



Non-targeted Metabolomics in Diverse Sorghum Breeding Lines Indicates Primary and Secondary Metabolite Profiles Are Associated with Plant Biomass Accumulation and Photosynthesis

Marie F. Turner^{1†}, Adam L. Heuberger^{2†}, Jay S. Kirkwood^{1,3}, Carl C. Collins¹, Edward J. Wolfrum⁴, Corey D. Broeckling³, Jessica E. Prenni³ and Courtney E. Jahn^{1*}

OPEN ACCESS

Edited by:

Ute Roessner,
The University of Melbourne, Australia

Reviewed by:

Mingshu Cao,
AgResearch, Grasslands Research
Center, New Zealand
Robert Henry,
The University of Queensland,
Australia

*Correspondence:

Courtney E. Jahn
courtney.jahn@colostate.edu

[†] These authors are co-first authors.

Specialty section:

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

Received: 27 April 2016

Accepted: 15 June 2016

Published: 11 July 2016

Citation:

Turner MF, Heuberger AL,
Kirkwood JS, Collins CC, Wolfrum EJ,
Broeckling CD, Prenni JE and
Jahn CE (2016) Non-targeted
Metabolomics in Diverse Sorghum
Breeding Lines Indicates Primary and
Secondary Metabolite Profiles Are
Associated with Plant Biomass
Accumulation and Photosynthesis.
Front. Plant Sci. 7:953.
doi: 10.3389/fpls.2016.00953

¹ Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO, USA, ² Department of Horticulture and Landscape Architecture, Colorado State University, Fort Collins, CO, USA, ³ Proteomics and Metabolomics Facility, Colorado State University, Fort Collins, CO, USA, ⁴ National Renewable Energy Laboratory, National Bioenergy Center, Golden, CO, USA

Metabolomics is an emerging method to improve our understanding of how genetic diversity affects phenotypic variation in plants. Recent studies have demonstrated that genotype has a major influence on biochemical variation in several types of plant tissues, however, the association between metabolic variation and variation in morphological and physiological traits is largely unknown. *Sorghum bicolor* (L.) is an important food and fuel crop with extensive genetic and phenotypic variation. Sorghum lines have been bred for differing phenotypes beneficial for production of grain (food), stem sugar (food, fuel), and cellulosic biomass (forage, fuel), and these varying phenotypes are the end products of innate metabolic programming which determines how carbon is allocated during plant growth and development. Further, sorghum has been adapted among highly diverse environments. Because of this geographic and phenotypic variation, the sorghum metabolome is expected to be highly divergent; however, metabolite variation in sorghum has not been characterized. Here, we utilize a phenotypically diverse panel of sorghum breeding lines to identify associations between leaf metabolites and morpho-physiological traits. The panel (11 lines) exhibited significant variation for 21 morpho-physiological traits, as well as broader trends in variation by sorghum type (grain vs. biomass types). Variation was also observed for cell wall constituents (glucan, xylan, lignin, ash). Non-targeted metabolomics analysis of leaf tissue showed that 956 of 1181 metabolites varied among the lines (81%, ANOVA, FDR adjusted $p < 0.05$). Both univariate and multivariate analyses determined relationships between metabolites and morpho-physiological traits, and 384 metabolites correlated with at least one trait (32%, $p < 0.05$), including many secondary metabolites such as glycosylated flavonoids and chlorogenic acids. The use of metabolomics to explain relationships between two or more morpho-physiological traits was explored and showed chlorogenic and shikimic acid to

be associated with photosynthesis, early plant growth and final biomass measures in sorghum. Taken together, this study demonstrates the integration of metabolomics with morpho-physiological datasets to elucidate links between plant metabolism, growth, and architecture.

Keywords: *Sorghum bicolor*, GC-MS, LC-MS, biomass, metabolomics, photosynthesis, chlorogenic acid, shikimic acid

INTRODUCTION

The increasing availability of tools such as transcriptomics, metabolomics, and proteomics has helped to facilitate the discovery of previously cryptic connections between genotype and phenotype. Metabolomics, in particular, provides an opportunity to examine a more nuanced phenotypic analysis in that it, while still removed from the visible phenotype, offers a chance at a more functional or mechanistic interpretation of species variation or plant responses at the molecular level (Feussner and Polle, 2015). Additionally, because small molecules are regulated by upstream gene expression and transcript formation, their measurement presents a complementary and functional approach to conventional genomics and transcriptomics (Bino et al., 2004). In this way, metabolomics offers benefits to both basic integrative fields such as systems biology as well as new ways to identify targets for improvement in applied disciplines such as plant breeding (Fernie and Schauer, 2009).

In recent studies, metabolomics has been used in such applications as characterization of biochemical variation within species (e.g., Kusano et al., 2015), discovery of potential metabolic engineering targets (Tsogtbaatar et al., 2015) and examination of plant responses to the whole environment (Steinfath et al., 2010; Heuberger et al., 2014a), as well as responses to individual biotic (e.g., Scandiani et al., 2015) and abiotic (Ganie et al., 2015; Sanchez-Martin et al., 2015) stressors. In addition, genome wide association mapping (GWAS) is being used to explicitly link the chemical diversity of metabolomic profiles with specific locations in the genome to help dissect quantitative traits (Riedelsheimer et al., 2012). Further, metabolic subsets are being explored as potential front-end tools such as biomarkers or for model-based prediction of traits that would otherwise require the high time and resource investments for one or more seasons of field trials (Meyer et al., 2007; Steinfath et al., 2010; Heuberger et al., 2014b).

Drought is the most costly abiotic stress for agriculture worldwide. Globally, many regions are expected to experience even more frequent and severe droughts in the coming century due to increasing atmospheric temperatures (Dai, 2011), making research on drought resistant crops a crucial component to improve food security or reduce costly inputs to biofuel production. Plants with a C₄ photosynthetic pathway are particularly valuable in crop production due to their physiological advantage under hot and dry conditions (Taylor et al., 2014).

Sorghum [*Sorghum bicolor* (L.) Moench] is an internationally important C₄ crop which produces grain, sugar syrup, and cellulosic biomass and can therefore be diverted to multiple markets, including food for human and animal consumption,

and feedstock for various methods of biofuel production. This market flexibility is due to extensive phenotypic variation for the ways in which sorghum accumulates and allocates biomass to its leaves, stems, and panicles. Sorghum is also increasingly used as a model for other C₄ species due to its small genome, available sequence, and annotation resources (Mace et al., 2013; Mullet et al., 2014). In addition, even within relatively limited breeding populations, sorghum is genetically diverse (Evans et al., 2013), with variation for agronomically important traits such as resistance to drought and tolerance of poor soils (Mace et al., 2013). Further, sorghum lines vary for photoperiod sensitivity, a foundational trait that enables breeders to shift carbon pools away from grain and toward vegetative tissues in plants well-suited for forage, biofuel feedstocks, or sugar (Rooney et al., 2007). Varieties that remain vegetative for longer periods of time maintain higher growth rates and can therefore accumulate up to 100% more biomass than grain-types that are quick to reach reproductive maturity (Mullet et al., 2014). Several morphological factors contribute to end biomass yield in sorghum, including variation in not only growth rate, but also allocation to different plant organs (leaves, stems, panicles). We define this collection of associated phenotypes (e.g., growth rate, harvest indices, final yield) as the process of “biomass accumulation.” Despite this morphological variation, sorghum can be broadly classified into two “types” based on allocation of carbon pools to major distinct tissues: (1) “grain type”: small plants bred for dense panicles, or (2) “biomass type”: large plants bred for total biomass (used as forage, sugar, or biofuels).

Because of the significant phenotypic variation in sorghum, it is reasonable to expect that metabolic variation among sorghum lines should also be high; however, this variation has yet to be characterized. This study described herein had two major objectives: (1) To examine and characterize the metabolic variation in an important set of sorghum breeding lines via non-targeted GC- and LC-MS analyses and (2) To explore the association of these metabolite profiles with an array of measured phenotypes (morphological, physiological, and structural carbohydrate content) expected to be relevant to plant growth, biomass accumulation, and biomass quality. Indeed, we found that both individual metabolites and profiles varied widely across lines and many small molecules had strong associations with morphological and physiological phenotypes.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Eleven diploid lines from both “grain” and “biomass” type sorghums were selected to represent available variation and genetic tools [e.g., BTx623 which is sequenced (Paterson et al.,

2009), or *RTx430* which can be transformed (Wu et al., 2014)], making them valuable in pursuit of improved crops. Major characteristics and select publications related to this panel are presented in **Table 1**.

Seeds were germinated (25°C, dark) in Petri dishes on filter paper with fungicide solution (Maxim XL, Syngenta) for 1 week prior to transplanting to pots (3.8 L) filled with Fafard 4P all-purpose, high-porosity potting mix. Plants were grown in a controlled greenhouse environment (mean 25/20°C day/night; 38/51% relative humidity day/night). Five replicates of each line were randomized to minimize effects of varying conditions across the bench. Supplemental lighting maintained a 16/8h light/dark photoperiod and mean daytime PAR was $323 \pm 52 \mu\text{mol m}^{-2}\text{s}^{-1}$. Plants were checked daily to maintain benign (unstressed) conditions and watered when the top 2 cm of medium became dry. Starter nutrients were present in the potting medium but to avoid limitation, plants were fertilized one additional time at 7 weeks post-germination with Osmocote time release (14-14-14) at a rate of 18 g per pot.

Phenotyping of Morphological, Physiological, and Structural Carbohydrate Traits

Plants were evaluated starting at 28 days post-germination and growth rates were calculated as the difference between plant heights in consecutive weeks. At reproductive maturity for each line, the following measurements were made: plant height was measured as distance from potting medium surface to tip of tallest leaf. Stem diameter was measured directly above the first node on the main tiller with a digital caliper (Neiko). Leaf length and width were measured on the five most recently fully expanded leaves, averaged, and used to estimate area. At harvest, biomass was partitioned into leaves, stems, and panicles before oven-drying at 93°C until samples reached constant weight. Harvest indices were calculated as proportions (dry biomass of particular tissue type: total dry biomass).

Chlorophyll extraction was performed on the youngest, fully expanded leaf of 11 week-old plants, which was cut into pieces and ground into a fine powder under liquid nitrogen. Chlorophyll was extracted from 300 mg ground tissue in 50 ml 80% acetone at 4°C, with shaking at 125 rpm for 30 min in the dark. Debris were removed with centrifugation at 1280 relative centrifugal force (rcf) for 15 min at 4°C. Absorbance of the chlorophyll solution was measured using a Synergy HT Multi-Detection Microplate Reader (BIO-TEK Instruments, Inc., Winooski, VT, U.S.A.) at 645 and 663 nm. Chlorophyll content was estimated using the formula of (Arnon, 1949).

Gas exchange measurements were made in the greenhouse on vegetative, 6-week-old plants, on the youngest, fully expanded leaf. Measurements were made in randomized order within a midday 4-h window on 2 consecutive days using the LI-6400XT portable photosynthesis system (LI-COR, Inc.) with the leaf chamber fluorometer attachment. The cuvette was placed at the midpoint of the leaf, avoiding the midrib. Measurements were taken under the following conditions: leaf temperature = 25°C, photosynthetically active radiation (PAR) = $1500 \mu\text{mol m}^{-2}\text{s}^{-1}$, $\text{CO}_2 = 400 \mu\text{mol mol}^{-1}$, and ambient RH = (38–42%). CO_2 fixation rate, stomatal conductance, and Ci:Ca (ratio of intercellular to ambient CO_2) were recorded and intrinsic water use efficiency (WUE) was calculated as a ratio of CO_2 fixation to stomatal conductance.

For NIRS analysis, the stem fraction (at harvest) of dry biomass was ground with a Model 4 Wiley Mill (Thomas Scientific) to pass through a 2 mm screen. Samples were and analyzed via NIRS and a multiple-feedstock multivariate calibration model (Wolfrum et al., 2013; Payne and Wolfrum, 2015) was used to generate predicted percentages and uncertainties for glucan, xylan, lignin, and ash. The calibration model was developed using primary compositional analysis data on a wide variety of biomass samples (including corn stovers, sorghum, miscanthus, cool-season grasses, and switchgrass) using standard wet chemical techniques (Sluiter et al., 2010; Templeton et al., 2010). All samples passed quality control,

TABLE 1 | Line, type, characteristics, and associated publications for lines in this study.

Line	Type/use	Resources and characteristics	Selected publications
BTx623	Grain	Sequenced; pre-flowering drought tolerant	Hart et al., 2001; Brown et al., 2006; Murray et al., 2008b; Paterson et al., 2009
RTx430	Grain	Used for transformation; post-flowering drought susceptible	Miller, 1984; MacKinnon et al., 1987; Howe et al., 2006; Liu and Godwin, 2012; Wu et al., 2014
IS3620C	Grain	Converted inbred	Brown et al., 2006; Burow et al., 2011
BTx642	Grain	Sequenced; post-flowering drought tolerant	Subudhi et al., 2000; Evans et al., 2013
Tx7000	Grain	Elite line; pre-flowering drought tolerant, Sequenced	Subudhi et al., 2000; Kebede et al., 2001; Evans et al., 2013
SC56	Grain	Stay green; pre-flowering drought susceptible	Kebede et al., 2001
SC170	Grain	Grain mold research cultivar	Little and Magill, 2009
M35-1	Biomass/grain	RIL parent	Reddy et al., 2012
100M	Biomass	Photoperiod sensitive NIL(Maturity)	Sorrells and Myers, 1982; Childs et al., 1992; Murphy et al., 2011
Rio	Biomass/sweet	Juicy-stalked; parent line	Broadhead, 1972; Murray et al., 2008a,b; Felderhoff et al., 2012
N598	Biomass/forage	Low-lignin mutant NIL (BMR)	Pedersen et al., 2006

NIL, Near Isogenic Line; RIL, Recombinant Inbred Line; BMR, Brown Mid Rib mutant (low lignin).

and uncertainties (UC) in predictions of compositional data were characterized using the empirical U-deviation method (Zhang and Garcia-Munoz, 2009), which calculates multivariate confidence intervals (CIs), similar in principal to 95% CIs for linear models. Larger UC values indicate higher uncertainties in predicted values. Mass of each constituent was calculated as the decimal percentage multiplied by final biomass.

Metabolite Extraction

Leaf samples were collected from 4-week-old plants. One leaf from each of $n = 5$ plant replicates per breeding line was used. Tissue was excised from leaves via a cork borer. The tissue was collected in the same spot on each leaf replicate, from the middle of the fully developed sixth leaf beside, but not including the midrib. For each sample, ~ 25 mg tissue (dry weight) was immediately placed in a 1.5 mL tube containing stainless steel grinding balls, frozen in liquid N_2 , and homogenized using a paint-shaker. Metabolites were extracted by adding 1 mL of methanol:water (70:30, v:v) and vortexing for 2 h at room temperature. Samples were centrifuged at high speed (13,500 rcf, 10 min, 4°C), and 800 μ L of supernatant was transferred to new tubes and stored at -80°C .

Metabolite Detection using GC-MS

Five hundred microliter of the metabolite extract was dried using a speedvac. Samples were derivatized by re-suspending the extract in 50 μ L of pyridine containing 15 mg/mL of methoxyamine hydrochloride, incubating at 60°C for 45 min, sonicating for 10 min, and incubating again at 60°C for an additional 45 min. Next, 50 μ L of N-methyl-N-trimethylsilyltrifluoroacetamide with 1% trimethylchlorosilane (MSTFA+1% TMCS, Thermo Scientific) was added and samples were incubated at 60°C for 30 min, centrifuged at $3000 \times g$ for 5 min at 4°C , cooled to room temperature, and 80 μ L of supernatant was transferred to a 150 μ L glass insert. Metabolites were detected with a Trace GC Ultra coupled to a Thermo DSQ II (Thermo Scientific), acquiring mass spectra of 50–650 m/z at 5 scans s^{-1} in electron impact mode after separation on a 30 m TG-5MS column (Thermo Scientific, 0.25 mm i.d., 0.25 μ m film thickness). Both inlet and transfer lines were set at 280°C . Samples were injected in a 1:10 split ratio twice in discrete randomized blocks with a 1.2 ml min^{-1} flow rate, following a program of 80°C for 30 sec, a ramp of 15°C per min to 330°C , and holding at 330°C for 8 min.

Metabolite Detection using Ultra Performance LC-MS

One microliter of metabolite extract was injected into an Acquity UPLC system (Waters Corporation). Separation was conducted with an Acquity UPLC T3 column (1.8 μ m, 1.0×100 mm; Waters Co.), using a gradient from solvent A (water, 0.1% formic acid) to solvent B (acetonitrile, 0.1% formic acid). Injections were made in 100% A, which was held for 1 min, a 12 min linear gradient to 95% B was then applied, and held at 95% B for 3 min, returned to starting conditions over 0.05 min, and allowed to re-equilibrate for 3.95 min. Flow rate was constant (200 μ L

min^{-1}) for the entire run duration. The column was held at 50°C with samples held at 5°C . Column eluent was coupled directly to a Xevo G2 Q-ToF MS (Waters Co.) fitted with an electrospray source. Data was collected in positive ion mode, scanning from 50–1200 at 5 scans s^{-1} , alternating between MS and MS^E mode. Collision energy was set to 6 V for MS mode, and ramped from 15–30 V for MS^E mode. Calibration was performed prior to analysis via infusion of sodium formate solution, with mass accuracy within 1 ppm. Capillary voltage was held at 2200 V, source temperature at 150°C , and desolvation temperature at 350°C at a nitrogen desolvation gas flow rate of 800 L h^{-1} .

Metabolomics Data Analysis

For each sample, a matrix of molecular features defined by retention time and mass (m/z) was generated using XCMS software (Smith et al., 2006) using settings as previously described (Broeckling et al., 2014), independently for GC- and LC-MS data sets. Samples were normalized to total ion current and relative quantity of each molecular feature was determined by mean area of the chromatographic peak among replicate injections ($n = 2$). For LC- and GC-MS, mass spectra, and metabolite quantities were generated using an algorithm that clusters masses into spectra (“spectral clusters” that represent “compounds”) based on co-variation and co-elution in the data set (Broeckling et al., 2014), and were annotated by searching against in-house and external metabolite databases including NIST v12 (<http://www.nist.gov>), Massbank (<http://www.massbank.jp>), Golm (gmd.mpimp-golm.mpg.de), and Metlin (metlin.scripps.edu). Annotations using high resolution mass spectrometry (LC-MS) involved the identification of precursor ions within ~ 5 ppm error of the expected positive ion adduct (e.g., H^+ or Na^+). Glycosides were identified by neutral loss of 162.05 m/z in LC-MS spectral clusters.

Statistical Analysis

Analysis of Variance (ANOVA) for morphological and physiological traits was performed in JMP Pro 10 (SAS Institute). Prior to analysis, data were Box-Cox transformed to improve normality. For metabolites, Pearson's and Spearman's correlations, ANOVA, and hierarchical clustering were conducted using cor, aov, and hclust functions in R, respectively (R Core Team, 2012). ANOVA p -values were adjusted for false discovery rate (FDR) using the p.adjust function in R (Benjamini and Hochberg, 1995). Heat maps were generated using the corrplot package in R (Wei, 2010). Z scores were calculated using the mean value of a metabolite compared to the mean and standard deviation of the metabolite for the control variety (BTx623). Principal component analysis (PCA) on morpho-physiological and structural traits was performed on data that was mean-centered and scaled to unit variance (UV). PCA on metabolites was performed on data that was mean-centered and Pareto-scaled using SIMCA v14.0 (Umetrics, Umea, Sweden). The O2PLS model was performed in SIMCA and was a regression of morphology and physiology traits (y variables, scaled to unit variance) against metabolites (x variables, scaled to unit variance).

RESULTS

Significant Metabolite Variation was Observed in Sorghum Leaf Extracts Among Sorghum Lines and between “Grain” and “Biomass” Types

Biochemical variation was investigated among the 11 lines of the sorghum panel in order to explore relationships of primary and secondary metabolites to morphological, physiological, and structural carbohydrate traits. Additionally, the sorghum panel was also analyzed by sorghum type, with respect to size/use as (a) grain sorghum (smaller stature) or (b) forage/sweet/biomass sorghum (larger; referred to as “biomass sorghum”).

A non-targeted metabolomics approach was used to evaluate the extent of variation among diverse genotypes without the need to assign compound names to mass spectra. For this study, metabolites were annotated if they correlated to a morpho-physiological trait, described below (for annotation procedures, see Materials and Methods Section “Metabolomics data analysis”). Overall, 6494 and 10,957 molecular features were detected in LC- and GC-MS datasets, respectively. After clustering features, 487 and 694 spectral clusters were generated for LC- and GC-MS datasets, respectively, for a total of 1181 estimated compounds in the sorghum leaf extracts. Of the 1181 compounds, 584 of 694 compounds detected by GC-MS (84.1%), and 372 of 487 (76.3%) detected by LC-MS varied among the 10 sorghum lines (ANOVA, FDR adjusted $p < 0.05$). When grouped by size type, (biomass or grain) 188 of 694 GC-MS-detected and 132 of 487 LC-MS-detected compounds (27.1% for both methods) were significantly different between types (Supplementary Table 1).

Z-score and Principal component analysis (PCA) were conducted on these compounds and characterized significant metabolic variation across lines. Nine principal components explained 69.8% of the variation (Table 2). The first two principle components (PCs) are depicted in Figure 1; Although the PCA does not appear to show larger metabolomics trends grain vs. biomass lines, this is likely to be due to a larger influence of “line” on metabolite composition than sorghum “type.” ANOVA was conducted on each PC to determine if metabolites differed among lines and/or sorghum types. All nine PCs varied among sorghum lines ($p < 0.05$) and PCs 2, 5, 6, and 7 varied between sorghum types (“grain” vs. “biomass”; Table 2). As an alternative method to characterize metabolic variation within this sorghum population, all spectral clusters were z-transformed using the breeding line BTx623 as the control. The z-scores were plotted as a heat map and organized with hierarchical clustering (Figure 2). The heat map provides an overview of the data and indicates metabolic similarity and variation among the breeding lines. Taken together, the metabolomic profiling data indicates significant variation in sorghum leaf metabolite profiles. Multivariate and univariate analyses demonstrate that most metabolites vary by sorghum line and type, and such variation allows for the discovery of associations to morpho-physiological traits.

TABLE 2 | ANOVA on metabolite principal components, by line, and by type.

Principal component	% Variation explained	p-value by line	p-value by type ^a
1	22.1	<0.0001	0.9181
2	11.4	<0.0001	0.0005
3	7.6	<0.0001	0.1732
4	6.5	<0.0001	0.9971
5	5.8	0.0008	0.0039
6	5.1	<0.0001	0.0044
7	4.2	0.0198	0.0168
8	3.7	<0.0001	0.6501
9	3.4	<0.0001	0.5176

^atype, “biomass” vs. grain.

Morphological, Physiological, and Cell-Wall Structural Traits Varied by Sorghum Line and Type

Multiple morphological, physiological, and structural traits were measured in the sorghum panel (Table 1). Significant variation among lines was identified for most traits, although effect sizes varied widely. A summary of the mean morphological and physiological trait values and ANOVA by line is available in Table 3. Variation in growth rate and reproductive maturity impact the duration of sorghum biomass accumulation; partial growth curves for the sorghum panel are presented in Figure 3. The tallest line (*M35-1*) had highest overall growth rates in early, but not later weeks, and also produced most total dry biomass (83.5 g/plant), most of which derived from stems. In contrast, *IS3620C* produced least biomass, though it was mid-range in height; however, it also had smallest stem diameters (Table 3).

Structural components of plant tissue have important relationships with biomass quality and quantity; therefore NIRS analysis and a calibration model based on wet chemistry were used to examine composition of stem tissue (Table 4). Composition of glucan, xylan, lignin, and ash ranged from 26.8 to 31.3, 16.4 to 19.7, 14.8 to 17.3, and 4.5 to 7.2%, respectively. Total mass/plant of each constituent ranged from 18.9 to 22.5, 9.6 to 15.2, 9.0 to 13.3, and 2.9 to 4.5 g/plant for xylan, glucan, lignin, and ash, respectively.

As expected, biomass varied between the two size types (Figure 4), although interestingly, there was no significant difference between panicle biomass of grain and biomass sorghums. However, it should be noted that the line *100M* influenced certain otherwise consistent trait trends between sorghum types; this is discussed at the end of this section. The remaining traits varied significantly between sorghum grain and biomass types, except for the measured or calculated chlorophyll metrics (Table 5). Biomass sorghums were taller and produced more mass; grain sorghums had greater stem diameters, but allocated less overall biomass to stems. Biomass sorghums allocated a large proportion of mass to stems (stem harvest index, SHI) and had highest absolute stem biomass. Grain sorghums had higher leaf harvest indices (LHI), but biomass sorghums

TABLE 3 | Physiological and morphological trait means, standard errors, and results of ANOVA by line.

Trait	100M	BTx623	BTx642	IS3620C	M35-1	N598	Rio	RTx430	SC170	SC56	Tx7000	F-Ratio	Prob > F (p)	Adj. R ²
Total dry biomass (g)	71.9 ± 2.9	66.3 ± 2.5	63.5 ± 3.0	58.0 ± 1.3	83.5 ± 2.4	72.6 ± 2.3	80.4 ± 2.0	58.5 ± 2.3	63.0 ± 1.4	59.2 ± 1.4	65.3 ± 1.9	15.55	<0.0001	0.66
Leaf dry biomass (g)	25.2 ± 3.0	14.8 ± 0.4	16.2 ± 0.5	13.9 ± 0.3	17.4 ± 0.8	17.3 ± 0.7	20.0 ± 0.3	14.7 ± 0.3	16.2 ± 0.4	15.2 ± 0.3	15.5 ± 0.3	13.76	<0.0001	0.62
Panicle dry biomass (g)	5.6 ± 5.6	27.3 ± 1.3	23.1 ± 1.9	23.3 ± 0.8	25.5 ± 2.2	25.0 ± 1.1	25.5 ± 0.7	25.0 ± 1.4	27.0 ± 0.4	24.1 ± 0.7	28.8 ± 1.2	14.71	<0.0001	0.65
Stem dry biomass (g)	37.0 ± 1.1	24.2 ± 1.0	24.2 ± 1.0	20.8 ± 0.5	40.1 ± 2.7	30.4 ± 1.3	34.9 ± 1.9	18.8 ± 0.8	18.7 ± 1.2	19.9 ± 0.7	21.0 ± 0.6	32.16	<0.0001	0.80
Plant height (cm)	158.3 ± 2.0	138.8 ± 3.0	109.0 ± 4.7	150.6 ± 2.5	231.5 ± 0.2	178.5 ± 6.6	183.1 ± 1.5	118.0 ± 3.0	94.2 ± 2.7	106.0 ± 1.5	124.3 ± 2.7	37.76	<0.0001	0.82
Stem diameter (mm)	13.7 ± 0.28	11.9 ± 0.40	18.2 ± 0.54	10.4 ± 0.25	11.5 ± 0.41	11.1 ± 0.28	12.8 ± 0.28	12.8 ± 0.43	17.8 ± 0.47	16.4 ± 0.53	14.7 ± 0.48	43.78	<0.0001	0.84
Avg. individual leaf area (mm ²)	510.0 ± 9.9	283.9 ± 13	330.3 ± 15	270.4 ± 11	300.6 ± 15	288.4 ± 17	260.4 ± 16	268.8 ± 14	286.2 ± 6.0	322.9 ± 18	330.0 ± 11	28.82	<0.0001	0.78
Leaf harvest index	0.41 ± 0.042	0.22 ± 0.005	0.26 ± 0.012	0.24 ± 0.006	0.21 ± 0.007	0.24 ± 0.005	0.25 ± 0.006	0.25 ± 0.006	0.26 ± 0.003	0.26 ± 0.005	0.24 ± 0.005	13.95	<0.0001	0.63
Panicle harvest index	0.07 ± 0.070	0.41 ± 0.007	0.36 ± 0.013	0.40 ± 0.005	0.31 ± 0.029	0.34 ± 0.010	0.32 ± 0.011	0.43 ± 0.006	0.43 ± 0.008	0.41 ± 0.007	0.44 ± 0.007	27.25	<0.0001	0.78
Stem harvest index	0.52 ± 0.030	0.36 ± 0.005	0.38 ± 0.007	0.36 ± 0.004	0.48 ± 0.026	0.42 ± 0.013	0.43 ± 0.015	0.32 ± 0.004	0.31 ± 0.006	0.34 ± 0.006	0.32 ± 0.004	37.72	<0.0001	0.83
Growth Rt. weeks 1-2 (cm/week)	18.4 ± 1.4	27.5 ± 1.6	13.2 ± 2.3	22.5 ± 2.1	37.6 ± 4.2	36.3 ± 2.5	27.0 ± 5.0	20.7 ± 3.1	11.9 ± 1.7	17.6 ± 1.9	14.8 ± 2.2	8.90	<0.0001	0.50
Growth Rt. weeks 3-4 (cm/week)	24.2 ± 1.2	6.4 ± 0.5	3.8 ± 1.3	21.6 ± 1.2	38.3 ± 4.9	13.6 ± 3.0	8.6 ± 0.9	10.3 ± 1.0	2.5 ± 1.1	6.6 ± 1.2	3.9 ± 0.8	25.85	<0.0001	0.76
ChIA (mg/g)	0.63 ± 0.05	0.66 ± 0.07	0.54 ± 0.05	0.64 ± 0.04	0.52 ± 0.03	0.69 ± 0.02	0.49 ± 0.03	0.72 ± 0.08	0.73 ± 0.06	0.69 ± 0.03	0.49 ± 0.04	4.12	0.0006	0.38
ChIB (mg/g)	0.17 ± 0.02	0.17 ± 0.02	0.15 ± 0.02	0.16 ± 0.01	0.12 ± 0.01	0.19 ± 0.01	0.15 ± 0.01	0.18 ± 0.02	0.22 ± 0.02	0.17 ± 0.01	0.13 ± 0.02	3.65	0.0015	0.34
ChITot (mg/g)	0.81 ± 0.07	0.83 ± 0.09	0.69 ± 0.06	0.81 ± 0.05	0.66 ± 0.03	0.89 ± 0.03	0.64 ± 0.04	0.91 ± 0.10	0.97 ± 0.08	0.88 ± 0.03	0.62 ± 0.06	4.06	0.0006	0.38
A/B	3.74 ± 0.18	3.91 ± 0.12	3.78 ± 0.30	3.97 ± 0.16	4.19 ± 0.08	3.70 ± 0.12	3.23 ± 0.08	3.95 ± 0.28	3.35 ± 0.14	3.96 ± 0.10	3.82 ± 0.26	2.6	0.0154	0.24
A/T	0.78 ± 0.008	0.79 ± 0.005	0.78 ± 0.015	0.79 ± 0.006	0.80 ± 0.003	0.78 ± 0.006	0.75 ± 0.005	0.79 ± 0.011	0.76 ± 0.007	0.79 ± 0.004	0.78 ± 0.011	2.51	0.0185	0.23
Cl/Ca	0.319 ± 0.054	0.273 ± 0.022	0.377 ± 0.007	0.360 ± 0.012	0.362 ± 0.030	0.240 ± 0.034	0.288 ± 0.012	0.268 ± 0.040	0.312 ± 0.015	0.455 ± 0.015	0.309 ± 0.023	6.30	<0.0001	0.19
gs (mmolH ₂ O m ⁻² s ⁻¹)	0.099 ± 0.014	0.133 ± 0.010	0.226 ± 0.004	0.179 ± 0.006	0.184 ± 0.008	0.097 ± 0.007	0.150 ± 0.010	0.193 ± 0.009	0.173 ± 0.011	0.282 ± 0.013	0.195 ± 0.009	30.18	<0.0001	0.57
A _l (μmolCO ₂ m ⁻² s ⁻¹)	13.13 ± 0.50	21.39 ± 1.03	31.83 ± 0.79	25.93 ± 0.68	26.07 ± 0.92	15.87 ± 0.48	24.06 ± 1.30	31.71 ± 0.38	27.70 ± 2.16	33.91 ± 1.62	29.883 ± 1.04	41.04	<0.0001	0.65
WUE	159.2 ± 13.4	168.5 ± 5.79	140.5 ± 1.30	146.4 ± 2.84	145.6 ± 7.11	177.9 ± 8.58	164.0 ± 3.27	167.9 ± 9.39	157.4 ± 2.93	121.2 ± 3.43	157.8 ± 5.67	7.02	<0.0001	0.22

Cl/Ca ratio of intercellular (leaf) [CO₂] to ambient [CO₂]; gs, stomatal conductance; A, photosynthesis; WUE, Water Use Efficiency; ChIA, B, Tot, Chlorophylls A, B, and Total, respectively; A/B and A/T, ratio of chlorophyll A to B and A to Total, respectively.

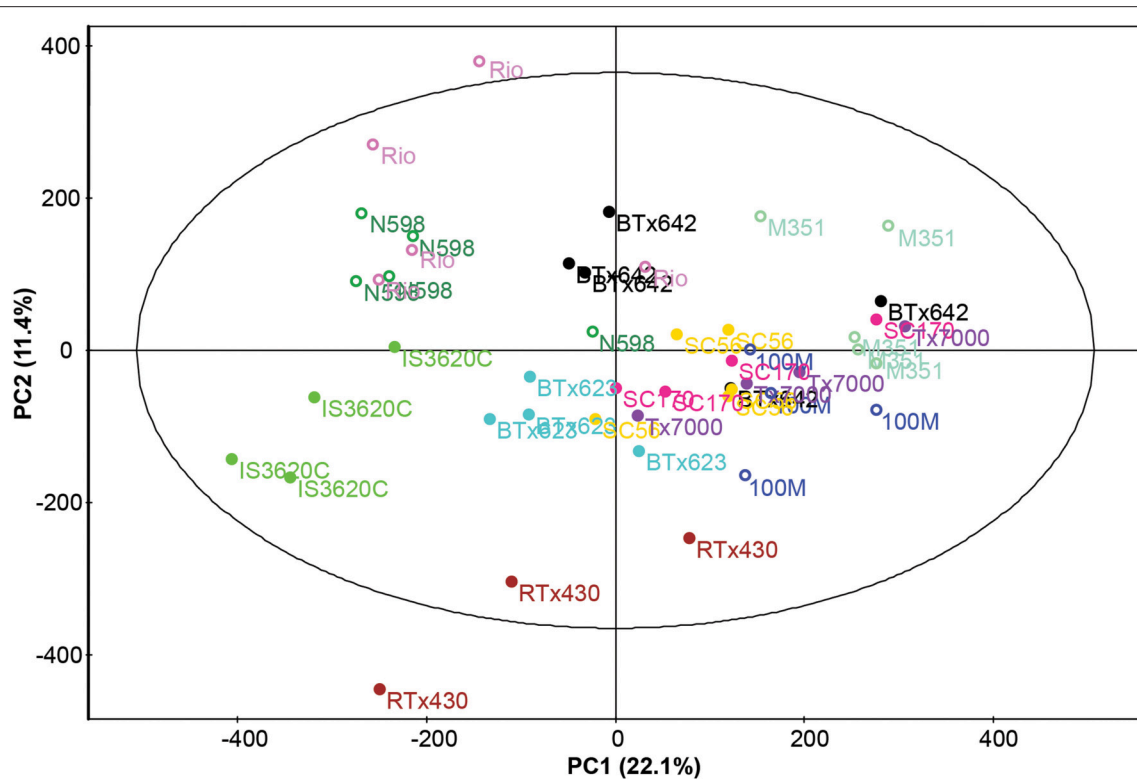


FIGURE 1 | Scores plot from PCA of the metabolomic analysis of 11 sorghum lines. Data from GC- and LC- MS analyses were combined. Biomass types are shown with open symbols and grain types with closed symbols.

TABLE 4 | Predicted percentages of biomass constituents, uncertainties, and estimated grams/plant for the stem fraction of each line.

	Glucan			Xylan			Lignin			Ash		
	(%) ^a	UC	(g/plant) ^b	(%) ^a	UC	(g/plant) ^b	(%) ^a	UC	(g/plant) ^b	(%) ^a	UC	(g/plant) ^b
100M	28.3	3.7	20.4	19.7	2.3	14.2	16.3	2.5	11.7	4.5	3.0	3.2
BTx623	30.0	3.3	21.6	16.9	2.1	11.2	16.0	2.2	10.6	5.9	2.7	3.9
BTx642	28.3	3.8	20.4	17.7	2.5	11.2	16.1	2.6	10.2	4.5	3.1	2.9
IS3620C	31.1	3.7	22.4	18.7	2.3	10.9	15.5	2.4	9.0	6.1	3.0	3.5
M35-1	26.8	4.1	19.3	18.2	2.6	15.2	15.9	2.7	13.3	4.6	3.4	3.8
N598	26.3	5.1	18.9	18.0	3.2	13.1	14.8	3.4	10.8	5.8	4.2	4.2
Rio	27.1	3.8	19.5	18.8	2.4	15.1	16.0	2.5	12.9	5.0	3.1	4.0
RTx430	29.2	3.4	21.0	16.4	2.2	9.6	16.7	2.3	9.8	6.4	2.8	3.7
SC170	31.3	2.9	22.5	16.6	1.9	10.5	15.4	2.0	9.7	7.2	2.4	4.5
SC56	30.4	3.7	21.9	19.5	2.3	11.5	17.3	2.4	10.2	4.9	3.0	2.9
Tx7000	29.8	3.9	21.4	17.0	2.5	11.1	15.0	2.6	9.8	6.5	3.2	4.3

UC, Uncertainty of predicted % Glucan, Xylan, Lignin, and Ash composition (see Materials and Methods). Higher UCs reflect greater uncertainty of the predicted percentage.

^aPredictions determined via NIRS and mixed feedstock calibration model.

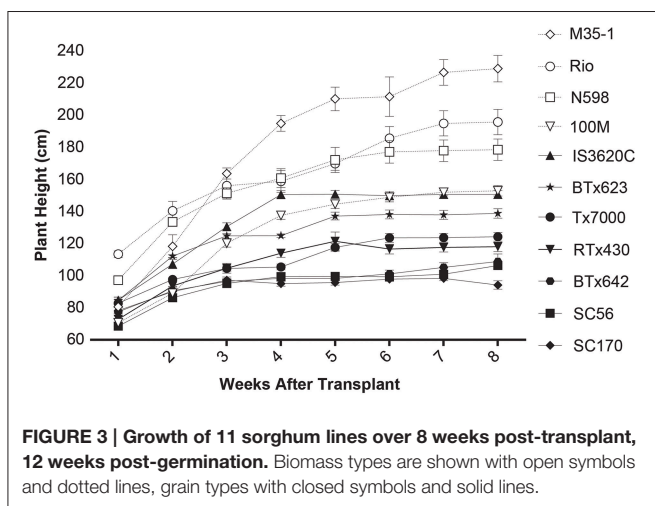
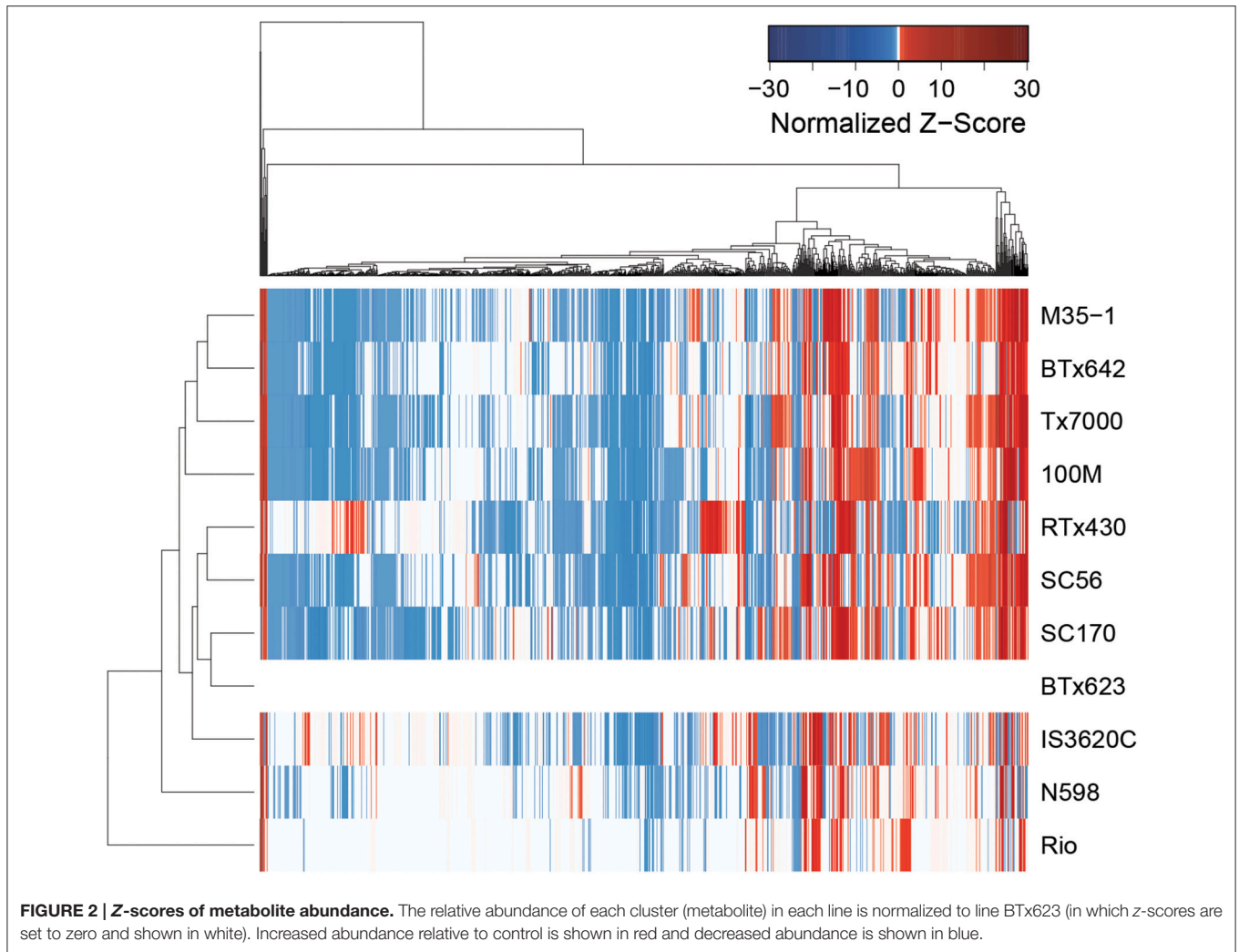
^bEstimations based on predicted percentage and actual dry biomass.

produced more overall leaf biomass; Grain sorghum had higher panicle harvest indices (PHI).

Significant differences in physiological traits were also found between types (Table 5). Smaller grain sorghums had higher rates of photosynthesis, stomatal conductance, and higher Ci:Ca. Biomass sorghums had higher intrinsic water use efficiency (WUE). In addition, PCA was conducted on NIR spectra for

stems, and the first two components explained 40 and 19% of variation, respectively. PC1 and PC2 largely separated grain and biomass (Figure 5), indicating cell wall composition profiles to be relatively consistent with sorghum type.

As noted above, line 100M displayed certain morphological characters which differed from its type ("biomass"). Therefore, for these traits, ANOVA was done both with and without



100M (Table 5). Without *100M*, biomass lines had smaller leaves than grain lines; inclusion of *100M* skewed the trend significantly in the opposite direction for both leaf area and

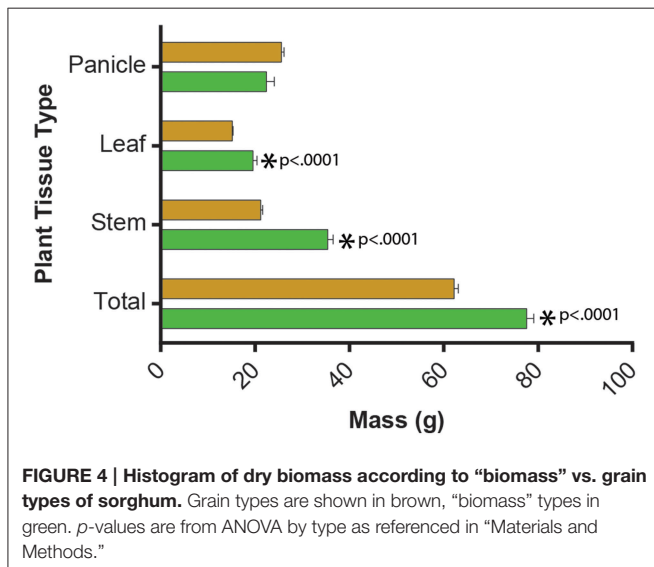
LHI. Additionally, only one of eight *100M* replicates produced a panicle. Therefore, when analysis included *100M*, panicle weight for biomass sorghum was significantly smaller. However, when *100M* was excluded, no difference existed between absolute panicle weights of biomass and grain-type sorghums, though PHI remained significantly different. Aside from these traits, *100M* was consistent with other biomass types.

The co-variation of morpho-physiological traits in this panel was investigated using Spearman's rank correlation (Figure 6, Supplementary Table 2). Photosynthesis and stomatal conductance were negatively associated with biomass and other growth traits and physiological processes were correlated with one another. Photosynthesis was positively correlated with stomatal conductance ($r_s = 0.95$), Ci/Ca ratios ($r_s = 0.48$), and negatively with WUE ($r_s = -0.66$). Plant height and growth rate were also negatively correlated with photosynthesis ($r_s = -0.69, -0.49$). Chlorophyll A (ChlA) and B (ChlB) were positively correlated with one another ($r_s = 0.90$) and ChlA and total chlorophyll (ChlT) were negatively correlated with stem biomass ($r_s = -0.66$ and -0.56 , respectively). Percent glucan was negatively correlated with total dry biomass ($r_s =$

TABLE 5 | Physiological and morphological trait means, standard errors, and results of ANOVA by type.

Trait	Biomass sorghum	Grain sorghum	F-ratio	p	Adj. R ²
	Mean ± SEM	Mean ± SEM			
Plant height (cm)	189.8 ± 6.828	121.6 ± 2.834	106.6	< 0.0001	0.57
Stem diameter (mm)	12.2 ± 0.239	14.6 ± 0.437	13.43	0.0004	0.14
Average Individual leaf area (mm ²)	323.2 ± 19.02 (276.5 ± 9.52)*	300.7 ± 5.997	4.99	0.0287	0.01
Panicle harvest index	0.29 ± 0.022	0.4 ± 0.004	85.13	< 0.0001	0.53
Leaf harvest index	0.26 ± 0.014 (0.23 ± 0.005)*	0.2 ± 0.003	4.28	0.0423	0.01
Stem harvest index	0.45 ± 0.012	0.3 ± 0.004	122.5	< 0.0001	0.62
Growth rate, weeks 1–2 (cm/week)	30.6 ± 2.249	18.4 ± 1.070	26.30	< 0.0001	0.24
Growth rate, weeks 3–4 (cm/week)	21.4 ± 2.697	8.0 ± 0.971	29.25	< 0.0001	0.27
Ci/Ca	0.297 ± 0.016	0.341 ± 0.009	6.44	0.0118	0.02
g _s (mmol H ₂ O m ⁻² s ⁻¹)	0.134 ± 0.006	0.196 ± 0.006	57.25	< 0.0001	0.20
A (μmol CO ₂ m ⁻² s ⁻¹)	20.19 ± 0.730	28.36 ± 0.606	78.76	< 0.0001	0.26
WUE	162.9 ± 4.083	150.3 ± 2.319	6.26	0.0131	0.02
ChlA (mg/g)	0.582 ± 0.025	0.632 ± 0.024	1.99	0.1648	–
ChlB (mg/g)	0.159 ± 0.008	0.168 ± 0.007	0.738	0.3942	–
ChlT (mg/g)	0.749 ± 0.033	0.810 ± 0.030	1.7173	0.196	–
A/T	0.777 ± 0.004	0.781 ± 0.004	0.4636	0.4991	–
A/B	3.716 ± 0.096	3.809 ± 0.080	0.5427	0.4647	–

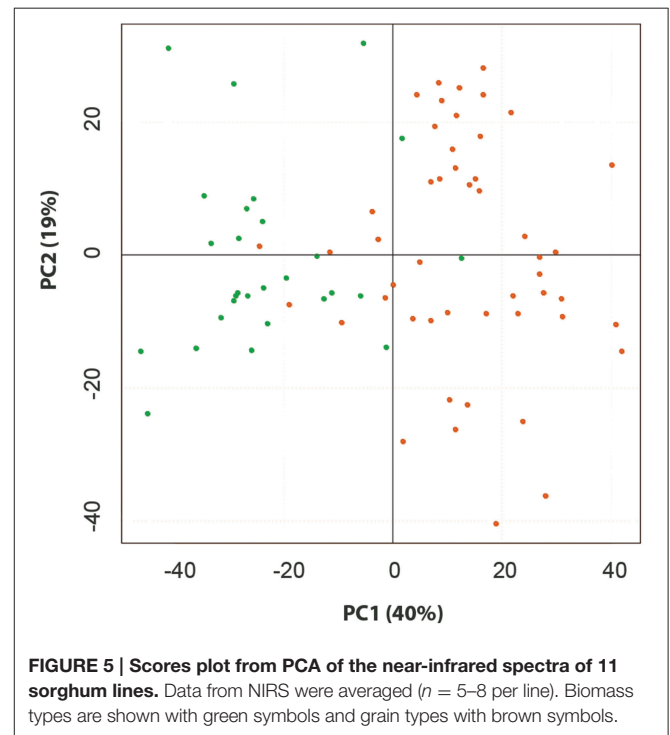
*Starred values calculated with 100M removed as an outlier. Ci/Ca = ratio of intercellular (leaf) [CO₂] to ambient [CO₂]; g_s, stomatal conductance; A, photosynthesis; WUE, Water Use Efficiency; ChlA, B, Tot, Chlorophylls A, B, and Total, respectively; A/B and A/T, ratio of chlorophyll A to B and A to Total, respectively.



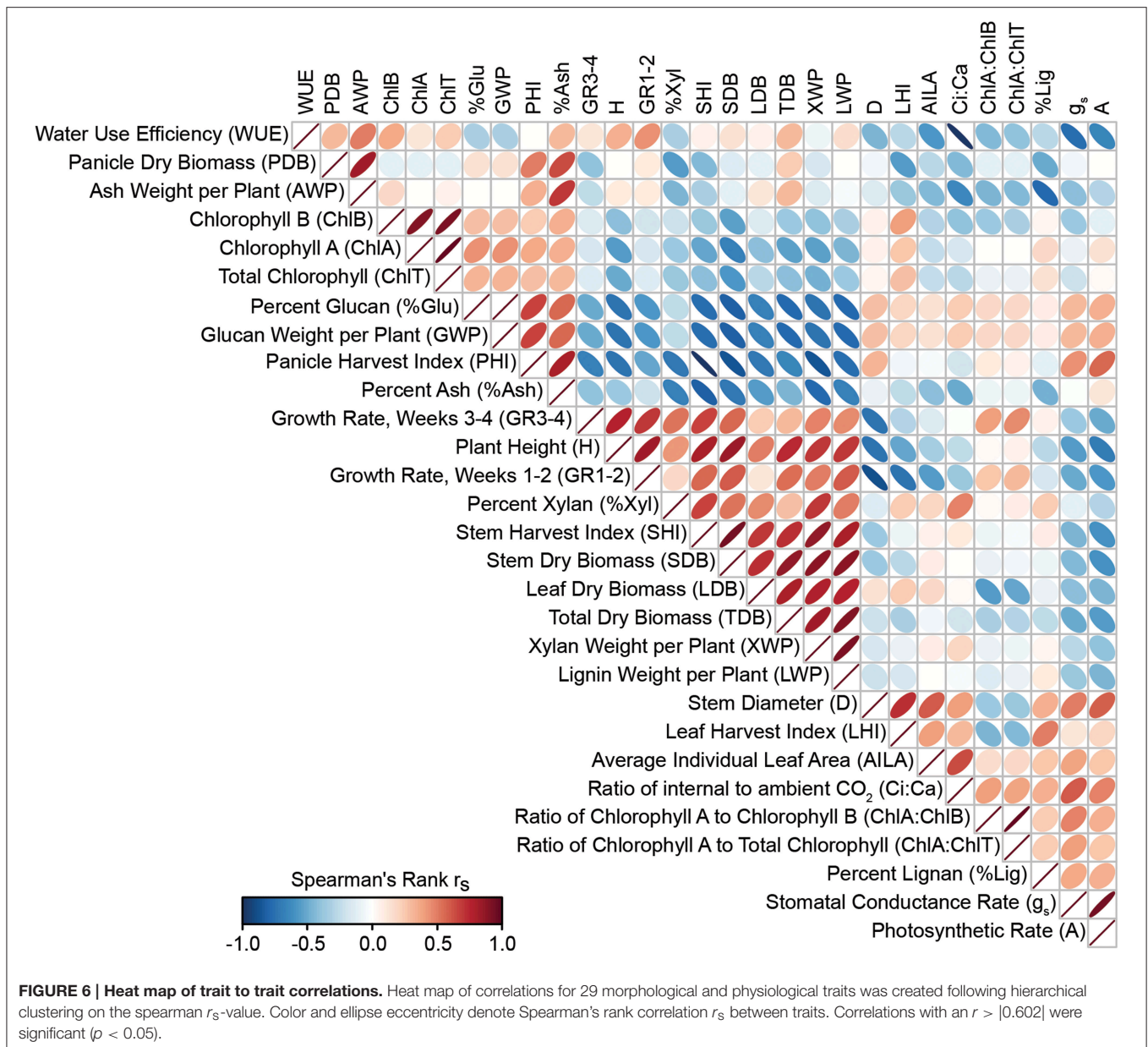
–0.78), plant height ($r_s = -0.72$), and total stem biomass ($r_s = -0.80$) and percent ash was negatively correlated with SHI ($r_s = -0.82$).

Leaf Metabolic Variation Is Associated with Morpho-Physiological Phenotypes

Associations among metabolites and morpho-physiological phenotypes were investigated using univariate and multivariate methods. Spearman’s rank correlation was used and characterized 384 of 1181 clusters (32.5%) as associated



with at least one morpho-physiological trait ($r_s > |0.5|$). Of the 386 metabolites, 36 (~10%) could be annotated by matching mass spectra to several metabolite databases, and are displayed as a heat map following hierarchical clustering with correlations to all measured traits (Figure 7A, Supplementary Table 3).



The data indicate sets of metabolites are associated with sets of morpho-physiological traits. The most notable trend was the positive association of glycosylated flavonoids with photosynthesis-related traits (e.g., photosynthesis, stomatal conductance). Organic acids, including erythronic/threonic acid (two compounds with identical mass spectra), lactic acid, and a pentose sugar acid were negatively correlated with photosynthesis. Further, mass spectra that matched to chlorogenic acid (CGA) were detected at three distinct elution times, indicating the presence of three structural isomers. CGA isomer 1 was correlated with both photosynthesis (positive), biomass (negative), and growth rates (negative for both weeks 1–2 and 3–4).

These data were further integrated using O2PLS, a multivariate regression technique based on orthogonal projection to latent structures (OPLS), is a method to integrate two different multivariate datasets (Bylesjo et al., 2007). Here, O2PLS regressed the 20 morpho-physiological traits against 1181 LC- and GC-MS spectral clusters (Trygg, 2002; **Figure 8**). The O2PLS model resulted in four components that explained 82% of the variation. The O2PLS (**Figure 8**) indicates associations among morpho-physiological traits, and highlights metabolites that co-vary with these trends. For example, a positive correlation exists between total dry biomass, stem biomass, and HI-stems, and these traits were negatively associated with HI-Panicles. This supports that sorghum biomass is largely driven by carbon

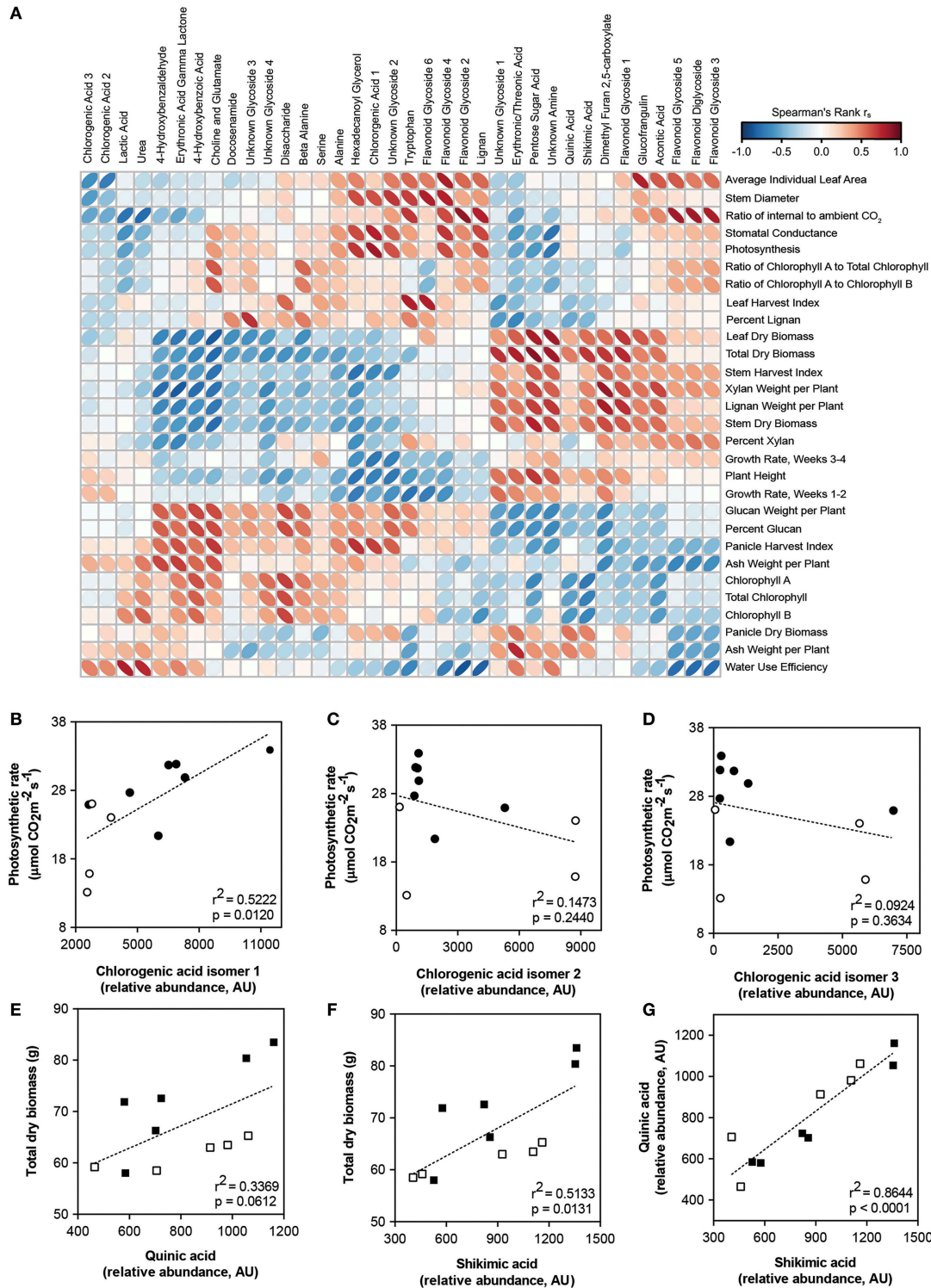
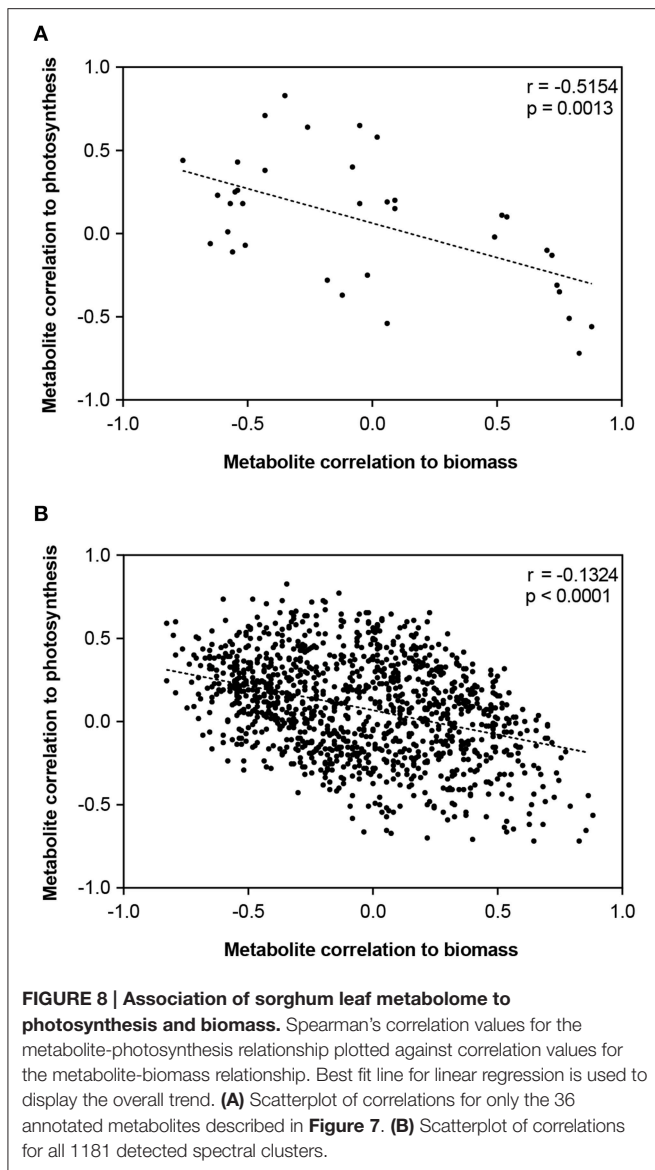


FIGURE 7 | Correlation of metabolites with morphological and physiological traits. (A) Heat map with hierarchical clustering of 36 metabolites associated with morphology or physiology in sorghum. Color and ellipse eccentricity denote Spearman's rank correlation r_s between metabolites and traits. Correlations with an $r > |0.602|$ are significant ($p < 0.05$). Numerical values, including r_s and p -values, are given in Additional file 4. **(B–D)** Scatterplots of mean relative abundance for three chlorogenic acid isomers plotted against mean photosynthetic rate for each sorghum line. **(E,F)** Scatterplots of mean abundance for shikimic and quinic acid plotted against mean total biomass for each sorghum line. **(G)** Scatterplot of mean abundance for shikimic acid plotted against mean abundance for quinic acid for each sorghum line. Closed circles, low biomass sorghum lines; open circles, high biomass sorghum lines; closed squares, low photosynthesis sorghum lines; open squares, high photosynthesis sorghum lines. Dashed lines indicate best fit line for linear regression performed on metabolite-trait relationships.



allocation toward stems (vs. panicles), and that there are metabolites that follow similar trends.

Shikimic and Chlorogenic Acid Are Potential Mediators of Carbon Assimilation, Growth Rate, and Final Dry Biomass Phenotypes in Sorghum

An interesting trend was observed between photosynthesis, growth rates, and final dry biomass. In this study, the lines with the highest photosynthesis (carbon assimilation rates measured at an early time point) had the lowest growth rates in weeks 1–2 and 3–4, as well as the lowest total dry biomass at the end of the experiment. While counterintuitive, several other studies have found negative associations between early photosynthetic measurements and plant growth phenotypes (Wassom et al., 2003; Jahn et al., 2011).

The metabolites that co-varied with these morpho-physiological traits included both primary and secondary metabolites. Of the three CGA isomers, one isomer was negatively correlated to growth rates (1–2 and 3–4 weeks) and total dry biomass (**Figure 7B**); the other two had slight negative correlations (**Figures 7C,D**). Shikimic and quinic acid were positively correlated with biomass (**Figures 7E,F**). In addition, abundances of shikimic and quinic acid within all sorghum lines were highly correlated (**Figure 7G**). Relative abundances of shikimic, quinic, and chlorogenic acid (isomer 1) among lines are presented in **Supplementary Figure 1**. Interestingly, metabolites positively correlated to photosynthesis also tended to be negatively correlated to biomass (**Figure 8A**). When this relationship was explored in the full metabolomics dataset (all 1181 spectral clusters), a significant trend was also observed (**Figure 8B**).

Taken together, these data indicate shikimate and CGA are associated with the relationship between carbon allocation (photosynthesis) and plant growth (early-developmental growth rates and final dry biomass). This is supported by the associations for these metabolites in both univariate and multivariate statistical models.

DISCUSSION

Recent research has shown increasing use of both targeted and non-targeted metabolomic approaches to characterize a molecular level of phenotypic variation in both model species such as *Arabidopsis* (e.g., Sulpice et al., 2010) as well in crops such as tomato (Tikunov et al., 2005) and rice (Kusano et al., 2015). In addition to defining a species metabolome, this characterization of variation is also used as a “top-down” approach to better comprehend the complexities of metabolic pathways (Fernie and Klee, 2011). Further, metabolomics are being used to find biomarkers that are indicative of valuable traits in crops such as potato (Steinfath et al., 2010) and barley (Heuberger et al., 2014a) and to examine changes in metabolite profiles as plant immune responses to biotic influence (e.g., the influence of *Fusarium* on soybean root profiles as in Scandiani et al., 2015) as well as in response to abiotic stresses such as drought, as in oats (Sanchez-Martin et al., 2015) or wheat (Bowne et al., 2012). The global profiles of small molecules and subsets therein are also being used in attempts to predict complex traits in the model, *Arabidopsis* (Meyer et al., 2007), although it is becoming increasingly clear that these profiles are likely to be subject to a great deal of environmental influence (Sulpice et al., 2013).

Because of existing genetic tools, the lines examined in this study represent avenues for further discovery and valuable trait exploitation [e.g., sequenced sorghum line *BTx623* (Paterson et al., 2009), re-sequenced lines *Tx7000* and *BTx642* (Evans et al., 2013), or *RTx430*, which is the only commonly transformed sorghum line in this recalcitrant species (Howe et al., 2006; Liu and Godwin, 2012; Wu et al., 2014)].

This study sought to (1) characterize the metabolome variation in this important set of sorghum breeding lines and (2) explore its relationship to morpho-physiological traits related to

plant growth and biomass accumulation. Indeed, the data showed that in a phenotypically and genetically diverse species such as sorghum, the metabolome follows suit, exhibiting great variation in small molecule profiles among lines (Figure 3). In addition, correlative relationships among morphological, physiological, and structural traits related to biomass, as well as between metabolite traits and biomass were identified (Figures 6, 7A). It is also of note that certain subsets and whole metabolome patterns demonstrated interesting opposing relationships with biomass, growth rates and photosynthesis (Figures 7B–G, 9) as well as relationships with other physiological and morphological traits (Figures 7A, 8).

The relationship between photosynthetic measurements and subsequent carbon allocation to the accumulation of biomass is notoriously complex. Alteration of photosynthesis is often discussed as a potential mode of crop improvement (Murchie et al., 2009; Zhu et al., 2010), however, relationships between physiological traits and biomass are not necessarily positive (Wassom et al., 2003; Jahn et al., 2011), nor straightforward.

It should also be noted there are limitations to measuring photosynthesis in a controlled environment and on a per-leaf area, and these relationships may change at the whole-plant and whole-canopy levels in the field. However, despite this complexity, the results of this study do demonstrate a negative relationship of early growth rates and final biomass to leaf-area based photosynthesis measurements (Figure 6) and are further supported by the relationships of these traits to the leaf metabolome (Figures 7–9). Therefore, our data indicate that the relationship between physiological rates and biomass accumulation may be mediated through both primary and secondary metabolism. Indeed, biomass and photosynthetic rate were among traits with strongest correlations to leaf metabolic profile, highlighting the utility of a metabolomics approach to understand the mechanisms behind biomass regulation.

Specific identified metabolites included Chlorogenic acid (CGA) which was found to be associated with photosynthesis (positive), growth rate (negative), and biomass (negative). CGA

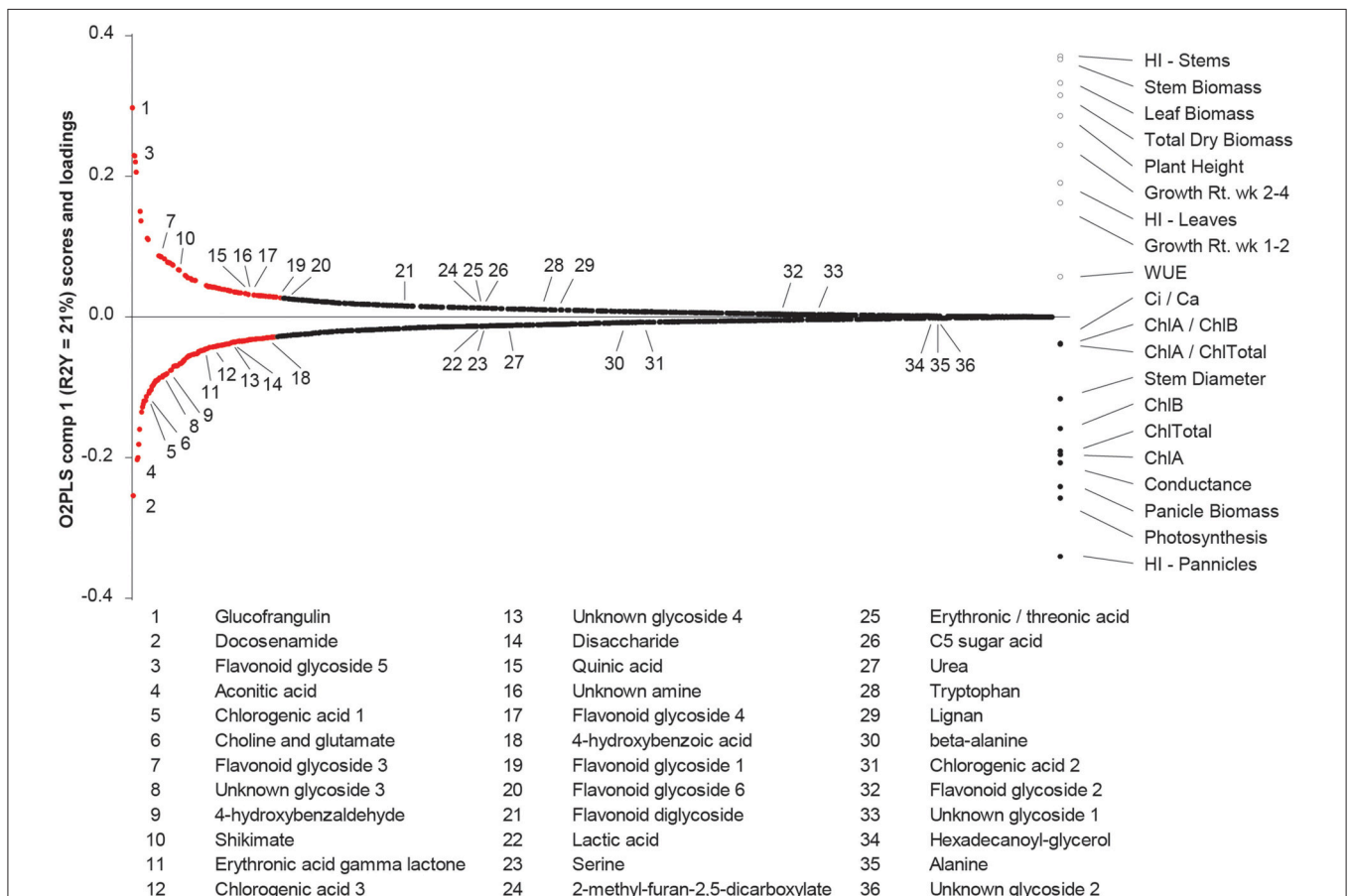


FIGURE 9 | Prediction of physiological and morphological traits based on metabolic variation. O2PLS integrated datasets and overlaid trends in metabolites and traits. The y-axis represents relative contribution of metabolites to the model (loadings), or relative contribution of traits to the model (scores). Higher or lower loadings or scores denote metabolites and traits with largest contributions to the model and loadings in red were considered to significantly contribute to the model (see z-thresholds described in Materials and Methods). The x-axis is an ordered list of the 1181 metabolites from highest (left side of x-axis) to lowest contributors (right side of x-axis). Traits are listed at the right side of the x-axis. The positive or negative direction of metabolites and traits indicate co-variation (e.g., stem and leaf biomass co-varies with metabolites 1, 3, 7, 10, etc., and negatively correlate to photosynthesis). The 36 metabolites correlated to at least one trait are denoted.

is a highly abundant compound and has been demonstrated to have important roles in multiple plant organs including leaves (Sheen, 1973; Mondolot et al., 2006; Clé et al., 2008; Leiss et al., 2009), roots, and root hairs, (Narukawa et al., 2009; Franklin and Dias, 2011), and has roles in diverse processes such as wound response (Ramamurthy et al., 1992; Campos-Vargas and Saltveit, 2002) and cell-wall building (Aerts and Baumann, 1994; Mondolot et al., 2006). In our study, three CGA isomers were identified in leaf extracts. These isomers have been previously described as containing identical MS/MS fragmentation (Xue et al., 2013), but were here separated chromatographically, allowing for independent measurements. One of these three detected isomers had the strongest positive correlation to photosynthesis (**Figure 7B**); the others were weakly negative. As in this work, other studies have also found isomer-specific metabolite relationships. For example, Xue et al. demonstrated that only one of three chlorogenic acid isomers responded to heat stress and subsequent changes to photosynthesis in *Arabidopsis* (Xue et al., 2013).

CGA is a product of the phenylpropanoid pathway (an intermediate in the lignin pathway), and is a fairly well-characterized scavenger of reactive oxygen species (ROS). In young plant leaves, CGA has been found to localize to chloroplasts (Mondolot et al., 2006); evidence for a protective role against light damage. Further, in green pepper, chlorogenic acid rescued photosynthesis from paraquat-induced inhibition (Laskay and Lakos, 2011). Interestingly, phenolics and specifically CGA, are often inhibitors of plant growth (Einhellig and Kuan, 1971; Li et al., 1993). Thus, these data support that sorghum lines exhibiting higher rates of carbon assimilation also have high abundances of particular CGA, isomers subsequently inhibiting early growth rates that affect final quantities of biomass.

Shikimic acid is another compound involved in important leaf and stem processes, as seen in many studies (e.g., Ossipov et al., 2003; Chaves et al., 2011; Dizengremel et al., 2012; Zhang et al., 2013). Of the metabolites identified in this study, shikimic acid also had one of the strongest positive correlations to biomass. Shikimic acid is a precursor for aromatic amino acids leading into the phenylpropanoid pathway, which is responsible for synthesis of lignin and other secondary metabolites and although it may seem intuitive that more lignin is necessary to support increased biomass, the relationship between them is variable (Hu et al., 1999; Chen and Dixon, 2007; Jahn et al., 2011). Interestingly, shikimic acid has also been observed to be an inhibitor of the Phosphoenolpyruvate Carboxylase (PEPC) enzyme (Colombo et al., 1998) which plays a critical role in the fixation of carbon in C4 organisms. While it remains unclear why a reduction in the amount of photosynthetically fixed carbon would sometimes be associated with an increase in biomass, the competitive inhibition of PEPC may also help to explain why larger sorghums (with more shikimic acid) exhibit lower leaf-level rates of photosynthesis.

Quinic acid, also a constituent of the phenylpropanoid pathway, had a similar correlation profile to shikimic acid (**Figure 7G**) and is a precursor to CGA. However, in high biomass

plants, higher quinic acid levels likely reflect lignin and not chlorogenic acid biosynthesis, as evidenced by lower levels of chlorogenic acid in these lines. A study in *Arabidopsis* found shikimic acid and several other phenylpropanoid metabolites to increase in mutants deficient in lignin biosynthetic enzymes, perhaps as an attempt to recover lignin content (Vanholme et al., 2012). In sorghum, it is therefore likely that increased synthesis of shikimic and quinic acid provide the lignin content necessary to support increased biomass.

This study examined the metabolic variation of an important set of sorghum breeding lines via non-targeted GC- and LC-MS analyses and explored the association of these metabolites and profiles with many measured phenotypes (morphological, physiological, and structural carbohydrate content) relevant to biomass and quality. The sorghum panel exhibited high metabolic variation, some of which co-varied with phenotypic traits; in particular, patterns of and specific metabolites that appear to influence the relationship between carbon assimilation, early stage growth, and final biomass. Because of these notable associations, future research should explore the potential utility of early phenotype prediction via the metabolome to circumvent resource-intensive greenhouse and field evaluation of phenotypic traits.

AUTHOR CONTRIBUTIONS

MT, AH, JK, CC, EW, CB, JP, and CJ wrote the manuscript and had primary responsibility for the final content. CC, CJ, JP conceived the studied and generated the experimental design. MT, CJ, EW collected and analyzed morpho-physiological data. AH, JK, CC, CB collected and analyzed the metabolomics data. MT, AH, JK integrated morpho-physiological and metabolomics data. All authors have read and approved the final manuscript.

FUNDING

The work of MT, JK, CC, and CJ was supported by the Colorado State University Agricultural Experiment Station and the Energy Institute at Colorado State University. The work of EW was supported by the U.S. Department of Energy under Contract No. DE-AC36-08GO28308 with the National Renewable Energy Laboratory with funding provided by US DOE Bioenergy Technologies Office (BETO). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Energy Institute at Colorado State University. We would like to acknowledge generous support by the Colorado State University Libraries Open Access Research and Scholarship Fund. The funding bodies did not participate in the design, collection, analysis and interpretation of data; or in preparation of the manuscript.

ACKNOWLEDGMENTS

We would like to thank John Mullet and William Rooney at Texas A&M University, who provided

guidance on the selection of important breeding populations for this study and the seed to initiate this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2016.00953>

REFERENCES

- Aerts, R. J., and Baumann, T. W. (1994). Distribution and utilization of chlorogenic acid in *Coffea* seedlings. *J. Exp. Bot.* 45, 497–503. doi: 10.1093/jxb/45.4.497
- Arnon, D. I. (1949). Copper enzymes in isolated chloroplasts - polyphenoloxidase in beta-vulgaris. *Plant Physiol.* 24, 1–15. doi: 10.1104/pp.24.1.1
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Methodol.* 57, 289–300.
- Bino, R. J., Hall, R. D., Fiehn, O., Kopka, J., Saito, K., Draper, J., et al. (2004). Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* 9, 418–425. doi: 10.1016/j.tplants.2004.07.004
- Bowne, J. B., Erwin, T. A., Juttner, J., Schnurbusch, T., Langridge, P., Bacic, A., et al. (2012). Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Mol. Plant* 5, 418–429. doi: 10.1093/mp/sss114
- Broadhead, D. M. (1972). Registration of Rio Sweet Sorghum1 (Reg. No. 113). *Crop Sci.* 12, 716–716. doi: 10.2135/cropsci1972.0011183X001200050068x
- Broeckling, C. D., Afsar, F. A., Neumann, S., Ben-Hur, A., and Prenni, J. E. (2014). RAMClust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data. *Anal. Chem.* 86, 6812–6817. doi: 10.1021/ac501530d
- Brown, P. J., Klein, P. E., Bortiri, E., Acharya, C. B., Rooney, W. L., and Kresovich, S. (2006). Inheritance of inflorescence architecture in sorghum. *Theor. Appl. Genet.* 113, 931–942. doi: 10.1007/s00122-006-0352-9
- Burrow, G. B., Klein, R. R., Franks, C. D., Klein, P. E., Schertz, K. F., Pederson, G. A., et al. (2011). Registration of the BTx623/IS3620C Recombinant inbred mapping population of Sorghum. *J. Plant Reg.* 5, 141–145. doi: 10.3198/jpr2010.04.0219crmp
- Bylesjo, M., Eriksson, D., Kusano, M., Moritz, T., and Trygg, J. (2007). Data integration in plant biology: the O2PLS method for combined modeling of transcript and metabolite data. *Plant J.* 52, 1181–1191. doi: 10.1111/j.1365-3113.2007.03293.x
- Campos-Vargas, R., and Saltveit, M. E. (2002). Involvement of putative chemical wound signals in the induction of phenolic metabolism in wounded lettuce. *Physiol. Plant.* 114, 73–84. doi: 10.1034/j.1399-3054.2002.114.0111.x
- Chaves, L., Passarinho, J. A., Capitão, C., Chaves, M. M., Fevereiro, P., and Ricardo, C. P. (2011). Temperature stress effects in *Quercus ruber* leaf metabolism. *J. Plant Physiol.* 168, 1729–1734. doi: 10.1016/j.jplph.2011.05.013
- Chen, F., and Dixon, R. A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* 25, 759–761. doi: 10.1038/nbt1316
- Childs, K. L., Cordonnierpratt, M. M., Pratt, L. H., and Morgan, P. W. (1992). Genetic-regulation of development in *Sorghum-bicolor*.7. Ma3r flowering mutant lacks a phytochrome that predominates in green tissue. *Plant Physiol.* 99, 765–770. doi: 10.1104/pp.99.2.765
- Clé, C., Hill, L. M., Niggeweg, R., Martin, C. R., Guisez, Y., Prinsen, E., et al. (2008). Modulation of chlorogenic acid biosynthesis in *Solanum lycopersicum*; consequences for phenolic accumulation and UV-tolerance. *Phytochemistry* 69, 2149–2156. doi: 10.1016/j.phytochem.2008.04.024
- Colombo, S. L., Andreo, C. S., and Chollet, R. (1998). The interaction of shikimic acid and protein phosphorylation with pep carboxylase from the C4 dicot *Amaranthus viridis*. *Phytochemistry* 48, 55–59. doi: 10.1016/S0031-9422(97)01100-X
- Dai, A. (2011). Drought under global warming: a review. *Wiley Interdiscipl. Rev. Clim. Change* 2, 45–65. doi: 10.1002/wcc.81
- Dizengremel, P., Vaultier, M.-N., Le Thiec, D., Cabané, M., Bagard, M., Gérant, D., et al. (2012). Phosphoenolpyruvate is at the crossroads of leaf metabolic responses to ozone stress. *New Phytol.* 195, 512–517. doi: 10.1111/j.1469-8137.2012.04211.x
- Einhellig, F. A., and Kuan, L.-Y. (1971). Effects of scopoletin and chlorogenic acid on stomatal aperture in tobacco and sunflower. *Bull. Torrey Bot. Club*, 98, 155–162. doi: 10.2307/2484031
- Evans, J., McCormick, R. F., Morishige, D., Olson, S. N., Weers, B., Hilley, J., et al. (2013). Extensive variation in the density and distribution of DNA polymorphism in Sorghum genomes. *PLoS ONE* 8:e79192. doi: 10.1371/journal.pone.0079192
- Felderhoff, T. J., Murray, S. C., Klein, P. E., Sharma, A., Hamblin, M. T., Kresovich, S., et al. (2012). QTLs for energy-related traits in a sweet x grain Sorghum *Sorghum bicolor* (L.) moench mapping population. *Crop Sci.* 52, 2040–2049. doi: 10.2135/cropsci2011.11.0618
- Fernie, A., and Klee, H. (2011). The use of natural genetic diversity in the understanding of metabolic organization and regulation. *Front. Plant Sci.* 2:59. doi: 10.3389/fpls.2011.00059
- Fernie, A. R., and Schauer, N. (2009). Metabolomics-assisted breeding: a viable option for crop improvement?. *Trends Genet.* 25, 39–48. doi: 10.1016/j.tig.2008.10.010
- Feussner, I., and Polle, A. (2015). What the transcriptome does not tell—proteomics and metabolomics are closer to the plants' patho-phenotype. *Curr. Opin. Plant Biol.* 26, 26–31. doi: 10.1016/j.pbi.2015.05.023
- Franklin, G., and Dias, A. C. P. (2011). Chlorogenic acid participates in the regulation of shoot, root and root hair development in *Hypericum perforatum*. *Plant Physiol. Biochem.* 49, 835–842. doi: 10.1016/j.plaphy.2011.05.009
- Ganie, A. H., Ahmad, A., Pandey, R., Aref, I. M., Yousuf, P. Y., Ahmad, S., et al. (2015). Metabolite profiling of low-P tolerant and low-P sensitive maize genotypes under phosphorus starvation and restoration conditions. *PLoS ONE* 10:e0129520. doi: 10.1371/journal.pone.0129520
- Hart, G. E., Schertz, K. F., Peng, Y., and Syed, N. H. (2001). Genetic mapping of *Sorghum bicolor* (L.) Moench QTLs that control variation in tillering and other morphological characters. *Theor. Appl. Genet.* 103, 1232–1242. doi: 10.1007/s001220100582
- Heuberger, A. L., Broeckling, C. D., Kirkpatrick, K. R., and Prenni, J. E. (2014a). Application of nontargeted metabolite profiling to discover novel markers of quality traits in an advanced population of malting barley. *Plant Biotechnol. J.* 12, 147–160. doi: 10.1111/pbi.12122
- Heuberger, A. L., Robison, F. M., Lyons, S. M. A., Broeckling, C. D., and Prenni, J. E. (2014b). Evaluating plant immunity using mass spectrometry-based metabolomics workflows. *Front. Plant Sci.* 5:291. doi: 10.3389/fpls.2014.00291
- Howe, A., Sato, S., Dweikat, I., Fromm, M., and Clemente, T. (2006). Rapid and reproducible *Agrobacterium*-mediated transformation of sorghum. *Plant Cell Rep.* 25, 784–791. doi: 10.1007/s00299-005-0081-6
- Hu, W. J., Harding, S. A., Lung, J., Popko, J. L., Ralph, J., Stokke, D. D., et al. (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nat. Biotechnol.* 17, 808–812. doi: 10.1038/11758

- Jahn, C. E., Mckay, J. K., Mauleon, R., Stephens, J., McNally, K. L., Bush, D. R., et al. (2011). Genetic variation in biomass traits among 20 diverse rice varieties. *Plant Physiol.* 155, 157–168. doi: 10.1104/pp.110.165654
- Kebede, H., Subudhi, P. K., Rosenow, D. T., and Nguyen, H. T. (2001). Quantitative trait loci influencing drought tolerance in grain sorghum (*Sorghum bicolor* L. Moench). *Theor. Appl. Genet.* 103, 266–276. doi: 10.1007/s001220100541
- Kusano, M., Yang, Z., Okazaki, Y., Nakabayashi, R., Fukushima, A., and Saito, K. (2015). Using metabolomic approaches to explore chemical diversity in rice. *Mol. Plant* 8, 58–67. doi: 10.1016/j.molp.2014.11.010
- Laskay, G., and Lakos, A. (2011). Counteracting the effects of paraquat on photosynthesis by chlorogenic acid. *Acta Biol. Szegediensis* 55, 101–103.
- Leiss, K. A., Maltese, F., Choi, Y. H., Verpoorte, R., and Klinkhamer, P. G. L. (2009). Identification of chlorogenic acid as a resistance factor for thrips in *Chrysanthemum*. *Plant Physiol.* 150, 1567–1575. doi: 10.1104/pp.109.138131
- Li, H.-H., Inoue, M., Nishimura, H., Mizutani, J., and Tsuzuki, E. (1993). Interactions of trans-cinnamic acid, its related phenolic allelochemicals, and abscisic acid in seedling growth and seed germination of lettuce. *J. Chem. Ecol.* 19, 1775–1787.
- Little, C. R., and Magill, C. W. (2009). The grain mold pathogen, *Fusarium thapsinum*, reduces caryopsis formation in *Sorghum bicolor*. *J. Phytopathol.* 157, 518–519. doi: 10.1111/j.1439-0434.2008.01530.x
- Liu, G. Q., and Godwin, I. D. (2012). Highly efficient sorghum transformation. *Plant Cell Rep.* 31, 999–1007. doi: 10.1007/s00299-011-1218-4
- Mace, E. S., Tai, S. S., Gilding, E. K., Li, Y. H., Prentis, P. J., Bian, L. L., et al. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 4:2320. doi: 10.1038/ncomms3320
- MacKinnon, C., Gunderson, G., and Nabors, M. W. (1987). High-efficiency plant-regeneration by somatic embryogenesis from callus of mature embryo explants of bread wheat (*Triticum aestivum*) and grain-sorghum (*Sorghum-bicolor*). *In Vitro* 23, 443–448. doi: 10.1007/bf02623861
- Meyer, R., Steinfath, M., Lisec, J., Becher, M., Witucka Wall, H. O., et al. and Altmann, T. (2007). The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4759–4764. doi: 10.1073/pnas.0609709104
- Miller, F. R. (1984). Registration of RTx430 sorghum parental line. *Crop Sci.* 24, 1224–1224. doi: 10.2135/cropsci1984.0011183X002400060074x
- Mondolot, L., La Fisca, P., Buatois, B., Talansier, E., De Kochko, A., and Campa, C. (2006). Evolution in caffeoylquinic acid content and histolocalization during *Coffea canephora* leaf development. *Ann. Bot.* 98, 33–40. doi: 10.1093/aob/mcl080
- Mullet, J., Morishige, D., McCormick, R., Truong, S., Hilley, J., Mckinley, B., et al. (2014). Energy Sorghum—a genetic model for the design of C4 grass bioenergy crops. *J. Exp. Bot.* 65, 3479–3489. doi: 10.1093/jxb/eru229
- Murchie, E. H., Pinto, M., and Horton, P. (2009). Agriculture and the new challenges for photosynthesis research. *New Phytol.* 181, 532–552. doi: 10.1111/j.1469-8137.2008.02705.x
- Murphy, R. L., Klein, R. R., Morishige, D. T., Brady, J. A., Rooney, W. L., Miller, F. R., et al. (2011). Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16469–16474. doi: 10.1073/pnas.1106212108
- Murray, S. C., Rooney, W. L., Mitchell, S. E., Sharma, A., Klein, P. E., Mullet, J. E., et al. (2008a). Genetic improvement of Sorghum as a biofuel feedstock: II. QTL for stem and leaf structural carbohydrates. *Crop Sci.* 48, 2180–2193. doi: 10.2135/cropsci2008.01.0068
- Murray, S. C., Sharma, A., Rooney, W. L., Klein, P. E., Mullet, J. E., Mitchell, S. E., et al. (2008b). Genetic improvement of Sorghum as a biofuel feedstock: I. QTL for stem sugar and grain nonstructural carbohydrates. *Crop Sci.* 48, 2165–2179. doi: 10.2135/cropsci2008.01.0016
- Narukawa, M., Kanbara, K., Tominaga, Y., Aitani, Y., Fukuda, K., Kodama, T., et al. (2009). Chlorogenic acid facilitates root hair formation in lettuce seedlings. *Plant Cell Physiol.* 50, 504–514. doi: 10.1093/pcp/pcp010
- Ossipov, V., Salminen, J.-P., Ossipova, S., Haukioja, E., and Pihlaja, K. (2003). Gallic acid and hydrolysable tannins are formed in birch leaves from an intermediate compound of the shikimate pathway. *Biochem. Syst. Ecol.* 31, 3–16. doi: 10.1016/S0305-1978(02)00081-9
- Paterson, A., Bowers, J., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., et al. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457, 551–556. doi: 10.1038/nature07723
- Payne, C., and Wolfrum, E. (2015). Rapid analysis of composition and reactivity in cellulosic biomass feedstocks with near-infrared spectroscopy. *Biotechnol. Biofuels* 8, 43. doi: 10.1186/s13068-015-0222-2
- Pedersen, J. F., Funnell, D. L., Toy, J. J., Oliver, A. L., and Grant, R. J. (2006). Registration of seven forage sorghum genetic stocks near-isogenic for the brown Midrib genes bmr-6 and bmr-12 joint contribution of the USDA, ARS and the University of Nebraska Agric. Exp. Stn. as Paper no. 14578, Journal Series, Nebraska Agric. Exp. Stn. Registration by CSSA. *Crop Sci.* 46, 490–491. doi: 10.2135/cropsci2005.06-0095
- R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org/>
- Ramamurthy, M. S., Maiti, B., Thomas, P., and Nair, P. M. (1992). High-performance liquid chromatography determination of phenolic acids in potato tubers (*Solanum tuberosum*) during wound healing. *J. Agric. Food Chem.* 40, 569–572. doi: 10.1021/jf00016a007
- Reddy, R. N., Madhusudhana, R., Mohan, S. M., Chakravarthi, D. V. N., and Seetharama, N. (2012). Characterization, development and mapping of Unigene-derived microsatellite markers in sorghum *Sorghum bicolor* (L.) Moench. *Mol. Breed.* 29, 543–564. doi: 10.1007/s11032-011-9571-0
- Riedelsheimer, C., Cziedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., et al. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat. Genet.* 44, 217–220. doi: 10.1038/ng.1033
- Rooney, W. L., Blumenthal, J., Bean, B., and Mullet, J. E. (2007). Designing sorghum as a dedicated bioenergy feedstock. *Biofuels Bioprod. Biorefining Bioppr* 1, 147–157. doi: 10.1002/bbb.15
- Sanchez-Martin, J., Heald, J., Kingston-Smith, A., Winters, A., Rubiales, D., Sanz, M., et al. (2015). A metabolomic study in oats (*Avena sativa*) highlights a drought tolerance mechanism based upon salicylate signalling pathways and the modulation of carbon, antioxidant and photo-oxidative metabolism. *Plant Cell Environ.* 38, 1434–1452. doi: 10.1111/pce.12501
- Scandiani, M. M., Luque, A. G., Razori, M. V., Casalini, L. C., Aoki, T., O'donnell, K., et al. (2015). Metabolic profiles of soybean roots during early stages of *Fusarium tucumaniae* infection. *J. Exp. Bot.* 66, 391–402. doi: 10.1093/jxb/eru432
- Sheen, S. J. (1973). Correlation between chlorophyll and chlorogenic acid content in tobacco-leaves. *Plant Physiol.* 52, 422–426. doi: 10.1104/pp.52.5.422
- Sluiter, J. B., Ruiz, R. O., Scarlata, C. J., Sluiter, A. D., and Templeton, D. W. (2010). Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *J. Agri. Food Chem.* 58, 9043–9053. doi: 10.1021/jf1008023
- Smith, C. A., Want, E. J., 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
- Sorrells, M. E., and Myers, O. (1982). Duration of developmental stages of 10 milo maturity genotypes. *Crop Sci.* 22, 310–314. doi: 10.2135/cropsci1982.0011183X002200020026x
- Steinfath, M., Strehmel, N., Peters, R., Schauer, N., Groth, D., Hummel, J., et al. (2010). Discovering plant metabolic biomarkers for phenotype prediction using an untargeted approach. *Plant Biotechnol. J.* 8, 900–911. doi: 10.1111/j.1467-7652.2010.00516.x
- Subudhi, P. K., Rosenow, D. T., and Nguyen, H. T. (2000). Quantitative trait loci for the stay green trait in sorghum (*Sorghum bicolor* L. Moench): consistency across genetic backgrounds and environments. *Theor. Appl. Genet.* 101, 733–741. doi: 10.1007/s001220051538
- Sulpice, R., Nikoloski, Z., Tschoep, H., Antonio, C., Kleessen, S., Larhlmi, A., et al. (2013). Impact of the carbon and nitrogen supply on relationships and connectivity between metabolism and biomass in a broad panel of *Arabidopsis* Accessions. *Plant Physiol.* 162, 347–363. doi: 10.1104/pp.112.210104
- Sulpice, R., Trenkamp, S., Steinfath, M., Usadel, B., Gibon, Y., Witucka-Wall, H., et al. (2010). Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of *Arabidopsis* accessions. *Plant Cell* 22, 2872–2893. doi: 10.1105/tpc.110.076653

- Taylor, S. H., Ripley, B. S., Martin, T., De-Wet, L.-A., Woodward, F. I., and Osborne, C. P. (2014). Physiological advantages of C4 grasses in the field: a comparative experiment demonstrating the importance of drought. *Glob. Chang. Biol.* 20, 1992–2003. doi: 10.1111/gcb.12498
- Templeton, D. W., Scarlata, C. J., Sluiter, J. B., and Wolfrum, E. J. (2010). Compositional analysis of lignocellulosic feedstocks. 2. Method uncertainties. *J. Agri. Food Chem.* 58, 9054–9062. doi: 10.1021/jf100807b
- Tikunov, Y., Lommen, A., De Vos, C. H. R., Verhoeven, H. A., Bino, R. J., Hall, R. D., et al. (2005). A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. *Plant Physiol.* 139, 1125–1137. doi: 10.1104/pp.105.068130
- Trygg, J. (2002). O2-PLS for qualitative and quantitative analysis in multivariate calibration. *J. Chemom.* 16, 283–293. doi: 10.1002/cem.724
- Tsogtbaatar, E., Cocuron, J. C., Sonera, M. C., and Alonso, A. P. (2015). Metabolite fingerprinting of pennycress (*Thlaspi arvense* L.) embryos to assess active pathways during oil synthesis. *J. Exp. Bot.* 66, 4267–4277. doi: 10.1093/jxb/erv020
- 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
- Wassom, J. J., Knepp, A. W., Tranel, P. J., and Wax, L. M. (2003). Variability in photosynthetic rates and accumulated biomass among greenhouse-grown common cocklebur (*Xanthium strumarium*) accessions. *Weed Technol.* 17, 84–88. doi: 10.1614/0890-037X(2003)017[0084:VIPRAA]2.0.CO;2
- Wei, T. (2010). *Corrplot: Visualization of a Correlation Matrix. R Package Version 0.2-0*. Available online at: <https://cran.r-project.org/web/packages/corrplot/index.html>
- Wolfrum, E., Payne, C., Stefaniak, T., Dighe, N., Bean, B., and Dahlberg, J. (2013). *Multivariate Calibration Models for Sorghum Composition using Near-Infrared Spectroscopy. National Renewable Energy Lab Technical Report NREL/TP-5100-56838*.
- Wu, E., Lenderts, B., Glassman, K., Berezowska-Kaniewska, M., Christensen, H., Asmus, T., et al. (2014). Optimized Agrobacterium-mediated sorghum transformation protocol and molecular data of transgenic sorghum plants. *In Vitro Cell. Dev. Biol. Plant* 50, 9–18. doi: 10.1007/s11627-013-9583-z
- Xue, L. J., Guo, W. B., Yuan, Y. A., Anino, E. O., Nyamdari, B., Wilson, M. C., et al. (2013). Constitutively elevated salicylic acid levels alter photosynthesis and oxidative state but not growth in transgenic *Populus*. *Plant Cell* 25, 2714–2730. doi: 10.1105/tpc.113.112839
- Zhang, L., and Garcia-Munoz, S. (2009). A comparison of different methods to estimate prediction uncertainty using Partial Least Squares (PLS): a practitioner's perspective. *Chemometrics Intell. Lab. Syst.* 97, 152–158. doi: 10.1016/j.chemolab.2009.03.007
- Zhang, X., Rerksiri, W., Liu, A., Zhou, X., Xiong, H., Xiang, J., et al. (2013). Transcriptome profile reveals heat response mechanism at molecular and metabolic levels in rice flag leaf. *Gene* 530, 185–192. doi: 10.1016/j.gene.2013.08.048
- Zhu, X. G., Long, S. P., and Ort, D. R. (2010). Improving photosynthetic efficiency for greater yield. *Annu. Rev. Plant Biol.* 61, 235–261. doi: 10.1146/annurev-arplant-042809-112206

Conflict of Interest Statement: 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.

Copyright © 2016 Turner, Heuberger, Kirkwood, Collins, Wolfrum, Broeckling, Prenni and Jahn. 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) or licensor 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.