青年英才 A 类附件材料

植物科学技术学院

陈 伟

2020-11-25

目录

1 论文

1.1 Shi et al., *Plant J*, 2020（共同通讯排最后）

the plant journal

The Plant Journal (2020) **103**, 279–292 doi: 10.1111/tpj.14727

Metabolomics analysis and metabolite-agronomic trait associations using kernels of wheat (*Triticum aestivum*) recombinant inbred lines

Taotao Shi[1,2,†], Anting Zhu[1,2,†], Jingqi Jia[1,2], Xin Hu[1,2], Jie Chen[1,2], Wei Liu[1,2], Xifeng Ren[2], Dongfa Sun[2], Alisdair R. Fernie[3], Fa Cui[4,*] and Wei Chen[1,2,*]

[1]*National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China,*

[2]*College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China,*

[3]*Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm 14476, Germany, and*

[4]*Wheat Molecular Breeding Innovation Research Group, Key Laboratory of Molecular Module-Based Breeding of High Yield and Abiotic Resistant Plants in Universities of Shandong, School of Agriculture, Ludong University, Yantai, China*

Received 30 September 2019; revised 17 January 2020; accepted 7 February 2020; published online 19 February 2020.
*For correspondence (e-mail chenwei0609@mail.hzau.edu.cn; sdaucf@126.com).
†These authors contributed equally to this work.

SUMMARY

Plants produce numerous metabolites that are important for their development and growth. However, the genetic architecture of the wheat metabolome has not been well studied. Here, utilizing a high-density genetic map, we conducted a comprehensive metabolome study via widely targeted LC-MS/MS to analyze the wheat kernel metabolism. We further combined agronomic traits and dissected the genetic relationship between metabolites and agronomic traits. In total, 1260 metabolic features were detected. Using linkage analysis, 1005 metabolic quantitative trait loci (mQTLs) were found distributed unevenly across the genome. Twenty-four candidate genes were found to modulate the levels of different metabolites, of which two were functionally annotated by *in vitro* analysis to be involved in the synthesis and modification of flavonoids. Combining the correlation analysis of metabolite-agronomic traits with the co-localization of methylation quantitative trait locus (mQTL) and phenotypic QTL (pQTL), genetic relationships between the metabolites and agronomic traits were uncovered. For example, a candidate was identified using correlation and co-localization analysis that may manage auxin accumulation, thereby affecting number of grains per spike (NGPS). Furthermore, metabolomics data were used to predict the performance of wheat agronomic traits, with metabolites being found that provide strong predictive power for NGPS and plant height. This study used metabolomics and association analysis to better understand the genetic basis of the wheat metabolism which will ultimately assist in wheat breeding.

Keywords: ***Triticum aestivum*** **L., mature seed, metabolic quantitative trait loci, agronomic trait, metabolic prediction.**

INTRODUCTION

Plants are highly enriched in specific metabolites that play important roles in the plant life cycle and mediate their interactions within the complex environments in which they live (Dixon and Strack, 2003; Saito and Matsuda, 2010; Peng *et al.*, 2017). Metabolomics aims to be the qualitative and quantitative analysis of all metabolites in a biological sample (Fiehn *et al.*, 2000), however current methodologies fall well short of this goal (Alseekh *et al.*, 2017). That said, combining metabolomics with genomics and transcriptomics has proven powerful in analyzing metabolic diversity and its underlying genetic variation, as well as in identifying numerous new genes and metabolic pathways (Tohge and Fernie, 2010; Fernie and Tohge, 2017; Alseekh and Fernie, 2018; Fang *et al.*, 2019). For example, hundreds of metabolic quantitative trait loci (mQTLs) have been detected in Arabidopsis, tomato, maize, and rice by linkage analysis (Lisec *et al.*, 2009; Matsuda *et al.*, 2012; Toubiana *et al.*, 2012; Alseekh *et al.*, 2015; Jin *et al.*, 2017), with the identification a large number of structural and regulatory genes involved in managing crop metabolite abundances.

1.2 Chen et al., *Plant Biotechnol J*, 2020（共同通讯排最后）

Plant Biotechnology Journal
Open Access

Plant Biotechnology Journal (2020) **18**, pp. 1722–1735 doi: 10.1111/pbi.13335

Metabolite-based genome-wide association study enables dissection of the flavonoid decoration pathway of wheat kernels

Jie Chen[1,2], Xin Hu[1,2], Taotao Shi[1,2], Huanran Yin[1,2], Dongfa Sun[2], Yuanfeng Hao[3], Xianchun Xia[3], Jie Luo[1], Alisdair R. Fernie[4], Zhonghu He[3,*] and Wei Chen[1,2,*]

[1]*National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan, China*
[2]*College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China*
[3]*National Wheat Improvement Center, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China*
[4]*Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm, Germany*

Received 10 September 2019; accepted 29 December 2019.
**Correspondence* (Tel +86-18995611853; fax +86-27-87285101; email chenwei0609@mail.hzau.edu.cn (W.C.)); (Tel +86-10-82108547; fax +86-10-82108547; email hezhonghu02@caas.cn (Z.H.))

Keywords: flavonoid decoration, metabolite-based genome-wide association study (mGWAS), metabolomics-associated breeding, pathway elucidation, wheat kernels.

Summary

The marriage of metabolomic approaches with genetic design has proven a powerful tool in dissecting diversity in the metabolome and has additionally enhanced our understanding of complex traits. That said, such studies have rarely been carried out in wheat. In this study, we detected 805 metabolites from wheat kernels and profiled their relative contents among 182 wheat accessions, conducting a metabolite-based genome-wide association study (mGWAS) utilizing 14 646 previously described polymorphic SNP markers. A total of 1098 mGWAS associations were detected with large effects, within which 26 candidate genes were tentatively designated for 42 loci. Enzymatic assay of two candidates indicated they could catalyse glucosylation and subsequent malonylation of various flavonoids and thereby the major flavonoid decoration pathway of wheat kernel was dissected. Moreover, numerous high-confidence genes associated with metabolite contents have been provided, as well as more subdivided metabolite networks which are yet to be explored within our data. These combined efforts presented the first step towards realizing metabolomics-associated breeding of wheat.

Introduction

Plant metabolites play crucial roles in the interaction of plants with their surrounding environments (Saito and Matsuda, 2010; Schwab, 2003) and are necessary for humans in that they directly or indirectly constitute our nutritional supply (De Luca *et al.*, 2012; Keurentjes, 2009; Saito and Matsuda, 2010; Wang *et al.*, 2009). One class of specialized metabolites (also called secondary metabolites), the flavonoids, have been proposed to possess a range of functional roles. For instance, different decorations of the basic flavonoid structure were associated with varied UV tolerance in both rice cultivars and *Arabidopsis* ecotypes dispersed in various latitudes (Peng *et al.*, 2017; Tohge *et al.*, 2016), and flavonoid metabolites are believed to confer, among other bioactivities, anti-inflammatory activity when provided in the diet (Kang *et al.*, 2016; Martin and Li, 2017; Zhou and Ibrahim, 2009). However, the enormous number of predicted metabolites (Dixon and Strack, 2003) and the severe variation in their abundance between species (Morohashi *et al.*, 2012) mean that vast majority of metabolic pathways remain to be fully unveiled. Indeed, unlike primary metabolites which are similarly present across the plant kingdom, secondary metabolic pathways are highly divergent between species. An elegant recent study revealed that part of the previously known pathway for flavonoid syntheses was reconstructed in rice, in which the naringenin-to-tricin route was redirected bypassing the formation of tricetin (Lam *et al.*, 2015). More recently, the progress in flavonoid biosynthesis pathways in model plants and several crop species has been reviewed (Tohge *et al.*, 2017). These studies demonstrated the essentiality of applying metabolomics as a systematic approach to study specialized plant metabolism. Specifically, metabolomic genome-wide association study (mGWAS) or metabolomic quantitative trait loci mapping (mQTL) has proven highly powerful in understanding the diversification of metabolites (Chen *et al.*, 2014; Gong *et al.*, 2013; Zhu *et al.*, 2018), as well as the association of these metabolites with biotic and abiotic stress defence processes (Chen *et al.*, 2018; Glauser *et al.*, 2011; Peng *et al.*, 2017) or with food quality and flavour (Peng *et al.*, 2016; Sharma *et al.*, 2016; Tieman *et al.*, 2017). However, to date only very few mGWAS or mQTL studies have been conducted in wheat (Hill *et al.*, 2013; Hill *et al.*, 2015; Matros *et al.*, 2017).

Wheat (*Triticum aestivum* L.) is a leading cereal crop ultimately accounting for approximately 20% of the calories consumed by humans (Simmonds *et al.*, 2016). To secure worldwide food supply, intense selection of high-yield and broad-adaptation wheat cultivars has been the primary breeding target in wheat breeding programmes (Dubcovsky and Dvorak, 2007). However, owing to linkage drag and relatively low recombination frequencies, conventional breeding processes are usually time-consuming and have low efficiency and predictability (Holland, 2007), which renders it necessary to develop and incorporate genomic tools to assist wheat breeding programmes (Simmonds *et al.*, 2016). Indeed, the available genomic information of hexaploid wheat (IWGSC, 2018) and its tetraploid (Avni *et al.*, 2017) and diploid progenitors (Ling *et al.*, 2018; Luo *et al.*, 2017) has greatly facilitated fundamental research in wheat (Cui *et al.*, 2019; Ju

1.3 Peng et al., *Front Plant Sci*, 2018（共同通讯排最后）

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 August 2018
doi: 10.3389/fpls.2018.01196

Genome-Wide Association Studies of Free Amino Acid Levels by Six Multi-Locus Models in Bread Wheat

Yanchun Peng[1†], Hongbo Liu[2†], Jie Chen[1], Taotao Shi[1], Chi Zhang[3], Dongfa Sun[1], Zhonghu He[4], Yuanfeng Hao[4] and Wei Chen[1]**

[1] College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [2] National Key Laboratory of Crop Genetic Improvement, National Center of Plant Gene Research, Huazhong Agricultural University, Wuhan, China, [3] School of Chemical Science and Engineering, Royal Institute of Technology, Stockholm, Sweden, [4] Institute of Crop Science, National Wheat Improvement Center, Chinese Academy of Agricultural Sciences, Beijing, China

Genome-wide association studies (GWAS) have been widely used to dissect the complex biosynthetic processes of plant metabolome. Most studies have used single-locus GWAS approaches, such as mixed linear model (MLM), and little is known about more efficient algorithms to implement multi-locus GWAS. Here, we report a comprehensive GWAS of 20 free amino acid (FAA) levels in kernels of bread wheat (*Triticum aestivum* L.) based on 14,646 SNPs by six multi-locus models (FASTmrEMMA, FASTmrMLM, ISISEM-BLASSO, mrMLM, pKWmEB, and pLARmEB). Our results showed that 328 significant quantitative trait nucleotides (QTNs) were identified in total (38, 8, 92, 45, 117, and 28, respectively, for the above six models). Among them, 66 were repeatedly detected by more than two models, and 155 QTNs appeared only in one model, indicating the reliability and complementarity of these models. We also found that the number of significant QTNs for different FAAs varied from 8 to 41, which revealed the complexity of the genetic regulation of metabolism, and further demonstrated the necessity of the multi-locus GWAS. Around these significant QTNs, 15 candidate genes were found to be involved in FAA biosynthesis, and one candidate gene (*TraesCS1D01G052500*, annotated as tryptophan decarboxylase) was functionally identified to influence the content of tryptamine *in vitro*. Our study demonstrated the power and efficiency of multi-locus GWAS models in crop metabolome research and provided new insights into understanding FAA biosynthesis in wheat.

Keywords: wheat, free amino acid (FAA), genome-wide association studies, multi-locus models, QTNs

OPEN ACCESS

Edited by:
Zhenyu Jia,
University of California, Riverside, United States

Reviewed by:
Jin-Ying Gou,
Fudan University, China
Yang Zhang,
Sichuan University, China

***Correspondence:**
Yuanfeng Hao
haoyuanfeng@caas.cn
Wei Chen
chenwei0609@mail.hzau.edu.cn
† These authors have contributed equally to this work

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *27 May 2018*
Accepted: *26 July 2018*
Published: *14 August 2018*

Citation:
Peng Y, Liu H, Chen J, Shi T, Zhang C, Sun D, He Z, Hao Y and Chen W (2018) Genome-Wide Association Studies of Free Amino Acid Levels by Six Multi-Locus Models in Bread Wheat. Front. Plant Sci. 9:1196. doi: 10.3389/fpls.2018.01196

INTRODUCTION

Genome-wide association studies (GWAS) have largely been applied to the genetic dissection of complex traits in plants. With the landmark GWAS study of 107 phenotypes in *Arabidopsis* (Atwell et al., 2010), numerous other studies have been successfully performed, including those addressing the flowering time and grain yield in rice (Huang et al., 2012; Yang W. et al., 2014), salinity tolerance in barley (Fan et al., 2016), male inflorescence size in maize (Wu et al., 2016), floret fertility in wheat (Guo et al., 2017), and the reducing levels of cucurbitacin in cucumber domestication (Shang et al., 2014). Of these studies, the mixed linear model (MLM) has been adopted most frequently owing to its effective control of spurious associations (Yu et al., 2006). However, as a single-locus GWAS

1.4 Chen et al., *Nat Commun*, 2016（共同一作排第一）

nature COMMUNICATIONS

ARTICLE

Received 11 Dec 2015 | Accepted 29 Jul 2016 | Published 4 Oct 2016

DOI: 10.1038/ncomms12767 OPEN

Comparative and parallel genome-wide association studies for metabolic and agronomic traits in cereals

Wei Chen[1,*], Wensheng Wang[1,*], Meng Peng[1,*], Liang Gong[1], Yanqiang Gao[1], Jian Wan[1], Shouchuang Wang[1], Lei Shi[1], Bin Zhou[1], Zongmei Li[1], Xiaoxi Peng[1], Chenkun Yang[1], Lianghuan Qu[2], Xianqing Liu[2] & Jie Luo[1]

The plant metabolome is characterized by extensive diversity and is often regarded as a bridge between genome and phenome. Here we report metabolic and phenotypic genome-wide studies (mGWAS and pGWAS) in rice grain that, in addition to previous metabolic GWAS in rice leaf and maize kernel, show both distinct and overlapping aspects of genetic control of metabolism within and between species. We identify new candidate genes potentially influencing important metabolic and/or morphological traits. We show that the differential genetic architecture of rice metabolism between different tissues is in part determined by tissue specific expression. Using parallel mGWAS and pGWAS we identify new candidate genes potentially responsible for variation in traits such as grain colour and size, and provide evidence of metabotype-phenotype linkage. Our study demonstrates a powerful strategy for interactive functional genomics and metabolomics in plants, especially the cloning of minor QTLs for complex phenotypic traits.

[1] National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China. [2] College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.L. (email: jie.luo@mail.hzau.edu.cn).

1.5 Peng et al., *Plant Cell*, 2016（共同一作排第三）

The Plant Cell, Vol. 28: 1533–1550, July 2016, www.plantcell.org

LARGE-SCALE BIOLOGY ARTICLE

Evolutionarily Distinct BAHD *N*-Acyltransferases Are Responsible for Natural Variation of Aromatic Amine Conjugates in Rice[OPEN]

Meng Peng,[a,1] Yanqiang Gao,[a,1] Wei Chen,[a,1] Wensheng Wang,[a] Shuangqian Shen,[a] Jian Shi,[a] Cheng Wang,[a] Yu Zhang,[a] Li Zou,[a] Shouchuang Wang,[a] Jian Wan,[a] Xianqing Liu,[b] Liang Gong,[a] and Jie Luo[a,2]

[a] National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China

[b] College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China

ORCID ID: 0000-0002-2903-9668 (Y.G.)

Phenolamides (PAs) are specialized (secondary) metabolites mainly synthesized by BAHD *N*-acyltransferases. Here, we report metabolic profiling coupled with association and linkage mapping of 11 PAs in rice (*Oryza sativa*). We identified 22 loci affecting PAs in leaves and 16 loci affecting PAs in seeds. We identified eight BAHD *N*-acyltransferases located on five chromosomes with diverse specificities, including four aromatic amine *N*-acyltransferases. We show that genetic variation in PAs is determined, at least in part, by allelic variation in the tissue specificity of expression of the BAHD genes responsible for their biosynthesis. Tryptamine hydroxycinnamoyl transferase 1/2 (Os-THT1/2) and tryptamine benzoyl transferase 1/2 (Os-TBT1/2) were found to be bifunctional tryptamine/tyramine *N*-acyltransferases. The specificity of Os-THT1 and Os-TBT1 for agmatine involved four tandem arginine residues, which have not been identified as specificity determinants for other plant BAHD transferases, illustrating the versatility of plant BAHD transferases in acquiring new acyl acceptor specificities. With phylogenetic analysis, we identified both divergent and convergent evolution of *N*-acyltransferases in plants, and we suggest that the BAHD family of tryptamine/tyramine *N*-acyltransferases evolved conservatively in monocots, especially in Gramineae. Our work demonstrates that omics-assisted gene-to-metabolite analysis provides a useful tool for bulk gene identification and crop genetic improvement.

INTRODUCTION

Plants produce a large number of specialized (secondary) metabolites that are crucial for their interactions with the ever-changing environment (Keurentjes, 2009; Saito and Matsuda, 2010). Much of the enormous diversity of specialized metabolites in plants comes from modifications or decorations of core structures by different tailing reactions such as acylation (Bontpart et al., 2015), glycosylation (Bowles et al., 2005), methylation (Lam et al., 2007), and so on (Schwab, 2003). In addition, many studies have revealed that plant specialized metabolites accumulate with tissue specificity (Schilmiller et al., 2010; McDowell et al., 2011; Kim et al., 2012; Toubiana et al., 2012; Watanabe et al., 2013; Dong et al., 2015; Wen et al., 2015).

Phenolamides (PAs), also referred to as hydroxycinnamic acid amides or phenylamides, are the hydroxycinnamoyl acylated products of the mono-, di-, or triphenolic acid (coumaric, caffeic, or ferulic acid) substitutions of polyamines (putrescine, spermidine, agmatine, tryptamine, and serotonin) or arylmonoamines such as tyramine, octopamine, and anthranilate (Figure 1; Bienz et al., 2005; Edreva et al., 2007; Bassard et al., 2010). PAs are a diverse group of specialized metabolites found in many plants, including wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), rice (*Oryza sativa*), and maize (*Zea mays*) (Martin-Tanguy et al., 1978; Martin et al., 1985; Facchini et al., 2002; Edreva et al., 2007; Bassard et al., 2010). Accumulating evidence shows that PAs play an important role in plant defense responses against pathogens and insect herbivores (Newman et al., 2001; Tanaka et al., 2003; Kaur et al., 2010; Park et al., 2014). They are also suggested to play roles in sulfur starvation, heat shock, and protection against UV irradiation (Klapheck, 1983; Edreva et al., 1998; Kaur et al., 2010), although clear evidence to support these roles is yet to be reported (Fellenberg and Vogt, 2015).

Condensation of the hydroxycinnamoyl and amine moieties is the key step for the biosynthesis of PAs and is catalyzed by a diversity of hydroxycinnamoyl transferases (Petersen, 2015). One of the first genes encoding this type of enzyme to be characterized was the gene encoding tyramine:*N*-hydroxycinnamoyl transferase (THT; Hohlfeld et al., 1995; Schmidt et al., 1999; Yu and Facchini, 1999; Back et al., 2001; Von Roepenack-Lahaye

[1] These authors contributed equally to this work.
[2] Address correspondence to jie.luo@mail.hzau.edu.cn.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Jie Luo (jie.luo@mail.hzau.edu.cn).
[OPEN] Articles can be viewed without a subscription.
www.plantcell.org/cgi/doi/10.1105/tpc.16.00265

1.6 Chen et al., *Plant Physiol*, 2018（参与）

Metabolome Analysis of Multi-Connected Biparental Chromosome Segment Substitution Line Populations[1]

Jie Chen,[a,b,2] Jilin Wang,[c,2] Wei Chen,[b] Wenqiang Sun,[a,b] Meng Peng,[a] Zhiyang Yuan,[a,b] Shuangqian Shen,[a] Kun Xie,[a,b] Cheng Jin,[a] Yangyang Sun,[a] Xianqing Liu,[d] Alisdair R. Fernie,[e,f] Sibin Yu,[a,b,3] and Jie Luo[a,g,3,4]

[a]National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China
[b]College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
[c]Rice Research Institute, Jiangxi Academy of Agricultural Sciences, Nanchang 330200, China
[d]College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
[e]Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm 144776, Germany
[f]Center of Plant System Biology and Biotechnology, Plovdiv 4000, Bulgaria
[g]Institute of Tropical Agriculture and Forestry of Hainan University, Haikou, Hainan 572208, China

ORCID IDs: 0000-0002-6282-7244 (J.C.); 0000-0001-7858-1683 (J.W.); 0000-0001-7225-3785 (W.C.); 0000-0003-4241-6242 (W.S.); 0000-0002-4374-5366 (M.P.); 0000-0002-8459-0831 (S.S.); 0000-0002-8755-7607 (C.J.); 0000-0003-3232-2939 (Y.S.); 0000-0002-6136-7098 (X.L.); 0000-0001-9000-335X (A.R.F.); 0000-0001-9508-4494 (J.L.)

Metabolomic analysis coupled with advanced genetic populations represents a powerful tool with which to investigate the plant metabolome. However, genetic analyses of the rice (*Oryza sativa*) metabolome have been conducted mainly using natural accessions or a single biparental population. Here, the flag leaves from three interconnected chromosome segment substitution line populations with a common recurrent genetic background were used to dissect rice metabolic diversity. We effectively used multiple interconnected biparental populations, constructed by introducing genomic segments into Zhenshan 97 from ACC10 (A/Z), Minghui 63 (M/Z), and Nipponbare (N/Z), to map metabolic quantitative trait loci (mQTL). A total of 1,587 mQTL were generated, of which 684, 479, and 722 were obtained from the A/Z, M/Z, and N/Z chromosome segment substitution line populations, respectively, and we designated 99 candidate genes for 367 mQTL. In addition, 1,001 mQTL were generated specifically from joint linkage analysis with 25 candidate genes assigned. Several of these candidates were validated, such as *LOC_Os07g01020* for the in vivo content of pyridoxine and its derivative and *LOC_Os04g25980* for cis-zeatin glucosyltransferase activity. We propose a novel biosynthetic pathway for *O*-methylapigenin *C*-pentoside and demonstrated that *LOC_Os04g11970* encodes a component of this pathway through fine-mapping. We postulate that the methylated apigenin may confer plant disease resistance. This study demonstrates the power of using multiple interconnected populations to generate a large number of veritable mQTL. The combined results are discussed in the context of functional metabolomics and the possible features of assigned candidates underlying respective metabolites.

A vast number of metabolites are produced by plants, many of which are essential for plants to interact with the environment (Schwab, 2003; Saito and Matsuda, 2010) and represent important constituents of the human diet and development (Keurentjes, 2009; Saito and Matsuda, 2010; De Luca et al., 2012). For example, the water-soluble B6 group of metabolites not only function in cellular defense against oxidative stress in plants (Herrero et al., 2011) but also may reduce the incidence of important human diseases, such as hypertension and diabetes (Hellmann and Mooney, 2010; Fitzpatrick et al., 2012). Similarly, it has been reported that specialized metabolites such as flavonoids are involved in biotic and abiotic stress tolerance in plants (Luo et al., 2009; Kaur et al., 2010; Saito et al., 2013) and confer health-promoting effects against chronic diseases and certain cancers in humans (Niggeweg et al., 2004; Butelli et al., 2008). Therefore, in view of the importance of metabolites, it is necessary both to further study their in planta functions and explore their value for humans. However, this is a somewhat daunting task, since it was estimated that more than 200,000 metabolites are produced in the plant kingdom (Dixon and Strack, 2003), and these metabolites exhibit severe

[1]This work was supported by the National Science Fund for Distinguished Young Scholars (No. 31625021), by The Ministry of Science and Technology of China (2016YFD0100500) to J.L., by the National Natural Science Foundation of China (No. 31671656) and the National High Technology Research and Development of China (No. 2014AA10A604) to S.Y., and by the National Natural Science Foundation of China (No. 31770328) to W.C.

[2]These authors contributed equally to the article.

[3] Author for contact: ysb@mail.hzau.edu.cn and jie.luo@mail.hzau.edu.cn.

[4]Senior author.

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) are: Sibin Yu (ysb@mail.hzau.edu.cn) and Jie Luo (jie.luo@mail.hzau.edu.cn).

J.L. and S.Y. designed the study; J.C., J.W., and W.C. performed data analysis; J.C., J.W., and W.C. conducted main experiments, with W.S., M.P., Z.Y., S.S., K.X., C.J., Y.S., and X.L. in assistance; J.C., J.W., A.R.F., and J.L. wrote the article.

www.plantphysiol.org/cgi/doi/10.1104/pp.18.00490

1.7 Li et al., *Plant J*, 2019（参与）

the plant journal

The Plant Journal (2019) **100**, 908–922 doi: 10.1111/tpj.14482

Comparative analysis of metabolome of rice seeds at three developmental stages using a recombinant inbred line population

Kang Li[1], Dehong Wang[1], Liang Gong[2], Yuanyuan Lyu[1], Hao Guo[1], Wei Chen[3], Cheng Jin[1], Xianqing Liu[4], Chuanying Fang[4,*] and Jie Luo[1,4,*]

[1]*National Key Laboratory of Crop Genetic Improvement and National Center of Plant Research, Huazhong Agricultural University, Wuhan 430070, China,*

[2]*The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA,*

[3]*College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China, and*

[4]*Hainan Key Laboratory for Sustainable Utilization of Tropical Bioresource, College of Tropical Crops, Hainan University, Haikou 570288, China*

Received 14 May 2019; revised 13 July 2019; accepted 23 July 2019; published online 29 July 2019.
*For correspondence (e-mails chuanyingfang@126.com; jie.luo@mail.hzau.edu.cn).

SUMMARY

Plants are considered an important food and nutrition source for humans. Despite advances in plant seed metabolomics, knowledge about the genetic and molecular bases of rice seed metabolomes at different developmental stages is still limited. Here, using Zhenshan 97 (ZS97) and Minghui 63 (MH63), we performed a widely targeted metabolic profiling in seeds during grain filling, mature seeds and germinating seeds. The diversity between MH63 and ZS97 was characterized in terms of the content of metabolites and the metabolic shifting across developmental stages. Taking advantage of the ultra-high-density genetic map of a population of 210 recombinant inbred lines (RILs) derived from a cross between ZS97 and MH63, we identified 4681 putative metabolic quantitative trait loci (mQTLs) in seeds across the three stages. Further analysis of the mQTLs for the codetected metabolites across the three stages revealed that the genetic regulation of metabolite accumulation was closely related to developmental stage. Using *in silico* analyses, we characterized 35 candidate genes responsible for 30 structurally identified or annotated compounds, among which *LOC_Os07g04970* and *LOC_Os06g03990* were identified to be responsible for feruloylserotonin and L-asparagine content variation across populations, respectively. Metabolite–agronomic trait association and colocation between mQTLs and phenotypic quantitative trait loci (pQTLs) revealed the complexity of the metabolite–agronomic trait relationship and the corresponding genetic basis.

Keywords: rice seeds, metabolome, seeds during grain filling, mature seeds, germinating seeds, metabolic quantitative trait loci.

INTRODUCTION

As readouts of the physiological or biochemical status of an organism, metabolites are essential for plant growth and plant–environment interactions, as well as for human health (Keurentjes, 2009; Saito and Matsuda, 2010; De Luca *et al.*, 2012; Wurtzel and Kutchan, 2016). Benefiting from the extreme diversity of metabolites, plants have become ideal models for dissecting the mechanism of metabolite biosynthesis and its regulation (Keurentjes *et al.*, 2006; Morohashi *et al.*, 2012; Luo, 2015; Fang *et al.*, 2019a,b; Fang and Luo, 2019). In plants, which are sessile in nature, the number of metabolites is estimated to be between 100 000 and 1 million (Dixon and Strack, 2003; Afendi *et al.*, 2012). Many metabolites display differential shifts during development. For instance, the contents of the majority of *C*-glycosylated and *O*-glycosylated flavonoids significantly increase in seedlings during the first 10 days after germination, which decrease slightly during later stages (Dong *et al.*, 2014). Furthermore, the accumulation of anthocyanins and most of the indole-derived glucosinolates in leaves largely increases continuously throughout leaf senescence (Watanabe *et al.*, 2013). Development-dependent accumulation is also observed concerning the content of primary metabolites (Mounet *et al.*, 2007; Hu *et al.*, 2016; Silva *et al.*, 2017). For instance, Hu *et al.* (2016)

1.8 Liu et al., *Plant Biotechnol J*, 2019（参与）

Plant Biotechnology Journal
Open Access

Plant Biotechnology Journal (2019) **17**, pp. 2038–2040 doi: 10.1111/pbi.13180

Editorial

An automatic UPLC-HRMS data analysis platform for plant metabolomics

Pingping Liu[1,†], Huina Zhou[1,†], Qingxia Zheng[1], Peng Lu[1], Yong-Jie Yu[2,*], Peijian Cao[1], Wei Chen[3] and Qiansi Chen[1,*]

[1]*Zhengzhou Tobacco Research Institute of CNTC, Zhengzhou, Henan, China*
[2]*College of Pharmacy, Ningxia Medical University, Yinchuan, Ningxia, China*
[3]*National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei, China*

Received 25 April 2019;
revised 25 May 2019;
accepted 29 May 2019.
*Correspondence (Tel 86-371-67672089 (QC)/86-951-6880693 (YY); fax 86-371-67672079 (QC)/86-951-6880693 (YY); emails yongjie.yu@163.com (YY)/ chen_qiansi@163.com (QC))
†These authors contributed equally to this paper.
Keywords: ion-clustering peak annotation, plant metabolomics, UPLC-HRMS, chemometrics.

Here, we want to introduce our new automatic data analysis platform for untargeted metabolomic analysis of complex plant samples. Many laboratories across the world have adopted ultra-high performance liquid chromatography-high resolution mass spectrometry (UPLC-HRMS) as they seek to thoroughly characterize metabolites in complex plant samples, with the larger aim of identifying compounds with impactful biological functions (Rinschen *et al.*, 2019; Shen *et al.*, 2019). Despite the widespread deployment of analytical hardware that is capable of capturing robust data sets for many types of biological samples, it is the data analysis steps for UPLC-HRMS – seeking to accurately extract qualitative and quantitative information for thousands of metabolites – that continues to be one of the most challenging tasks and has become a frustrating bottleneck for metabolomics and lipidomics (Domingo-Almenara *et al.*, 2018).

A number of proprietary and freely available methodologies have been developed for such analyses, including AntDAS (Fu *et al.*, 2017), Mzmine2 (Pluskal *et al.*, 2010), XCMS (Smith *et al.*, 2006) and MS-DIAL (Tsugawa *et al.*, 2015). These methods typically integrate the entire data analysis workflow for untargeted metabolomics (i.e. EIC construction, peak detection, peak annotation and peak alignment) and connect to chemometrics methods for screening out functionally impactful metabolites that exhibit significant differences amongst various experimental groups.

In practical applications, especially for complex plant sample analysis, however, researchers are still faced with the challenge that many identified signals can correspond to the same biological metabolite, because hundreds of ions can screen out based on analysis of variance (ANOVA) or popular algorithms in chemometrics like partial least square. Aiming to address this problem, researchers seeking to perform compound identification must manually identify ions that putatively originate from a single metabolite (e.g. neutral loss and fragment ions), which is certainly a very time-consuming task. Aiming to address this problem, we have developed an ion clustering-based fragment identification algorithm. We combine this new algorithm with our previously developed data analysis methods (Fu *et al.*, 2016a; Yu *et al.*, 2019), including peak detection and time shift correction and registration modules, to provide an integrated data analysis platform for UPLC-HRMS-based untargeted metabolomics. The platform comprises five modules, including one for EIC peak extraction, one for time shift correction, one for peak registration across samples, one for peak screening module and finally our ion clustering-based peak annotation module (a full flow chart of the platform is presented in Figure 1).

EIC peak extraction module

The acquired UPLC-HRMS data files from an instrument are transformed into *mzxml* file format using ProteoWizard (http://proteowizard.sourceforge.net/). EICs are constructed using an ion density clustering algorithm(Yu *et al.*, 2019), which extracts EICs based on the fact that ions from a metabolite exhibit almost identical *m/z* values so that, within a small *m/z* tolerance like 0.01 Da, the specific ion density will be higher than any particular background noise signal. Peak detection is performed for each EIC. First, a local-minimal value-based baseline drift correction algorithm (Fu *et al.*, 2016b) is introduced for baseline correction.

Next, chromatographic peaks are extracted using a Gaussian smoothing-based strategy (Fu *et al.*, 2016a), which extracts peaks by smoothing the EIC under different smoothing scales; it then searches the ridge lines across successively increased smoothing scales. An example of peak extraction is provided in Figure 1a1-a3, where three chromatographic peaks in the EIC have been extracted. Then, retention time and *m/z* value of the ion acquired at the peak apex are used to characterize each EIC peak (Figure 1a4). Quantitative information based on peak area and peak height is extracted, respectively, as the sum of responses in the elution EIC peak and as the intensity at the peak apex.

Time shift correction module

Time shift correction is performed for each EIC. First, the total number of peaks in each EIC is counted, and the EIC with the maximum number of components is selected as the reference sample (Figure 1b1). For each peak in the reference, its candidate traces in a test sample must satisfy both *m/z* and retention tolerances, which are set as 0.01 Da and 0.5 min, respectively. The similarity between a test peak and a reference peak is assessed based on the Pearson correlation coefficient values of EIC curves. A similarity matrix can thus be constructed based on the test and reference peaks (Figure 1b2). We align all reference

2 基金项目

2.1 国家自然科学基金重大研究计划，100 万元，子课题主持

国家自然科学基金
合作任务书

项目名称： 小麦产量与品质性状关键基因的图位克隆和遗传网络解析

项目编号： 91935304

依托单位（甲方）： 中国农业大学

项目负责人： 倪中福 所在学院： 农学院

合作研究单位（乙方）：华中农业大学

合作研究任务负责人： 陈伟

合作研究任务名称：利用代谢组解析小麦特殊营养品质调控网络

起止年限：2020 年 1 月 1 日至 2021 年 12 月 31 日

签订日期： 2019 年 11 月 19 日

中国农业大学科学技术发展研究院

二〇一七年制

2.2 国家自然科学基金面上项目，59 万元，主持

国家自然科学基金面上项目，31770328，整合多亲本染色体片段代换系群体研究水稻代谢数量性状位点，2018/01-2021/12，59 万元，在研，主持；

关于国家自然科学基金资助项目批准及有关事项的通知

陈伟　先生/女士：

根据《国家自然科学基金条例》的规定和专家评审意见，国家自然科学基金委员会（以下简称自然科学基金委）决定批准资助您的申请项目。项目批准号：31770328，项目名称：整合多亲本染色体片段代换系群体研究水稻代谢数量性状位点，直接费用：59.00万元，项目起止年月：2018年01月至2021年12月，有关项目的评审意见及修改意见附后。

请尽早登录科学基金网络信息系统（https://isisn.nsfc.gov.cn），获取《国家自然科学基金资助项目计划书》（以下简称计划书）并按要求填写。对于有修改意见的项目，请按修改意见及时调整计划书相关内容；如对修改意见有异议，须在计划书电子版报送截止日期前提出。**注意：请严格按照《国家自然科学基金资助项目资金管理办法》填写计划书的资金预算表，其中，劳务费、专家咨询费科目所列金额与申请书相比不得调增。**

计划书电子版通过科学基金网络信息系统（https://isisn.nsfc.gov.cn）上传，由依托单位审核后提交至自然科学基金委进行审核。审核未通过者，返回修改后再行提交；审核通过者，打印为计划书纸质版（一式两份，双面打印），由依托单位审核并加盖单位公章后报送至自然科学基金委项目材料接收工作组。计划书电子版和纸质版内容应当保证一致。

向自然科学基金委提交和报送计划书截止时间节点如下：

1、提交计划书电子版截止时间为**2017年9月11日16点**（视为计划书正式提交时间）；

2、提交计划书电子修改版截止时间为**2017年9月18日16点**；

3、报送计划书纸质版截止时间为**2017年9月26日16点**。

请按照以上规定及时提交计划书电子版，并报送计划书纸质版，未说明理由且逾期不报计划书者，视为自动放弃接受资助。

附件：项目评审意见及修改意见表

国家自然科学基金委员会
生命科学部
2017年8月17日

3 奖励荣誉

3.1 2017 年湖北省优秀博士学位论文


陈 伟：

你的博士学位论文《水稻代谢组的生化及遗传基础研究》（指导老师：罗杰）被评为湖北省优秀博士学位论文。

湖北省人民政府学位委员会 湖北省教育厅

二〇一七年三月二日

3.2 2016 年湖北省自然科学优秀学术论文特等奖


湖北省自然科学优秀学术论文

證 書

罗 杰
陈兴明
陈 伟 同志撰写的 Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism

论文被评为湖北省第十六届自然科学优秀学术论文特等奖。

湖北省科学技术协会 湖北省人力资源和社会保障厅 湖北省科学技术厅

二〇一六年十二月