THE CROP JOURNAL 9 (2021) 181–191

Available online at www.sciencedirect.com

ScienceDirect

Kernel metabolites depict the diversity of relationship between maize hybrids and their parental lines

Yang Xu[a], *Ying Ma*[a], *Xin Wang*[a], *Cheng Li*[a], *Xuecai Zhang*[b], *Pengcheng Li*[a], *Zefeng Yang*[a], *Chenwu Xu*[a,*]

[a]*Key Laboratory of Plant Functional Genomics of the Ministry of Education, Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops, College of Agriculture, Yangzhou University, Yangzhou 225009, Jiangsu, China*

[b]*International Maize and Wheat Improvement Center (CIMMYT), 06600 Mexico, D.F., Mexico*

ARTICLE INFO

Article history:
Received 3 January 2020
Received in revised form 25 April 2020
Accepted 28 May 2020
Available online 24 July 2020

Keywords:
Maize
Kernel
Metabolite
Reciprocals
Heterosis

ABSTRACT

As the end products of cellular regulatory processes, metabolites provide the link between genotypes and phenotypes. Although metabolites have been widely applied for functional gene detection and phenotype prediction in maize, there is little research focusing on the genetic information of metabolites per se. Here, we performed genetic analyses for the kernel metabolites of 11 parental inbred lines of six representative maize varieties, including Zhongdan 2, Danyu 13, Yedan 13, Zhengdan 958, Xianyu 355, and Suyu 16, as well as their 26 reciprocal hybrids. We identified a total of 208 metabolites in maize kernels using untargeted metabolite profiling technology. Both cluster analysis and principal component analysis indicated that kernel metabolites could distinguish hybrids from their parents. Analysis of variance further revealed that 163 metabolites exhibited significant differences between parents and hybrids, and 40 metabolites showed significant differences between reciprocal crosses. We also investigated the genetic effects and heterosis for each metabolite. By taking all hybrids into consideration, about two-thirds of all metabolites displayed overdominant with 36.8% and 31% of them displaying positive overdominant and negative overdominant, respectively. Besides, 27.5% and 20.4% of all hybrid combinations showed significant mid-parent heterosis and over-parent heterosis, respectively. Our findings revealed that kernel metabolites exhibited the diversity of relationship between maize hybrids and their parental lines. Additionally, we identified 25 significant metabolic markers related to 11 agronomic traits using the LASSO method. Seven metabolic markers were associated with more than one trait simultaneously. These results provide a genetic basis for further utilization of metabolites in the genetic improvement of maize.

* *Corresponding author.*
E-mail address: cwxu@yzu.edu.cn. (C. Xu).

Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

https://doi.org/10.1016/j.cj.2020.05.009

1. Introduction

Maize is the world's most important crop for food, feed and industrial materials. Hybrid breeding is favorable for increasing maize yield by taking advantage of heterosis, which results in about 35% increase in yield. Since the 1950s, maize varieties in China have been renewed six times. Breeders have selected and bred a large number of superior inbred lines, such as Zi330, Ye478, Dan340, Chang7-2, and Mo17, and then used these inbred lines to successively produce several elite varieties, including Zhongdan 2, Danyu 13, Yedan 13, Zhengdan 958, and so on. These maize varieties greatly contributed to maize production and global food security.

In recent years, the rapid development of next generation sequencing techniques coupled with quantitative trait loci (QTL) mapping or genome-wide association study (GWAS) enable dissection of the genetic architecture underlying complex quantitative traits in maize, such as oil biosynthesis and leaf architecture [1–5]. Despite the progress of genetic mapping of quantitative traits at the genomic level, the gene sequence is limited in capturing effects of gene interactions and downstream regulation [6]. As the final products in the cellular regulatory processes, the metabolome might be expected to integrate all previous genetic processes and gene interactions as well as reflect the physiological status [7]. Analyzing metabolome is thus of fundamental importance for a deeper understanding of the genetic regulatory mechanism of phenotypic traits.

There are more than 200,000 metabolites in the plant kingdom, including the necessary primary metabolites for the maintenance of plant life activities as well as growth and development, and secondary metabolites for protecting plants against pathogens and herbivores as well as abiotic stresses [8,9]. The rapid development of high-throughput metabolite profiling technologies enables accurate quantification of a fraction of metabolites [10,11]. Previous studies suggested that metabolites are heritable and controlled by multiple loci and can thus be regarded as quantitative traits to detect novel genes [12]. Recently, metabolome genome-wide association study (mGWAS) has received increasing attention. In plants, mGWAS is conducted from model specie *Arabidopsis* to many major crops, such as rice and maize [13–15]. Several studies indicate that mGWAS is a powerful complementary tool to traditional GWAS, and even able to directly identify the corresponding functional genes [16,17]. Wen et al. [18] identified 1459 SNPs significantly associated with 983 metabolites in maize kernel using mGWAS and validated two genes by mutant and transgenic analysis. Chen et al. [12] detected a total of 2947 lead SNPs for 840 metabolites in rice leaves via mGWAS and functionally annotated five candidate genes.

Metabolites can not only be treated as quantitative traits but also as biomarkers to predict crop phenotypes [6]. Meyer et al. [19] first used 181 metabolites to predict biomass in *Arabidopsis* based on a recombinant inbred line (RIL) population and found that the correlation coefficient between the predicted and real biomass reached 0.58. Gärtner et al. [20] proposed that combining genetic markers and metabolites could increase the prediction accuracy of biomass heterosis. Riedelsheimer et al. [21] crossed 285 inbred lines with two testers and predicted combining abilities for seven biomass traits, and demonstrated that prediction accuracies ranged from 0.60 to 0.80 by using 130 metabolites. Xu et al. [22,23] predicted the yield of 21,945 potential hybrids derived from 210 RILs using 1000 metabolites and proposed that selection of top ten crosses might bring about an approximately 30% increase in yield. Dan et al. [24] reported high prediction accuracies of metabolomic prediction for three agronomic traits in rice using a complete diallel cross population derived from 18 inbred lines.

Given the above, numerous studies of metabolites have been carried out for gene detection or phenotype prediction. However, there remain limited studies focusing on the genetic information of metabolites per se. In this study, we identified 208 kernel metabolites from 11 parental inbred lines of six representative maize varieties (Zhongdan 2, Danyu 13, Yedan 13, Zhengdan 958, Xianyu 355, and Suyu 16) and their 26 reciprocal crosses. The objectives of this study were to (1) test whether metabolite levels display significant differences between parents and hybrids as well as between reciprocal crosses, (2) evaluate the genetic effect of each metabolite in each hybrid, (3) calculate the mid-parent heterosis and over-parent heterosis for each metabolite, and (4) identify metabolic markers associated with agronomic traits.

2. Materials and methods

2.1. Plant materials and field trials

In this study, the parents of six popular maize varieties (Zhongdan 2, Danyu 13, Yedan 13, Zhengdan 958, Xianyu 355, and Suyu 16) were selected as parental inbred lines, which are E28 (P1), Zheng58 (P2), Chang7-2 (P3), Mo17 (P4), Zi330 (P5), PH6CW (P6), PH4CV (P7), Ye478 (P8), Dan340 (P9), JB (P10) and Y53 (P11).

According to the mating design of P1 × P2, P1 × P3, P4 × P1 (Danyu 13), P1 × P5, P2 × P3 (Zhengdan 958), P2 × P4, P2 × P5, P3 × P4, P3 × P5, P4 × P5 (Zhongdan 2), P6 × P7 (Xianyu 335), P8 × P9 (Yedan 3), and P10 × P11 (Suyu 16), 11 parental lines were crossed to produce 13 F_1 hybrid seeds and 13 reciprocal seeds in Sanya, Hainan province, China, in 2016. The 26 hybrids include six popular maize varieties and their reciprocal hybrids as well as hybrids derived from the complete diallel cross among five foundation parents including E28, Chang7-2, Zheng58, Mo17, and Zi330. All 37 maize materials including 11 parents and 26 hybrids were planted with a randomized block design of four repetitions in the experimental field of Yangzhou University, Jiangsu province, China, in 2017. For each line of each replication, 13 plants were grown in a row of 3.0 m, with 0.5 m space between rows.

For each line, ears from six randomly selected plants were harvested at the consistent maturity for phenotypic data acquisition. The 11 agronomic traits including ear weight (EW), ear length (EL), ear diameter (ED), ear row number (ERN), kernel number per row (KNR), cob weight (CW), cob diameter (CD), kernel length (KL), kernel width (KW), kernel thickness (KT), and 100-kernel weight (HKW) were measured. The average phenotypic values were used for data analysis.

2.2. Sample preparation and metabolite profiling

Kernel samples of maize material were collected 30 days after pollination (DAP). For each replication, about 20 well growth kernels were randomly selected from five plants of each material. After collection, the samples were rapidly frozen in liquid nitrogen and placed in an ultra-low temperature refrigerator at –80 °C. Then, about 50 mg of samples were applied to extraction procedure, and extracted with 800 μL of methanol (with internal standard 2.8 mg mL^{-1}, DL-o-Chlorophenylalanine). Afterwards, the samples were ground to fine powder using Grinding Mill at 65 HZ for 120 s, and were ultra sonicated at 40 kHz on ice bath for 30 min, and then centrifuged at 12,000 r min^{-1} and 4 °C for 15 min. Finally, 200 μL of supernatant was transferred to vial for liquid chromatography-mass spectrometry analysis (LC-MS) analysis using an ACQUITY UPLC system. The raw metabolite data from the LC-MS analysis was transformed to CDF files by Masslynx 4.1 software and input into XCMS software for peak picking, peak alignment, peak filtering and peak filling. The data include retention time, observations, peak area and mass-to-charge ratio, etc. Then, the peak areas of metabolites were normalized via dividing them by the sum of all peak areas in the corresponding sample for metabolite data analysis.

2.3. Statistical analysis

We performed principal component analysis (PCA), hierarchical cluster analysis, analysis of variance (ANOVA) for metabolite data. The metabolite data were further standardized to zero mean and unit variance for PCA and cluster analysis. PCA was performed using R function prcomp, hierarchical cluster analysis was performed using the R function hclust, and ANOVA was implemented using our own R program. The heatmap was plotted using R package called R/pheatmap.

2.4. Heterosis analysis and genetic effect of each metabolite

Mid-parent heterosis (%) was calculated as $(F_1 - MP)/MP \times 100$, positive over-parent heterosis (%) was calculated as $(F_1 - HP)/HP \times 100$, and negative over-parent heterosis (%) was calculated as $(F_1 - LP)/LP \times 100$, where F_1 represents the level of each metabolite in each hybrid, MP represents the mid-parent value in metabolite level of corresponding parental lines, HP represents the high-parent value of the corresponding parent, and LP represents the low-parent value of the corresponding parent. The *t*-test was applied to test the significance of heterosis.

According to d/a ratio, we classified metabolic effects into seven categories including negative overdominance ($d/a \leq -1.2$), negative dominance ($-1.2 < d/a \leq -0.8$), negative partially dominance ($-0.8 < d/a \leq -0.2$), additive ($-0.2 < d/a < 0.2$), positive partially dominance ($0.2 \leq d/a < 0.8$), positive dominance ($0.8 \leq d/a < 1.2$), and positive overdominance ($d/a \geq 1.2$). Here, a is defined as 1/2(HP – LP), and d is defined as $F_1 - MP$.

2.5. Detection of metabolic markers

We combined least absolute shrinkage and selection operator (LASSO) and linear regression method to identify metabolic markers significantly associated with 11 agronomic traits. LASSO is a popular method in variable selection, but it does not have a default strategy to calculate the *P*-values for markers [25]. After variable selection, the number of the non-zero effects is always smaller than the sample size. Thus, the ordinary least squares can be used to calculate standard errors for estimated effects. In this study, we used R package called R/glmnet to implement LASSO for variable selection, and then used R function called lm to calculate the *P*-values of selected metabolic markers. The significant metabolic markers were determined by the *P*-values less than 0.05.

3. Results

3.1. Annotation of metabolites

Using untargeted LC-MS metabolite profiling technology, a total of 208 metabolites were identified and annotated from maize kernels of all lines (Table S1). The 208 metabolites contained primary metabolites and secondary metabolites, which could be classified into ten categories, including amino acids, lipids, hormones, and so on (Fig. 1). Among these metabolites, the amino acids were the most, accounting for 28.4%, followed by the lipids with 17.3% and the phytohormones with 11.5%, and the terpenoids were the least.

3.2. Hierarchical cluster analysis and PCA

We performed hierarchical cluster analysis and PCA for 208 metabolites. As shown in Fig. 2, most metabolite levels exhibited variations in different materials. The 37 maize lines were classified into two clusters, which separated the hybrids from their parental lines except the cross P1 × P5, suggesting that there exist large differences between parents and corresponding hybrids. Also, 208 metabolites were classified into three clusters, and ten categories of metabolites were all observed in each cluster. The first cluster presented obvious variation for hybrids, while the second and the third clusters presented diversity for parental lines. Additionally, the enrichment of some metabolites appeared to be line specific. For example, β-carotene (m74) was only abundant in P11, histamine (m30) was only abundant in P4, and gibberellin A3 (m160) was only abundant in P6 × P7. Most metabolites of the third cluster displayed high levels in P11.

The PCA result indicated that PC1 could sufficiently distinguish parents from hybrids, which was consistent with the result of cluster analysis (Fig. 3). We further selected ten metabolites with the greatest contribution to PC1 and found that eight metabolites of them belonged to amino acids, all of which fell into the first cluster from the cluster analysis.

3.3. Analysis of variance for metabolite data

To further understand the genetic characteristics of kernel metabolites, we performed analysis of variance for each metabolite from 37 maize materials and the results are presented in Table S2. Totally, 163 metabolites showed significant differences between parents and hybrids, 144 metabolites displayed significant differences among parents, and 109 metabolites displayed significant differences among hybrids. Between reciprocal

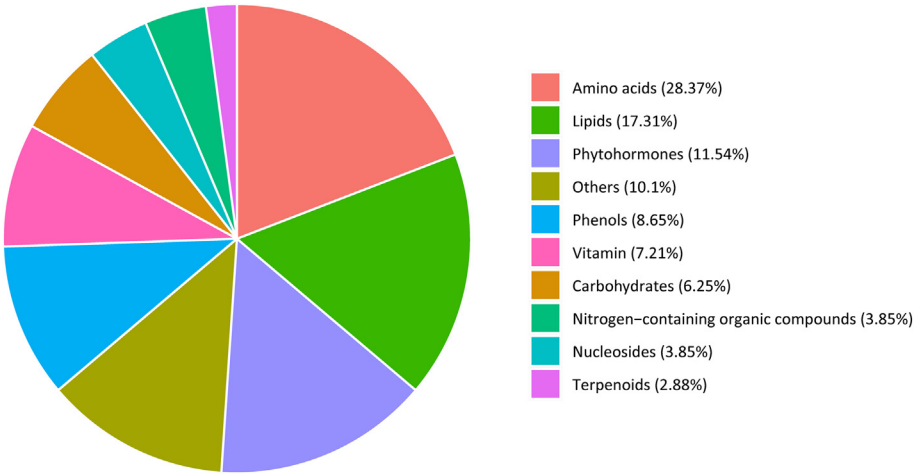

Fig. 1 – Ten categories of metabolites identified from maize kernels. The identified 208 metabolites were divided into ten categories, which are amino acids and their derivatives, lipids, plant hormones, phenols, vitamins, carbohydrates, nitrogen-containing organic compounds, nucleotides, terpenes, and else.

crosses, 40 metabolites showed significant differences ($P < 0.05$), and 16 metabolites had highly significant differences ($P < 0.01$). Five of the 16 metabolites belonged to amino acids, and four of them were hormones. To test the differences between reciprocal crosses for each combination, the degree freedom between reciprocal crosses were further decomposed into 13 single degree (Table 1). In the 16 metabolites, nine metabolites exhibited highly significant differences between reciprocal cross combinations of P1 × P5 and P5 × P1; five metabolites displayed significant differences between Zhengdan 958 (P2 × P3) and its reciprocal hybrid, including l-histidinol (m10), D-ribose (m85), pentadecanoic acid (m126), gibberellin A53 (m162), and p-coumaroylquinic acid (m197); four metabolites showed significant differences between Xianyu 355 (P6 × P7) and its reciprocal cross, which were l-histidinol (m10), L-glutamine (m110), gibberellin A12 (m152), and gibberellin A3 (m160), respectively. Pentadecanoic acid (m126) and p-coumaroylquinic acid (m197) showed significant differences in five reciprocal combinations, followed by and L-Glutamine (m110) and gibberellin A53 (m162), displaying significant differences in four reciprocal combinations.

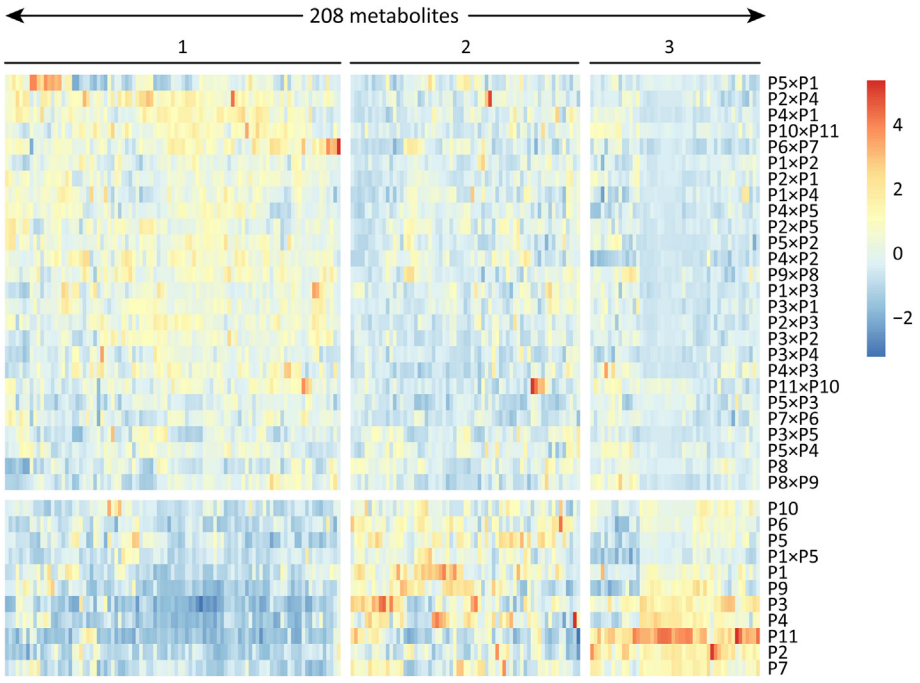

Fig. 2 – Cluster analysis of 208 metabolites from all 37 maize lines. Each column represents a metabolite, and each row represents a maize line. The 208 metabolites were clustered into three clusters, and the 37 maize lines were divided into two clusters. The red color indicates high metabolite levels and blue color indicates low metabolite levels.

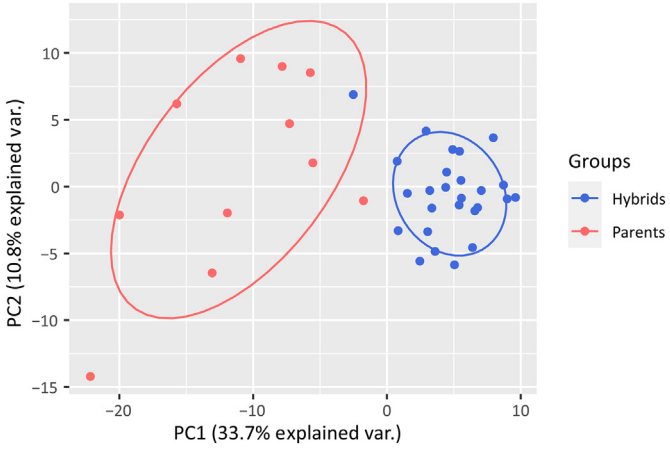

Fig. 3 – Principal component analysis (PCA) of the metabolites from all maize lines. The blue block represents hybrids, and the red block represents parents.

3.4. *Relationship between parents and hybrids in metabolite levels*

In order to know the relationship between parents and their hybrids in metabolite levels, we further investigated the genetic effects of each metabolite (Table S3). Seven types of genetic effects for all metabolites, including negative overdominance (1), negative dominance (2), negative partially dominance (3), additive (4), positive partially dominance (5), positive dominance (6), and positive overdominance (7) are illustrated in Fig. 4. Cluster analysis of genetic effects of metabolites showed that most metabolites of the first type presented positive overdominance, while the third type of metabolites mostly presented negative overdominance. By considering all hybrids, about two-thirds of all metabolites (67.8%) displayed overdominance with 36.8% exhibiting positive overdominance and 31.0% exhibiting negative overdominance. All seven metabolic effects were observed, and the positive dominance only accounted for 3.6%. Genetic effects of certain metabolites seemed to be lines specific. For example, ethanol (m33), homovanillic acid (m46), and syringic acid (m54) exhibited positive overdominance in all 26 hybrids, while apigenin (m66), malvidin (m69), and phytol (m90) exhibited negative overdominance in all 26 hybrids. The metabolites a-L-rhamnose (m80)

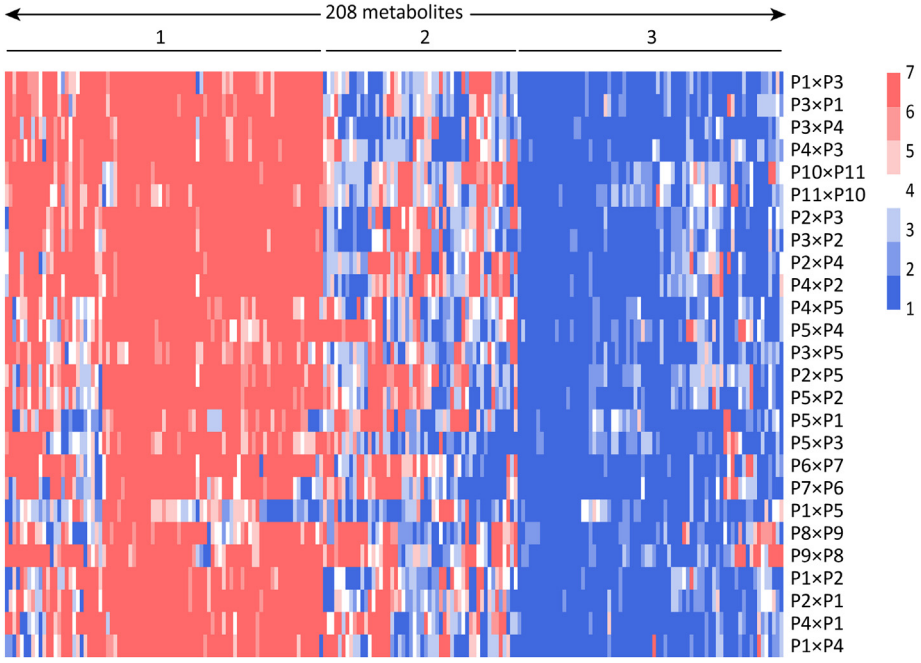

Fig. 4 – Genetic effects for metabolites in each hybrid. Each column represents a metabolite, and each row represents a hybrid. Each color represents a kind of effect, and 1–7 represent negative overdominance, negative dominance, negative partially dominance, additive, positive partially dominance, positive dominance, and positive overdominance, respectively.

and pyruvate (m200) exhibited negative overdominance in P1 × P5, while positive overdominance in all other hybrids. For individual hybrid, large differences in genetic effects could be observed in different metabolites. Taking Zhengdan 958 (P2 × P3) as an example, 90 and 64 metabolites presented positive and negative overdominance, and only six metabolites displayed positive dominance. Overall, the genetic effects of different metabolites displayed large diversity in the same hybrid, while the genetic effects of the same metabolite varied little over different hybrids.

3.5. *Heterosis analysis*

We calculated mid-parent heterosis and over-parent heterosis of each metabolite in 26 hybrids (Tables S4–S6). The mid-parent heterosis and the significance test for 208 metabolites are illustrated in Fig. 5. For all hybrids, mid-parent heterosis ranged from –73.46% (oleic acid) to 599.88% (physostigmine), indicating considerable variation among metabolites in mid-parent heterosis (Table S4). Out of all 5408 combinations, 810 combinations exhibited significant negative mid-parent heterosis and 676 exhibited significant positive mid-parent heterosis, which accounted for 15% and 12.5%, respectively. Several metabolites displayed mid-parent heterosis in most hybrid combinations. For example, l-threonine (m19), sucrose (m87), and chlorogenic acid (m182) showed significant positive mid-parent heterosis in 21 hybrid combinations with the mean values of 114.64%, 53.21%, and 46%, respectively. Phytol (m90), skimmianine (m204), and vitamin D (m104) exhibited significant positive mid-parent heterosis in 24, 23, and 22 combinations with the mean values of –67.95%, –62.41%, and –50.82%, respectively.

The over-parent heterosis and the significance test for 208 metabolites are illustrated in Fig. 6A–B. Out of all combinations, 544 combinations exhibited significant positive over-parent heterosis and 558 combinations exhibited significant negative

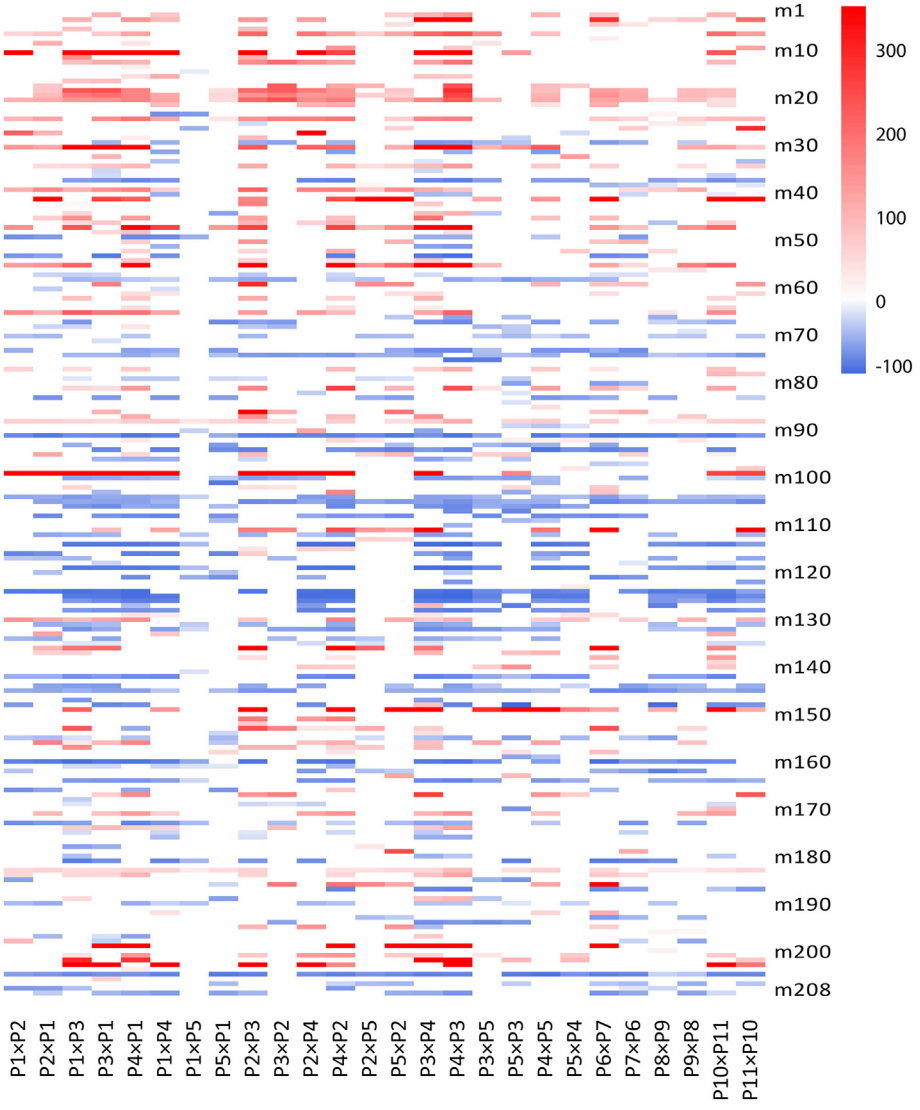

Fig. 5 – Significant mid-parent heterosis of 208 metabolites. Each block represents the mid-parent heterosis of each metabolite in the corresponding hybrid. The red and blue blocks represent positive and negative significance mid-parent heterosis, respectively.

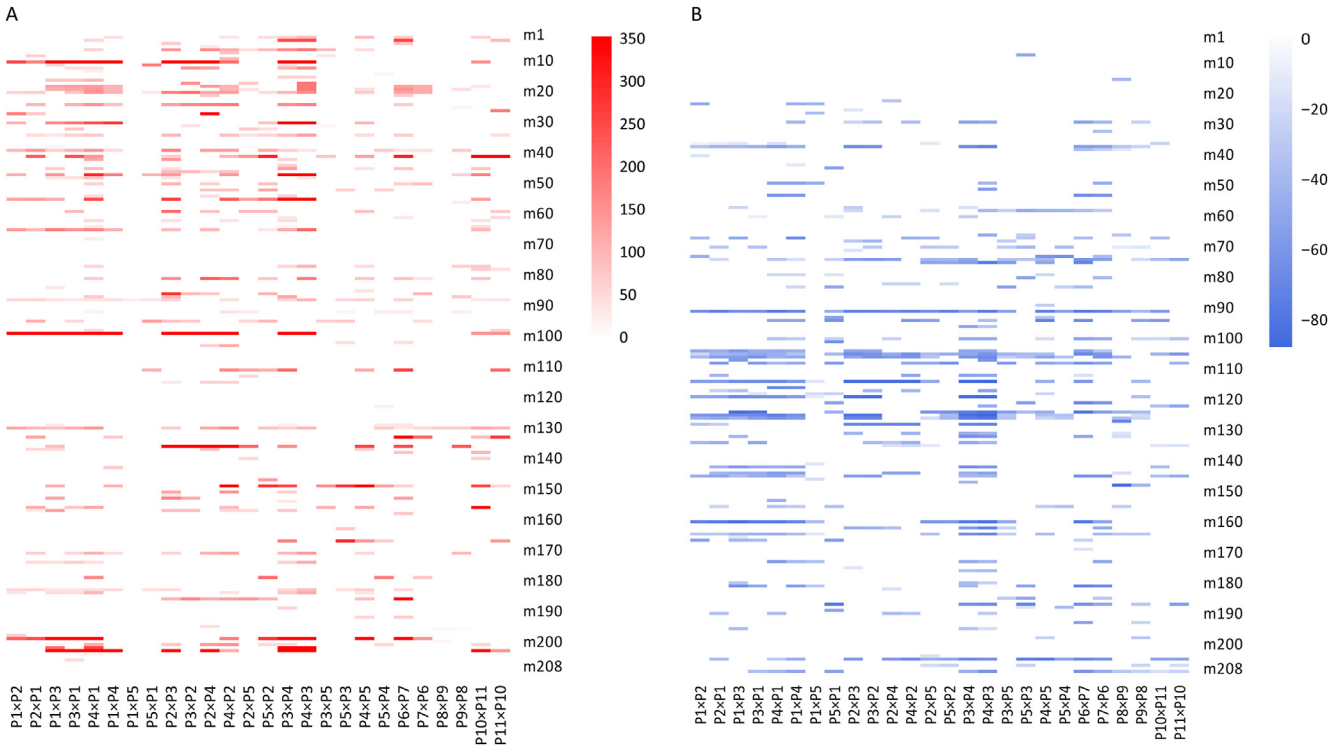

Fig. 6 – Significant positive over-parent heterosis and negative over-parent heterosis of 208 metabolites. A) Significant positive over-parent heterosis of 208 metabolites. B) Significant negative over-parent heterosis of 208 metabolites.

over-parent heterosis, accounting for 10.1% and 10.3%, respectively. The positive over-parent heterosis in l-Histidine (m9), pyridoxine (m98), physostigmine (m198), and quinic acid (m202) were relatively high with mean values of 200.43%, 364.11%, 408.25%, and 254.9% over all hybrids, respectively (Fig. 6A, Table S5). Propionic acid (m129) and sucrose (m87) exhibited significant positive over-parent heterosis in 18 and 17 combinations, respectively, while phytol (m90), vitamin D3 (m104), and α-tocotrienol (m105) exhibited significant negative over-parent heterosis in 23, 22, and 19 combinations, respectively (Fig. 6B).

3.6. Significant metabolic markers associated with agronomic traits

A total of 25 significant metabolic markers were detected for 11 agronomic traits (Table 2). Thirteen metabolites were associated with 100-kernel weight (HKW), indicating that the HKW might be related with multiple metabolic pathways. Seven metabolites were simultaneously related with two traits, wherein abscisic acid (m146) and gibberellin A3 (m160) simultaneously influenced EW and CW, l-arginine (m3) influenced ED and CD simultaneously, and stigmasterol (m72) affected ED and KL simultaneously. Based on the correlation analysis among agronomic traits, we found that the correlations between EW and CW, ED and CD as well as between ED and KL were significant (Fig. 7). Among these markers, gibberellin A3 (m160) associated with EW and CW could explain 23.59% of the phenotypic variation. Malvidin (m69) was significantly related with ERN, and gentisic acid (m45) was significantly related with HKW, which accounted for 19.33% and 18.90% of phenotypic variations, respectively.

4. Discussion

Understanding the genetic characteristics of metabolites per se will facilitate the application of metabolites for crop genetic improvement. In this study, we performed quantitative genetic analyses for 208 metabolites detected from 11 parental inbred lines of six representative maize varieties and their 26 reciprocal hybrids crosses. We found that metabolite levels varied a lot across different lines and different metabolites. However, these metabolites could separate parental lines from their corresponding hybrids based on cluster analysis and PCA. The conclusion is consistent with Lisec et al. [26] who revealed metabolic profiles of maize root could separate six parental inbred lines and their 14 hybrids. Among the quantified 208 metabolites, 59 metabolites are amino acids. The PCA result indicates that they play significant roles in distinguishing parents from hybrids, and 46 of 59 amino acids exhibit significant differences between parents and hybrids. Amino acids are not only the basic unit to proteins, but also form an important raw material for energy metabolism and material metabolism. Great processes have been achieved with regard

Table 1 – Metabolites with significant differences between reciprocal crosses.

Metabolite	Number	P1 × P2	P1 × P3	P4 × P1	P1 × P5	P2 × P3	P2 × P4	P2 × P5	P3 × P4	P3 × P5	P4 × P5	P6 × P7	P8 × P9	P10 × P11
		P2 × P1	P3 × P1	P1 × P4	P5 × P1	P3 × P2	P4 × P2	P5 × P2	P4 × P3	P5 × P3	P5 × P4	P7 × P6	P9 × P8	P11 × P10
L-Histidinol	m10				**	*						**		
L-Methionine	m14				**		*				*			
Pyroglutamic acid	m25				**									
Sphagnum acid	m51				**			*						
cAMP	m63				**						*			
D-Ribose	m85				**	*							*	
L-Glutamine	m110			*	**						*	*		
cis-gondoic acid	m116						**							
Pentadecylic acid	m126				**	*	**		*				*	
MG(0:0/20:5(5Z,8Z,11Z,14Z,17Z)/0:0)	m145													**
Gibberellin A12	m152		**									**		
Gibberellin A3	m160											*		
Gibberellin A53	m162					*	**			**				**
(–)-Jasmonic acid	m170		*	**									**	
p-Coumaroyl quinic acid	m197	**	*	*		*				*				
Skimmianine	m204				**					**				

* Indicates significant difference at the 0.05 probability level.
** Indicates significant difference at the 0.01 probability level.

Table 2 – Significant metabolites identified for agronomic traits using the LASSO method.

Trait	Metabolite	Number	Effect	R^2 (%)	P-value
EW	(+)-Abscisic acid	m146	38.0290	6.58	0.0438
EW	Gibberellin A3	m160	50.8681	11.77	0.0165
EL	L-Tyrosine	m22	–7.1012	9.52	0.0161
ED	L-Arginine	m3	1.4308	7.18	0.0309
ED	Stigmasterol	m72	–1.5632	8.57	0.0200
ERN	p-cresol	m48	0.6085	7.25	0.0270
ERN	Malvidin	m69	–0.9936	19.33	0.0029
ERN	Kynurenic acid	m192	–0.5271	5.44	0.0311
KNR	Sulfuric acid	m108	–1.5072	6.04	0.0477
CW	Spermidine	m31	–4.6262	6.26	0.0122
CW	(+)-Abscisic acid	m146	5.2630	8.11	0.0084
CW	Gibberellin A3	m160	8.9774	23.59	0.0009
CD	L-Arginine	m3	1.0510	11.75	0.0458
CD	(+)-α-Tocopherol	m92	–1.2085	15.53	0.0131
KL	Stigmasterol	m72	–0.2309	7.37	0.0274
KW	p-Cresol	m48	–0.2155	13.25	0.0035
KW	Sulfuric acid	m108	–0.1777	9.02	0.0132
KW	Kynurenic acid	m192	0.1595	7.26	0.0165
KT	Pantoic acid	m77	–0.1094	12.23	0.0210
HKW	L-Tryptophan	m20	0.6532	3.02	0.0002
HKW	Gentisic acid	m45	–1.6337	18.90	1.53E-07
HKW	Sphagnum acid	m51	–0.5148	1.88	0.0059
HKW	β-Carotene	m74	–1.1230	8.93	0.0009
HKW	β-Sitosterol	m91	0.8933	5.65	0.0004
HKW	Folic acid	m93	–0.6820	3.29	0.0407
HKW	Pyridoxal phosphate	m96	0.4132	1.21	0.0150
HKW	Docosanoic acid	m117	–0.2898	0.59	0.0119
HKW	Gibberellin A12	m153	–0.3700	0.97	0.0095
HKW	(–)-Caaverine	m168	–0.5503	2.14	0.0220
HKW	Allantoic acid	m177	–1.0285	7.49	0.0001
HKW	Pyruvate	m200	0.6903	3.37	0.0077
HKW	Vanillin	m206	0.2927	0.61	0.0170

R^2 indicates proportion of the total phenotypic variation explained by the metabolite. EW, ear weight; EL, ear length; ED, ear diameter; ERN, ear row number; KNR, kernel number per row; CW, cob weight; CD, cob diameter; KL, kernel length; KW, kernel width; KT, kernel thickness; HKW, 100-kernel weight.

[35]. We will further take multiple tissues and sampling time into account in the following studies.

The results of ANOVA indicate that metabolites exhibit significant differences among parents, hybrids, as well as parents and hybrids. A small number of metabolites exhibit significant differences between reciprocal crosses, suggesting that nuclear inheritance is dominant. However, 16 metabolites exhibit significant differences ($P < 0.01$) between reciprocal crosses, indicating the existence of cytoplasm effects for these metabolites. Cytoplasm effects have also been found for agronomic traits and quality traits in maize and other crops. Tang et al. [36] revealed that the cytoplasm contributed to about 8.75% and 39.91% of the phenotypic variance for plant height and ear height in maize. It is well known that genotype is identical between reciprocal crosses, which may limit the application of genomic data for hybrid prediction. Xu et al. [23] found that when using parental metabolites to predict the yield of hybrids, the predictability is much higher than genomic prediction. However, metabolite data tend not to be

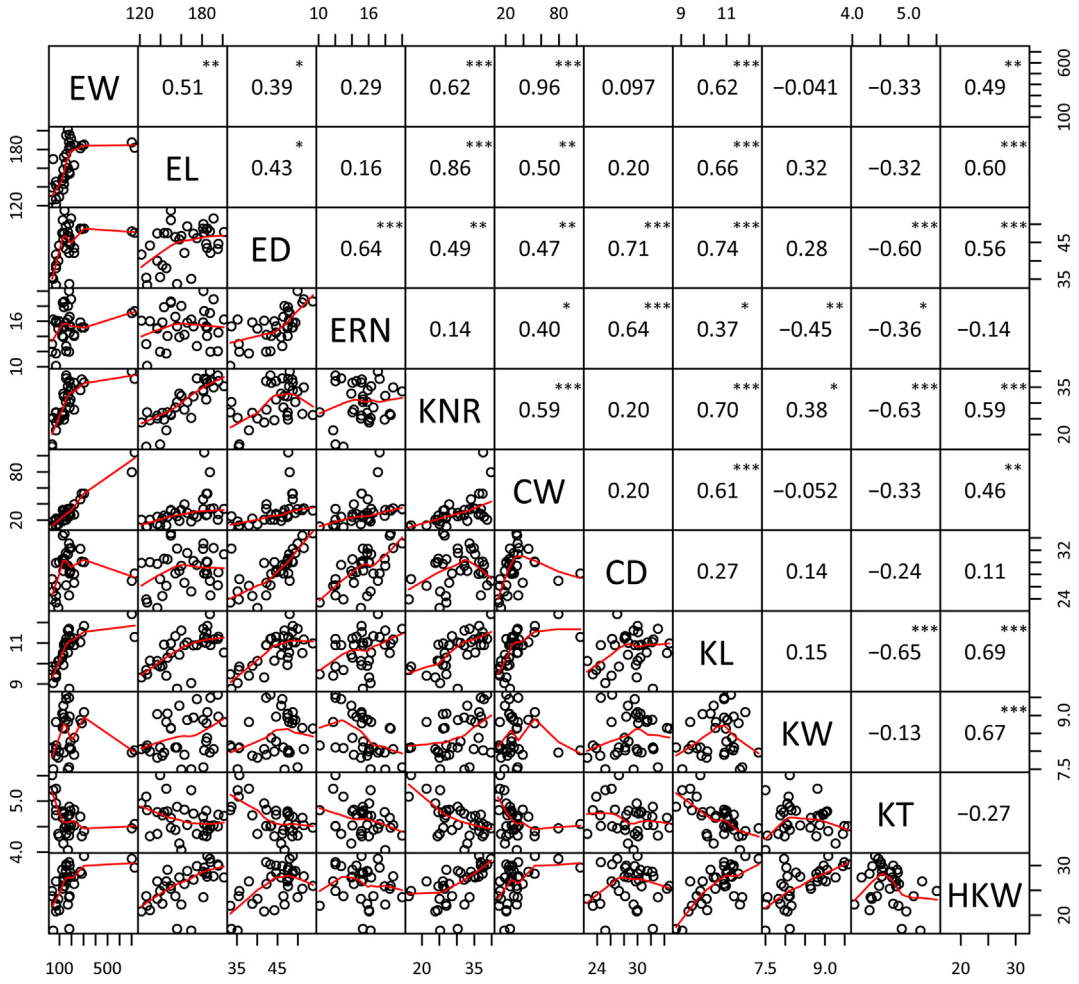

Fig. 7 – The pairwise correlation analysis among 11 agronomic traits. The lower left shows a bivariate scatter plot with a fitted line, and the upper right shows the correlation coefficient and significance level. *, **, and * indicate significant difference at the 0.05, 0.01, and 0.001 probability levels, respectively. EW, ear weight; EL, ear length; ED, ear diameter; ERN, ear row number; KNR, kernel number per row; CW, cob weight; CD, cob diameter; KL, kernel length; KW, kernel width; KT, kernel thickness; HKW, 100-kernel weight.**

collected from hybrids but indirectly inferred from their parents. The relationship between metabolite levels of hybrids and those of their parents is unclear, which affects the predictability of hybrid prediction based on parental metabolites. Dan et al. [24] used sums, differences, ratios of the parental metabolite levels to construct predictive variables of hybrids. Xu et al. [23] proposed to use the additive-dominance model for metabolite levels to predict hybrids. Here, we find that there are no uniform inheritance patterns for all metabolites. Seven categories of genetic effects were all observed for the 208 metabolites. Taking all combinations into account, most metabolites followed an overdominance inheritance pattern, accounting for 67.8%. The result is consistent with the previous study in maize root metabolome, which also demonstrated that two-thirds of all metabolites displayed an overdominance inheritance pattern [26]. We also observed large differences in inheritance patterns of different metabolites. Therefore, it seems inappropriate to construct a uniform model for all metabolites when using parental metabolite data to predict the performance of hybrids. The optimal model for hybrid prediction based on parental metabolite data need to be further validated.

Supplementary data for this article can be found online at https://doi.org/10.1016/j.cj.2020.05.009.

Declaration of competing interest

Authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (31801028, 41801013, 31902101), the National Key Research and Development Program of China (2016YFD0100303), Natural Science Foundation of Jiangsu Province (BK20180939), Qing Lan Project of Yangzhou University, and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Author contributions

Chenwu Xu and Yang Xu designed the research; Yang Xu, Ying Ma, and Xin Wang performed the research; Cheng Li, Xuecai Zhang, Pengcheng Li, and Zefeng Yang analyzed the data; and Yang Xu wrote the paper. All authors read and approved the final manuscript.

REFERENCES

[1] Y. Xiao, H. Liu, L. Wu, M. Warburton, J. Yan, Genome-wide association studies in maize: praise and stargaze, Mol. Plant 10 (2017) 359–374.

[2] M.D. McMullen, S. Kresovich, H.S. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, P. Brown, C. Browne, M. Eller, K. Guill, C. Harjes, D. Kroon, N. Lepak, S.E. Mitchell, B. Peterson, G. Pressoir, S. Romero, M. Oropeza Rosas, S. Salvo, H. Yates, M. Hanson, E. Jones, S. Smith, J.C. Glaubitz, M. Goodman, D. Ware, J.B. Holland, E.S. Buckler, Genetic properties of the maize nested association mapping population, Science 325 (2009) 737–740.

[3] F. Tian, P.J. Bradbury, P.J. Brown, H. Hung, Q. Sun, S. Flint-Garcia, T.R. Rocheford, M.D. McMullen, J.B. Holland, E.S. Buckler, Genome-wide association study of leaf architecture in the maize nested association mapping population, Nat. Genet. 43 (2011) 159–162.

[4] H. Li, Z. Peng, X. Yang, W. Wang, J. Fu, J. Wang, Y. Han, Y. Chai, T. Guo, N. Yang, J. Liu, M.L. Warburton, Y. Cheng, X. Hao, P. Zhang, J. Zhao, Y. Liu, G. Wang, J. Li, J. Yan, Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels, Nat. Genet. 45 (2013) 43–50.

[5] X. Wang, H. Wang, S. Liu, A. Ferjani, J. Li, J. Yan, X. Yang, F. Qin, Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings, Nat. Genet. 48 (2016) 1233–1241.

[6] T.A. Schrag, M. Westhues, W. Schipprack, F. Seifert, A. Thiemann, S. Scholten, A.E. Melchinger, Beyond genomic prediction: combining different types of omics data can improve prediction of hybrid performance in maize, Genetics 208 (2018) 1373–1385.

[7] B. Peng, H. Li, X.X. Peng, Functional metabolomics: from biomarker discovery to metabolome reprogramming, Protein Cell 6 (2015) 628–637.

[8] D. Yang, X. Du, Z. Yang, Z. Liang, Z. Guo, Y. Liu, Transcriptomics, proteomics, and metabolomics to reveal mechanisms underlying plant secondary metabolism, Eng. Life Sci. 14 (2014) 456–466.

[9] J. Luo, Metabolite-based genome-wide association studies in plants, Curr. Opin. Plant Biol. 24 (2015) 31–38.

[10] O. Fiehn, J. Kopka, P. Dörmann, T. Altmann, R.N. Trethewey, L. Willmitzer, Metabolite profiling for plant functional genomics, Nat. Biotechnol. 18 (2000) 1157–1161.

[11] C. Kristensen, M. Morant, C.E. Olsen, C.T. Ekstrom, D.W. Galbraith, B.L. Moller, S. Bak, Metabolic engineering of dhurrin in transgenic Arabidopsis plants with marginal inadvertent effects on the metabolome and transcriptome, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 1779–1784.

[12] W. Chen, Y. Gao, W. Xie, L. Gong, K. Lu, W. Wang, Y. Li, X. Liu, H. Zhang, H. Dong, Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism, Nat. Genet. 46 (2014) 714.

[13] C. Riedelsheimer, J. Lisec, A. Czedik-Eysenberg, R. Sulpice, A. Flis, C. Grieder, T. Altmann, M. Stitt, L. Willmitzer, A.E. Melchinger, Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 8872–8877.

[14] E.K. Chan, H.C. Rowe, J.A. Corwin, B. Joseph, D.J. Kliebenstein, Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in Arabidopsis thaliana, PLoS Biol. 9 (2011), e1001125.

[15] F. Matsuda, R. Nakabayashi, Z. Yang, Y. Okazaki, J. Yonemaru, K. Ebana, M. Yano, K. Saito, Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism, Plant J. 81 (2015) 13–23.

[16] W. Wen, K. Li, S. Alseekh, N. Omranian, L. Zhao, Y. Zhou, Y. Xiao, M. Jin, N. Yang, H. Liu, Genetic determinants of the network of primary metabolism and their relationships to plant performance in a maize recombinant inbred line population, Plant Cell 27 (2015) 1839–1856.

[17] C. Fang, J. Luo, Metabolic GWAS-based dissection of genetic bases underlying the diversity of plant metabolism, Plant J. 97 (2019) 91–100.

[18] W. Wen, D. Li, X. Li, Y. Gao, W. Li, H. Li, J. Liu, H. Liu, W. Chen, J. Luo, Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights, Nat. Commun. 5 (2014) 3438.

[19] R.C. Meyer, S. Matthias, L. Jan, B. Martina, W.W. Hanna, T.R. Ottó, F. Oliver, E. Anne, W. Lothar, S. Joachim, The metabolic signature related to high plant growth rate in Arabidopsis thaliana, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 4759–4764.

[20] G. Tanja, S. Matthias, A. Sandra, L. Jan, R.C. Meyer, A. Thomas, W. Lothar, S. Joachim, Improved heterosis prediction by combining information on DNA- and metabolic markers, PLoS One 4 (2009), e5220.

[21] C. Riedelsheimer, A. Czedik-Eysenberg, C. Grieder, J. Lisec, F. Technow, R. Sulpice, T. Altmann, M. Stitt, L. Willmitzer, A.E. Melchinger, Genomic and metabolic prediction of complex heterotic traits in hybrid maize, Nat. Genet. 44 (2012) 217–220.

[22] S. Xu, D. Zhu, Q. Zhang, Predicting hybrid performance in rice using genomic best linear unbiased prediction, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 12456–12461.

[23] S. Xu, Y. Xu, L. Gong, Q. Zhang, Metabolomic prediction of yield in hybrid rice, Plant J. 88 (2016) 219–227.

[24] Z. Dan, J. Hu, W. Zhou, G. Yao, R. Zhu, Y. Zhu, W. Huang, Metabolic prediction of important agronomic traits in hybrid rice (*Oryza sativa* L.), Sci. Rep. 6 (2016) 21732.

[25] Y. Xu, C. Xu, S. Xu, Prediction and association mapping of agronomic traits in maize using multiple omic data, Heredity 119 (2017) 174–184.

[26] J. Lisec, L. Römisch-Margl, Z. Nikoloski, H.P. Piepho, P. Giavalisco, J. Selbig, A. Gierl, L. Willmitzer, Corn hybrids display lower metabolite variability and complex metabolite inheritance patterns, Plant J. 68 (2011) 326–336.

[27] R. Muralla, C. Sweeney, A. Stepansky, T. Leustek, D. Meinke, Genetic dissection of histidine biosynthesis in Arabidopsis, Plant Physiol. 144 (2007) 890–903.

[28] R.A. Azevedo, C. Damerval, J. Landry, P.J. Lea, C.M. Bellato, L. W. Meinhardt, M. Le Guilloux, S. Delhaye, A.A. Toro, S.A. Gaziola, Regulation of maize lysine metabolism and endosperm protein synthesis by opaque and floury mutations, Eur. J. Biochem. 270 (2003) 4898–4908.

[29] V. Tzin, G. Galili, New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants, Mol. Plant 3 (2010) 956–972.

[30] M. Nam, E. Bang, T. Kwon, Y. Kim, E. Kim, K. Cho, W. Park, B.G. Kim, I. Yoon, Metabolite profiling of diverse rice germplasm and identification of conserved metabolic markers of rice roots in response to long-term mild salinity stress, Int. J. Mol. Sci. 16 (2015) 21959–21974.

[31] J. Wei, A. Wang, R. Li, H. Qu, Z. Jia, Metabolome-wide association studies for agronomic traits of rice, Heredity 120 (2018) 342–355.

[32] S. Yamaguchi, Gibberellin metabolism and its regulation, Annu. Rev. Plant Biol. 59 (2008) 225–251.
[33] K. Shu, W. Zhou, F. Chen, X. Luo, W. Yang, Abscisic acid and gibberellins antagonistically mediate plant development and abiotic stress responses, Front. Plant Sci. 9 (2018) 416.
[34] O. Fernandez, M. Urrutia, S. Bernillon, C. Giauffret, F. Tardieu, J.L. Gouis, N. Langlade, A. Charcosset, A. Moing, Y. Gibon, Fortune telling: metabolic markers of plant performance, Metabolomics 12 (2016) 158.
[35] M. Westhues, T.A. Schrag, C. Heuer, G. Thaller, H.F. Utz, W. Schipprack, A. Thiemann, F. Seifert, A. Ehret, A. Schlereth, Omics-based hybrid prediction in maize, Theor. Appl. Genet. 130 (2017) 1927–1939.
[36] Z. Tang, Z. Yang, Z. Hu, D. Zhang, X. Lu, B. Jia, D. Deng, C. Xu, Cytonuclear epistatic quantitative trait locus mapping for plant height and ear height in maize, Mol. Breed. 31 (2013) 1–14.