



OPEN ACCESS

**Edited by:**

Breeanna Urbanowicz,  
University of Georgia, United States

**Reviewed by:**

Jaime Barros,  
University of North Texas,  
United States  
Anna Katherine Block,  
Center for Medical, Agricultural and  
Veterinary Entomology, United States  
Xingxing Li,  
Michigan State University,  
United States  
Hiroshi A. Maeda,  
University of Wisconsin-Madison,  
United States

**\*Correspondence:**

Brian Dilkes  
bdilkes@purdue.edu  
Clint Chapple  
chapple@purdue.edu

**†ORCID:**

Jeffrey P. Simpson  
orcid.org/0000-0003-4019-0681  
Jacob Olson  
orcid.org/0000-0003-4479-575X  
Brian Dilkes  
orcid.org/0000-0003-2799-954X  
Clint Chapple  
orcid.org/0000-0002-5195-562X

**Specialty section:**

This article was submitted to  
Plant Metabolism and Chemodiversity,  
a section of the journal  
Frontiers in Plant Science

**Received:** 24 May 2021

**Accepted:** 23 July 2021

**Published:** 14 September 2021

**Citation:**

Simpson JP, Olson J, Dilkes B and  
Chapple C (2021) Identification of the  
Tyrosine- and Phenylalanine-Derived  
Soluble Metabolomes of Sorghum.  
Front. Plant Sci. 12:714164.  
doi: 10.3389/fpls.2021.714164

# Identification of the Tyrosine- and Phenylalanine-Derived Soluble Metabolomes of Sorghum

Jeffrey P. Simpson<sup>1†</sup>, Jacob Olson<sup>1†</sup>, Brian Dilkes<sup>1,2\*†</sup> and Clint Chapple<sup>1,2\*†</sup>

<sup>1</sup> Department of Biochemistry, Purdue University, West Lafayette, IN, United States, <sup>2</sup> Purdue University Center for Plant Biology, West Lafayette, IN, United States

The synthesis of small organic molecules, known as specialized or secondary metabolites, is one mechanism by which plants resist and tolerate biotic and abiotic stress. Many specialized metabolites are derived from the aromatic amino acids phenylalanine (Phe) and tyrosine (Tyr). In addition, the improved characterization of compounds derived from these amino acids could inform strategies for developing crops with greater resilience and improved traits for the biorefinery. Sorghum and other grasses possess phenylalanine ammonia-lyase (PAL) enzymes that generate cinnamic acid from Phe and bifunctional phenylalanine/tyrosine ammonia-lyase (PTAL) enzymes that generate cinnamic acid and *p*-coumaric acid from Phe and Tyr, respectively. Cinnamic acid can, in turn, be converted into *p*-coumaric acid by cinnamate 4-hydroxylase. Thus, Phe and Tyr are both precursors of common downstream products. Not all derivatives of Phe and Tyr are shared, however, and each can act as a precursor for unique metabolites. In this study, <sup>13</sup>C isotopic-labeled precursors and the recently developed Precursor of Origin Determination in Untargeted Metabolomics (PODIUM) mass spectrometry (MS) analytical pipeline were used to identify over 600 MS features derived from Phe and Tyr in sorghum. These features comprised 20% of the MS signal collected by reverse-phase chromatography and detected through negative-ionization. Ninety percent of the labeled mass features were derived from both Phe and Tyr, although the proportional contribution of each precursor varied. In addition, the relative incorporation of Phe and Tyr varied between metabolites and tissues, suggesting the existence of multiple pools of *p*-coumaric acid that are fed by the two amino acids. Furthermore, Phe incorporation was greater for many known hydroxycinnamate esters and flavonoid glycosides. In contrast, mass features derived exclusively from Tyr were the most abundant in every tissue. The Phe- and Tyr-derived metabolite library was also utilized to retrospectively annotate soluble MS features in two *brown midrib* mutants (*bmr6* and *bmr12*) identifying several MS features that change significantly in each mutant.

**Keywords:** sorghum, phenylpropanoids and phenolics, tyrosine and derivatives, metabol/nomics, specialized (secondary) metabolite

## INTRODUCTION

*Sorghum bicolor* (L.) is a C4 cereal crop known for its adaptation to, and exceptional productivity in, hot and water-limited environments. In addition to being a food and forage crop, especially in arid climates, there is interest in improving the properties of sorghum for the biorefinery (Brenton et al., 2016). Sorghum is genetically diverse and has a high potential for genetic improvement through breeding (Brenton et al., 2016). One potential path for improvement is the formulation of new or modified profiles of metabolites to enhance the resilience of the plant to biotic and abiotic stress.

Untargeted mass spectrometry (MS) can detect and quantify many of the metabolites contained in plant tissues. The output of these analyses is a list of mass features, each associated with a mass to charge ratio ( $m/z$ ), column retention time, and ion count. The term mass feature rather than compound is appropriate because features can include the parental ions of true metabolites and in-source-generated molecular fragments or adducts. Untargeted MS datasets are rich in information concerning the metabolic repertoires of plants but are insufficient to predict compound structures. Although some metabolites can be identified through comparison to authentic standards and interpretation of fragment ions generated by higher-order MS, in untargeted MS, there are still hundreds to thousands of “unknown” metabolite features that have no structural information associated with them.

To increase the informational value of untargeted MS datasets we developed an R-based program called “Precursor of Origin Determination in Untargeted Metabolomics” (PODIUM) (Simpson et al., 2021). In this procedure, a heavy mass isotope representing a precursor to a class of metabolites (e.g., ring labeled  $^{13}\text{C}_6$  phenylalanine (Phe) to identify phenylpropanoids) is fed to plants, and MS features that incorporate the isotope are identified from mass spectrometry data that was processed by XCMS (Smith et al., 2006). The labeled and unlabeled features are then annotated as co-eluting isotopologues and the natural abundance isotopolog (i.e., light or  $^{12}\text{C}_6$ ) is identified. PODIUM was first applied to identify Phe-derived MS features produced in stems of wild-type *Arabidopsis* and 10 genotypes with defects in phenylpropanoid metabolism (Simpson et al., 2021). Over 2,000 MS features, which account for almost 30% of the negative-ion signals collected, were Phe-derived across the different genotypes. This library of metabolites produced from Phe was then used to annotate Phe-derived mass features within an untargeted metabolome of 440 *Arabidopsis* ecotypes. Genome-wide association mapping of those Phe-derived metabolites uncovered multiple loci contributing to natural variation in phenylpropanoid metabolism.

The objective of this study was to catalog specialized metabolites in sorghum derived from the aromatic amino acids Phe and Tyr. Phe and Tyr provide the core structure to a variety of ubiquitous, lineage-, and species-specific specialized metabolites in plants, including soluble phenylpropanoids, lignin, and cyanogenic glycosides, which are compounds that are particularly relevant to this study (Widhalm and Dudareva, 2015; Tohge and Fernie, 2017; Schenck and Maeda, 2018). All

plants synthesize phenylpropanoids from Phe via Phe ammonia lyase (PAL), but grasses can also produce phenylpropanoids from Tyr through the Tyr ammonia-lyase (TAL) activity of some PAL enzymes. Thus, this bifunctional enzyme, called Phe/Tyr ammonia-lyase (PTAL), can deaminate Phe and Tyr into cinnamic acid and *p*-coumaric acid, respectively (Barros and Dixon, 2020). In the Phe-derived pathway, cinnamic acid can be hydroxylated by cinnamate 4-hydroxylase (C4H) to generate *p*-coumaric acid, which could combine into a pool of *p*-coumaric acid that is derived from the TAL activity of PTAL (Figure 1). However, in the model grass *Brachypodium*, Barros et al. (2016) identified differences in the proportions of lignin and other phenylpropanoids derived from Phe or Tyr and through PAL or PTAL, suggesting that the pools of Phe- and Tyr-derived *p*-coumaric acid are not entirely shared or equivalent and that each pool contributes to different end products.

In this study, PODIUM was applied to identify Phe- and Tyr-derived metabolomes (FDM and YDM, respectively), in three tissues of the sorghum genotype BTx623. Ring-labeled  $^{13}\text{C}_6$ -Phe and  $^{13}\text{C}_6$ -Tyr were fed independently to sorghum seedlings, and metabolite features derived from these amino acids were then identified in the untargeted liquid chromatography-MS (LC-MS) dataset. These data allowed us to estimate the relative incorporation of labeled Phe and Tyr into features labeled with both precursors. Finally, the FDM and YDM were also annotated in two brown midrib mutants (*bmr6* and *bmr12*) that have defects in the phenylpropanoid pathway enzymes cinnamyl alcohol dehydrogenase (CAD) and caffeic acid/5-hydroxyferulic acid *O*-methyltransferase (COMT), respectively.

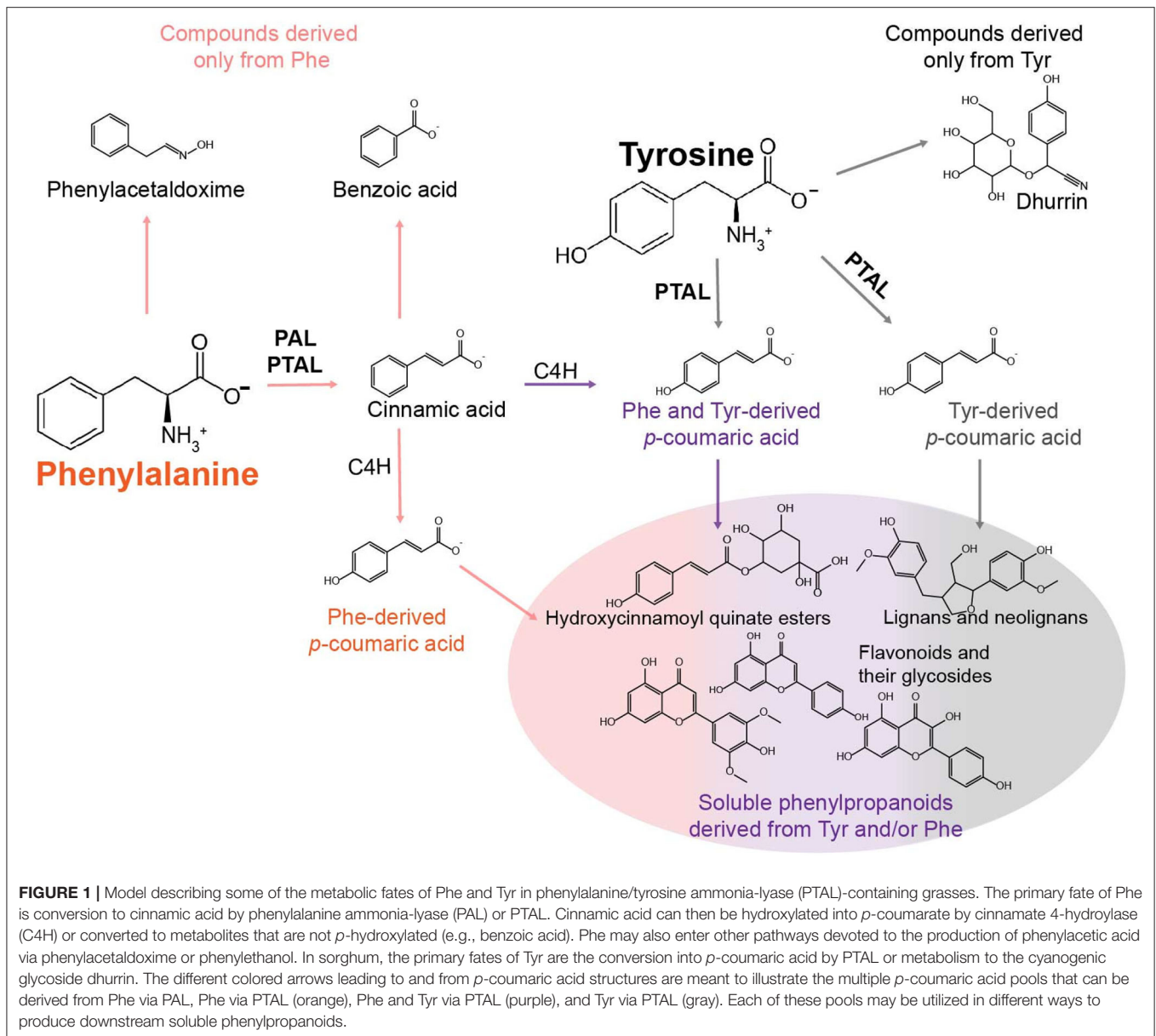
## METHODS

### Plant Material and Growth Conditions

*S. bicolor* (L.) (BTx623 genotype) plants used for Phe and Tyr feeding were grown in Redi-Earth Plug and Seedling Mixture (Sun Gro Horticulture, Quincy MI) augmented with Scotts Osmocote Plus controlled-release fertilizer (Hummert International, Earth City, MO). Plants were grown to their three leaf stage (~10 days post emergence) in a greenhouse under long-day conditions [25°C/23°C (day/night) 16 h light/8 h dark photoperiod] with natural light supplemented with high-pressure sodium lamps (Philips, MASTER GreenPower T400W, Belgium). The *bmr6* and *bmr12* mutants and their wild-type (BTx623) controls were grown under the same conditions.

### Phenylalanine and Tyrosine Feeding

Phe and Tyr feeding was performed similarly to the methods of Wang P. et al. (2018) and Simpson et al. (2021). Plants were removed from the soil, washed with water, and placed into a 120 ml beaker containing the labeling medium along with two other replicate plants. The labeling medium consisted of 30 ml of a Murashige and Skoog medium supplemented with either 1 mM L-phenylalanine (Sigma) or L-tyrosine, or 1 mM ring- $^{13}\text{C}_6$ -labeled L-phenylalanine or ring- $^{13}\text{C}_6$ -labeled L-tyrosine [Cambridge Isotope Laboratories, Tewksbury, MA. Cat No. CLM-1055 (Phe) or CLM-1542 (Tyr)]. Plants were fed for 24 h in the greenhouse under the conditions described above,



after which, each plant was rinsed with water and patted dry. The roots, basal 2 cm of the plant (representing the stem and 2 cm of leaf sheath and developing leaves), and 5 cm of the non-fully expanded leaves were dissected from the plant, flash-frozen in liquid nitrogen, and stored at  $-70^{\circ}\text{C}$  until metabolite extraction.

## Metabolite Extraction and LC-MS Analysis of Soluble Metabolites

Soluble metabolites were extracted from frozen tissue in 50% aqueous methanol (v/v) at a concentration of  $100\text{ mg fresh mass ml}^{-1}$  at  $65^{\circ}\text{C}$  for 2 h, with vortexing every 30 min. Samples were then centrifuged for 5 min at  $13,000 \times g$ , and the soluble fraction was transferred to a new tube. Samples were concentrated in a speed vacuum at  $30^{\circ}\text{C}$  and the dried extract was then re-dissolved

in 50% aqueous methanol (v/v) at 20% of the original volume. All extracts were stored at  $-20^{\circ}\text{C}$  prior to LC-MS analysis.

Chromatographic separations were performed using an Agilent 1100 HPLC system (Agilent Technologies, Palo Alto, CA, USA) with a Shimadzu Shim-pack XR-ODS ( $3 \times 75\text{ mm} \times 2.2\text{ }\mu\text{m}$ , Shimadzu Corp., Japan) separation column and a  $5\text{ }\mu\text{l}$  injection volume. The LC and MS running conditions are described in Simpson et al. (2021). Following data collection, Agilent MassHunter raw data (.d directories) were converted to mzXML using the msconvert tool in ProteoWizard. The mzXML files were grouped into different folders based on tissue and treatment and processed with PODIUM (available at <https://github.com/chapple-lab/podium>) or XCMS (Smith et al., 2006) for the unfed *bmr6*, *bmr12*, and wild-type samples. Peak picking was performed using centWave using

the following parameters: ppm = 15, peakwidth = c(3, 27), snthresh = 15, and prefilter = c(2, 15). Peaks were aligned using “obiwarp,” grouped with bw = 2, mzwid = 0.02, and minsamp = 3, and fillChromPeaks was used to integrate peaks across all samples that were not detected during the initial peak picking.

## RESULTS

### Phenylalanine and Tyrosine Are Used to Produce Unique and Common MS Features

The FDM and YDM libraries were established on three leaf stage greenhouse-grown sorghum. Plants from the inbred line BTx623 were fed  $^{13}\text{C}_6$  ring-labeled Phe and Tyr by removing whole seedlings from their pots and placing them in beakers for 24 h with enough  $^{12}\text{C}$  or  $^{13}\text{C}_6$ -Phe or  $^{13}\text{C}_6$ -Tyr feeding solution to cover the roots. Following feeding, plants were harvested into three portions: (1) roots, (2) the 2 cm section above the roots encompassing the stem and portions of developing leaves and leaf bases (called “base” here), and (3) a 5 cm section above the base containing developing and unexpanded leaf blades and leaf sheaths (called “leaf” here). Semi-polar metabolites were extracted from each sample and analyzed by LC-MS using reverse-phase chromatography and negative-mode ionization. The resulting MS data were run through the XCMS-based PODIUM pipeline (Smith et al., 2006; Simpson et al., 2021) to identify the FDM and YDM. Only MS features that eluted between 100 and 1,100 s in the LC-MS gradient were included in this FDM and YDM analysis, and labeled and unlabeled Phe and Tyr were removed from downstream analyses.

In total, the PODIUM pipeline identified 668 metabolite features containing one to three phenyl rings derived from Phe or Tyr. The  $^{13}\text{C}_6$ -containing metabolite features derived from Phe or Tyr represented 5–13% of the total ion counts in  $^{13}\text{C}$ -fed samples (Figure 2A). In some cases, the signal derived from a mass feature with a +6, +12, or +18 mass in the samples that were fed with the  $^{12}\text{C}$  isotope represented either co-chromatographing features and/or a signal value derived from the “peak-filling” step employed by XCMS integration. As expected for *bona fide* metabolite features, but not for such artifacts, the majority of features identified by PODIUM were more abundant in the heavy-isotope label fed samples. Of the three samples, the base- and root-derived extracts incorporated the most  $^{13}\text{C}$ -Phe and -Tyr, which may reflect better uptake of the label due to the closer proximity to the feeding solution.

Identification of these Phe- and Tyr-derived features in the  $^{12}\text{C}$ -fed samples allowed us to calculate the contribution of these two precursors. Taken together, the FDM and YDM were between 20 and 45% of the total ion count across the three tissues (Figure 2B). Aggregated ion counts for predicted Phe- and Tyr-derived metabolite features in the base and leaf portions were similar, while the ion signal from root samples was approximately 50% lower. There was no linear relationship between the abundance of specific Phe- and Tyr-derived  $^{12}\text{C}$

metabolite features between tissues, indicating that each tissue produces distinct metabolite profiles (Figure 2C).

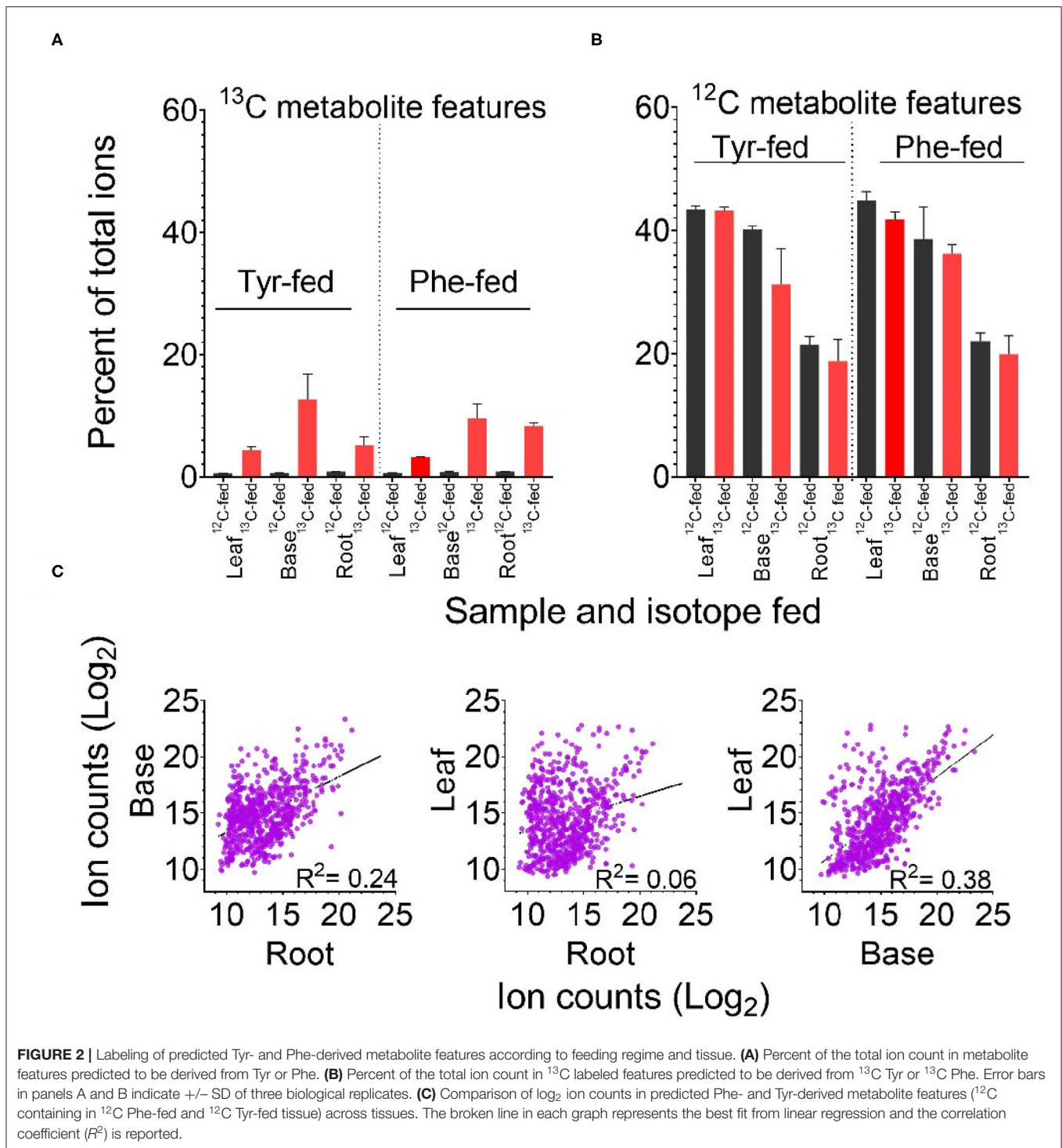
Replicated samples from the  $^{12}\text{C}$ -fed plants grouped with each other along the principal components, whereas replicated samples from the different tissues were separated (Figure 3A). Some effects of precursor supplementation were also visible in these plots. Notably, within the root and base tissues, Phe- and Tyr-fed samples were separated in the score plot (Figure 3A), indicating that the administration of Phe, Tyr, or both altered the synthesis of some  $^{12}\text{C}$  features in their respective FDMs and YDMs. However, comparing the accumulation of individual  $^{12}\text{C}$ -containing features between the Phe-fed and Tyr-fed samples by linear regression revealed a very strong linear relationship between most metabolite features (Figure 3C). This demonstrates that the administration of the  $^{12}\text{C}$  amino acids had only a minor effect on the abundance of  $^{12}\text{C}$  metabolites in the FDM and YDM.

Unlike the analysis of the  $^{12}\text{C}$  features, principal component analysis (PCA) and linear regression using the abundances of the  $^{13}\text{C}_{6,12,18}$ -containing features is a direct measure of the *de novo* production of metabolites during the 24 h feeding. Thus, this analysis can detect metabolites produced from Phe and Tyr via PTAL activity and identify the features that are differentially or uniquely derived from either amino acid. Principal components calculated from these data showed clear separation of the root, base, and leaf samples (Figure 3B). Heavy isotope-labeled Phe or Tyr feeding also resulted in clear separation for all three tissues (Figure 3B). Taken together, this suggested that Tyr and Phe feeding had different effects on compound *de novo* biosynthesis (Figure 3B). Linear regression comparing the abundances of the  $^{13}\text{C}$  metabolite features between feedings also demonstrated the presence of some metabolites derived primarily from either Phe or Tyr, although the majority were of mixed origin (Figure 3D). Most metabolite features were labeled from both Phe and Tyr, indicating that sorghum employs both precursors to synthesize many of its soluble phenyl-ring-containing metabolites. The existence of metabolite features that fall closer to each axis (Figure 3D) indicates that some PAL or PTAL-dependent metabolites and were preferentially derived from one amino acid.

### Phe- and Tyr-Derived Metabolites Have Distinct Patterns of $^{13}\text{C}$ Incorporation, Chromatographic Properties, and Abundances

Tyr and Phe can be metabolized into phenyl ring-containing nitrogenous metabolites or compounds resulting from amino group removal by transamination, oxidative deamination, or, most prominently, deamination through the ammonia-lyase activity of PAL or PTAL (Widhalm and Dudareva, 2015; Schenck and Maeda, 2018). Some of these metabolite pathways that utilize Phe and Tyr are uniquely fed by one amino acid (i.e., Tyr-derived cyanogenic glycosides, Phe-derived cinnamoyl, or phenylacetaldoxime derivatives). In grasses, both amino acids can act as precursors for phenylpropanoids produced downstream of *p*-coumaric acid. In this study, the incorporation

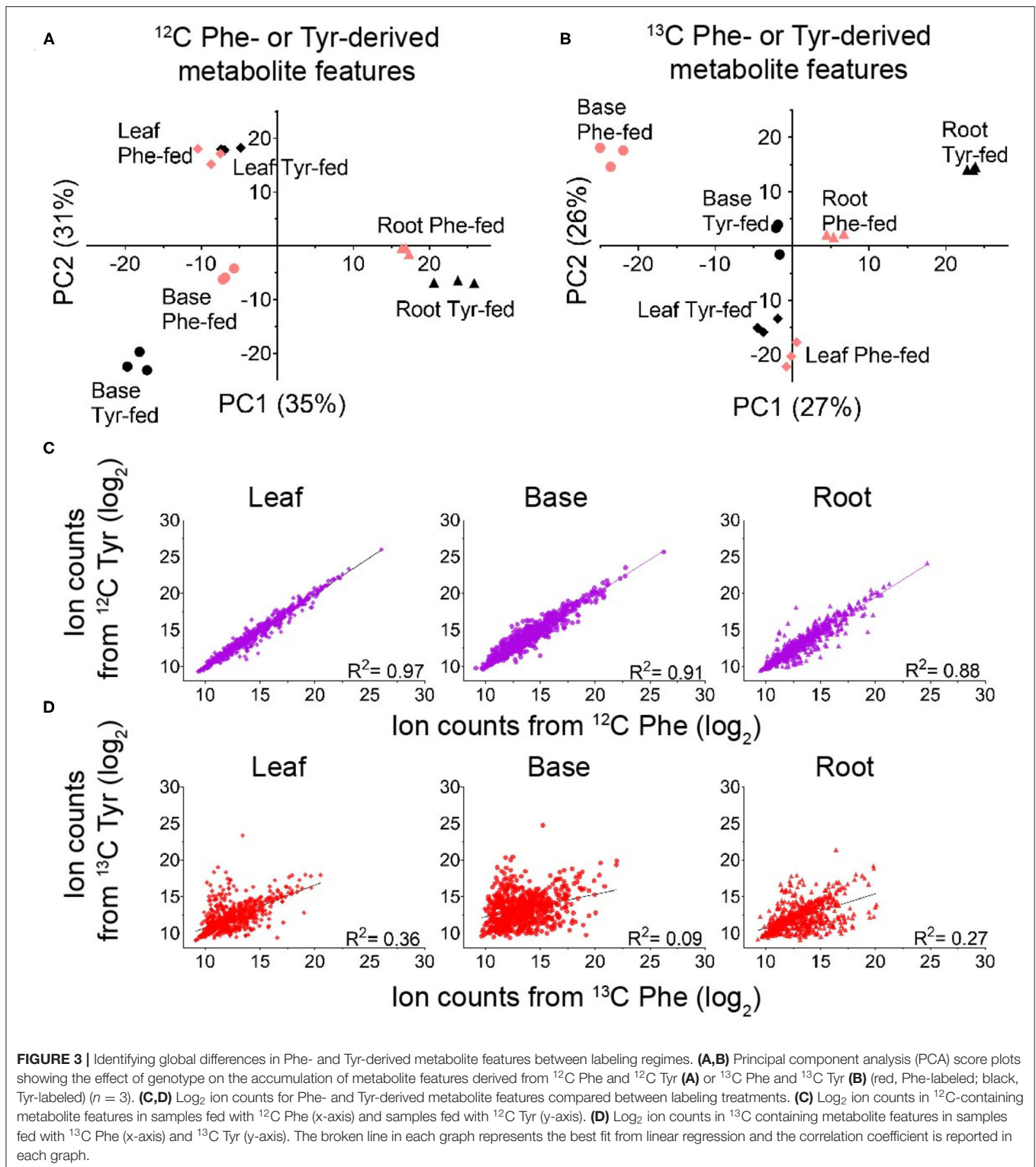




of labeled Phe and Tyr into the FDM and YDM were compared to estimate the relative contribution of Phe and Tyr into metabolites derived through PAL and/or PTAL.

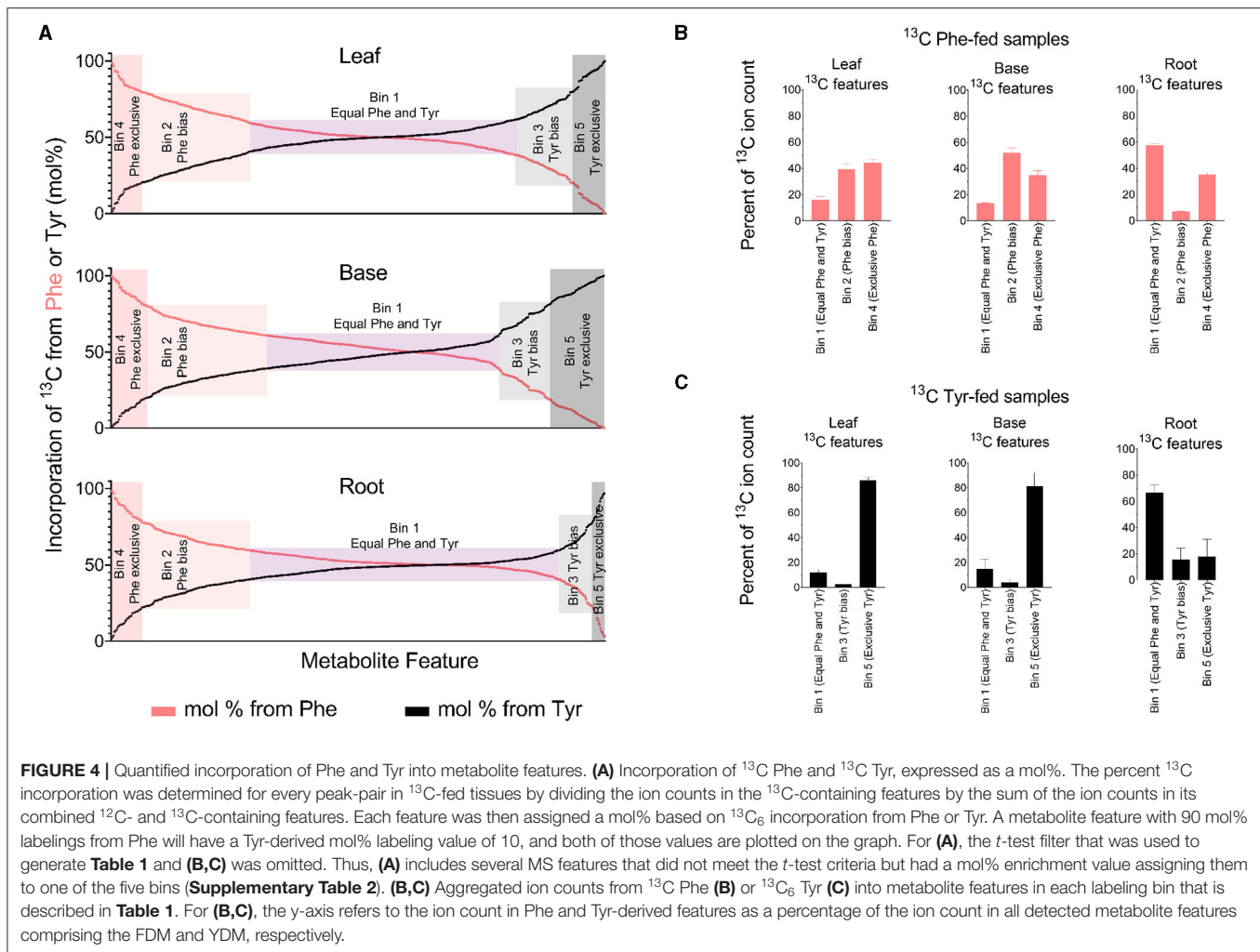
Label incorporation into metabolites was determined by dividing the <sup>13</sup>C<sub>6</sub> ion signal (and ion signal from <sup>13</sup>C<sub>12</sub> and

<sup>13</sup>C<sub>18</sub> features, when present) for each labeled metabolite feature by the ion signal from the entire pool of the metabolite feature [i.e., <sup>13</sup>C<sub>6,12,18</sub> ion signals/(<sup>13</sup>C<sub>6,12,18</sub> + <sup>12</sup>C ion signals) for a labeled metabolite feature]. The resulting values were also then converted into a mol% to directly compare the



capacity of Phe and Tyr to label each metabolite feature. For example, at the end of the labeling period, 25% of a chlorogenic acid (i.e., caffeoylquinic acid) pool was isotopically

labeled in the  $^{13}\text{C}$ -Phe-fed leaves, and 4.2% was labeled in  $^{13}\text{C}$ -Tyr-fed leaves. Then, calculating the relative *de novo* contributions (i.e., mol%) indicated that 86 mol% of leaf



chlorogenic acid was derived from  $^{13}\text{C}$ -Phe, and 13 mol% was derived from  $^{13}\text{C}$ -Tyr.

Metabolite features were sorted into five bins based on mol% enrichment from Phe and Tyr in each tissue and a *t*-test that compared  $^{13}\text{C}$  incorporation into metabolite features between Phe-fed and Tyr-fed samples (**Figure 4**, **Table 1**). Bin 1 contained metabolite features that were labeled at 50  $\pm$  10 mol% from each of the precursor amino acids (i.e., considered derived equally from both amino acids). Metabolite features in this bin represented the majority of labeled features in each tissue. Bins 2 and 3 contained metabolite features with a mol% enrichment  $>60\%$  and  $<80\%$  from Phe (Bin 2) and Tyr (Bin 3), respectively ( $p < 0.05$ ). Metabolite features with a Phe bias were more numerous in all three tissues. Bins 4 and 5 contained metabolite features with an average enrichment of 80 mol% or higher from Phe (Bin 4) and Tyr (Bin 5), respectively ( $p < 0.05$ ), and thus were considered derived exclusively from a single amino acid. The Tyr-derived features were more numerous

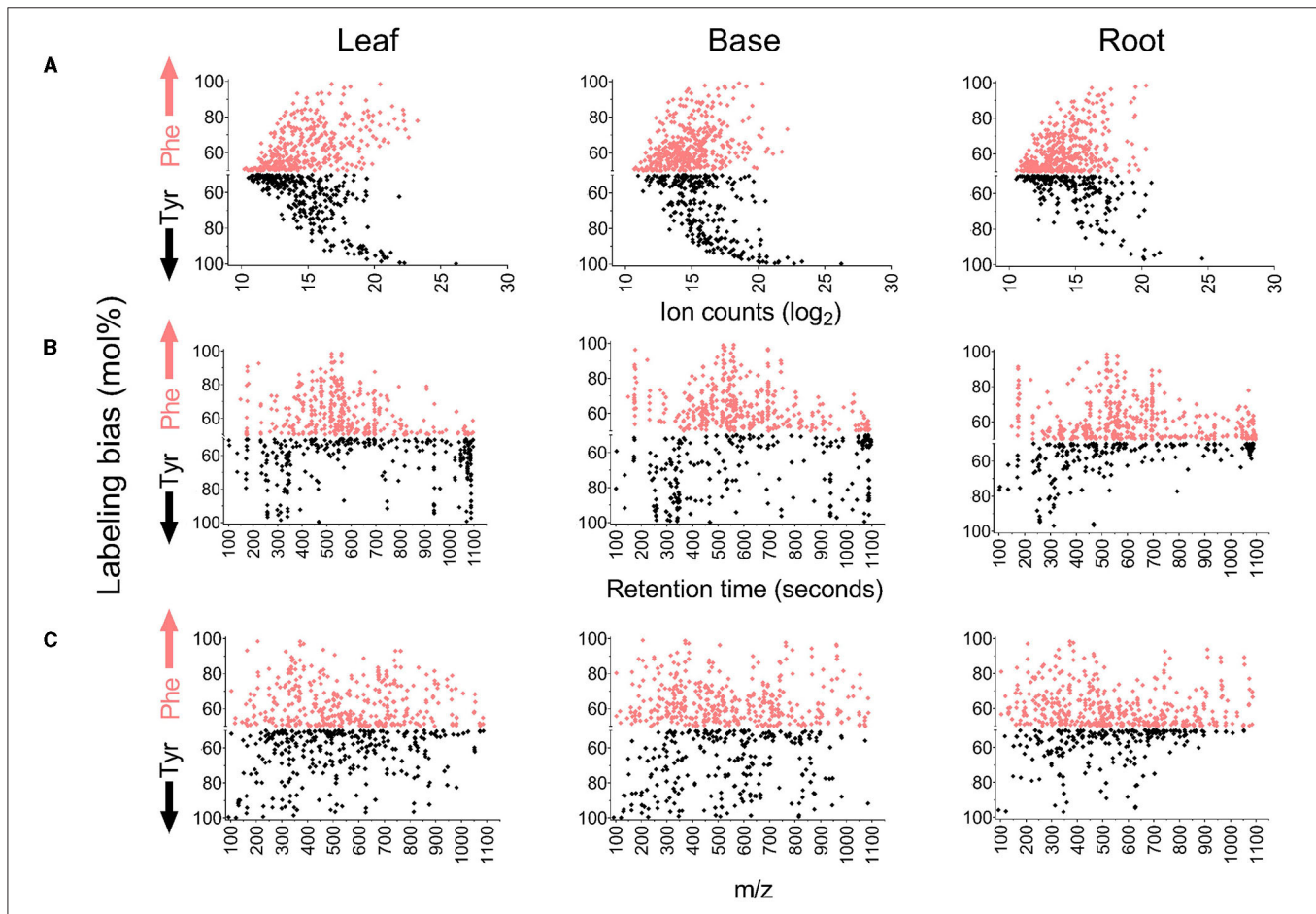
than the Phe-derived features, but Bins 4 and 5 contained the fewest features.

The  $^{13}\text{C}$ -Tyr and  $^{13}\text{C}$ -Phe flux into metabolites were estimated by summing the  $^{13}\text{C}$  ion signal for the respective bins (**Figures 4B,C**). The total ion count from Tyr-derived metabolites (Bin 5) was several-fold higher than for metabolites derived exclusively from Phe (Bin 4), but only in above-ground tissues. Thus, more of the fed  $^{13}\text{C}$ -Tyr went into metabolites produced exclusively from that amino acid, despite there being more total PTAL-derived metabolite features produced. Along the same lines, in above-ground tissues, there was a relatively even distribution of the  $^{13}\text{C}$  signal from metabolites derived exclusively from Phe (Bin 4) and PTAL-derived metabolites with a Phe-bias (Bin 2) as compared to metabolites of approximately equal Phe and Tyr labeling.

Differences in the abundance, retention time, and *m/z* for the labeled metabolite features were also compared to their preferred amino acid of origin (**Figure 5**). Many of the most abundant

**TABLE 1** | Categorizing labeled MS features derived from Phe and Tyr.

| Bin and amino acid bias                                      | Percent of labeled MS features in each bin |      |      |
|--|--|------|------|
|  | Leaf                                       | Base | Root |
| Bin 1: Neither Phe or Tyr bias 50–59 mol % toward Phe or Tyr | 44   | 34   | 55   |
| Bin 2: 60–80 mol % toward Phe                                | 14   | 22   | 16   |
| Bin 3: 60–80 mol % toward Tyr                                | 6  | 7    | 4    |
| Bin 4: 80–100 mol % toward Phe                               | 5  | 7    | 5    |
| Bin 5: 80–100 mol % toward Tyr                               | 5  | 10   | 1    |



**FIGURE 5** | Comparison of properties of metabolite features based on their incorporation of  $^{13}\text{C}$  Phe or  $^{13}\text{C}$  Tyr. The y-axis on each graph reports mol% labeling from Phe or Tyr if the bias toward an amino acid was above 50%. Features that fall above the middle axis were more heavily derived from Phe, and features that fall below were more heavily derived from Tyr. For example, an MS feature with a 90 mol% labeling from Phe and 10 mol% labelings from Tyr will be represented on the y-axis by a single point at 90 in the Phe-bias portion of the graph. On the contrary, an MS feature with a 10 mol% labeling from Phe and a 90 mol% labeling from Tyr will be represented by 90 mol% in the Tyr-bias portion of the graph. **(A)** Comparison of labeling bias to the ion counts in each metabolite feature ( $^{12}\text{C}$ -fed samples). **(B)** Comparison of labeling bias to the retention time of metabolite features. **(C)** Comparison of labeling bias to m/z value for metabolite features.

features in all tissues were predominantly or exclusively derived from either Phe or Tyr (Bins 4 and 5), whereas features derived more evenly from either amino acid were less abundant. Most of the principally Tyr-derived features eluted between 250 and 350 s and approximately 1,100 s, indicating that they are relatively polar or non-polar molecules, respectively. In contrast, most

of the features that were more strongly Phe-derived or had a mixed Phe and Tyr origin eluted between 400 and 700 s and are thus intermediate in hydrophobicity. Consistent with this result, it was verified by tandem MS (MS/MS) that the Tyr-derived metabolite dhurrin eluted at 310 s and known flavonoids and hydroxycinnamate esters (HCEs) eluted between 500 and 900 s.



**TABLE 2** | Tyr-derived metabolites and their MS/MS fragmentation ratios.

| Feature name | <sup>12</sup> C feature m/z | <sup>13</sup> C feature m/z | Peak pair retention time (seconds) | Fragment m/z and % of spectrum represented by fragment* |   | Tentative identification                              |
|--------------|-----------------------------|-----------------------------|------------------------------------|---|---|---|
|              |                             |                             |                                    | Phenyl ring-containing fragments                        | Non phenyl ring fragments   |   |
| M121T468     | 121.0298                    | 127.0496                    | 468                                | 121 (62); 92 (23); 93 (3)                               | None detected   | <i>p</i> -hydroxybenzaldehyde                         |
| M130T746     | 130.0299                    | 136.0498                    | 746                                | 131 (35); 130 (5); 102 (13)                             | None detected   | Unknown   |
| M131T1041    | 131.0376                    | 137.0579                    | 1,041                              | 131 (51); 87 (10); 85 (3)                               | None detected   | Unknown   |
| M162T1088    | 162.0017                    | 168.0218                    | 1,088                              | 162 (73)  | None detected   | Unknown   |
| M227T1070    | 227.0464                    | 233.0663                    | 1,070                              | 227 (42); 168 (12); 167 (6); 142 (2); 141 (15)          | 66 (12)   | Unknown   |
| M263T938     | 263.0827                    | 269.1026                    | 938                                | 263 (1); 219 (1); 153 (4); 132 (20); 131 (64)           |   | Unknown   |
| M310T325     | 310.0935                    | 316.1138                    | 325                                | 310 (1); 148 (12); 121 (2); 113 (3)                     | 179 (4); 119 (7); 89 (15); 71 (6); 59 (28)                            | Dhurrin   |
| M313T258     | 313.0932                    | 319.1129                    | 258                                | 313 (1); 151 (11); 107 (22); 106 (3)                    | 159 (1); 113 (4); 101 (6); 85 (13); 83 (1)                            | <i>p</i> -hydroxyphenylacetic acid glucoside          |
| M346T312     | 346.0709                    | 352.0907                    | 312                                | 346 (2); 101 (4)  | 179 (12); 161 (3); 131 (1); 119 (5); 113(4); 89 (20); 71 (8); 59 (19) | Unknown, non-phenyl ring is consistent with a glucose |
| M445T311     | 445.1357                    | 451.1558                    | 311                                | 445 (17); 401 (6)                                       | 161 (9); 101 (5); 113 (3); 71 (8); 59 (3)                             | Unknown, non-phenyl ring is consistent with a glucose |
| M814T340     | 814.2629                    | 820.2822                    | 340                                | 814 (1); 472 (2); 683 (5); 503 (1); 341 (23)            | 179 (9); 161 (23); 113 (17); 89 (4)                                   | Unknown, non-phenyl ring is consistent with a glucose |

\*MS/MS was performed after the initial untargeted MS<sup>1</sup> analysis, and fragmentation spectra were collected from metabolite features identified by PODIUM as being derived from Phe or Tyr.

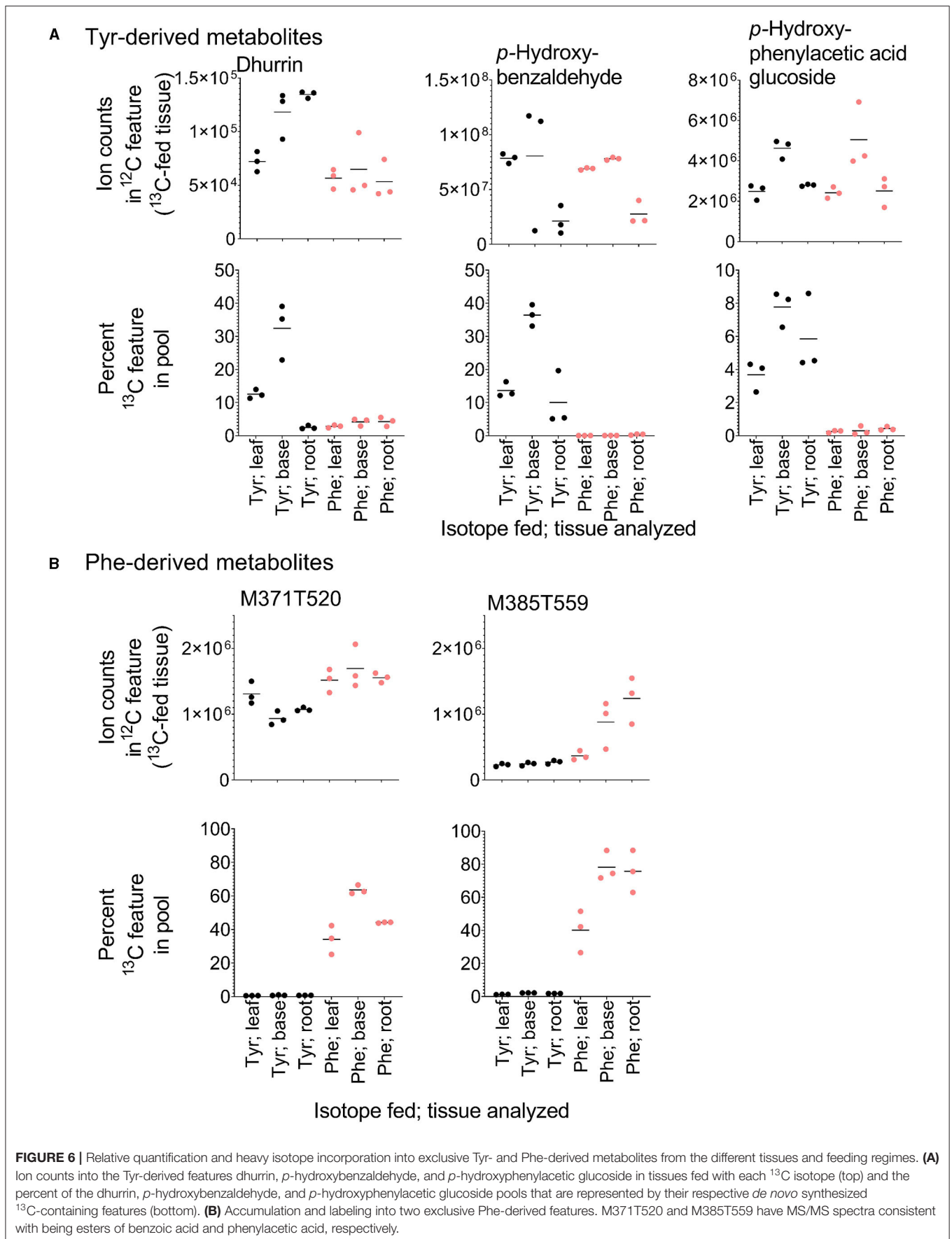
The comparison of m/z showed no clear difference based on Tyr or Phe origins.

## A Large Proportion of the YDM Is Derived Independently of PTAL

Metabolite features derived exclusively from a single amino acid represented only 1–10% of the labeled features in each library (Bins 4 and 5 in **Table 1**), but constituted a large sink for the <sup>13</sup>C-Phe and <sup>13</sup>C-Tyr that was fed to the tissue (**Figures 4B,C**). Over 80% of <sup>13</sup>C-Tyr flowed into metabolites that were not produced from Phe, consistent with the fact that sorghum produces large quantities of the cyanogenic glycoside dhurrin (Gleadow and Møller, 2014). In this study, dhurrin was identified in the untargeted MS data as feature M310T325 (which denotes a feature with an m/z M-H value of 310 and retention time of 325 s) and *post-hoc* MS/MS verified this identification and its exclusive Tyr origin (**Table 2**). <sup>12</sup>C Dhurrin was detected in all tissues and feeding regimes, but there was a small increase in <sup>12</sup>C dhurrin in <sup>13</sup>C Tyr-fed samples and not Phe-fed samples (**Figure 6A**, top), indicating that the synthesis or catabolism of <sup>12</sup>C dhurrin was affected by the application of the <sup>13</sup>C Tyr isotope. *De novo* synthesized <sup>13</sup>C dhurrin represented between 10 and 40% of the total dhurrin pool (i.e., the proportion of <sup>13</sup>C dhurrin relative to the sum of <sup>12</sup>C and <sup>13</sup>C dhurrin) in the base and leaf portions, but not

in the roots (**Figure 6A**, bottom). The apparent lack of dhurrin biosynthesis in roots has been reported previously (Halkier and Møller, 1989), but some studies have identified dhurrin and its catabolic products, including hydrogen cyanide, in roots (Starr et al., 1984, Adewusi, 1990). There were 39 additional Tyr-derived metabolite features, representing 15 likely parental metabolites, that fell into the exclusive Tyr-derived group (Bin 5), most of which were found only in above-ground tissues. The identities for 11 of these metabolites were also assessed by MS/MS performed subsequent to the initial untargeted analysis. In addition to dhurrin, putative products of dhurrin catabolism were identified, including *p*-hydroxybenzaldehyde (M121T468) and *p*-hydroxyphenylacetic acid glycoside (M313T258) and several other unidentified glycosylated metabolites (**Table 2**, **Figure 6**) (Gleadow and Møller, 2014). Although <sup>13</sup>C dhurrin was not detected in the roots, the presence of labeled catabolic products suggests that they, or <sup>13</sup>C dhurrin, were transported to the roots from the site of synthesis but, apparently, at a rate too slow to lead to a significant pool of <sup>13</sup>C dhurrin. Alternatively, the labeled *p*-hydroxybenzaldehyde could originate from transported *p*-hydroxymandelonitrile, which is both a dhurrin biosynthetic intermediate and a breakdown product (Halkier and Møller, 1989).

A similar analysis identified 16 exclusively Phe-derived metabolite features, representing 10 likely parental metabolites



**TABLE 3** | Phe-Derived metabolites and their MS/MS fragmentation ratios.

| Feature name | <sup>12</sup> C feature m/z | <sup>13</sup> C feature m/z | Peak pair retention time (seconds) | Fragment m/z and % of spectrum represented by fragment* |  | Tentative identification                                     |
|--------------|-----------------------------|-----------------------------|------------------------------------|---|--|--|
|              |                             |                             |                                    | Phenyl ring-containing fragments                        | Phenyl ring-containing fragments                             |  |
| M206T561     | 206.0821                    | 212.1021                    | 561                                | 206 (3); 164 (17); 147 (4); 103 (1); 91 (11)            | 72 (3); 70 (4); 58 (13)                                      | Acetylated phenylalanine                                     |
| M371T520     | 371.0986                    | 377.1187                    | 746                                | 371 (16); 12 (11); 77 (3)                               | 249 (3); 113 (3); 99 (2); 85 (2); 75 (2); 71 (2); 59 (3)     | 1-0-4-benzoyl-3-O-glucuronosyl glycerol (Stark et al., 2017) |
| M385T559     | 385.1131                    | 391.1327                    | 1041                               | 385 (17); 135 (2); 117 (2); 111 (1); 91 (8)             | 267 (17); 249 (4); 175 (1); 161 (1); 113 (6); 99 (1); 89 (1) | 1-0-4-phenylacetyl-3-O-glucuronosyl glycerol                 |

\*MS/MS was performed after the initial untargeted MS<sup>1</sup> analysis, and fragmentation spectra were collected from metabolite features identified by PODIUM as being derived from Phe or Tyr.

common to all three libraries with an additional metabolite found only in above-ground tissues (Table 3). This list included acetylated phenylalanine and metabolites M371T520 and M385T559 that were tentatively identified as conjugates of benzoic acid and phenylacetic acid, which are molecules that cannot be made from Tyr (Table 3, Figure 6B). Phenylacetic acid has recently been shown to be produced from Phe through a PAL/PTAL-independent pathway involving phenylacetaldoxime (Perez et al., 2021).

## Phe and Tyr Are Differentially Incorporated Into Common Hydroxycinnamates and Flavonoids

The relative proportions of <sup>13</sup>C incorporation into putative PTAL-dependent metabolites differed between amino acid treatment and between tissues (Figure 4A). Because the metabolism of Phe and Tyr both produce *p*-coumaric acid, these differences in the incorporation into downstream metabolites suggested more than one *p*-coumaric acid pool exists, with the contributions of Phe and Tyr to these pools not being equivalent. This possibility was examined for individual metabolites by determining label incorporation into *p*-coumaric acid and 25 other abundant HCEs and flavonoids whose structures were predicted by *post-hoc* MS/MS (Figure 7). For *p*-coumaric acid, there was a <10% difference in <sup>13</sup>C labeling from Tyr and Phe. This measurement represents a tissue-wide average and does not shed light on potentially distinct cellular or subcellular pools. However, downstream of *p*-coumaric acid, the flavonoids kaempferol glucoside, apigenin, apigenin glycoside, and an isomer of an apigenin C-glycoside were enriched in label from <sup>13</sup>C-Phe. Label from <sup>13</sup>C-Tyr was enriched in triclin, apigenin glucoside rhamnoside, and an isomer of apigenin C-glycoside. For HCEs, as was seen with *p*-coumaric acid, there was no significant difference in Tyr and Phe incorporation into *p*-coumaroylshikimate and one isomer of *p*-coumaroylquininate in leaves (Figure 7A), although *p*-coumaroylshikimate in the base and root portions were slightly more enriched from Tyr (Figures 7B,C). In contrast, two of the three detected isomers of

*p*-coumaroylquininate and most caffeoyl and feruloyl esters were labeled more from Phe than Tyr in the leaf and base portions. In roots, only one isomer of feruloylquininate was more enriched in label from Phe, while the other HCEs were labeled evenly from both amino acids. Taken together, the lack of a consistent labeled Phe-to-Tyr incorporation ratio between metabolites derived through PAL and/or PTAL implies that there is more than one *p*-coumaric acid pool within each tissue and that they are differentially derived from Phe and Tyr.

## The FDM and YDM Include Soluble Phenylpropanoids Affected in *Brown midrib (bmr6 and bmr12) Mutants of Sorghum*

The library of Tyr and Phe-derived metabolite features was used to characterize the FDM and YDM of two genetically characterized *bmr* mutants, *bmr6* and *bmr12*, which carry loss of function mutations in CAD and COMT, respectively (Palmer et al., 2008). Based on the positions of *bmr6* and *bmr12* in the phenylpropanoid pathway, loss of CAD could negatively affect the production of monolignols and soluble derivatives, such as lignans and neolignans; whereas, loss of COMT activity in *bmr12* could affect the methylation of 5-hydroxyferulic acid, 5-hydroxyconiferaldehyde, and 5-hydroxyconiferyl alcohol into sinapate, sinapaldehyde, and sinapyl alcohol, respectively, and the methylation of flavonoids (Eudes et al., 2017). In total 469 Phe- and Tyr-derived metabolite features were annotated in *bmr* and wild-type leaves, representing 21–28% of the ion counts in those samples (Figure 8A). A PCA using ion abundances of the predicted Phe- and Tyr-derived features separated each genotype along with both components, indicating that the *bmr* mutants have different metabolic phenotypes from each other and those from the wild type (Figure 8B). The total aggregated ion counts for Phe- and Tyr-derived metabolite features in the mutants were not different from those in the wild-type plants.

In total, 163 labeled features were significantly altered in abundance ( $p < 0.05$ ) in one or both mutants, compared to

those from the wild type. Of these, most features significantly decreased in one or both mutants relative to those from the wild type (47 in *bmr6* only, 49 in *bmr12* only, and 39 in *bmr6* and *bmr12*). Those that increased relative to these wild-type plants included 24 metabolite features in *bmr6*, 5 mass features in *bmr12*, and 2 features in both mutants. Zero mass features were significantly affected in both mutants in opposing directions. Most *bmr*-affected features were preferentially derived from Phe in wild type, a result consistent with many other known phenylpropanoids detected through heavy isotope labeling in this study (Figure 8C). Of the MS/MS corroborated metabolites, feruloylquinic acid hyperaccumulated in *bmr6*, but *p*-coumaric acid, an ester of *p*-coumaroylquinic acid, kaempferol, and *p*-coumaroylglycerate were reduced in the same mutant (Figure 8D). In *bmr12*, there was also a significant reduction in *p*-coumarate, *p*-coumaroylquinic acid, *p*-coumaroylglycerate, and the flavonoid tricetin (Figure 8D). The reduction in tricetin is consistent with a previous report suggesting that the COMT disrupted in *bmr12* also methylates some flavonoids (Eudes et al., 2017). Furthermore, neolignans and lignans are phenylpropanoid dimers produced from monolignols, thus, their biosynthesis is CAD-dependent (Dima et al., 2015). These neolignans and lignans can be distinguished in isotopic labeling data if the metabolite features incorporated more than one labeled phenyl ring and, indeed, the dual-labeled features that negatively affected in *bmr6*, but not in *bmr12*, were metabolites that were tentatively identified as lariciresinol hexose (G (8-8) G) (M521T562), a formate adduct of G (8-O-4) caffeoyl hexose (M565T604) (Vanholme et al., 2012), and G(8-5)H-hexoside (M535T598 and M535T623) (Sundin et al., 2014) (Figure 8D).

## DISCUSSION

Although specialized metabolites can be detected and measured by tools such as LC-MS and GC-MS, the identities of most of the mass features collected are not known (Perez De Souza et al., 2017). The PODIUM isotopic labeling protocol and computational pipeline helps solve this by providing precursor-of-origin annotations to metabolites and their features collected by untargeted MS (Simpson et al., 2021). In our previous study, the ability to sort metabolite features into precursor-of-origin groups helped in metabolite identification and the merging of metabolomic data with genomic data to identify genes influencing specific aspects of phenylpropanoid metabolism. Notably, the PODIUM-derived metabolite annotations identified the metabolic consequences of mutating various phenylpropanoid pathway enzymes and uncovered potentially biologically relevant gene-to-metabolite associations through genome-wide association.

Here, PODIUM was applied to identify the soluble FDM and YDM of the reference sorghum genotype BTx623. Despite the emerging importance of sorghum to the bioeconomy, there is considerably less information about its metabolic space, especially compared to *Arabidopsis* and even other monocots, such as rice and maize. The determination of these two sorghum

metabolomes in this study sets the stage to identify metabolite-to-gene associations in re-sequenced mutant populations and diversity panels through genome-wide association. In addition, the determination of both the FDM and YDM is particularly relevant in grasses because they possess dual PAL and PTAL activity; thus, this data allows us to compare the relative capacity of sorghum to incorporate each aromatic amino acid into common soluble phenylpropanoids.

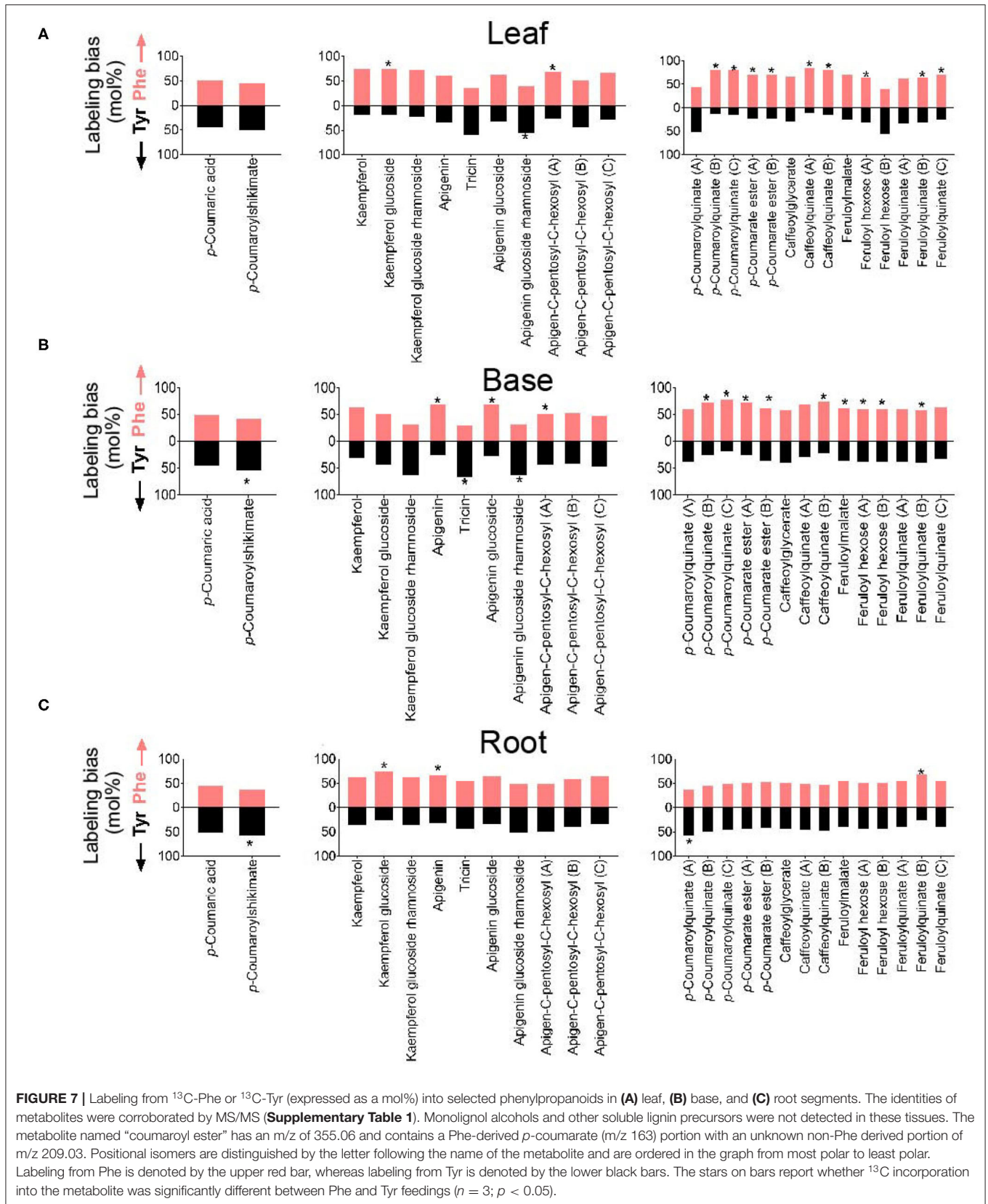
The primary focus of this work was to annotate Phe- and Tyr-derived soluble metabolomes of young sorghum plants. Many of the so-called specialized metabolites that confer resistance to biotic and abiotic stresses are contained within this subset of the metabolome. In sorghum, this includes the cyanogenic glycoside dhurrin that is derived from Tyr, and various hydroxycinnamic esters and flavonoid glycosides that are produced by the phenylpropanoid pathway following the deamination of Phe or Tyr. Although our labeling approach captured many specialized metabolites derived from the amino acids, including *p*-coumaric and ferulic acids, the monolignol alcohols that are polymerized into lignin were not detected. Previously, the feeding of  $^{13}\text{C}_6$ -Phe to *Arabidopsis* stems resulted in the label incorporation into lignin and the detection of monolignols and free hydroxycinnamic acids (Wang P. et al., 2018). In contrast, the young sorghum plants used in this study were likely not synthesizing secondary cell walls at comparable rates, so the pools of soluble lignin precursors may have been too low to be detected.

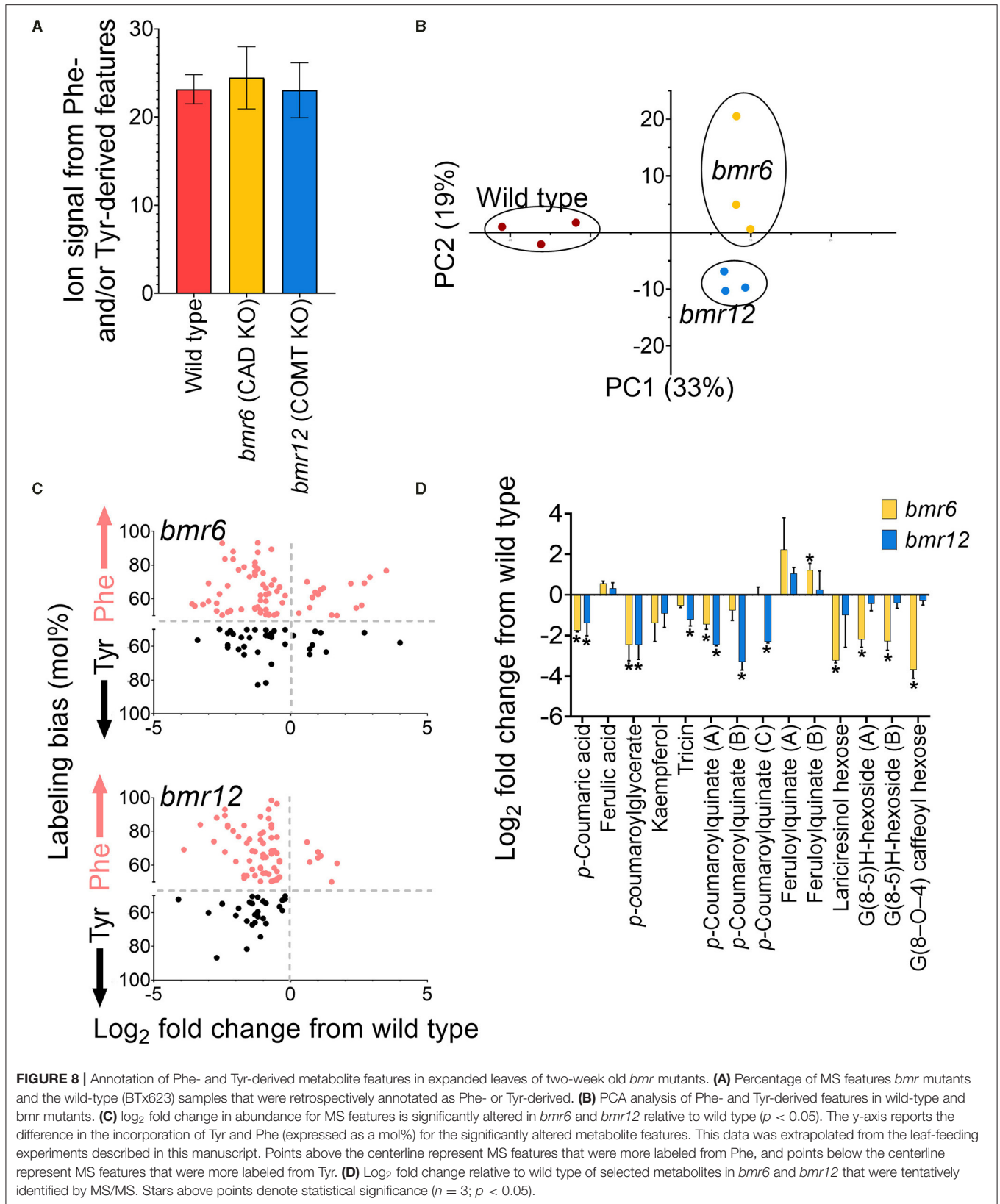
The total ion count from Phe-derived metabolites in sorghum was between 25 and 45% of the total ion count, which is a value comparable to what was observed in the Phe-derived soluble metabolome of *Arabidopsis* stems, a plant that does not have PTAL activity (Simpson et al., 2021). Phe-derived metabolite features detected in both species include free hydroxycinnamic acids *p*-coumarate and ferulate, some kaempferol-derived flavonoids, and coniferyl alcohol coupling products. However, for the most part, the two FDMs are very different. The most abundant metabolite features in *Arabidopsis* stems were hydroxycinnamic acids, notably sinapic acid, esterified to glucose and malate, which were not detected in sorghum. In contrast, coumarate, caffeate, and ferulate esters to quinic acid, and apigenin-containing flavonoids that are produced in sorghum were largely absent from *Arabidopsis* stems.

## Phe and Tyr Contribute Differently to Phenylpropanoid Biosynthesis in Sorghum

Phenylalanine/tyrosine ammonia-lyase activity in plants has been known since the initial elucidation of the phenylpropanoid pathway by radioisotope labeling (Brown and Neish, 1956; Neish, 1961; Jangaard, 1974). Dual PAL and TAL activity in plants appears to only exist in grasses, where PTAL activity and the corresponding genes have been functionally characterized in maize (Rösler et al., 1997), *Brachypodium* (Cass et al., 2015; Barros et al., 2016), and sorghum (Jun et al., 2018), among others (Barros and Dixon, 2020). Sequence analysis and mutagenesis indicated that PTALs can use Tyr as a substrate because of a Phe-to-His substitution in the active







site of the enzyme (Watts et al., 2006), although additional substitutions may contribute to substrate promiscuity (Jun et al., 2018). Sorghum contains eight PAL isoenzymes, and sequence analysis and *in vitro* characterization identified two of them as PTALs: Sb04g026510 (SbPAL1) and Sb06g022740 (Jun et al., 2018). In this study, precursor labeling identified 668 metabolite features derived from Phe and/or Tyr across three tissues, representing over 20% of the ion signal acquired in each tissue by reverse-phase chromatography and negative-mode ionization. Most labeled metabolite features were labeled from both Phe and Tyr, providing *in planta* evidence that sorghum can use both amino acids to synthesize soluble *p*-coumaric acid-derived phenylpropanoids.

The analysis of the flow of  $^{13}\text{C}$  into metabolites identified clear differences in the relative incorporation of Phe and Tyr into *p*-coumaric acid-derived phenylpropanoids. The incorporation of labeled substrates is influenced by the size of preexisting precursor pools, the capacity of the substrate to access metabolically active pools, and differences in substrate uptake. The incorporation of Phe and Tyr into metabolites presented here provided an estimation of the capacity of the plant to channel exogenous Phe and Tyr into downstream metabolites. Most labeled metabolite features did not exhibit any precursor-of-origin bias; however, half of the structurally characterized HCEs and flavonoids were more strongly labeled from Phe. This bias was strongest in the two above-ground tissues. These observations are consistent with multiple independent *p*-coumaric acid pools that are differentially utilized to synthesize some PAL and/or PTAL-derived metabolites. Distinct *p*-coumaric acid pools and the idea that PAL and PTAL are responsible for different products were previously suggested after silencing the single PTAL gene of *Brachypodium* (Barros et al., 2016). In the mentioned study, labeled Phe and Tyr were incorporated relatively equally into lignin in wild type *Brachypodium*. However, in the PTAL RNAi line, syringyl monomers and cell-wall-bound *p*-coumarate were more affected than other lignin or cell wall components. Because PTAL affinity and activity toward Tyr were greater than for Phe, the authors proposed that the TAL functionality of PTAL is preferentially utilized for making those cell wall and lignin components. They also showed that the downregulation of PTAL causes a significant reduction in soluble *p*-coumarate and some flavonoids, whereas ferulic acid and chlorogenic acid increased over 2-fold in that mutant. This data indicated that *Brachypodium* PTAL should be used along with PAL to also synthesize soluble products of the phenylpropanoid pathway, and, similar to the data in sorghum, suggested that PAL and PTAL may have different roles for soluble phenylpropanoid biosynthesis. How Phe and Tyr are metabolized into different *p*-coumaric pools to produce different products could be influenced by several factors including PTAL vs. PAL expression, the kinetic properties of PALs and PTALs, and the compartmentalization of enzymes and substrates within plant cells and tissues. Localized expression differences may contribute significantly to the differences in Phe and Tyr incorporation into phenylpropanoids and/or lignin in *Brachypodium*, where PTAL and PAL genes are co-expressed

with different suites of genes (Barros and Dixon, 2020, Barros et al., 2016). Publicly available expression data from sorghum of equivalent age and tissue to the study (accessed from gramene.org) (Davidson et al., 2012; Shakoor et al., 2014; Wang et al., 2018) showed that the PALs with characterized PTAL functionalities (Jun et al., 2018) were more highly expressed than the six other annotated sorghum PAL isozymes. *In vitro* kinetic analysis of *Brachypodium* PTAL revealed a preference for Tyr, although sorghum and maize PTALs had similar activities toward both amino acids (Barros et al., 2016; Jun et al., 2018). Distinct *p*-coumaric acid pools that are utilized differentially could also come about through metabolic channeling between PAL and the ER-bound P450s, such as C4H, that modify phenylpropanoid pathway intermediates (Achnine et al., 2004; Barros and Dixon, 2020). Barros et al. (2016) showed that the lignin from  $^{13}\text{C}$ -Tyr (and thus requiring its PTAL enzyme) was more sensitive to isotope dilution from exogenously fed *p*-coumaric acid compared with the lignin produced from  $^{13}\text{C}$ -Phe. This result demonstrates that the endogenous  $^{13}\text{C}$ -Phe-derived *p*-coumaric acid pool is unable to mix with exogenous *p*-coumaric acid and, thus, points to the physical separation of PAL and PTALs and/or the *p*-coumarate pools to which they contribute (Barros et al., 2016). Therefore, in this model, Tyr would be excluded from PAL-C4H metabolons because it is not a PAL substrate. Alternatively, if Tyr enters a PTAL-C4H metabolon, then the *p*-coumarate generated may be eliminated from the channel because it cannot be acted upon by C4H (Barros et al., 2016; Dixon and Barros, 2019). How these different strategies of *p*-coumarate production would then impact activation by 4-coumarate coenzyme A ligase (4CL) is unclear. Tyr-derived *p*-coumarate acid could also bypass this metabolon *via* the recently discovered cytosolic bifunctional coumarate 3-hydroxylase/ascorbate peroxidase (C3H/APX) that can hydroxylate *p*-coumarate acid directly to caffeic acid (Barros et al., 2019). It was proposed that this pathway may be favored to produce grass-type lignin (i.e., S-type) and the lignin made during stress because of the association of C3H/APX with detoxifying reactive oxygen species (Barros et al., 2019; Barros and Dixon, 2020).

Besides producing *p*-coumaric acid, Phe and Tyr can enter several alternate pathways leading to unique metabolites (Widhalm and Dudareva, 2015; Schenck and Maeda, 2018). Sorghum produces large amounts of the cyanogenic glycoside dhurrin derived only from Tyr (Halkier and Møller, 1989; Gleadow and Møller, 2014) and, accordingly, substantial  $^{13}\text{C}$ -Tyr incorporation into dhurrin, mass features consistent with dhurrin breakdown products, and other metabolites derived only from Tyr were observed. Metabolites derived exclusively from Phe were present, but  $^{13}\text{C}$  incorporation into them was considerably less compared to Tyr-only metabolites. Because dhurrin biosynthesis places a high demand on Tyr pools, the observed Phe-labeling bias for many PTAL-derived phenylpropanoids may be due to the reduced availability of Tyr for PTAL. If so, dhurrin accumulating grasses may exhibit a shared pattern of the proportioning of Phe and Tyr into phenylpropanoids.

## Identification of Soluble Phenylpropanoids Affected in *Brown midrib* Mutants

*Brown midrib* mutants (*bmr*) are distinguished by an altered lignin composition and the accumulation of red or brown pigmentation in their leaf and stem vasculature (Sattler et al., 2014). Three *bmr* loci in sorghum encode the known phenylpropanoid pathway enzymes 4CL (*bmr2*) (Saballos et al., 2012), CAD (*bmr6*), and OMT (*bmr12*) (Palmer et al., 2008). The altered and/or reduced lignin phenotypes in *bmr* mutants have been characterized extensively because of their utility in saccharification and forage applications (Sattler et al., 2010). There are contradictory reports on the susceptibility of *bmr* maize and sorghum to diseases, with some studies reporting an increased susceptibility due to impaired cell wall function and others reporting increased resistance, possibly due to the increased levels of phenylpropanoid specialized metabolites (Sattler et al., 2010). Previous work examining soluble hydroxycinnamic acids from mature stems of two different sorghum lines containing *bmr6* or *bmr12* mutations showed increased levels of free ferulic acid, but opposing effects on *p*-coumaric acid where it was decreased in *bmr6* and increased in *bmr12* (Palmer et al., 2008; Tetreault et al., 2020). In this study, a similar effect on *p*-coumarate, and *p*-coumarate and ferulate esters in *bmr6* were noted, but reduced levels of *p*-coumarate esters in *bmr12* were observed as well.

The synthesis of phenylpropanoid dimers, such as neolignans and lignans, requires the production of cinnamyl alcohols by CAD. Accordingly, some of the most negatively affected metabolites in *bmr6*, and not *bmr12*, were dual-labeled metabolites. This result also indicates that the CAD locus affected in *bmr6* is required for lignan synthesis, in addition to its role in lignin deposition. The *bmr6* and *bmr12* mutants of sorghum are defective at the same respective loci as the *cadc*, *cadd*, and *omt1* mutants of *Arabidopsis* (Fraser and Chapple, 2011). In *Arabidopsis*, the FDMs of *cadc*, *cadd*, and *omt1* were distinct from each other and those from the wild type, and a similar result was found in sorghum. However, distinct alterations in phenylpropanoid compounds were found in each species. Notably, the loss of *omt1* *Arabidopsis* resulted in the accumulation of the glycosylated substrate of the enzyme

(5-hydroxyferuloyl hexose), but this metabolite was not detected in sorghum. Both the loss of CAD/CADD in *Arabidopsis* and the loss of an equivalent CAD in sorghum caused a reduction in known coniferyl alcohol-derived dimers; however, the loss of these enzymes in *Arabidopsis* was also accounted with a general increase in upstream hydroxycinnamic esters, namely, sinapoyl and feruloyl malate, which was a trend not observed in sorghum.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

JS, BD, and CC conceived of the project. JS and CC wrote the manuscript. All authors contributed to the conducting of the experiments and the analysis and discussion of the results.

## FUNDING

This work was supported by the US Department of Energy, Office of Science (BER), Grant DE-SC0020368 (CC and BD), and JS were supported by a United States Department of Agriculture National Institute of Food and Agriculture postdoctoral grant 2018-08121/1019231.

## ACKNOWLEDGMENTS

We thank Dr. Bruce Cooper (Bindley Bioscience Center, Purdue University) for assistance in the acquisition of the LC-MS metabolite profiling data. We also thank Wilfred Vermerris (University of Florida) for donating the *bmr6* and *bmr12* seeds.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.714164/full#supplementary-material>

## REFERENCES

- Achnine, L., Blancaflor, E. B., Rasmussen, S., and Dixon, R. A. (2004). Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell* 16, 3098–3109. doi: 10.1105/tpc.104.024406
- Adeyemi, S. R. (1990). Turnover of dhurrin in green sorghum seedlings. *Plant Physiol.* 94, 1219–1224. doi: 10.1104/pp.94.3.1219
- Barros, J., and Dixon, R. A. (2020). Plant Phenylalanine/tyrosine ammonia-lyases. *Trends Plant Sci.* 25, 66–79. doi: 10.1016/j.tplants.2019.09.011
- Barros, J., Escamilla-Trevino, L., Song, L., Rao, X., Serrani-Yarce, J. C., Palacios, M. D., et al. (2019). 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nat. Commun.* 10:1994. doi: 10.1038/s41467-019-10082-7
- Barros, J., Serrani-Yarce, J. C., Chen, F., Baxter, D., Venables, B. J., and Dixon, R. A. (2016). Role of bifunctional ammonia-lyase in grass cell wall biosynthesis. *Nat. Plants* 2:16050. doi: 10.1038/nplants.2016.50
- Brenton, Z. W., Cooper, E. A., Myers, M. T., Boyles, R. E., Shakoore, N., Zielinski, K. J., et al. (2016). A genomic resource for the development, improvement, and exploitation of sorghum for bioenergy. *Genetics* 204, 21–33. doi: 10.1534/genetics.115.183947
- Brown, S. A., and Neish, A. C. (1956). Studies of lignin biosynthesis using isotopic carbon. V. Comparative studies on different plant species. *Can. J. Biochem. Phys.* 34, 769–778. doi: 10.1139/y56-081
- Cass, C. L., Peraldi, A., Dowd, P. F., Mottiar, Y., Santoro, N., Karlen, S. D., et al. (2015). Effects of phenylalanine ammonia lyase (PAL) knockdown on cell wall composition, biomass digestibility, and biotic and abiotic stress responses in *Brachypodium*. *J. Exp. Bot.* 66, 4317–4335. doi: 10.1093/jxb/erv269
- Davidson, R. M., Gowda, M., Moghe, G., Lin, H., Vaillancourt, B., Shiu, S. H., et al. (2012). Comparative transcriptomics of three poaceae species



- reveals patterns of gene expression evolution. *Plant J.* 71, 492–502. doi: 10.1111/j.1365-313X.2012.05005.x
- Dima, O., Morreel, K., Vanholme, B., Kim, H., Ralph, J., and Boerjan, W. (2015). Small glycosylated lignin oligomers are stored in *Arabidopsis* leaf vacuoles. *Plant Cell* 27, 695–710. doi: 10.1105/tpc.114.134643
- Dixon, R. A., and Barros, J. (2019). Lignin biosynthesis: old roads revisited and new roads explored. *Open Biol.* 9:190215. doi: 10.1098/rsob.190215
- Eudes, A., Dutta, T., Deng, K., Jacquet, N., Sinha, A., Benites, V. T., et al. (2017). SbCOMT (Bmr12) is involved in the biosynthesis of triclin-lignin in sorghum. *PLoS ONE* 12:e0178160. doi: 10.1371/journal.pone.0178160
- Fraser, C. M., and Chapple, C. (2011). The phenylpropanoid pathway in *Arabidopsis*. *Arabidopsis Book* 9:e0152. doi: 10.1199/tab.0152
- Gleadow, R. M., and Möller, B. L. (2014). Cyanogenic glycosides: synthesis, physiology, and phenotypic plasticity. *Annu. Rev. Plant Biol.* 65, 155–185. doi: 10.1146/annurev-arplant-050213-040027
- Halkier, B. A., and Möller, B. L. (1989). Biosynthesis of the cyanogenic glucoside dhurrin in seedlings of *Sorghum bicolor* (L.) Moench and partial purification of the enzyme system involved. *Plant Physiol.* 90, 1552–1559. doi: 10.1104/pp.90.4.1552
- Jangaard, N. O. (1974). The characterization of phenylalanine ammonia-lyase from several plant species. *Phytochemistry* 13, 1765–1768. doi: 10.1016/0031-9422(74)85086-7
- Jun, S. Y., Sattler, S. A., Cortez, G. S., Vermerris, W., Sattler, S. E., and Kang, C. (2018). Biochemical and structural analysis of substrate specificity of a phenylalanine ammonia-lyase. *Plant Physiol.* 176, 1452–1468. doi: 10.1104/pp.17.01608
- Neish, A. C. (1961). Formation of *m*- and *p*- coumaric acids by enzymatic deamination of the corresponding isomers of tyrosine. *Phytochemistry* 1, 1–24. doi: 10.1016/S0031-9422(00)82806-X
- Palmer, N. A., Sattler, S. E., Saathoff, A. J., Funnell, D., Pedersen, J. F., and Sarath, G. (2008). Genetic background impacts soluble and cell wall-bound aromatics in *brown midrib* mutants of sorghum. *Planta* 229, 115–127. doi: 10.1007/s00425-008-0814-1
- Perez De Souza, L., Naake, T., Tohge, T., and Fernie, A. R. (2017). From chromatogram to analyte to metabolite. How to pick horses for courses from the massive web resources for mass spectral plant metabolomics. *Gigascience* 6, 1–20. doi: 10.1093/gigascience/gix037
- Perez, V. C., Dai, R., Bai, B., Tomiczek, B., Askey, B. C., Zhang, Y., et al. (2021). Aldoximes are precursors of auxins in *Arabidopsis* and maize. *New Phytol.* 231, 1449–1461. doi: 10.1111/nph.17447
- Rösler, J., Krekel, F., Amrhein, N., and Schmid, J. (1997). Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. *Plant Physiol.* 113, 175–179. doi: 10.1104/pp.113.1.175
- Saballos, A., Sattler, S. E., Sanchez, E., Foster, T. P., Xin, Z., Kang, C., et al. (2012). Brown midrib2 (Bmr2) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (*Sorghum bicolor* (L.) Moench). *Plant J.* 70, 818–830. doi: 10.1111/j.1365-313X.2012.04933.x
- Sattler, S. E., Funnell-Harris, D. L., and Pedersen, J. F. (2010). Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Science* 178, 229–238. doi: 10.1016/j.plantsci.2010.01.001
- Sattler, S. E., Saballos, A., Xin, Z., Funnell-Harris, D. L., Vermerris, W., and Pedersen, J. F. (2014). Characterization of novel Sorghum brown midrib mutants from an EMS-mutagenized population. *G3* 4, 2115–2124. doi: 10.1534/g3.114.014001
- Schenck, C. A., and Maeda, H. A. (2018). Tyrosine biosynthesis, metabolism, and catabolism in plants. *Phytochemistry* 149, 82–102. doi: 10.1016/j.phytochem.2018.02.003
- Shakoor, N., Nair, R., Crasta, O., Morris, G., Feltus, A., and Kresovich, S. (2014). A *Sorghum bicolor* expression atlas reveals dynamic genotype-specific expression profiles for vegetative tissues of grain, sweet and bioenergy sorghums. *BMC Plant Biol.* 14:35. doi: 10.1186/1471-2229-14-35
- Simpson, J. P., Wunderlich, C., Li, X., Svedin, E., Dilkes, B., and Chapple, C. (2021). Metabolic source isotopic pair labeling and genome-wide association are complementary tools for the identification of metabolite-gene associations in plants. *Plant Cell* 33 492–510. doi: 10.1093/plcell/koaa046
- Smith, C. A., E. J. W., O'maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* 78, 779–787. doi: 10.1021/ac051437y
- Stark, T. D., Lösch, S., Balemba, O. B., and Hofmann, T. (2017). Two new benzoyl glucuronosyl glycerols from the leaves of *Garcinia buchananii* baker. *Phytochem. Lett.* 19, 187–190. doi: 10.1016/j.phytol.2016.12.033
- Starr, J. L., Newton, R. J., and Miller, F. R. (1984). Presence of dhurrin in sorghum root tissue and the effect of pathogenesis on hydrogen cyanide potential. *Crop Sci.* 24, 739–742. doi: 10.2135/cropsci1984.0011183X002400040026x
- Sundin, L., Vanholme, R., Geerinck, J., Goeminne, G., Höfer, R., Kim, H., et al. (2014). Mutation of the inducible *ARABIDOPSIS THALIANA CYTOCHROME P450 REDUCTASE2* alters lignin composition and improves saccharification. *Plant Physiol.* 166, 1956–1971. doi: 10.1104/pp.114.245548
- Tetreault, H. M., Gries, T., Palmer, N. A., Funnell-Harris, D. L., Sato, S., Ge, Z., et al. (2020). Overexpression of ferulate 5-hydroxylase increases syringyl units in sorghum bicolor. *Plant Mol. Biol.* 103, 269–285. doi: 10.1007/s11103-020-00991-3
- Tohge, T., and Fernie, A. R. (2017). Leveraging natural variance towards enhanced understanding of phytochemical sunscreens. *Trends Plant Sci.* 22, 308–315. doi: 10.1016/j.tplants.2017.01.003
- Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J. H., Goeminne, G., et al. (2012). A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. *Plant Cell* 24, 3506–3529. doi: 10.1105/tpc.112.102574
- Wang, B., Regulski, M., Tseng, E., Olson, A., Goodwin, S., McCombie, W. R., et al. (2018). A comparative transcriptional landscape of maize and sorghum obtained by single-molecule sequencing. *Genome Res.* 28, 921–932. doi: 10.1101/gr.227462.117
- Wang, P., Guo, L., Jaini, R., Klempien, A., McCoy, R. M., Morgan, J. A., et al. (2018). A <sup>13</sup>C isotope labeling method for the measurement of lignin metabolic flux in *Arabidopsis* stems. *Plant Methods* 14:51. doi: 10.1186/s13007-018-0318-3
- Watts, K. T., Mijts, B. N., Lee, P. C., Manning, A. J., and Schmidt-Dannert, C. (2006). Discovery of a substrate selectivity switch in tyrosine ammonia-lyase, a member of the aromatic amino acid lyase family. *Chem. Biol.* 13, 1317–1326. doi: 10.1016/j.chembiol.2006.10.008
- Widhalm, J. R., and Dudareva, N. (2015). A familiar ring to it: biosynthesis of plant benzoic acids. *Mol. Plant* 8, 83–97. doi: 10.1016/j.molp.2014.12.001

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Simpson, Olson, Dilkes and Chapple. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.