), whereas other Caryophyllales including non-core Caryophyllales produce anthocyanins (blue). A previous study (Lopez-Nieves et al., 2018) found that betalain-producing core Caryophyllales have TyrA $\alpha$  enzymes exhibiting relaxed sensitivity to tyrosine inhibition (filled red dots), whereas other plants have tyrosine-sensitive TyrA $\alpha$  (filled blue dots). The present study characterized the functionality of TyrA $\alpha$  enzymes from non betalain-producing core Caryophyllales (open red dots). A TyrA $\alpha$  gene duplicated into TyrA $\alpha$  and TyrA $\beta$  (black arrow) early in the core Caryophyllales before the origin of betalains (Sheehan et al., 2020).

(c) Aroenate dehydrogenase activity measured in the presence of different tyrosine concentrations using recombinant TyrA $\alpha$  from *Simmondsia chinensis* (ScTyrA $\alpha$ ), *Microtea debilis* (MdTyrA $\alpha$ ), and *Macarthuria australis* (MaTyrA $\alpha$ ). The data for tyrosine-sensitive TyrA $\alpha$  enzymes from *Nepenthes alata* (NaTyrA $\beta$ , previously reported in Lopez-Nieves et al., 2018) are shown as the control. The  $IC_{50}$  values of the enzymes are shown in Table 1. The data are shown as a percentage of the relative respective 0  $\mu$ M value ( $n = 3$ ). N.D., not detectable. N.T., not tested.

and 26% of activity, respectively, even at 1 mM tyrosine (Figure 1c), whereas TyrA $\beta$  from *Nepenthes alata* (NaTyrA $\beta$ ), a member of non-core Caryophyllales, showed no detectable activity (Lopez-Nieves et al., 2018). The  $IC_{50}$  values of ScTyrA $\alpha$ , MaTyrA $\alpha$ , and MdTyrA $\alpha$  were in the range 371–670  $\mu$ M, being 22- to 39-fold higher than that of NaTyrA $\beta$  from the non-core Caryophyllales (17  $\mu$ M)

(Table 1). Despite the lack of detectable betalain pigmentations, *Simmondsia*, *Macarthuria*, and *Microtea*, which diverged early in the core Caryophyllales, contain deregulated TyrA $\alpha$  enzymes. Taken together with previous reconstruction of the origin of betalain pigmentation (Brockington et al., 2015; Sheehan et al., 2020; Walker et al., 2018), these results indicate that deregulated TyrA $\alpha$

**Table 1** The  $IC_{50}$  values of TyrA wild-type and mutant enzymes, as calculated using PRISM ( $n = 3$ )

Enzyme abbreviations	Species	$IC_{50}$ value ( $\mu\text{M}$ )
ScTyrA $\alpha$	<i>Simmondsia chinensis</i>	371.0 $\pm$ 20.5
MdTyrA $\alpha$	<i>Microtea debilis</i>	582.9 $\pm$ 17.9
MaTyrA $\alpha$	<i>Macarthuria australis</i>	670.2 $\pm$ 22.0
SoTyrA $\alpha$	<i>Spinacia oleracea</i>	683.1 $\pm$ 34.2
SoTyrA $\alpha$ <sup>D208E</sup>	<i>Spinacia oleracea</i>	282.7 $\pm$ 14.1
SoTyrA $\alpha$ <sup>ALL13</sup>	<i>Spinacia oleracea</i>	600.1 $\pm$ 100.0
SoTyrA $\alpha$ <sup>D208E_ALL13</sup>	<i>Spinacia oleracea</i>	79.14 $\pm$ 18 <sup>a</sup>
NaTyrA $\beta$	<i>Nepenthes alata</i> $\times$ <i>ventricosa</i>	17.0 $\pm$ 8.3
NaTyrA $\beta$ <sup>E199D</sup>	<i>Nepenthes alata</i> $\times$ <i>ventricosa</i>	111.0 $\pm$ 8.0
AtTyrA2	<i>Arabidopsis thaliana</i>	72.1 $\pm$ 21.7
AtTyrA2 <sup>E179D</sup>	<i>Arabidopsis thaliana</i>	148.0 $\pm$ 10.4
AtTyrA2 <sup>D241N</sup>	<i>Arabidopsis thaliana</i>	ND <sup>b</sup>
AtTyrA2 <sup>D241N_E179D</sup>	<i>Arabidopsis thaliana</i>	ND <sup>b</sup>

<sup>a</sup>Value reported previously in Lopez-Nieves et al., 2018.

<sup>b</sup>ND, the  $IC_{50}$  values could not be determined because only limited inhibition was observed at 5 mM tyrosine effector, representing the highest concentration that can be achieved as a result of its limited solubility.

emerged early in core Caryophyllales and before the origin of betalain biosynthesis (Figure 1b).

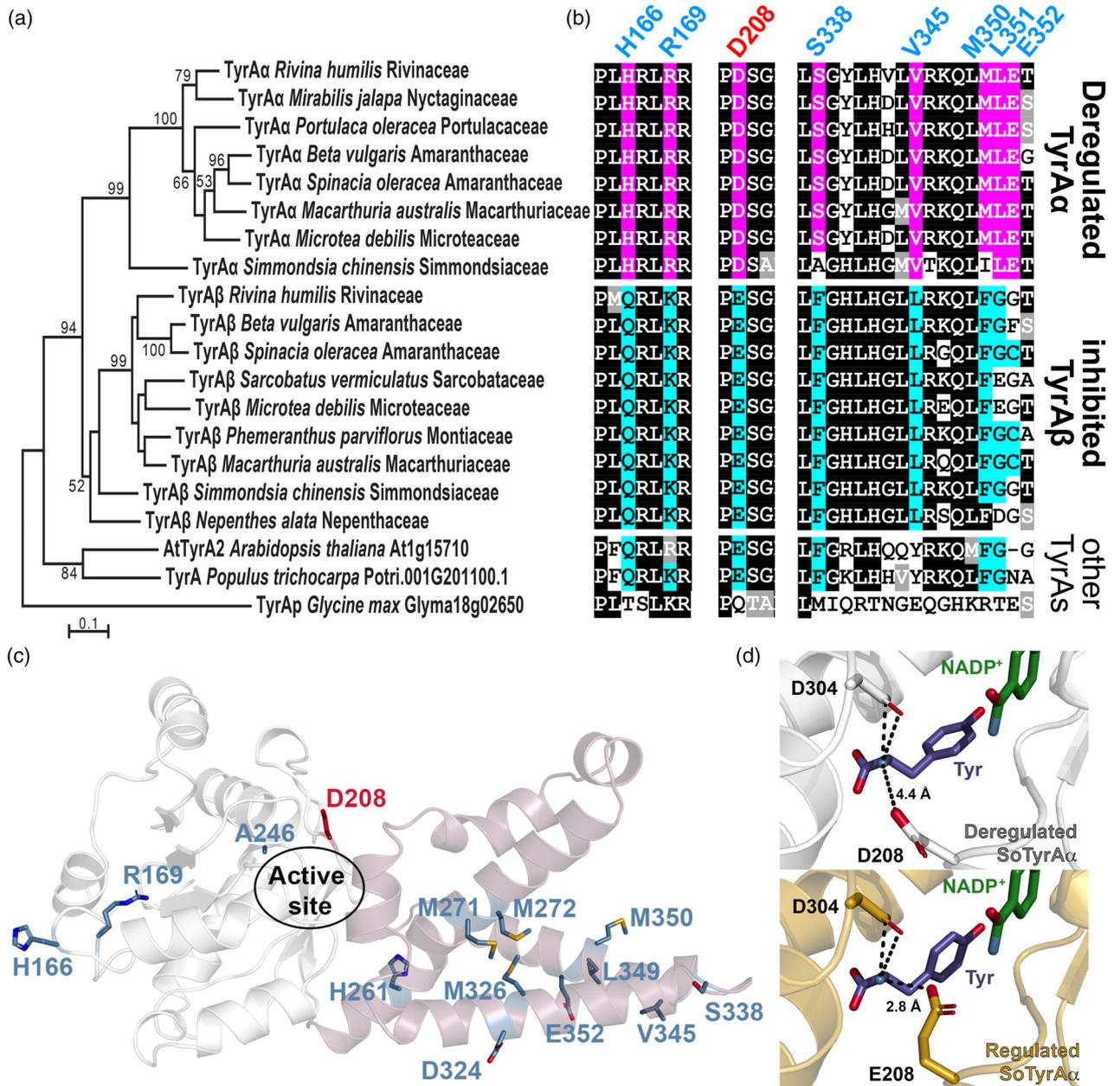
#### The active site residue D208 is critical for the relaxed regulation of TyrA $\alpha$ against tyrosine feedback inhibition

Using the timing of the emergence of functionally deregulated TyrA $\alpha$  enzymes that we established above (Figure 1), we searched for residues that might be responsible for their functionalization and tyrosine insensitivity. Amino acid sequence alignments were conducted using TyrA protein coding sequences that were obtained in the present study and from previously reported transcriptome data (Brockington et al., 2015; One Thousand Plant Transcriptomes Initiative, 2019), along with TyrA enzymes from non-Caryophyllales species (Figure S1). The phylogenetic analyses of representative TyrA sequences (i.e. functionally characterized TyrA $\alpha$  enzymes and corresponding TyrA $\beta$  from the same or closely related species) showed that Caryophyllales TyrA $\alpha$  and TyrA $\beta$  formed separate and well-supported monophyletic groups (Figure 2a), consistent with a previous study (Lopez-Nieves et al., 2018). The comparison of their amino acid sequences showed that, in total, 14 residues are generally conserved among but different between TyrA $\alpha$  and TyrA $\beta$  (Figure 2b for representative residues, Figure S1 for all residues). These residues were numbered based on TyrA $\alpha$  from *S. oleracea* (SoTyrA $\alpha$ ) (Table S1), which was used as a representative Caryophyllales TyrA $\alpha$  in the present study because the recombinant SoTyrA $\alpha$  expressed better than BvTyrA $\alpha$  (Lopez-Nieves et al., 2018).

Mapping of these 14 residues in a structural model of SoTyrA $\alpha$ , which was built based on the X-ray crystal structure of soybean TyrA $\beta$ /PDH as a template (Protein Data Bank ID: 5WHX) (Schenck et al., 2017a), revealed that only one residue, aspartate at the 208 position (D208), is located at the active site (Figure 2c), which corresponds to a glutamate (E) in all TyrA $\beta$  sequences (Figure 2b). Interestingly, the corresponding residue in the legume TyrA $\beta$  is a glutamine (Q) (Figure 2b), which was previously shown to be important for catalysis (Holland and Jez, 2018). Based on the comparison of the model with the three-dimensional structure of soybean TyrA $\beta$ /PDH, D208 likely interacts with the side-chain amine of tyrosine and argenine and is located opposite from D304, which is equivalent to D222 and substituted with asparagine (D222N) in legume TyrA $\beta$  (Figure 2d) (Schenck et al., 2017a). To experimentally test whether D208 is involved in the relaxed regulation of TyrA $\alpha$  enzymes, we performed site-directed mutagenesis and analyzed the SoTyrA $\alpha$ <sup>D208E</sup> protein. Substitution of D208 with a glutamate reduced the  $IC_{50}$  with tyrosine by approximately two-fold compared to the SoTyrA $\alpha$  wild-type enzyme (283 versus 683  $\mu\text{M}$ ) (Figure 3, Table 1). These results indicate that the active site D208 residue plays a role in the deregulation of TyrA $\alpha$ .

#### Multiple non-active site residues further contribute to the deregulation of TyrA $\alpha$ together with D208

Because the  $IC_{50}$  value of SoTyrA $\alpha$ <sup>D208E</sup> (283  $\mu\text{M}$ ) was still more than 10-fold higher compared to TyrA $\beta$  (approximately 20–60  $\mu\text{M}$ ) (Table 1), we considered that the additional 13 non-active site residues (Figure 2c, Figure S1) may also contribute to the relaxed regulation of TyrA $\alpha$ . To test this hypothesis, the 13 non-active site residues were mutated all together with and without the D208E mutation (SoTyrA $\alpha$ <sup>D208E\_ALL13</sup> and SoTyrA $\alpha$ <sup>ALL13</sup>, respectively) and their effects on tyrosine inhibition were examined. The SoTyrA $\alpha$ <sup>D208E\_ALL13</sup> mutant was much more strongly inhibited by tyrosine than SoTyrA $\alpha$  wild-type or SoTyrA $\alpha$ <sup>D208E</sup> mutant enzymes (Figure 3). The  $IC_{50}$  value of SoTyrA $\alpha$ <sup>D208E\_ALL13</sup> was 79  $\mu\text{M}$  (Table 1), approaching the values reported for other TyrA $\beta$  enzymes (Lopez-Nieves et al., 2018). When seven of the 13 residues, which are highly conserved among TyrA $\beta$  (or TyrA $\alpha$ ) (Figure S1, Table S1), were mutated to SoTyrA $\alpha$  together with D208E, the resulting SoTyrA $\alpha$ <sup>D208E\_PARTIAL7</sup> mutant was less inhibited by tyrosine than SoTyrA $\alpha$ <sup>D208E\_ALL13</sup> (Figure S2), suggesting that these selected seven residues are not sufficient and multiple residues are likely involved. Interestingly, mutating these 13 residues without D208E (i.e. SoTyrA $\alpha$ <sup>ALL13</sup>) had little effect on TyrA regulation (Figure 3) because SoTyrA $\alpha$ <sup>ALL13</sup> and SoTyrA $\alpha$  wild-type had comparable  $IC_{50}$  values (600 versus 683  $\mu\text{M}$ , respectively) (Table 1). This was also the case for mutating the seven residues without D208E



**Figure 2.** Phylogeny-guided prediction of amino acid residues involved in TyrA enzyme regulation in Caryophyllales.

(a, b) The comparison of amino acid sequences between deregulated TyrA $\alpha$  and tyrosine-sensitive TyrA $\beta$  enzymes identified an active site aspartate (D208, red letter), as well as 13 non-active site residues (blue letters), potentially involved in the deregulation of TyrA $\alpha$  (Table S1; for the complete alignment and residues, see also Figure S1). Amino acids conserved among TyrA $\alpha$  and TyrA $\beta$  sequences are highlighted in pink and cyan, respectively. The phylogenetic tree was generated by the maximum likelihood method (see Experimental Procedures and Data S1) with *Glycine max* TyrA $\beta$  as an outgroup.

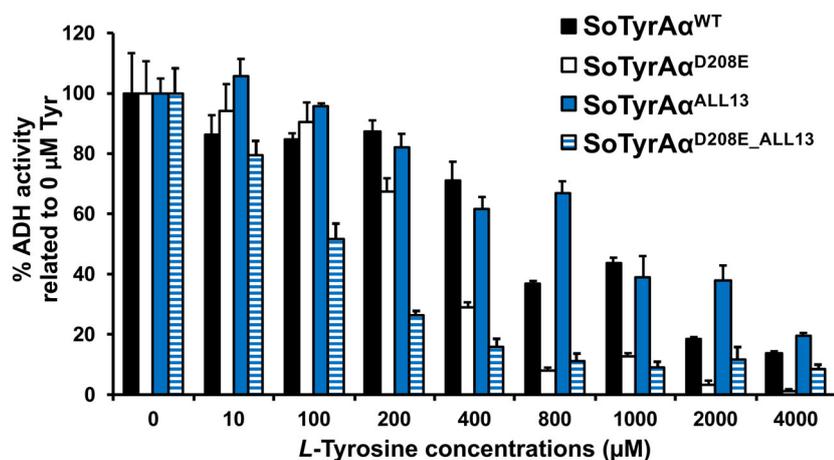
(c) Structural model of the SoTyrA $\alpha$  monomer with D208 (red) and the 13 non-active site residues shown (blue). The catalytic domain is shown in white, whereas the dimerization domain is shown in rose.

(d) Docking of tyrosine (purple) into the active site of the deregulated (top; white) and regulated (bottom; gold) SoTyrA $\alpha$  shows the potential contributions of aspartate and glutamate, respectively. Predicted electrostatic interactions between the backbone amine of tyrosine and active site residues are depicted using black dashed lines.

(SoTyrA $\alpha$ <sup>PARTIAL7</sup>) (Figure S2). Together, these results revealed that multiple non-active site residues contribute to the TyrA $\alpha$  deregulation, although only in the presence of D208, further highlighting the essential role of D208 in the relaxed regulation of TyrA $\alpha$ .

### The reciprocal E208D mutation can partially relax strongly-regulated TyrA $\beta$ and Arabidopsis TyrA enzymes

To test whether the acquisition of D208 can indeed confer relaxed regulation to TyrA enzymes against tyrosine



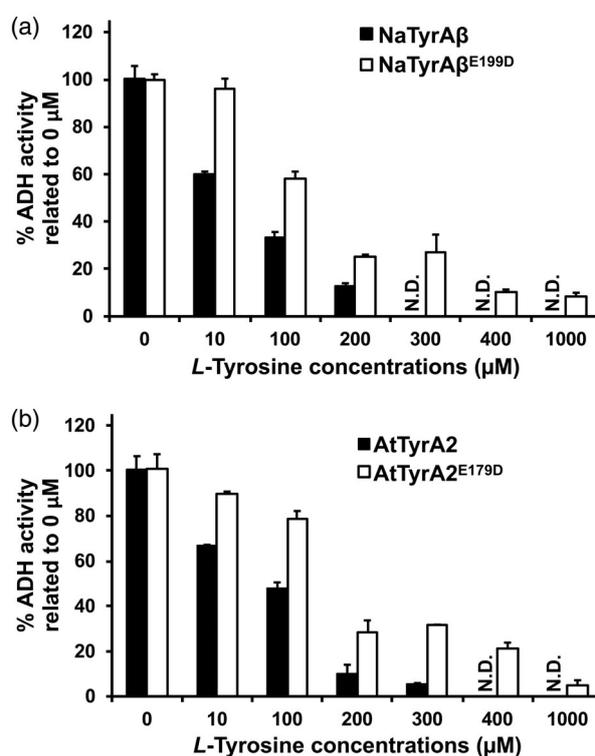
**Figure 3.** The active site aspartate 208 (D208), along with multiple non-active site residues, is involved in the relaxed regulation of TyrA $\alpha$  in Caryophyllales. The SoTyrA $\alpha$  was mutated at D208 and the other 13 residues either separately or together into specific residues highly represented in TyrA $\beta$  to generate SoTyrA $\alpha$ <sup>D208E</sup>, SoTyrA $\alpha$ <sup>ALL13</sup>, and SoTyrA $\alpha$ <sup>D208E\_ALL13</sup> mutant enzymes. These 13 mutated residues are shown in Table S1 and Figure S1. Their arogenate dehydrogenase activity was measured at different tyrosine concentrations. Data are expressed as a percentage of the respective control activity without tyrosine (0  $\mu$ M) and as the mean  $\pm$  SEM of three independent experiments. The corresponding  $IC_{50}$  values are shown in Table 1.

inhibition, the reciprocal E208D mutation was introduced to the tyrosine-sensitive TyrA $\beta$  enzyme from non-core Caryophyllales (i.e. NaTyrA $\beta$ ) (Lopez-Nieves et al., 2018) (Figure 1c). The NaTyrA $\beta$  enzyme with the E199D mutation (NaTyrA $\beta$ <sup>E199D</sup>), which is equivalent to E208D of SoTyrA $\alpha$ , became more resistant to tyrosine inhibition than NaTyrA $\beta$  wild-type (Figure 4a) with its  $IC_{50}$  value being increased from 17  $\mu$ M to 111  $\mu$ M (Table 1).

The protein sequence alignment of TyrA enzymes from diverse plant species showed that the E208 residue is well conserved among TyrA $\alpha$  enzymes even in non-Caryophyllales species (Figure 2b, Figure S1). Therefore, the effect of E208D mutation may also be extrapolated outside of the order Caryophyllales. To test this possibility, the E179D mutation, which corresponds to E208D of SoTyrA $\alpha$ , was introduced to *A. thaliana* TyrA2 (AtTyrA2<sup>E179D</sup>) and its effect on tyrosine sensitivity was assessed. The AtTyrA2<sup>E179D</sup> mutant exhibited enhanced resistance to tyrosine inhibition compared to the corresponding wild-type enzyme (Figure 4b) with a two-fold increase in  $IC_{50}$  (72 versus 148  $\mu$ M) (Table 1). These results further support the key role of D208 in TyrA enzyme regulation and suggest that the E208D mutation can serve as a useful tool for deregulating TyrA enzymes even in plant species beyond Caryophyllales.

#### Combined introduction of E208D and D222N converts *Arabidopsis* TyrA2 into a highly deregulated TyrA $\alpha$ enzyme that can produce higher levels of tyrosine *in planta*

Our previous study showed that D222N, the critical mutation in legume TyrA $\beta$  enzymes (Schenck et al., 2017a), can also introduce tyrosine insensitivity to AtTyrA2. Because D208 and N222, which evolved independently in core Caryophyllales and legumes, respectively, are located at different positions in the active site but both likely interact with the side-chain of tyrosine (Figure 2d), we considered whether these mutations could be combined to more effectively deregulate TyrA enzymes. To test this hypothesis,

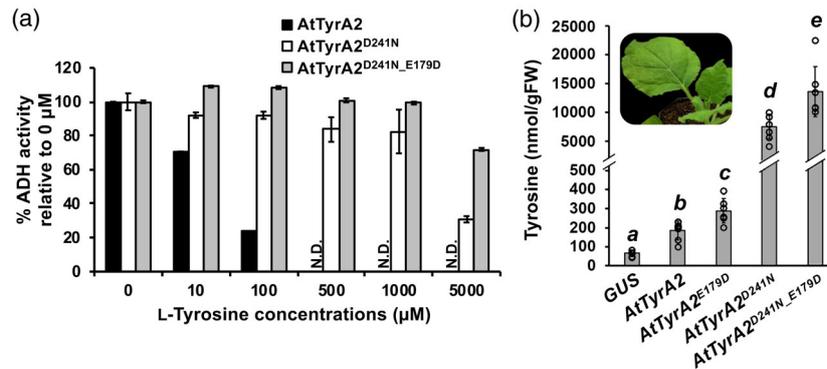


**Figure 4.** Introduction of the D208 mutation confers relaxed tyrosine sensitivity of highly-regulated TyrA enzymes. Arogenate dehydrogenase activity was measured at different tyrosine concentrations.

(a) Introducing the E199D mutation, analogous to the E208D mutation into *Nepenthes alata* TyrA $\beta$  (NaTyrA $\beta$ <sup>E199D</sup>) partially deregulated this tyrosine-sensitive TyrA enzyme.

(b) Introducing the E179D mutation, analogous to the E208D mutation, into the distantly-related *Arabidopsis thaliana* TyrA2 enzyme (AtTyrA2<sup>E179D</sup>) also partially deregulated this tyrosine-sensitive TyrA. Data are expressed as a percentage of the respective control activity without tyrosine (0  $\mu$ M) and as the mean  $\pm$  SEM of three independent experiments. The  $IC_{50}$  values are shown in Table 1. N.D., not detectable.

the E179D mutation (equivalent to E208D) was introduced to the background of AtTyrA2 with the D241N mutation (equivalent to D222N, AtTyrA2<sup>D241N</sup>) (Schenck et al., 2017a)



**Figure 5.** Introduction of two naturally occurring but independently evolved mutations additively deregulate TyrA.

(a) When the E179D mutation was introduced into the AtTyrA<sup>D241N</sup> mutant background, the AtTyrA<sup>D241N\_E179D</sup> double mutant became more resistant to tyrosine-mediated feedback inhibition than the AtTyrA<sup>D241N</sup> mutant. Arogenate dehydrogenase activity was measured at different tyrosine concentrations. Data are expressed as a percentage of the respective control activity without tyrosine (0 μM) and as the mean ± SEM of three independent experiments.

(b) Tyrosine content per gram of fresh weight (FW) upon transient expression of wild-type and mutant AtTyrA genes in *Nicotiana benthamiana* leaves. Letters (a to e) indicate groups that are statistically different in a pairwise comparison within the whole set of mean values by Student's two-sample *t*-test ( $\alpha = 0.05$ ). Data are the mean ± SD ( $n = 6$ ).

to generate the AtTyrA2<sup>D241N\_E179D</sup> double mutant. The single AtTyrA2<sup>D241N</sup> mutant was much more resistant to tyrosine inhibition than the wild-type, consistent with a previous study (Schenck et al., 2017a), but the AtTyrA2<sup>D241N\_E179D</sup> double mutant exhibited even more resistance to tyrosine than the single AtTyrA2<sup>D241N</sup> mutant (Figure 5a). AtTyrA2<sup>D241N\_E179D</sup> maintained approximately 80% activity even at 5 mM of tyrosine, the highest concentration that can be dissolved under physiological pH; hence its  $IC_{50}$  value could not be determined.

To further investigate the effects of these mutations, singly or simultaneously, on *in vivo* function of TyrA, the wild-type and the single and double mutants of AtTyrA2 were subcloned into a binary vector as C-terminal 3xFLAG-tag fusion proteins, and transiently expressed in *N. benthamiana* leaves through *Agrobacterium* infiltration. The infiltrated portion of the leaves was harvested after 3 days to determine the levels of tyrosine, the product of AtTyrA2. The results showed that the expression of the single mutant enzymes AtTyrA2<sup>D241N</sup> and, to a much lesser extent, AtTyrA2<sup>E179D</sup> resulted in significantly increased accumulation of tyrosine compared to wild-type AtTyrA2 (Figure 5b). Interestingly, the expression of the AtTyrA2<sup>D241N\_E179D</sup> double mutant further increased tyrosine accumulation compared to its respective single mutant enzymes (Figure 5b). Western blot analysis was also conducted using specific antibodies against the FLAG-tag, which was fused to the expressed AtTyrA2 proteins. Because the protein levels of AtTyrA2 wild-type and all three AtTyrA2 mutants were comparable with each other (Figure S3), none of these mutations impacted the overall expression and stability of these enzymes. Similar results were obtained in an independent experiment when two different concentrations (i.e. OD<sub>600</sub> of 0.5 and 0.25) of the *Agrobacterium* culture were infiltrated (Figure S4). These

results demonstrate that the combined introduction of the E208D and D222N mutations (E179D and D241N of AtTyrA2, respectively) effectively deregulates TyrA and can strongly elevate the production of tyrosine *in planta*.

## DISCUSSION

Amino acid biosynthetic pathways are highly regulated by the effector-mediated feedback regulation of enzymes that are often involved in branch point reactions; this balances the production of all 20 amino acids required for protein synthesis (Tzin and Galili 2010; Maeda and Dudareva, 2012; Galili et al., 2016; Xing and Last, 2017). Tyrosine biosynthesis is also tightly controlled by regulation of TyrA, so that other aromatic amino acids, such as phenylalanine, can be effectively produced when sufficient tyrosine is present (Rippert and Matringe, 2002a; Schenck and Maeda, 2018) (Figure 1a). However, deregulated TyrA enzymes were previously found in two plant lineages: the legume family and the order Caryophyllales (Lopez-Nieves et al., 2018; Schenck et al., 2015). In the deregulated TyrA<sub>p</sub> of legumes, the D222 residue, which is highly conserved among all plant and some bacterial TyrA arogenate dehydrogenases, was converted to non-acidic residue (e.g. asparagine, D222N) (Schenck et al., 2017a, 2017b). By contrast, the D222 residue (D304 in SoTyrA $\alpha$ ) (Figure 2d, Figure S1) was not altered in Caryophyllales TyrA $\alpha$  (Figure S1). Through phylogeny-guide structure-function analyses, the present study found that alteration of a second active site residue, E208D, played a critical role in the evolution of deregulated TyrA $\alpha$  enzymes in core Caryophyllales. Thus, deregulated TyrA enzymes evolved independently in legumes and Caryophyllales.

Based on the X-ray crystal structure of the N222D mutant of soybean TyrA<sub>p</sub> (GmPDH1), the side-chain carboxylate of D222 was shown to form an electrostatic

interaction with the backbone amine of the tyrosine effector, and likely the arogenate substrate, and thus was responsible for both tyrosine sensitivity and arogenate substrate specificity of TyrA<sub>α</sub> (Schenck et al., 2017a). Although repeated attempts to crystallize Arabidopsis and Caryophyllales TyrA enzymes have not been successful so far, structural modeling of TyrA<sub>α</sub> suggests that E208 is also located within the active site, opposite from D222 (D304 of SoTyrA<sub>α</sub>), and likely forms a charge–charge interaction with the backbone amine of tyrosine and, presumably, arogenate (Figure 2d). When tyrosine was docked into homology models of the SoTyrA<sub>α</sub> active site with an aspartate at position 208, the backbone amine of tyrosine was approximately 4.4 Å from the carboxylate oxygen of the aspartate side-chain. When tyrosine was docked into the active site of the SoTyrA<sub>α</sub><sup>D208E</sup>, the side-chain carboxylate of E208 was positioned 2.8 Å from the backbone amine of tyrosine (Figure 2d). Although these results are from molecular docking using homology models and not from proteins co-crystallized with the tyrosine effector, they suggest that the introduction of a slightly longer side-chain in glutamate, compared to aspartate, may position the residue closer to the effector, which in turn may enhance tyrosine regulation. Legume TyrA<sub>p</sub> enzymes, which are localized in cytosol, are completely insensitive to tyrosine inhibition (Schenck et al., 2015; 2017), whereas Caryophyllales TyrA<sub>α</sub> are only partially deregulated (Lopez-Nieves et al., 2018) (Figure 1c). Thus, in Caryophyllales, the subtle change from glutamate to aspartate, both comprising acidic residues, might have allowed partial deregulation of TyrA without depleting phenylalanine within the plastids where TyrA<sub>α</sub> is located (Lopez-Nieves, et al., 2018).

Although the introduction of D241N (equivalent to legume D222N) in distantly-related, highly-regulated Arabidopsis AtTyrA2 drastically reduced its sensitivity to tyrosine (Schenck et al., 2017a) (Figure 5a), the addition of E179D (equivalent to E208D) to the AtTyrA2<sup>D241N</sup> mutant further enhance its resistance to tyrosine inhibition; the AtTyrA2<sup>D241N\_E179D</sup> double mutant became completely insensitive to tyrosine concentrations of up to 1 mM (Figure 5a). Heterologous expression of these mutated enzymes and wild-type AtTyrA2 in *N. benthamiana* leaves further demonstrated that the expression of the AtTyrA2<sup>D241N\_E179D</sup> double mutant enhances tyrosine accumulation *in planta* (Figure 5b). This result suggests that two mutations, which evolved in two independent plant lineages, have additive effects on the regulation of TyrA enzymes.

The minor effects of D208 on the regulation of Caryophyllales TyrA<sub>α</sub> (Figures 3 and 4), in comparison with the drastic effect of N222 on legume TyrA<sub>p</sub> (Schenck et al., 2017a), suggested that additional residues are involved in the deregulation of Caryophyllales TyrA<sub>α</sub>. We further identified 13 additional mutations, which are conserved among

TyrA<sub>α</sub> but not in TyrA<sub>β</sub>, or vice versa (Figure 2a, Figure S1). These 13 residues were all located outside of the active site with 10 of the 13 residues located in the C-terminal dimerization domain (H261, M271, M272, D324, M326, S338, V345, L349, M350, and E352) (Figure 2c). The other three residues (H166, R169, and A246) were found in the catalytic domain but were not part of the active site. The introduction of these 13 mutations together with D208E essentially converted deregulated TyrA<sub>α</sub> into highly regulated TyrA<sub>β</sub>-like enzyme with an *IC*<sub>50</sub> value below 100 μM (Figure 3, Table 1), whereas the introduction of seven out of 13 mutations, together with D208E, was not sufficient to alter sensitivity to tyrosine (Figure S2). Therefore, in addition to the critical D208 residue, multiple non-active site residues were altered during the evolution of deregulated TyrA<sub>α</sub> after the duplication and divergent from TyrA<sub>β</sub> in core Caryophyllales. It is interesting to note that the introduction of these 13 or seven residues in TyrA<sub>α</sub> without D208E resulted in increased resistance to tyrosine inhibition at high tyrosine concentrations (Figure 3; Figure S2). Therefore, together with the subtle effect of D208 in the active site, these additional residues might have allowed for fine adjustment of the critical regulation between tyrosine and phenylalanine production and between secondary metabolite and protein synthesis. The contrasting mechanisms for the evolution of the deregulated TyrA enzymes between legumes and Caryophyllales also highlight diverse mutational pathways that were exploited by natural selection in these two different plant lineages.

The evolution of deregulated TyrA enzymes appears to be associated with the production of lineage-specific tyrosine-derived specialized metabolism (Schenck and Maeda, 2018). In young leaves of tropical legume trees from the genus *Inga*, extremely high accumulations of tyrosine and tyrosine-derived defense compounds (e.g. tyramine-gallate) are associated with the overexpression of the gene encoding the deregulated TyrA<sub>p</sub> (Coley et al., 2019), although the role of TyrA<sub>p</sub> enzymes in other legumes remains elusive (Schenck et al., 2020). The expression of the deregulated TyrAs also positively associates with elevated production of the tyrosine-derived betalain pigments across different families of core Caryophyllales (Lopez-Nieves et al., 2018). Notably, however, the present study revealed that deregulated TyrA enzymes are also present in non betalain-producing core Caryophyllales species (Figure 1c). This finding certainly raises an interesting future question about the potential function of these deregulated TyrA<sub>α</sub> enzymes for the production of yet unknown tyrosine-derived compounds, beyond betalain pigments.

Plants naturally produce diverse and abundant natural products from CO<sub>2</sub> and soil nutrients. The rapid development of synthetic biology and gene editing tools, along

with the discovery of novel plant specialized metabolic pathway genes, is providing exciting opportunities to use plant synthetic biology for the effective production of a variety of chemicals in a sustainable manner (Maeda, 2019b; Owen et al., 2017; Shih, 2018). Unlike specialized metabolic pathways, primary metabolism is typically highly conserved and constrained as a result of its essentiality (Maeda and Fernie, 2021; Moghe and Last, 2015; Pichersky and Lewinsohn, 2011; Weng et al., 2012) and often is difficult to manipulate (Maeda, 2019b). The present study demonstrated that the phylogeny-guided structure–function analyses of primary metabolic enzymes provides a powerful approach for uncovering the basis of primary metabolic innovations and generating tools to improve primary metabolism. For example, the elevated tyrosine, resulting from the expression of deregulated TyrAs, can be further converted into other high value products, such as betalain pigments, vitamins, and isoquinoline, as well as other alkaloids (Nett et al., 2020; Schenck and Maeda, 2018; Timoneda et al., 2018). The identified key mutations can be also introduced by base editing technology (Mishra et al., 2020; Rees and Liu, 2018), without the use of a transgenic approach, to enhance the endogenous production of tyrosine and tyrosine-derived natural products *in planta*.

## EXPERIMENTAL PROCEDURES

### Cloning, recombinant protein expression, purification, and site-directed mutagenesis

RNA extraction of *S. chinensis* was performed as described by Wang et al., 2011 from a new leaf collected from a specimen at the University of Wisconsin-Madison Botany Green House. The full-length or enzyme coding sequences of *S. chinensis*, *M. australis*, and *M. debilis* TyrA $\alpha$  were provided by the laboratory of Samuel Brockington in University of Cambridge, UK. For *M. australis* TyrA $\alpha$ , full-length coding sequences were recovered from the RNA-sequencing data (Matasci et al., 2014); for *S. chinensis* and *M. debilis* TyrA $\alpha$ , partial coding sequence were recovered from the RNA-sequencing, and then rapid amplification of cDNA ends PCR and inverse PCR (Ren et al., 2005) were used to obtain the full-length coding sequences. Using cDNA and gene-specific primers (Table S2), *ScTyrA $\alpha$*  was amplified and cloned into the pGEX-2T vector at the *Bam*HI and *Eco*RI sites with the In-Fusion HD cloning kit (Clontech, Madison, WI, USA). *MaTyrA $\alpha$*  and *MdTyrA $\alpha$*  were gene synthesized (General Biosystem, Morrisville, NC, USA) and directly cloned into pGEX-4T vector at the *Bam*HI and *Eco*RI sites. *SoTyrA $\alpha$ <sup>ALL13</sup>* and *SoTyrA $\alpha$ <sup>PARTIAL7</sup>* were initially gene synthesized into pET-28a vector at the *Nde*I and *Bam*HI sites and later subcloned into pGEX-2T at the the *Bam*HI and *Eco*RI sites. The protein expression constructs of *S. oleracea* (*SoTyrA $\alpha$* , KY207378), *B. vulgaris* (*BvTyrA $\beta$* , KY207366), *N. alata* (*NaTyrA $\beta$* , KY207377), and *A. thaliana* (*AtTyrA2*, At1g15710) were previously described in Lopez-Nieves et al. (2018).

For site-directed mutagenesis, the *SoTyrA $\alpha$* , *BvTyrA $\beta$* , *NaTyrA $\beta$* , *AtTyrA2*, *SoTyrA $\alpha$ <sup>ALL</sup>*, and *SoTyrA $\alpha$ <sup>PARTIAL7</sup>* plasmids were diluted 1:100 in water in accordance with the protocol described by Schenck et al., 2017a and using the mutagenesis primers as listed in Table S2. The double mutants were generated in the

background of the respective single mutants. All recombinant protein expression was performed as described previously, with the exception that protein cultures were grown at 28°C instead of 18°C (Lopez-Nieves et al., 2018).

### Phylogeny and identification of residues involved in relaxed sensitivity to tyrosine inhibition in Caryophyllales TyrAs

Sequences of cloned, transcriptome (Brockington et al., 2015) or publicly available TyrA sequence data were used to generate amino acid sequence alignment using CLUSTALW (<http://www.genome.jp/tools-bin/clustalw>) and color shaded with BOXSHADE ([https://embnet.vital-it.ch/software/BOX\\_form.html](https://embnet.vital-it.ch/software/BOX_form.html)). Representative sequences were also used to create a phylogenetic tree by the maximum likelihood method using MEGA, version 7 (Kumar et al., 2016) (Figure 2 and Figure S1) under the Jones–Taylor–Thornton model with 500 bootstrap replications and partial deletion, where all positions with less than 90% site coverage were eliminated.

### Arogenate dehydrogenase activity and tyrosine inhibition assays

The arogenate dehydrogenase activity and the enzyme inhibition by tyrosine were performed as described previously (Lopez-Nieves et al., 2018). The  $IC_{50}$  values of the enzymes were calculated using non-linear regression followed by the Dose–Response Inhibition function in PRISM, version 6 (GraphPad Software Inc., La Jolla, CA, USA).

### Molecular modeling and docking of *S. oleracea* TyrA arogenate dehydrogenase (SoTyrA)

The predicted three-dimensional protein structure of *SoTyrA $\alpha$*  was generated using I-TASSER (Roy et al., 2010; Yang et al., 2015; Zhang, 2008) and the X-ray crystal structure of soybean TyrA<sub>p</sub>/PDH (GmPDH1, Protein Data Bank ID: 5WHX) (Scheck et al., 2017a). The models were visualized in Pymol and CueMol (Figure 2c and d). For docking, structural models of *SoTyrA $\alpha$*  and *SoTyrA $\alpha$ <sup>D208E</sup>* were generated using Phyre2. NADP<sup>+</sup> from the active site of GmPDH1 was overlaid with the *Spinacia* proteins and added to the models before docking. Ligands were computational fit into the active site using AUTODOCKTOOLS, version 1.5.6 (Trott and Olson, 2010) with a grid box of 40 × 40 × 40 Å and the exhaustiveness set to 8. The alignment of key amino acid residues across TyrA enzymes was consistently observed between primary amino acid sequence alignment and structure modeling (Figure 2).

### Transient expression of TyrAs and tyrosine quantification in *N. benthamiana*

The regions corresponding to the putative mature enzyme (without the predicted transit peptide) of *AtTyrA2* wild-type and its mutant variants *AtTyrA2<sup>E179D</sup>*, *AtTyrA2<sup>D241N</sup>*, and *AtTyrA2<sup>D241N\_E179D</sup>* were amplified from their corresponding bacterial protein expression constructs by high-fidelity PCR and subcloned into the binary vector pAGM4673 (Addgene plasmid #48014, courtesy of Sylvestre Marillonnet; Weber et al., 2011) under control of CaMV 35S promoter. The first 216 nucleotides of the 5-enolpyruvylshikimate-3-phosphase synthase gene from *Petunia × hybrida*, which correspond to the predicted plastid transit peptide, were fused to the N-terminus of *TyrA* genes for plastidial targeting *in planta*. The 3xFLAG tag was added to the C-terminus of the *TyrA* in frame for immunodetection of the protein product with specific antibodies (see below). As terminator, we used the double terminator

EU+NbHSP, which has been reported to drastically increase the expression level of recombinant genes in *N. benthamiana* compared to most used terminators such as *NOS* (Diamos and Mason, 2018).

The constructs were transformed into *Agrobacterium tumefaciens* strain GV3101. Prior to infiltration, 10 ml of *A. tumefaciens* culture was grown at 28°C for approximately 24 h in Luria Bertani media with antibiotics and centrifuged at 3000 *g* for 5 min at room temperature. The pellets were washed twice with 3 ml of induction media [10 mM 2-(*N*-morpholine)-ethanesulphonic acid (MES) buffer at pH 5.6, 0.5% glucose, 2 mM NaH<sub>2</sub>PO<sub>4</sub>, 20 mM NH<sub>4</sub>Cl, 1 mM MgSO<sub>4</sub>, 2 mM KCl, 0.1 mM CaCl<sub>2</sub>, 0.01 mM FeSO<sub>4</sub>, and 0.2 mM acetosyringone] and incubated in induction media for 2–3 h at room temperature. After the incubation step, the cells were pelleted at 3000 *g* for 5 min at room temperature and resuspended into 3 ml of 10 mM MES buffer (pH 5.6) with 0.2 mM acetosyringone. The OD<sub>600</sub> was adjusted to 0.25 (or 0.5 for preliminary experiments, see Figure S4) prior to infiltration using the same MES buffer plus acetosyringone. *Nicotiana benthamiana* plants used for the experiment consisted of 4-weeks-old plants grown at 24°C, under approximately 140 μE of light intensity and a 12:12 h light/dark photocycle. Each plant was infiltrated in four different leaf spots with alternative constructs or controls, at approximately 1 h before the end of the light period. In total, each construct was infiltrated as six independent replicates into different plants with a randomized pattern, as well as preventing the infiltration of the same construct twice within different leaves of the same plant.

Samples were collected at 72 h after infiltration and were frozen immediately in liquid nitrogen. Leaf midrib and major veins were excluded from the sampling. Frozen plant samples were ground using liquid nitrogen and 15–25 mg of frozen powder were extracted into 400 μl of 2-amino-2-methyl-1-propanol 0.5% (pH 10.0) buffer in ethanol 75%, as described by Maeda et al., (2010). Plant extracts were analyzed by HPLC (Infinity 1260; Agilent, Santa Clara, CA, USA) equipped with a Atlantis T3 C18 column (Waters Corp., Milford, MA, USA) (3 μm, 2.1 × 150 mm) using mobile phases of A (water with 0.1% formic acid) and B (acetonitrile with 0.1% formic acid) in a 20-min gradient of the mobile phase B: 0–5 min, 1% isocratic; 5–10 min, linear increase from 1% to 76%; 10 to 12 min, linear decrease from 76% to 1%; 12 to 20 min, 1% isocratic. The tyrosine peak was detected at a retention time of approximately 3.5 min using fluorescence (excitation 274 nm, emission 303 nm) and was quantified using the authentic tyrosine standard (AAA1114118; Alfa Aesar, Tewksbury, MA, USA).

### Protein extraction and TyrA quantification by western blotting

Total proteins were extracted from around 10 mg of frozen leaf powder that were resuspended into 75 μl of 1X denaturing protein sample buffer (Tris buffer 60 mM, pH 6.8, SDS 2%, glycerol 10%, β-mercaptoethanol 3% and bromophenol blue 0.01%) with vigorous vortexing for 30 sec and boiled immediately at 95°C for 7 min. Samples were centrifuged at 15 000 *g* for 5 min and 5 μl of the supernatant was applied to the SDS-PAGE gel. 3xFLAG tagged TyrA enzymes were detected by quantitative immunoblotting using a monoclonal specific antibody conjugated to horseradish peroxidase (OctA-Probe HRP conjugated antibody clone H-5; Santa Cruz Biotechnology, Santa Cruz, CA, USA). Upon transfer of the SDS-PAGE gel, the membrane was blocked in 5% powdered skim milk prepared in TBS-Tween (50 mM Tris buffer, pH 7.5, 150 mM NaCl, and 0.05% of Tween-20) for 1 h, washed twice for 5 min with TBS-Tween, incubated in the antibody solution (prepared at a 1:1000 dilution in TBS-Tween) for 2 h at room

temperature, and washed three times for 10 min in TBS-Tween. Chemiluminescence was developed using SuperSignal™ West Pico PLUS Chemiluminescent substrate (Thermo Fisher, Waltham, MA, USA) and the signal was registered using in a ChemiDoc™ Imaging System (Bio-Rad, Hercules, CA, USA). The level of transgenic TyrA enzymes in the plant samples was estimated using IMAGEJ, version 1.52a (NIH, Bethesda, MD, USA) comparing the TyrA-FLAG chemiluminescence signal of the plant extracts with a FLAG standard (pure recombinant *Sorghum bicolor* SbTyrA1-3xFLAG protein produced in *E. coli*).

### ACCESSION NUMBERS

The sequences used for cloning have been deposited in the NCBI under Genbank accession numbers MG681186 (*S. chinensis*, ScTyrA $\alpha$ ), MG681187 (*M. australis*, MaTyrA $\alpha$ ), and MG681189 (*M. debilis*, MdTyrA $\alpha$ ). In addition to the sequences reported in Lopez-Nieves et al., 2018, the following sequences were obtained to build the phylogenetic tree and sequence alignment in Figure 2 and Figure S1: MG681193 (*S. chinensis*, ScTyrA $\beta$ ), MG681192 (*M. debilis*, MdTyrA $\beta$ ), and MG681191 (*M. australis*, MaTyrA $\beta$ ).

### ACKNOWLEDGEMENTS

We thank Sarah Friedrich for the pictures of *S. chinensis* and *M. debilis*, the authors in Wikimedia Commons, Attenboroughii and Margaret R Donald, for the pictures of *Nepenthes alata* and *Macarthuria australis*, respectively, and Dr. Ray Coolier (Wisconsin Crop Innovation Center, University of Wisconsin-Madison) for his assistance in designing the constructs used for transient expression of TyrA enzymes in *Nicotiana benthamiana*. The work was supported by grants from the US National Science Foundation (MCB-1818040 and DEB-1938597) and the US Department of Agriculture (NIFA-AFRI 2015-67013-22955 and 2020-67013-30898).

### AUTHOR CONTRIBUTIONS

SLN and HAM designed the research. JE performed the *Nicotiana* expression experiments. YM performed the biochemical characterization of AtTyrA2 mutant enzymes. CKH performed the molecular modeling and docking. TF and SFB obtained ScTyrA $\alpha$ , MaTyrA $\alpha$ , and MdTyrA $\alpha$  sequences. SLN performed the remainder of the experiments. SLN, JE, YM, CKH, JMJ, and HAM analyzed data. SLN and HAM wrote the paper.

### CONFLICT OF INTEREST

HAM has patents for legume TyrA enzymes. HAM and SLN have a pending patent for deregulated TyrA arogenate dehydrogenases. All of the other authors declare no conflict of interest.

### DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials.

### SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Table S1.** Residues potentially involved in the relaxed regulation of TyrA in Caryophyllales. The numbering of residues is in reference to SoTyrA $\alpha$ . The active site aspartate (D208, shown in bold) plays a critical role in the relaxed regulation of TyrA $\alpha$ . The 13 residues were mutated in the SoTyrA $\alpha$ <sup>ALL13</sup> mutant, whereas the seven out of these 13 residues mutated in the SoTyrA $\alpha$ <sup>PARTIAL7</sup> mutant are marked with an asterisk. Their positions in the protein sequence alignment are shown in Figure S1.

**Table S2.** Primers used in the present study.

**Figure S1.** Alignment of TyrA $\alpha$  and TyrA $\beta$  protein from Caryophyllales. The amino acid sequences of TyrA $\alpha$  and TyrA $\beta$  enzymes from different Caryophyllales species were aligned to predict residues that are involved in the relaxed regulation of TyrA $\alpha$ . TyrA $\alpha$  sequences are found between two horizontal red lines. The key D208 residue that is required for deregulation of TyrA $\alpha$  is shown in red. Thirteen residues that also contribute to TyrA regulation in Caryophyllales are shown in blue. The seven out of the 13 residues are marked by asterisks. These 13 and seven residues were mutated with and without D208E in the SoTyrA $\alpha$  background (Figure 3 and Figure S2, respectively). The yellow arrow and residue highlight the N222 residue critical for tyrosine insensitivity of legume TyrA $\beta$  enzyme (Schenck et al., 2017a). The letters are shaded in black or gray if the residues are present in > 50% of the sequences. The complete gene names, sequences, and their IDs are provided in Data S1.

**Figure S2.** Seven out of the thirteen residues are not sufficient to introduce strong feedback regulation by tyrosine in SoTyrA $\alpha$ . These seven out of the 13 residues (see Table S1) were mutated with and without the D208E mutation in the SoTyrA $\alpha$  enzyme. The resulting SoTyrA $\alpha$ <sup>D208E\_PARTIAL7</sup>, but not SoTyrA $\alpha$ <sup>PARTIAL7</sup> mutant enzyme was more inhibited by tyrosine than SoTyrA $\alpha$ <sup>WT</sup> but less than the SoTyrA $\alpha$ <sup>D208E\_ALL13</sup> mutant. As comparisons, the same data from Figure 3 are presented for SoTyrA $\alpha$ <sup>WT</sup>, SoTyrA $\alpha$ <sup>ALL13</sup>, and SoTyrA $\alpha$ <sup>D208E\_ALL13</sup> mutant enzymes. Their ADH activity was measured at different tyrosine concentrations using NADP<sup>+</sup> cofactor and purified recombinant enzymes. Data are expressed as the percentage of respective control activity without tyrosine (0  $\mu$ M) and the mean  $\pm$  SEM of three independent experiments.

**Figure S3.** (A) Anti-FLAG western blot in *Nicotiana benthamiana* samples expressing AtTyrA2 wild-type and its mutant variants. (B) FLAG standard, using recombinant *Sorghum bicolor* SbTyrA1-3xFLAG. (C) Determination of TyrA content from the image shown in (A) normalized by tissue fresh weight (bars from left to right in the same order than corresponding gel lanes). One-way ANOVA found no significant difference in normalized TyrA mass across the alternative enzymes ( $P = 0.253$ ).

**Figure S4.** Tyrosine quantification in *Nicotiana benthamiana* leaves transiently expressing AtTyrA2 and its mutant versions E179D, D241N and E179D+D41N, upon side-by-side infiltration of two alternative *Agrobacterium* culture densities (OD at 600 nm). Letters (a to e) indicate groups that are statistically different according to Student's two-sample  $t$ -test ( $\alpha = 0.05$ ,  $n = 4$ ). No significant differences were found between the two alternative ODs for any of the enzymes.

**Data S1.** The complete TyrA peptide sequences used for the alignment shown in Figure 2b and Figure S1.

## REFERENCES

- Brockington, S.F., Yang, Y., Gandia-Herrero, F., Covshoff, S., Hibberd, J.M., Sage, R.F. et al. (2015) Lineage-specific gene radiations underlie the evolution of novel betalain pigmentation in Caryophyllales. *New Phytologist*, **207**, 1170–1180.
- Coley, P.D., Endara, M.J., Ghabash, G., Kidner, C.A., Nicholls, J.A., Pennington, R.T. et al. (2019) Macroevolutionary patterns in overexpression of tyrosine: an anti-herbivore defense in a speciose tropical tree genus, *Inga* (Fabaceae). *Journal of Ecology*, **107**, 1620–1632. <https://doi.org/10.1111/1365-2745.13208>.
- Connelly, J.A. & Conn, E.E. (1986) Tyrosine biosynthesis in *Sorghum bicolor*: isolation and regulatory properties of arogenate dehydrogenase. *Zeitschrift Für Naturforschung C: A Journal of Biosciences*, **41**, 69–78. <https://doi.org/10.1515/znc-1986-1-212>.
- Cravens, A., Payne, J. & Smolke, C.D. (2019) Synthetic biology strategies for microbial biosynthesis of plant natural products. *Nature Communications*, **10**, 2142. <https://doi.org/10.1038/s41467-019-09848-w>.
- de Oliveira, M.V.V., Jin, X., Chen, X., Griffith, D., Batchu, S. & Maeda, H.A. (2019) Imbalance of tyrosine by modulating TyrA arogenate dehydrogenase impacts growth and development of *Arabidopsis thaliana*. *Plant Journal*, **97**, 901–922.
- Diamos, A.G. & Mason, H.S. (2018) Chimeric 3' flanking regions strongly enhance gene expression in plants. *Plant Biotechnology Journal*, **16**, 1971–1982. <https://doi.org/10.1111/pbi.12931>.
- Galili, G., Amir, R. & Fernie, A.R. (2016) The regulation of essential amino acid synthesis and accumulation in plants. *Annual Review of Plant Biology*, **67**, 153–178. <https://doi.org/10.1146/annurev-arplant-043015-112213>.
- Hatlestad, G., Sunnadeniya, R., Akhavan, N., Gonzalez, A., Goldman, I., McGrath, M. et al. (2012) The beet *R* locus encodes a new cytochrome P450 required for red betalain production. *Nature Genetics*, **44**, 816–820. <https://doi.org/10.1038/ng.2297>.
- Holland, C.K. & Joseph, J.M. (2018) Reaction mechanism of prephenate dehydrogenase from the alternative tyrosine biosynthesis pathway in plants. *Chem. Bio. Chem*, **19**, 1132–1136. <https://doi.org/10.1002/cbic.201800085>.
- Irmisch, S., Jancsik, S., Yuen, M.M.S., Madilao, L.L. & Bohlmann, J. (2020) Complete biosynthesis of the anti-diabetic plant metabolite montbretin A. *Plant Physiology*, **184**, 97–109. <https://doi.org/10.1104/pp.20.00522>.
- Jacobowitz, J.R. & Weng, J.-K. (2020) Exploring uncharted territories of plant specialized metabolism in the postgenomic era. *Annual Review of Plant Biology*, **71**, 631–658. <https://doi.org/10.1146/annurev-arplant-081519-035634>.
- Jozwiak, A., Sonawane, P.D., Panda, S., Garagounis, C., Papadopoulou, K.K., Abebie, B. et al. (2020) Plant terpenoid metabolism co-opts a component of the cell wall biosynthesis machinery. *Nature Chemical Biology*, **16**, 740–748. <https://doi.org/10.1038/s41589-020-0541-x>.
- Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870–1874.
- Lopez-Nieves, S., Yang, Y., Wang, M., Timoneda, A., Feng, T., Smith, S.A. et al. (2018) Relaxation of tyrosine pathway regulation underlies the evolution of betalain pigmentation in Caryophyllales. *New Phytologist*, **217**, 896–908.
- Maeda, H. (2019a) Evolutionary diversification of primary metabolism and its contribution to plant chemical diversity. *Frontiers in Plant Science*, **10**, 881. <https://doi.org/10.3389/fpls.2019.00881>.
- Maeda, H. (2019b) Harnessing evolutionary diversification of primary metabolism for plant synthetic biology. *The Journal of Biological Chemistry*, **294**, 16549–16566. <https://doi.org/10.1074/jbc.REV119.006132>.
- Maeda, H. & Dudareva, N. (2012) The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annual Review of Plant Biology*, **63**, 73–105.
- Maeda, H. & Fernie, A.R. (2021) Evolutionary history of plant metabolism. *Annual Review of Plant Biology*, **72**, 185–216. <https://doi.org/10.1146/annurev-arplant-080620-031054>.
- Maeda, H., Shasany, A.K., Schnepf, J., Orlova, I., Taguchi, G., Cooper, B.R. et al. (2010) RNAi suppression of arogenate dehydratase1 reveals that phenylalanine is synthesized predominantly via the arogenate pathway in petunia petals. *The Plant Cell*, **22**, 832–849.
- Matasci, N., Hung, L.-H., Yan, Z., Carpenter, E.J., Wickett, N.J., Mirarab, S. et al. (2014) Data access for the 1,000 Plants (1KP) project. *GigaScience*, **3**, 17. <https://doi.org/10.1186/2047-217X-3-17>.
- Mishra, R., Joshi, R.K. & Zhao, K. (2020) Base editing in crops: current advances, limitations and future implications. *Plant Biotechnology Journal*, **18**, 20–31. <https://doi.org/10.1111/pbi.13225>.
- Moghe, G.D. & Last, R.L. (2015) Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiology*, **169**, 1512–1523. <https://doi.org/10.1104/pp.15.00994>.

- Moore, B.M., Wang, P., Fan, P., Leong, B., Schenck, C.A., Lloyd, J.P. *et al.* (2019) Robust predictions of specialized metabolism genes through machine learning. *Proceedings of the National Academy of Science of the United States of America*, **116**, 2344–2353. <https://doi.org/10.1073/pnas.1817074116>.
- Nett, R.S., Lau, W. & Sattely, E.S. (2020) Discovery and engineering of colchicine alkaloid biosynthesis. *Nature*, **584**, 148–153. <https://doi.org/10.1038/s41586-020-2546-8>.
- Nishida, K. & Kondo, A. (2021) CRISPR-derived genome editing technologies for metabolic engineering. *Metabolic Engineering*, **63**, 141–147. <https://doi.org/10.1016/j.ymben.2020.12.002>.
- One Thousand Plant Transcriptomes Initiative (2019) One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, **574**, 679–685. <https://doi.org/10.1038/s41586-019-1693-2>.
- Owen, C., Patron, N.J., Huan, A. & Osbourn, A. (2017) Harnessing plant metabolic diversity. *Current Opinion in Chemical Biology*, **40**, 24–30. <https://doi.org/10.1016/j.cbpa.2017.04.015>.
- Pichersky, E. & Lewinsohn, E. (2011) Convergent evolution in plant specialized metabolism. *Annual Review of Plant Biology*, **62**, 549–566. <https://doi.org/10.1146/annurev-arplant-042110-103814>.
- Pyne, M.E., Narcross, L. & Martin, V.J.J. (2019) Engineering Plant secondary metabolism in microbial systems. *Plant Physiology*, **179**, 844–861. <https://doi.org/10.1104/pp.18.01291>.
- Rees, H.A. & Liu, D.R. (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. *Nature Reviews Genetics*, **12**, 770–788. <https://doi.org/10.1038/s41576-018-0059-1>.
- Ren, M., Chen, Q., Li, L., Zhang, R. & Guo, S. (2005) Successive chromosome walking by compatible ends ligation inverse PCR. *Molecular Biotechnology*, **30**, 95–101. <https://doi.org/10.1385/MB:30:2:095>.
- 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. <https://doi.org/10.1046/j.1432-1033.2002.03172.x>.
- Rippert, P. & Matringe, M. (2002b) Molecular and biochemical characterization of an *Arabidopsis thaliana* arogenate dehydrogenase with two highly similar and active protein domains. *Plant Molecular Biology*, **48**, 361–368.
- Roy, J., Kucukural, A. & Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, **5**, 725–738.
- Schenck, C.A., Chen, S., Siehl, D.L. & Maeda, H.A. (2015) Non-plastidic, tyrosine-insensitive prephenate dehydrogenases from legumes. *Nature Chemical Biology*, **11**, 52–57.
- Schenck, C.A., Holland, C.K., Schneider, M.R., Men, Y., Lee, S.G., Jez, J.M. *et al.* (2017a) Molecular Basis of the evolution of alternative tyrosine biosynthetic routes in plants. *Nature Chemical Biology*, **13**, 1029–1035.
- Schenck, C.A. & Maeda, H.A. (2018) Tyrosine biosynthesis, metabolism, catabolism in plants. *Phytochemistry*, **149**, 82–102. <https://doi.org/10.1016/j.phytochem.2018.02.003>.
- Schenck, C.A., Men, Y. & Maeda, H.A. (2017b) Conserved molecular mechanism of TyrA dehydrogenase substrate specificity underlying alternative tyrosine biosynthetic pathways in plants and microbes. *Frontiers in Molecular Biosciences*, **4**, 1–10. <https://doi.org/10.3389/fmolb.2017.00073>.
- Schenck, C.A., Westphal, J., Jayaraman, D., Garcia, K., Wen, J., Mysore, K.S.A. *et al.* (2020) Role of cytosolic, tyrosine-insensitive prephenate dehydrogenase in *Medicago truncatula*. *Plant Direct*, **4**, 1–15. <https://doi.org/10.1002/pld3.218>.
- Sheehan, H., Feng, T., Walker-Hale, N., Lopez-Nieves, S., Pucker, B., Guo, R. *et al.* (2020) Evolution of L-DOPA 4,5-dioxygenase activity allows for recurrent origins of betalain biosynthesis in Caryophyllales. *New Phytologist*, **227**, 914–929.
- Shih, P.M. (2018) Towards a sustainable bio-based economy: redirecting primary metabolism to new products with plant synthetic biology. *Plant Science*, **273**, 84–91. <https://doi.org/10.1016/j.plantsci.2018.03.012>.
- Thulin, M., Moore, A.J., El-Seedi, H., Larsson, A., Christin, P.A. & Edwards, E.J. (2020) Phylogeny and generic delimitation in Molluginaceae, new pigment data in Caryophyllales, and the new family Corbichoniaceae. *Taxon*, **65**, 775–793.
- Timoneda, A., Feng, T., Sheehan, H., Walker-Hale, N., Pucker, B., Lopez-Nieves, S. *et al.* (2020) The evolution of betalain biosynthesis in Caryophyllales. *New Phytologist*, **224**, 71–85.
- Timoneda, A., Sheehan, H., Feng, T., Lopez-Nieves, S., Guo, R. & Brockington, S.F. (2018) Redirecting primary metabolism to boost production of tyrosine-derived specialised metabolites in planta. *Scientific Report*, **8**, 17256. <https://doi.org/10.1038/s41598-018-33742-y>.
- Trott, O. & Olson, A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, **31**, 455–461.
- Tzin, V. & Galili, G. (2010) The biosynthetic pathways for shikimate and aromatic amino acids in *Arabidopsis thaliana*. *The Arabidopsis Book*, **8**, e0132. <https://doi.org/10.1199/tab.0132>.
- Walker, J.F., Yang, Y., Fen, T., Timoneda, A., Mikenas, J., Hutchison, V. *et al.* (2018) From cacti to carnivores: Improved phylotranscriptomic sampling and hierarchical homology inference provide further insight into the evolution of Caryophyllales. *American Journal of Botany*, **431**, 446–462.
- 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. <https://doi.org/10.1007/s12033-010-9357-3>.
- Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. (2011) A modular cloning system for standardized assembly of multinene constructs. *Plos One*, **6**, e16765. <https://doi.org/10.1371/journal.pone.0016765>.
- Weng, J.K., Philippe, R.N. & Noel, J.P. (2012) The rise of chemodiversity in plants. *Science*, **336**, 1667–1670. <https://doi.org/10.1126/science.1217411>.
- Xing, A. & Last, R.L. (2017) A regulatory hierarchy of the Arabidopsis branched-chain amino acid metabolic network. *The Plant Cell*, **29**, 1480–1499.
- Xu, X. & Qi, L.S. (2019) A CRISPR-dCas Toolbox for genetic engineering and synthetic biology. *Journal of Molecular Biology*, **431**, 34–47. <https://doi.org/10.1016/j.jmb.2018.06.037>.
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. & Zhang, Y. (2015) The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, **12**, 7–8.
- Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.