Exploring the genic resources underlying metabolites through mGWAS and mQTL in wheat: From large-scale gene identification and pathway elucidation to crop improvement

Jie Chen1,2, Mingyun Xue1,2, Hongbo Liu1, Alisdair R. Fernie3 and Wei Chen1,2,*

1National Key Laboratory of Crop Genetic Improvement and National Center of Plant Gene Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China
2College of Plant Science and Technology, Huazhong Agricultural University, Wuhan 430070, China
3Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm 14476, Germany
*Correspondence: Wei Chen (chenwei0609@mail.hzau.edu.cn)

https://doi.org/10.1016/j.xplc.2021.100216

ABSTRACT

Common wheat (Triticum aestivum L.) is a leading cereal crop, but has lagged behind with respect to the interpretation of the molecular mechanisms of phenotypes compared with other major cereal crops such as rice and maize. The recently available genome sequence of wheat affords the pre-requisite information for efficiently exploiting the potential molecular resources for decoding the genetic architecture of complex traits and identifying valuable breeding targets. Meanwhile, the successful application of metabolomics as an emergent large-scale profiling methodology in several species has demonstrated this approach to be accessible for reaching the above goals. One such productive avenue is combining metabolomics approaches with genetic designs. However, this trial is not as widespread as that for sequencing technologies, especially when the acquisition, understanding, and application of metabolic approaches in wheat populations remain more difficult and even arguably underutilized. In this review, we briefly introduce the techniques used in the acquisition of metabolomics data and their utility in large-scale identification of functional candidate genes. Considerable progress has been made in delivering improved varieties, suggesting that the inclusion of information concerning these metabolites and genes and metabolic pathways enables a more explicit understanding of phenotypic traits and, as such, this procedure could serve as an -omics-informed roadmap for executing similar improvement strategies in wheat and other species.

Key words: wheat, metabolomics, mGWAS, mQTL, pathway elucidation

Chen J., Xue M., Liu H., Fernie A.R., and Chen W. (2021). Exploring the genic resources underlying metabolites through mGWAS and mQTL in wheat: From large-scale gene identification and pathway elucidation to crop improvement. Plant Comm. 2, 100216.

INTRODUCTION

Metabolomics is an -omics tool that emerged following several other -omics, such as genomics, transcriptomics, and proteomics. This -omics approach has been followed for more than 20 years (Alseekh and Fernie, 2018), and techniques have become mainstream and protocols feasible even at a commercial level. Therefore, we will not detail the basic protocols or techniques, since they have been extensively described previously (Lisec et al., 2006; Vos et al., 2007; Kruger et al., 2008; Cajka and Fiehn, 2014). Common wheat (Triticum aestivum L.), also known as bread wheat, is a leading cereal crop ultimately accounting for approximately 20% of the calories consumed by humans (He et al., 2018). However, the genomic reference for this hexaploid crop has only recently become available (IWGSC, 2018), a fact that has significantly constrained fundamental research in this species. For example, it took several years from mapping the wheat vernalization gene VRN1 onto a chromosomal interval (Galiba et al., 1995; Dubcovsky et al., 1998; Barrett et al., 2002; Iwaki et al., 2002) to finally obtaining the gene by positional cloning (Yan et al.,...
Plant Communications

2003). Benefiting from advanced technologies, multiple -omics applications, including genomics, transcriptomics, proteomics, and metabolomics (Bilykioğlu et al., 2018; IWGSC, 2018; Xiang et al., 2019; Shi et al., 2020c; Ma et al., 2020), are available in the -omics toolbox to interpret the molecular alterations, and to contribute the genic resources, of wheat plants encountering various environmental conditions. Among these approaches, metabolomics was considered to be the bridge between genotypes and phenotypes (Fiehn, 2002) and has been widely applied to profile numerous wheat samples (Saia et al., 2019). However, early reports largely depicted metabolic adjustments rather than effectively pinpointing the underlying responsible genes, as has long been achieved for other crops (Albinsky et al., 2010; Gong et al., 2013; Wen et al., 2014; Chen et al., 2018). One way to accomplish this at large scale would be the combination of metabolomics and genetics, providing numerous candidate loci that may affect the relative contents of metabolites, and the potential responsible genes have thus been postulated and validated (Fernie and Tohge, 2017). In this review, we intend to introduce this predominant application of metabolomics in common wheat. The subsequently unveiled candidate genes and delineated metabolic pathways will help us to achieve the ultimate goal of wheat crop improvement.

THE APPLICATION OF METABOLOMICS IN WHEAT SHOULD DELIVER THE SWIFT DETECTION OF A VAST CATEGORY OF METABOLITES

The term “metabolomics” was coined at the end of the last century in parallel to the terms like “genomics” and “transcriptomics” (Oliver, 1998), and was subsequently defined as the investigation of all metabolites in an organism that requires simultaneous measurement of all metabolites in a given biological system (Dixon and Strack, 2003). However, it has somehow been estimated that there are in excess of 200 000 metabolites in the plant kingdom (Windsor et al., 2005), with some recent estimates suggesting that the number is close to a million (Wang et al., 2019). These metabolites are, however, distributed in a species-specific manner, with any given species predicted to contain over 5000 metabolites (Fernie et al., 2004). The uncertainty concerning specificity, numbers, and contents of metabolites within a given species/organism at a certain developmental stage renders comprehensive detection and identification a tough task. For instance, it has long been recognized that glucosinolates were mainly found in the Brassicaceae species (or the “cabbage clade” of Brassicales) (Halkier and Du, 1997); however, the total number of glucosinolate metabolites remains uncertain (Blazević et al., 2020). Likewise, the varied chemo-decorative forms of benzoazinoids, which represent an important class of plant defense metabolites that are distributed mainly in the Poaceae species, including wheat, maize, and rye, require extensive efforts to unveil their metabolic routes within these species (Bruijn et al., 2018).

Ideally, metabolomics should be rapid, unbiased, and comprehensive to fulfill the complete detection and identification of the metabolome from a given species. Regrettably, no single current analytical technique meets all these criteria at the same time. Of the two main technological platforms, NMR-based methodologies require the metabolites to be highly abundant or, alternatively, to be extracted from a large amount of tissue samples (Alseekh and Fernie, 2018). By contrast, mass spectrometry (MS)-based techniques can more sensitively detect the analytes. In addition, it can be coupled to either gas chromatography (GC) or liquid chromatography (LC), adding another dimension to differentiating the metabolite signals (Alseekh and Fernie, 2018). By employing the above-mentioned technological platforms, one could acquire hundreds of metabolic analytes from wheat samples using the NMR-based (Byeon et al., 2020) or MS-based techniques (Ameye et al., 2020; Wang et al., 2020). Although this number is far from covering the whole metabolome of a given species (hundreds detected versus thousands predicted), this is likely reflective of what can be currently achieved in wheat. Indeed, numerous experiments have been conducted to profile the wheat metabolomic landscape in recent years (for review, see Saia et al., 2019); these surveys may be inspired by pre-existing metabolomics reports. For instance, the Arabidopsis case revealed the detailed gene clusters responsible for root-specific metabolic pathways, wherein the specific metabolites involved selectively influence the root rhizosphere microbes (Huang et al., 2019). Following this clue, the interactions of metabolites in the wheat root rhizosphere exudates with soil microbes were respectively inspected using an LC–MS-based technique (Rieusset et al., 2021) and GC–MS platform (Iannucci et al., 2021). Although these wheat metabolomics examinations seldom provide candidate genes for swift functional validation, as commented elsewhere (Saia et al., 2019), recognizing the metabolic profiles of wheat samples under various environmental conditions or developmental stages provides a requisite knowledge for parallel and in-depth experimental designs. In the current review, we will not simply compile the wheat metabolomics reports to date, since this has been recently accomplished (Saia et al., 2019). Instead, we intend to probe how the underlying genes responsible for the content variation of metabolites were unveiled, and introduce some details of efficiently revealing the gene–metabolite linkage by combining the metabolomics approaches with genetic designs.

METABOLITE GENOME-WIDE ASSOCIATION STUDIES AND METABOLITE QUANTITATIVE TRAIT LOCI ARE POWERFUL WAYS TO DISSECT THE GENETIC ARCHITECTURES UNDERLYING A HIGHLY VARIED METABOLOME

Plant metabolites are indispensable for the plant itself, both for supporting growth and for its interactions with the surrounding environment (Saito and Matsuda, 2010; Huang et al., 2019). Some of these metabolites are also necessary for humans, as they constitute our nutritional supply (Luca et al., 2012; Martin and Li, 2017). Hence, it is of vital importance to elucidate the functional genes associated with plant metabolism. In comparison with the complexity of phenotyping only a small number of field phenotypes (Yang et al., 2020), one can easily profile the relative contents of multiple metabolites, which may
Exploring genic resources underlying metabolites

directly or indirectly relate to the phenotypes of interest. Therefore, identifying candidate functional genes for content variation of metabolites would lead us to a better understanding of important crop phenotypes. For instance, the metabolite trigonelline (N-methyl nicotinic acid) and the biosynthetic gene underlying its production were found to be responsible for rice grain width (Chen et al., 2016). Similarly, in tomato, a key flavonoid-regulating MYB transcription factor that is responsible for the fruit color (Adato et al., 2009) was found to be indirectly selected during the tomato improvement process (Zhu et al., 2018). A wheat MYB transcription factor not only positively regulates the flavonoid pathway that forms red kernels (Himi and Noda, 2005), but also positively elevates the relative content of abscisic acid that normally suppresses seed germination (Ju et al., 2019). The work of Lang et al. (2021) added an explanation for the long-questioned association between red-kernel and pre-harvest sprouting traits of wheat and in doing so provided a new potential target for breeding.

To unveil the candidate genes underlying metabolic routes, a reverse genetic strategy via homologous sequence alignment would be a straightforward scheme. An example is that of the benzoazoxinoids, a class of plant defense metabolites that are distributed mainly in the Poaceae species. Several Bx genes responsible for the biosynthesis of benzoazoxinoids have been cloned in wheat (Nomura et al., 2002, 2003; Sue et al., 2006, 2011; Li et al., 2018a), mainly via a homologous sequence alignment strategy against their respective maize orthologs. However, this procedure is not workable when wheat Bx genes have low sequence homology to their maize orthologs (Li et al., 2018a), nor is it likely to be applicable if the Poaceae species have specialized benzoazoxinoid metabolites (Brujin et al., 2018). Meanwhile, the homologous sequence alignment-based procedure would likely cause an unintended increase in labor. For instance, a total of 29 potential SNAT genes from grapevine that contain the key structural domain (Pfam00583) were generated, and of these, the authors tested five, and only one of the gene products was validated to have the designated enzymatic activity (Yu et al., 2019). To this end, we may get an even smaller chance in common wheat, given that it is a hexaploid species harboring an approximately 16 Gb genome. One way to potentially circumvent such an obstacle is the combination of metabolomics and forward genetics designs, as the enormous tested samples may potentially link genic alterations to metabolic variations of interest. Indeed, combinations of genetic approaches such as quantitative trait loci (QTL) and genome-wide association study (GWAS) with metabolomic profiling have been widely applied in order to identify the functional genes underlying the content variation of metabolites (Table 1). These methodologies are termed as mQTL (metabolite QTL) or mGWAS (metabolite GWAS), respectively. mQTL tools were first utilized in tomato (Schauer et al., 2006) and Arabidopsis (Keurentjes et al., 2006). Following these pioneering studies, the mQTL approach was broadly adapted to address the diverse aspects of genetics toward plant metabolomes (Rowe et al., 2008; Gong et al., 2013; Joseph et al., 2015), providing valuable insights into the genetic and biochemical bases of metabolic traits.

However, given that linkage mapping populations are usually derived from merely two parents, the results from their study are clearly not scalable to investigate the tremendous chemical and content diversity of metabolites from diverse germplasm resources of a given species. This prompted researchers to utilize mGWAS to probe the genetic contributions to metabolic diversity by simultaneously evaluating vast numbers of accessions. Pioneers in this field inspected the quantitative and qualitative variations in glucosinolates in Arabidopsis (Kliebenstein et al., 2001), finding that variations in 34 glucosinolate metabolites occurred in 39 Arabidopsis ecotypes. A subsequent, follow-up study further examined over 40 glucosinolates in 96 Arabidopsis accessions, revealing additional genes that may control natural variation in these traits (Chen et al., 2011). In recent years, mGWAS methodology has been widely utilized in major crops, including tomato (Tieman et al., 2017), maize (Wen et al., 2014), rice (Matsuda et al., 2015), barley (Zeng et al., 2020), and tea (Zhang et al., 2020). These studies provided numerous candidates underlying substantial taste and nutritional traits.

ANNOTATED GENOMES, GENOTYPED POPULATIONS, AND REQUISITE KNOWLEDGE ARE KEY ELEMENTS TO IDENTIFY CANDIDATE GENES FROM mGWAS/mQTL OUTPUT

Technically, the direct output generated from mQTL/mGWAS is merely linkages/associations between chromosomal locations and metabolite contents, and as such, further steps should be executed to possibly identify the potential candidate genes underlying their respective linkages/associations. The first and most facile element for candidate gene identification is the availability of genomic information, which is an essential, if not the only, way to sift through the possible genes included within the genomic interval (in the case of mQTL results) or adjacent to the associated marker loci (in the case of mGWAS results). For example, wheat mQTL surveys have been deployed for several years (Hill et al., 2013, 2015). In these two cases, 205 compounds (93 of which were unknown metabolites) (Hill et al., 2013) and 558 mass features (197 were identified) (Hill et al., 2015) were respectively profiled in 179 doubled haploid lines using GC–MS and LC–MS, respectively. However, these two investigations did not yield any candidate genes regarding the content variations of profiled metabolites, largely due to the absence of reference genomic information, which was publicly available only afterward (IWGSC, 2018). Following these pioneering wheat mQTL studies, a wheat mGWAS revealed possible candidate genes for the involved metabolic traits (Matros et al., 2017), benefiting from partially pre-released wheat genomic data. Recently, a wheat mGWAS successfully identified 26 candidate genes with high confidence, of which candidates the authors enzymatically validated two and accordingly unveiled the flavonoid metabolism pathway of wheat (Chen et al., 2020). In parallel, this team conducted a wheat mQTL investigation (Shi et al., 2020c), yielded and substantiated candidates, and probed the possibility of linkage between the wheat yield and the contents of certain metabolites. The brief history and progress of wheat mGWAS/mQTL studies (Hill et al., 2013, 2015; Matros et al., 2017; Chen et al., 2020; Shi et al., 2020c) clearly revealed the indispensable role of the availability of genomic reference information for assigning possible candidates. However, access to a reference genome does not
necessarily render the candidate gene assignment process facile. It is important to note that when we assign potential candidate genes to mQTL/mGWAS outcomes, we are postulating the probable genes encoding products that may enzymatically catalyze, translocate, or in some other way regulate the linked/associated metabolites. For this purpose, genes located within or around the mapped intervals/associated linkage disequilibrium blocks were evaluated, including their annotations and reports regarding their orthologs in other species, along with the chemical structures and pathway architectures of designated metabolites (Chen et al., 2014, 2020). For instance, trigonelline has long been thought to be synthesized by methylation from nicotinic acid (Joshi and Handler, 1960), indicating that the biosynthetic enzyme for this key step is probably a methyl transferase. In a rice mGWAS (Chen et al., 2014), LOC_Os02g57760 was annotated as a methyl-transferase-encoding gene and was located within a linkage disequilibrium block of the leading SNP associated with the content variation of trigonelline. Thereafter, this gene was designated and proved as a candidate for trigonelline (Chen et al., 2014, 2016). Likewise, a flavonoid glycosyl-transferase-encoding gene located within the QTL interval was assigned and enzymatically validated to be the candidate for apigenin 7-O-rutinoside, a glycosyl-decorated flavonoid, in a recent wheat mQTL study (Shi et al., 2020c). Two homoeologous wheat genes (TraesCS4B01G371700 and TraesCS4D01G365800) were selected as candidates behind the accumulation of the metabolite sucrose (Chen et al., 2020). Their ortholog in Arabidopsis AT3G19940, which is also known as STP10, encodes a high-affinity hexose transporter (Rottmann et al., 2016). These examples thus collectively demonstrate that the approach enables us to identify possible candidate genes from mQTL/mGWAS outputs, once we have a reference genome, a genotyped population, and requisite biochemical knowledge concerning the target metabolites. For wheat, the reference genomes of hexaploid wheat (IWGSC, 2018) and its progenitors (Avni et al., 2017; Ling et al., 2018; Maccaferri et al., 2019) have been released; numerous genotyped segregating populations/natural accessions are available (He et al., 2016; Guo et al., 2017; Chen et al., 2019, 2020); and an enormous range of functional genes for diverse metabolites can be postulated via reference to species such as Arabidopsis (http://www.arabidopsis.org), rice (http://rice.plantbiology.msu.edu), and maize (http://www.maizegdb.org). Thus, the stage is clearly set to perform mQTL analyses and mGWASs in wheat (Figure 1).

DELINEATING THE METABOLIC PATHWAYS IS ESSENTIAL TO EXPLAIN OR EXPLORE THE CONNECTIONS BETWEEN METABOLITES AND THE END PHENOTYPES

As stated above, numerous candidate genes underlying the content variation of metabolite abundance could be obtained. This raises the question of how to systemically evaluate these candidates/metabolites and finally link them to end phenotypes. For phenotype-driven studies, on one hand, we need to explain how metabolite contents are linked to the designated traits. For example, copious Arabidopsis colorless seed mutants have been generated, suggesting their seed coat pigment abundance is probably impaired. These mutants were named transparent testa (tt), and the responsible mutated genes (TT genes) were demonstrated to encode candidates in the flavonoid/anthocyanin pathways (Lepiniec et al., 2006). The indication that flavonoids/anthocyanins are one major class of plant chromogenic compounds is probably why researchers tended to conduct targeted metabolomics measurements covering the flavonoids and/or anthocyanins in plant-color-related cases (Shi et al., 2020b; Jiao et al., 2020; Qiu et al., 2020). Subsequently, differential metabolites and corresponding responsible candidate genes were collected to explain the molecular mechanisms of color variation at the metabolic and,

| Species     | Method | Trait                              | Candidate                                      | Associated metabolites                                | Reference                  |
|-------------|--------|------------------------------------|-----------------------------------------------|-----------------------------------------------------|----------------------------|
| Blueberry   | mGWAS  | volatile organic compounds         | linalool synthase and α-terpineol synthase genes | linalool, limonene, and eucalyptol                     | (Ferrão et al., 2020)     |
| Barley      | mGWAS  | grain oligosaccharide content      | acid β-fructofuranosidase genes                | fructan                                              | (Matros et al., 2021)     |
| Arabidopsis | mGWAS  | resistance to Ralstonia solanacearum | CYP735A1, CKX2, and CKX4                       | cytokinin                                            | (Alonso-Díaz et al., 2021) |
| Maize       | mGWAS  | salt-induced osmotic stress        | ZmCS3, ZmUGT, and ZmCYP709B2                   | citrate and flavonoids                               | (Liang et al., 2021)      |
| Tomato      | mGWAS  | vitamin E content                  | genes involved in the chorismate–tyrosine pathway | tocochromanols                                      | (Burgos et al., 2021)     |
| Soybean     | mQTL   | seed oil content                   | diacylglycerol lipase, phospholipase, and acyl-CoA dehydrogenase genes | fatty acids                                          | (Li et al., 2021b)        |
| Rice        | mQTL   | insect resistance                  | Glucosyl-transferase genes                     | flavonoid O-glucosides                               | (Yang et al., 2021)       |
| Carrot      | mQTL   | flavor-associated sabinene         | Terpene synthase genes                         | sabinene, α-thujene, α-terpinene, γ-terpinene, terpinen-4-ol, and 4-carene | (Reichardt et al., 2020)  |

Table 1. Recent examples of functional genes identified from mGWAS and mQTL.
probably, the transcriptomic levels (Li et al., 2020; Zhan et al., 2020).

For phenotype-independent metabolomics surveys, on the other hand, we need to establish the pathway through experimental output and further explore the possible biological functions of the metabolites themselves. The establishment of a pathway is essentially achieved in the same manner as that for identifying candidate genes illustrated above. We need to take the existing knowledge of pathway architectures, chemical structures of metabolites, and reported candidate genes into consideration. In a recent wheat mGWAS (Chen et al., 2020), the authors assigned a glycosyltransferase-coding gene as a candidate gene regulating tricin abundance, and an acyl-transferase-coding gene for tricin O-malonyl hexoside abundance. The pathway was primarily established to be the transformation from tricin to tricin...
Exploring genic resources underlying metabolites

Plant Communications

7-O-glucoside and finally to tricin 7-O-malonyl glucoside, which was sequentially catalyzed by the two designated gene products. The authors also explored the possible involvement of this pathway in lignin formation through the by-product tricin 4'-O-glucoside (Chen et al., 2020). Similarly, Chen et al. (2018) established a flavonoid decoration pathway in a rice mQTL survey and explored the O-methylated apigenin that may, similar to its structural analog sakuranetin (Kodama et al., 1992), confer disease-resistance activity in rice (Chen et al., 2018). Collectively, the metabolic pathways that were established from the specified metabolites and underlying candidates could be used to explain the phenotypic differences, or to explore the potential connections between comprised metabolites and seemingly relevant traits.

GENETIC STATISTICAL POWER AND METABOLIC DETECTION CAPACITY NEED TO BE IMPROVED IN FUTURE mGWAS/mQTL STUDIES

Although stand-alone mQTL and mGWAS can be powerful tools in the large-scale identification of candidate genes and are almost instantly applicable in wheat, it would be imprudent to neglect their drawbacks or the use of other multi-omics approaches. The main disadvantages of mQTL and mGWAS, at the level of genetic design, are the insufficient variant coverage beyond the two parents from biparental populations (Chen et al., 2018), hidden population structures (Flint-Garcia et al., 2003), and inadequate power to detect rare variations (Brescghello and Sorrells, 2006) within the natural accessions. Accordingly, utilizing multi-parental populations such as the NAM (nested association mapping) population (Tian et al., 2011; Jordan et al., 2018) or the MAGIC (multi-parent advanced generation intercross) population (Dell’Acqua et al., 2015) could compensate for the deficient variant source of biparental populations at some degree. Alternatively, one could combine the mGWAS and mQTL (Shi et al., 2020a) to enable a more powerful dissection of candidate genes.

Aside from utilizing varied combinations of populations to maximize the statistical power that can be leveraged in gene identification, the detection capacity of metabolomics tools affords the overall signals to be analyzed. Two strategies are usually available, including the targeted metabolomics that normally detects hundreds of known metabolites and the untargeted metabolomics that can generate more signals, yet most of the mass features cannot be assigned to exact chemical structures (Roepenack-Lahaye et al., 2004). Given that knowledge concerning the chemical skeletons of metabolites and the underlying pathway architectures is indispensable for identifying responsible candidate genes, the low level of identity of the detected analytes leads to a relatively low overall efficacy when it comes to assigning responsible candidates. Meanwhile, the low coverage (hundreds detected versus thousands predicted) of targeted metabolomics over the metabolome of given plant organisms has lessened the efficacy relative to uncovering the metabolic pathways and underlying genes. Given this, we have developed a widely targeted metabolomics method that combines merits from both targeted and untargeted metabolomics (Chen et al., 2013), which has been successfully utilized in numerous species (Chen et al., 2014; Wen et al., 2014; Song et al., 2015; Geng et al., 2016; Li et al., 2017, 2021a; Wang et al., 2017; Shareeb et al., 2018; Cao et al., 2019; Lee et al., 2020; Liu et al., 2020; Zeng et al., 2020; Fu et al., 2021; Nie et al., 2021). By applying this protocol, we are now able to detect some 5000 metabolites from bread wheat samples within a single run that typically takes 20 min (data not published). However, two major drawbacks are hindering the instant implementation of this methodology to other species. The first would be the nature of the species or organism specificity of metabolites (Fernie et al., 2004). Accordingly, we have constructed the MS2T (MS2 spectral tag) library covering a collection of wheat samples spanning the whole growth period to constitute a maximum metabolic detection capacity, which is time consuming and profession dependent. Misuse of MS2T libraries will likely cause a significant decrease in metabolites detected. The other issue that needs to be improved on is deciphering the chemistry. We are capable of identifying some 2000 metabolites within the above-mentioned wheat MS2T library (data not published), indicating that the remaining 3000 are currently unknown analytes. However, this does not indicate that we should pre-exclude all machine reads that could not be allocated to exact chemical structures in wheat metabolomics studies. Indeed, genetics can be a powerful tool by which to identify unknown metabolites. For example, in a rice mGWAS report (Chen et al., 2016), the authors showed that the identification of unknown metabolites and the assignment of candidate genes could be complementary to one another. The principal logic to instruct this is that the different decorations derived from the same chemical skeletons (tryptamine derivatives in this case) could be present in proportional levels and comprise the same major ion fragments, which enables us to link the unknown metabolites to the known ones. In accordance, we have discovered several sub-networks, including flavonoids, amino acids, nucleotides, and indole skeleton-involved metabolites, in a recent wheat mGWAS (Chen et al., 2020). Further downstream efforts could be paid to these clusters of metabolites in order to fully delineate their positions within metabolic networks.

BREEDING WHEAT TO BE A BETTER CROP ASSISTED BY METABOLOMICS APPROACHES AS INDICATION, PREDICTION, AND NAVIGATION TOOLS

Wheat, as an indispensable food crop, is facing the eternal breeding goals concerning yield and quality. However, the conventional breeding processes of wheat have been constrained by linkage drags and relatively low recombination frequencies rendering them time consuming and of low efficiency and predictability (Holland, 2007). Metabolomics has long been regarded as a bridge between genotypes and phenotypes (Fiehn, 2002). One form of the underlying contact is that certain metabolites contribute largely to phenotypes. In this regard a widely accepted concept is the pivotal position of phytohormones in plant growth and in response to stress (Yu et al., 2020), wherein the different decorative forms of phytohormones would lead to role transition (active or not) of these metabolites (Korasick et al., 2013). To a lesser extent, trehalose 6-phosphate has recently been implicated to play roles in a wide range of developmental processes (Figueroa and Lunn, 2016; Li et al., 2019), while a specific
has long been observed (Groos et al., 2002). At this scope, the association between these two agronomic traits phenylacylated flavonoids (Tohge et al., 2016) confers UV profiles of maize testcrosses (Riedelsheimer et al., 2012). The prediction was revealed following the analysis of the metabolic metabolomics data. For instance, a good performance machine reads, which is bridged by proper modeling of the complex phenotypes through the numerous and less discernable Zhu et al., 2018) enable the prediction of discernable yet pivotal metabolites and the underlying metabolic pathways could be subjected to direct detection and selection at early stages of plant development instead of tedious physiological tests when breeders intend to acquire the related agronomic traits.

Theoretically, the metabotype–phenotype ties (Adato et al., 2009; Zhu et al., 2018) enable the prediction of discernable yet complex phenotypes through the numerous and less discernable machine reads, which is bridged by proper modeling of the metabolomics data. For instance, a good performance prediction was revealed following the analysis of the metabolic profiles of maize testcrosses (Riedelsheimer et al., 2012).

Similarly, a recent study showed that metabolic profiling of the wheat leaf and spike bracts metabolomes independently revealed that they were good predictors of grain yield (Vergara-Diaz et al., 2020b). Another study showed that remote hyperspectral imaging of wheat allows one to estimate metabolite content, potentially rendering the development of biomarkers even more powerful (Vergara-Diaz et al., 2020a). That said, the prediction power of metabolomics data in wheats is still controversial (Zhao et al., 2015; Shi et al., 2020c) and it is thus perhaps too early to discuss now.

Aside from contributing to various agronomic traits, the contents of metabolites themselves could be direct breeding goals (Martin and Li, 2017). One pre-eminent example is the introduction of provitamin A into the carotenoid-free rice endosperm via metabolic engineering (Ye et al., 2000; Paine et al., 2005). Following this, several crop species, such as cassava, maize, and potato, have been targeted for vitamin biofortification (Bai et al., 2011; Sayre et al., 2011). The improvement of vitamin contents has also been a concern in wheat breeding (Table 2). Indeed, for instance, wheat germ oils contain the highest concentration of vitamin E of all species tested (Trela and Szymańska, 2019). Although substantial efforts have been made in elucidating the metabolic pathway of vitamin metabolism in other crops (Zhou et al., 2012;}

| Vitamin | Related metabolites | Key candidate genes | Orthologs |
|---------|---------------------|---------------------|-----------|
| VA      | β-carotene, provitamin A, β-cryptoxanthin, retinol derivatives | RALDH, retinal dehydrogenase; REH, retinyl ester hydrolase; PSY*, phytoene synthase | 7A1357000; 7B1296800; 7D1285000 |
| VB1     | thiamine, thiamin pyrophosphate | TMP-PPase, thiamin-phosphate pyrophosphorylase; THI*, thiamine thiazole synthase | 7A0916800; 7B0760700; 7D0879600 |
| VB2     | riboflavin, flavin adenine dinucleotide, flavin mononucleotide | RibA, GTP cyclohydrolase II; RibB, 3,4-dihydroxy-2-butane 4-phosphate synthase; PyrR*, pyrimidine reductase | 6A0988100; 6B1211700; 6D0873600 |
| VB3     | niacin, niacinamide, nicotinamide adenine dinucleotide (phosphate) | NadA, quinolinate synthase; nitrate reductase* | 6A0038200; 6B0056200; 6D0042700 |
| VB5     | pantothenic acid, pantetheine, pantetheine | PanB, 3-methyl-2-oxobutanoate hydroxymethyltransferase; PanK*, pantothenate kinase | 5A0779300; 5B0811000; 5D0737600 |
| VB6     | pyridoxine/pyridoxal/pyridoxamine 5'-phosphate | PDX*, pyridoxal 5'-phosphate synthase | 2A0661600; 2B0746500; 2D0617200 |
| VB7     | biotin, biocytin | OTBS, desthiobiotin synthetase; BIO*, biotin synthase | 6A0389100; 6B0491900; 6D0338900 |
| VB9     | folic acid, tetrahydrofolic acid derivatives | FPGS, folylpolyglutamate synthetase; DHFR*, dihydrofolate reductase; DHFS, dihydrofolate synthase; DHPS, dihydropteroteroate synthase | 2A1204700; 2B1374200; 2D1159900 |
| VC      | ascorbic acid | PMI, phosphomannose isomerase; Ala, aldonolactonase; GGP*, GDP-L-galactose phosphorylase | 4A0537700; 4B0239000; 4D0202300 |
| VE      | tocopherols and tocotrienols | VTE1*, tocopherol cyclase; VTE2, homogentisate phytol transferase; VTE3, dimethyl-phytylquinol methyl transferase; VTE4, γ-tocopherol C-methyl transferase | 1A0584500; 1B0677200; 1D055800 |

Table 2. Key genes underlying the vitamin contents that may subjected to wheat grain nutritional improvement.

The wheat candidate genes are generated by sequence alignment against reported genes marked by asterisks. Wheat gene IDs are abbreviations based on the IWGSC Chinese Spring genome v.2.1 annotation. For instance, 7A1357000 denotes TraesCS07G7A1357000.
Quadrana et al., 2014; Liu et al., 2017; Wang et al., 2018; Schuy et al., 2019; Zhang et al., 2019), less has been achieved in wheat (Li et al., 2018b; Watkins et al., 2019). In view of this, key candidate genes could be swiftly selected and validated through multi-omics approaches including metabolomics, providing molecular resources for breeding wheat to be a more nutritional food. Alternatively, the introgression of elevated vitamin contents (or other less discernable metabolites as immediate breeding targets) could be directly conducted, without knowing the precise genes or markers in advance (Figure 2). Following this route, combinations of widely planted cultivars as acceptors and high-vitamin accessions as donors are first established. Subsequently, several backcrosses to restore the genetic background and finally self-crossing to gain homozygosity are conducted (Wing et al., 2018). During these steps, metabolic measurements represent a type of functional selection that can be applied to half of the wheat kernels, ensuring that we can maintain the high-vitamin phenotypes. Meanwhile, the counterparts of the wheat kernels are grown for the next phase of breeding. The whole process can be further accelerated by speed breeding techniques (Ghosh et al., 2018), wherein the duration of a generation is merely 3 months rather than a full year. Finally, simultaneous utilization of wheat SNP chips (Cavanagh et al., 2013; Wang et al., 2014; Allen et al., 2016; Boeven et al., 2016; Winfield et al., 2016; Rasheed and Xia, 2019) to genotype the acceptors, donors, and progeny would enable the development of functional markers for the desired metabolic traits, and thereby allow the underlying candidate genes to be subsequently cloned (Figure 2).

CONCLUSIONS AND PERSPECTIVES

Fundamental wheat research has been considerably boosted since the availability of a reference genomic sequence, which greatly benefited candidate gene identification. In view of the mQTL/mGWAS protocols, the inherent flaws of segregating populations or natural collections require a more comprehensive genetic design, particularly in the case of a species with a complex genome such as bread wheat, as well as a more powerful statistical method, to enable a more accurate detection of inherited loci for metabolites. The demands awaiting improved computational capacity toward large-scale deciphering of the unknown analytes into possible chemical structures (Blazenović et al., 2018; Kind et al., 2018), which are also pivotal for candidate gene identification. Meanwhile, varied wheat samples encompassing different spatiotemporal organisms should be included to cover the tissue-specific metabolites and the underlying candidate genes. Finally, new metabolic technologies such as metabolome-scale labeling (Tsugawa et al., 2019) or single-cell metabolomics (Souza et al., 2020), which has been similarly applied in other -omics fields, will likely improve our understanding of the metabolic pathways of more specialized organisms. Thus, two major challenges lie ahead. The first is the improvement of the coverage of the metabolome. Second, we need to move toward functional metabolomics that delimits the biological roles of the molecules themselves. The framework presented here should allow advances toward both of these. Indeed, collectively, one could anticipate an acceleration in candidate gene identification and comprehensive metabolic pathway construction and that these molecular resources will ultimately assist in wheat crop improvement.

FUNDING

This work was supported by the National Natural Science Foundation of China (91935304, 31770329, and 32001541), the Huazhong Agricultural University Scientific & Technological Self-Innovation Foundation (2017RC006), the China Postdoctoral Science Foundation.
Exploring genic resources underlying metabolites (2018M642866 and 2021T140246), and the Hubei Provincial Natural Science Foundation (2020CFB149).

ACKNOWLEDGMENTS

The authors declare no conflict of interest.

REFERENCES

Adato, A., Mandel, T., Mintz-Oron, S., Venger, I., Levy, D., Yativ, M., Dominguez, E., Wang, Z., Vos, R.C.H. de., Jetter, R., et al. (2009). Fruit-surface flavonoid accumulation in tomato is controlled by a SIMYB12-regulated transcriptional network. PLoS Genet. 5:e1000777.

Albinsky, D., Kusano, M., Higuchi, M., Hayashi, N., Kobayashi, M., Fukushima, A., Mori, M., Ichikawa, T., Matsu, K., Kuroda, H., et al. (2010). Metabolomic screening applied to rice FOX Arabidopsis lines leads to the identification of a gene-changing nitrogen metabolism. Mol. Plant 3:135–142.

Allen, A.M., Winfield, M.O., Burridge, A.J., Downie, R.C., Bennow, H.R., Barker, G.L.A., Wilkinson, P.A., Coghill, J., Waterfall, C., Davarsi, A., et al. (2016). Characterization of a Wheat Breeders’ Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (Triticum aestivum). Plant Biotechnol. J. 15:390–401.

Alseekh, S., and Ferrone, A.R. (2018). Metabolomics 20 years on: what have we learned and what hurdles remain? Plant J. 94:933–942.

Ameye, M., van Meulebroek, L., Meuninck, B., Vanhaecke, L., Smagghe, G., Haesaert, G., and Audenaert, K. (2020). Metabolomics reveal induction of ROS production and glycosylation events in wheat upon exposure to the green leaf volatile z-3-hexenyl acetate. Front. Plant Sci. 11:586271.

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S.O., Gundlach, H., Hale, I., Mascher, M., Spannagl, M., Wiebe, K., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science 357:93–97.

Bai, C., Twyman, R.M., Farré, G., Sanahuja, G., Christou, P., Capell, T., and Zhu, C. (2011). A golden era—pro-vitamin A enhancement in diverse crops. In Vitro. Cell. Dev. Biol. Plant. 47:205–221.

Barrett, B., Bayram, M., and Kidwell, K. (2002). Identifying AFLP and microsatellite markers for vernalization response gene Vrn-B1 in hexaploid wheat using reciprocal mapping populations. Plant Breeding 121:400–406.

Burgos, E., Benen De Luca, M., Diouf, I., Haro, L.A. de, Albert, E., Sauvage, C., Tao, Z.J., Bermudez, L., Asis, R., Nesi, A.N., et al. (2021). Validated MAGIC and GWAS population mapping reveals the link between vitamin E content and natural variation in choline metabolism in tomato. Plant J. 105:907–923.

Biyiklioglu, S., Alpektin, B., Akpinar, B.A., Varella, A.C., Hofland, M.L., Weaver, D.K., Bothner, B., and Budak, H. (2018). A large-scale multimetics analysis of wheat stem solidity and the wheat stem sawfly feeding response, and syntenic associations in barley. Brachypodium, and rice. Funct. Integr. Genomics 18:241–259.

Blazević, I., Montaut, S., Burčul, F., Olsen, C.E., Burow, M., Rollin, P., and Agerbirk, N. (2020). Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. Phytochemistry 169:112100.

Blaženović, I., Kind, T., Ji, J., and Fiehn, O. (2018). Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. Metabolites 8:31.

Boeve, P.H.G., Longin, C.F.H., Leiser, W.L., Kollers, S., Eibmeyer, E., and Wurschum, T. (2016). Genetic architecture of male floral traits required for hybrid wheat breeding. Theor. Appl. Genet. 129:2343–2357.

Breseghello, F., and Sorrels, M.E. (2006). Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. Genetics 172:1165–1177.

Buijink, W.J.C. de., Gruppen, H., and Vincken, J.P. (2018). Structure and biosynthesis of benzoxazinoids: plant defence metabolites with potential as antimicrobial scaffolds. Phytochemistry 155:233–243.

Byeon, Y.S., Lee, D., Hong, Y.-S., Lim, S.-T., Kim, S.S., and Kwak, H.S. (2020). Comparison of physicochemical properties and metabolite profiling using 1H NMR spectroscopy of Korean wheat malt. Foods 9:1436.

Cai, T., and Fiehn, O. (2014). Comprehensive analysis of lipids in biological systems by liquid chromatography-mass-spectrometry. Trend. Anal. Chem. 61:192–206.

Cao, K., Li, Y., Deng, C.H., Gardiner, S.E., Zhu, G., Fang, W., Chen, C., Wang, X., and Wang, L. (2019). Comparative population genomics identified genomic regions and candidate genes associated with fruit domestication traits in peach. Plant Biotechnol. J. 17:1954–1970.

Cavanagh, C.R., Chao, S., Wang, S., Huang, B.E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G.L., Akhunova, A., et al. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc. Natl. Acad. Sci. U S A 110:8057–8062.

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

Chen, J., Hu, X., Shi, T., Yin, H., Sun, D., Hao, Y., Xia, Y., Luo, J., Ferrine, A.R., He, Z., et al. (2020). Metabolite-based genome-wide association study enables dissection of the flavonoid decoration pathway of wheat kernels. Plant Biotechnol. J. 18:1722–1735.

Chen, J., Wang, J., Chen, W., Sun, W., Peng, M., Yuan, Z., Shen, S., Xie, K., Jin, C., Sun, Y., et al. (2018). Metabolome analysis of multi-connected biparental chromosome segment substitution line populations. Plant Physiol. 178:612–625.

Chen, J., Zhang, F., Zhao, C., Lv, G., Sun, C., Pan, Y., Guo, X., and Chen, F. (2019). Genome-wide association study of six quality traits reveals the association of the TaRPP13L1 gene with flour color in Chinese bread wheat. Plant Biotechnol. J. 17:2106–2122.

Chen, W., Gao, Y., Xie, W., Gong, L., Liu, Y., Luo, J., Xia, X., Zhang, H., Dong, H., et al. (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. Nat. Genet. 46:714–721.

Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., Yu, S., Xiong, L., and Luo, J. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. Mol. Plant 6:1769–1780.

Chen, W., Wang, W., Peng, M., Gong, L., Gao, Y., Wan, J., Wang, S., Shi, L., Zhou, B., Li, Z., et al. (2016). Comparative and parallel genome-wide association studies for metabolic and agronomic traits in cereals. Nat. Commun. 7:12767.

Dell’Acqua, M., Gatti, D.M., Pea, G., Cavallaro, F., Coppens, F., Magris, G., Hlaing, A.L., Aung, H.H., Nelissen, H., Baute, J., et al.
Plant Communications

(2015). Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in Zea mays. Genome Biol. 16:167.

Dixon, R.A., and Strack, D. (2003). Phytochemistry meets genome analysis, and beyond. Phytochemistry 62:815–816.

Dubcovsky, J., Lijavetzky, D., Appendino, L., and Tranquilli, G. (1998). Comparative RFLP mapping of Triticum monococcum genes controlling vernalization requirement. Theor. Appl. Genet. 97:968–975.

Fernie, A.R., and Tohge, T. (2017). The genetics of plant metabolism. Annu. Rev. Genet. 51:287–310.

Fernie, A.R., Trehewey, R.N., Krotzky, A.J., and Willmitzer, L. (1997). The biosynthesis of glucosinolates. Mol. Cell Biol. 5:763–769.

Ferrão, L.F.V., Johnson, T.S., Benevenuto, J., Edger, P.P., Colquhoun, T.A., and Munoz, P.R. (2020). Genome-wide association of volatiles reveals candidate loci for blueberry flavor. New Phytol. 226:1725–1737.

Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. Plant Mol. Biol. 48:155–171.

Figueroa, C.M., and Lunn, J.E. (2016). A tale of two sugars: trehalose 6-phosphate and sucrose. Plant Physiol. 172:7–27.

Flint-Garcia, S.A., Thornsberry, J.M., and Buckler, E.S. (2003). Structure of linkage disequilibrium in plants. Annu. Rev. Plant Biol. 54:357–374.

Fu, A., Wang, Q., Mu, J., Ma, L., Wen, C., Zhao, X., Gao, L., Li, J., Shi, K., Wang, Y., et al. (2021). Combined genomic, transcriptomic, and metabolomic analyses provide insights into chayote (Sechium edule) evolution and fruit development. Hortic. Res. 8:36.

Galiba, G., Quarrie, S.A., Sutka, J., Morgounov, A., and Snape, J.W. (1995). RFLP mapping of the vernalization (Vrn1) and frost resistance (Fr1) genes on chromosome 5A of wheat. Theor. Appl. Genet. 90:1174–1179.

Geng, S., Misra, B.B., Armas, E. de, Huhman, D.V., Alborn, H.T., Sumner, L.W., and Chen, S. (2016). Jasmonate-mediated stomatal closure under elevated CO2 revealed by time-resolved metabolomics. Plant J. 88:947–962.

Ghareeb, M.A., Mohamed, T., Saad, A.M., Refahy, L.A.-G., Sobeh, M., Geng, S., Misra, B.B., Armas, E. de, Huhman, D.V., Alborn, H.T., Ghosh, S., Watson, A., Gonzalez-Navarro, O.E., Ramirez-Gonzalez, R.H., Yanes, L., Mendoza-Suárez, M., Simmonds, J., Wells, R., Rayner, T., Green, P., et al. (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. Nat. Protoc. 13:2944–2963.

Gong, L., Chen, W., Gao, Y., Liu, X., Zhang, H., Xu, C., Yu, S., Zhang, Q., and Luo, J. (2013). Genetic analysis of the met abolome exemplified using a rice population. Proc. Natl. Acad. Sci. U S A 110:20329–20335.

Gros, C., Gay, G., Perretant, M.-R., Gervais, L., Bernard, M., Dedryver, F., and Charmet, G. (2002). Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a white × red grain bread-wheat cross. Theor. Appl. Genet. 104:39–47.

Guo, Z., Chen, D., Alqudah, A.M., Roder, M.S., Ganal, M.W., and Schnurbusch, T. (2017). Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. New Phytol. 214:257–270.

Halkier, B.A., and Du, L. (1997). The biosynthesis of glucosinolates. Trends. Plant Sci. 2:425–431.

He, X., Lillemo, M., Shi, J., Wu, J., Bjornstad, Å., Belova, T., Dreisigacker, S., Duveiller, E., and Singh, P. (2016). QTL characterization of fusarium head blight resistance in CIMMYT bread wheat line Soru1. PLoS One 11:e0158052.

He, Z., Zhuang, Q., Cheng, S., Yu, Z., Zhao, Z., and Liu, X. (2018). Wheat production and technology improvement in China (in Chinese). J. Agric. 8:99–106.

Hill, C.B., Taylor, J.D., Edwards, J., Mather, D., Bacic, A., Langridge, P., and Roessner, U. (2013). Whole-genome mapping of agronomic and metabolic traits to identify novel quantitative trait Loci in bread wheat grown in a water-limited environment. Plant Physiol. 162:1266–1281.

Hill, C.B., Taylor, J.D., Edwards, J., Mather, D., Langridge, P., Bacic, A., and Roessner, U. (2015). Detection of QTL for metabolic and agronomic traits in wheat with adjustments for variation at genetic loci that affect plant phenology. Plant Sci. 233:143–154.

Himi, E., and Noda, K. (2005). Red grain colour gene (R) of wheat is a Myb-type transcription factor. Euphytica 143:239–242.

Holland, J.B. (2007). Genetic architecture of complex traits in plants. Curr. Opin. Plant Biol. 10:156–161.

Huang, A.C., Jiang, T., Liu, Y.-X., Bai, Y.-C., Reed, J., Qu, B., Goossens, A., Nutzmann, H.W., Bai, Y., and Osbourn, A. (2019). A specialized metabolic network selectively modulates Arabidopsis root microbiota. Science 364:eaau6389.

Iannucci, A., Canfora, L., Nigro, F., Vita, P., de, and Beleggia, R. (2021). Relationships between root morphology, root exudate compounds and rhizosphere microbial community in durum wheat. Appl. Soil Ecol. 158:103781.

Iwaki, K., Nishida, J., Yanagisawa, T., Yoshida, H., and Kato, K. (2002). Genetic analysis of Vm-B1 for vernalization requirement by using linked dCAPS markers in bread wheat (Triticum aestivum L.). Theor. Appl. Genet. 104:571–576.

IWGSC (The International Wheat Genome Sequencing Consortium). (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361:eaar7191.

Jiao, F., Zhao, L., Wu, X., Song, Z., and Li, Y. (2020). Metabolome and transcriptome analyses of the molecular mechanisms of flower color mutation in tobacco. BMC. Genomics. 21:611.

Jordon, K.W., Wang, S., He, F., Chao, S., Lun, Y., Paux, E., Sourdille, P., Sherman, J., Akhunova, A., Blake, N.K., et al. (2018). The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. Plant J. 95:1039–1054.

Joseph, B., Lau, L., and Kliebenstein, D.J. (2015). Quantitative variation in responses to root spatial constraint within Arabidopsis thaliana. Plant Cell 27:2227–2243.

Joshi, J.G., and Handler, P. (1960). Biosynthesis of trigonelline. J. Biol. Chem. 235:2981–2983.

Ju, L., Jing, Y., Shi, P., Liu, J., Chen, J., Yan, J., Chu, J., Chen, K.M., and Sun, J. (2019). JAZ proteins modulate seed germination through interacting with ABI5 in bread wheat and Arabidopsis. New Phytol. 223:246–260.

Keurentjes, J.J.B., Fu, J., Vos, C.H.R de., Lommen, A., Hall, R.D., Bino, R.J., van der Plas, L.H.W., Jansen, R.C., Vreugdenhil, D., and Koornneef, M. (2006). The genetics of plant metabolism. Nat. Genet. 38:842–849.

Kind, T., Tsugawa, H., Ćajka, T., Ma, Y., Lai, Z., Mehta, S.S., Wohlgemuth, G., Barupal, D.K., Showalter, M.R., Arita, M., et al. (2018). Identification of small molecules using accurate mass MS/MS search. Mass. Spectrom. Rev. 37:513–532.

Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., and Mitchell-Olds, T. (2001). Genetic control of
Exploring genic resources underlying metabolites

natural variation in Arabidopsis glucosinolate accumulation. Plant Physiol. 126:811–825.

Kodama, O., Miyakawa, J., Akatsuka, T., and Kiyosawa, S. (1992). Sakuranetin, a flavanone phytoalexin from ultraviolet-irradiated rice leaves. Phytochemistry 31:3807–3809.

Korasick, D.A., Enders, T.A., and Strader, L.C. (2013). Auxin biosynthesis and storage forms. J. Exp. Bot. 64:2541–2555.

Kruger, N.J., Troncoso-Ponce, M.A., and Ratcliffe, R.G. (2008). 1H NMR metabolite fingerprinting and metabolic analysis of perchloric acid extracts from plant tissues. Nat. Protoc. 3:1001–1012.

Lang, J., Fu, Y., Zhou, Y., Cheng, M., Deng, M., Li, M., Zhu, T., Yang, J., Guo, X., Gui, L., et al. (2021). Myb10-D confers PHS-3D resistance to pre-harvest sprouting by regulating NCED in ABA biosynthesis pathway of wheat. New Phytol. 230:1940–1952.

Lepiniec, L., Debeaujon, I., Routaboul, J.M., Baudry, A., Pourcel, L., Nesl, N., and Caboche, M. (2006). Genomics and biochemistry of seed flavonoids. Annu. Rev. Plant Biol. 57:405–430.

Lee, S., Oh, D.G., Singh, D., Lee, J.S., Lee, S., and Lee, C.H. (2021). Metabolomics-driven gene mining and natural variation in Arabidopsis glucosinolate accumulation. Plant Commun. 2:230–235.

Lepiniec, L., Debeaujon, I., Routaboul, J.M., Baudry, A., Pourcel, L., Nesl, N., and Caboche, M. (2006). Genomics and biochemistry of seed flavonoids. Annu. Rev. Plant Biol. 57:405–430.

Lee, S., Oh, D.G., Singh, D., Lee, J.S., Lee, S., and Lee, C.H. (2020). Exploring the metabolomic diversity of plant species across spatial (leaf and stem) components and phylogenic groups. BMC. Plant Biol. 20:39.

Li, B., Forster, C., Robert, C.A.M., Zust, T., Hu, L., Machado, R.A.R., Berset, J.D., Handrick, V., Knauer, T., Hensel, G., et al. (2018a). Convergent evolution of a metabolic switch between aphid and caterpillar resistance in cereals. Sci. Adv. 4:eaat6797.

Li, J., Liu, J., Wen, W., Zhang, P., Wan, Y., Xia, X., Zhang, Y., and He, Z. (2018b). Genome-wide association mapping of vitamins B1 and B2 in common wheat. New Phytol. 216:699–719.

Li, S., Deng, B., Tian, S., Guo, M., Liu, H., and Zhao, X. (2021a). Metabolic and transcriptomic analyses reveal different metabolite biosynthesis profiles between leaf buds and mature leaves in Ziziphus jujuba mill. Food Chem. 347:129005.

Li, T., Fan, Y., Qin, H., Dai, G., Li, G., Li, Y., Wang, J., Yin, Y., Chen, F., Qin, X., et al. (2020). Transcriptome and flavonoid metabolites metabolic analysis identifies regulatory networks and hub genes in black and white fruits of Triticum aestivum (leaf and stem) components and phylogenic groups. BMC. Plant Biol. 20:39.

Liu, Y., Lv, J., Liu, Z., Wang, J., Yang, B., Chen, W., Ou, L., Dai, X., Zhang, Z., and Zou, X. (2020). Integrative analysis of metabolome and transcriptome reveals the mechanism of color formation in pepper fruit (Capsicum annuum L.). Food Chem. 306:125629.

Luca, V. de., Salim, V., Atsumi, S.M., and Yu, F. (2012). Mining the biodiversity of plants: a revolution in the making. Science 336:1658–1661.

Ma, Q., Shi, C., Su, C., and Liu, Y. (2020). Complementary analyses of the transcriptome and iTRAQ proteome revealed mechanism of ethylene dependent salt response in bread wheat (Triticum aestivum L.). Food Chem. 326:126866.

Maccari, M., Harris, N.S., Twardziok, S.O., Pasam, R.K., Gundlach, H., Spannagl, M., Ormanbekova, D., Lux, T., Prade, V.M., Milner, S.G., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. Nat. Genet. 51:885–895.

Martin, C., and Li, J. (2017). Medicine is not health care, food is health care: plant metabolic engineering, diet and human health. New Phytol. 216:699–719.

Masato, A., Houston, K., Tucker, M.R., Schreiber, M., Berger, B., Aubert, M.K., Wilkinson, L.G.W., Witzel, K., Waugh, R., Seiffert, U., et al. (2021). Genome-wide association study reveals the genetic complexity of fructan accumulation patterns in barley grain. J. Exp. Bot. 72:2383–2402.

Matros, A., Liu, G., Hartmann, A., Jiang, Y., Zhao, Y., Wang, H., Ebmeyer, E., Korzun, V., Schachtschneider, R., Kazman, E., et al. (2017). Genome-metabolite associations revealed low heritability, high genetic complexity, and causal relations for leaf metabolites in winter wheat (Triticum aestivum). J. Exp. Bot. 68:415–428.

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

Nie, H., Chen, H., Li, G., Su, K., Song, M., Duan, Z., Li, X., Cao, X., Huang, J., Huang, S., et al. (2021). Comparison of flavonoids and phenylpropanoids compounds in Chinese water chestnut processed with different methods. Food Chem. 335:127662.

Nomura, T., Ishihara, A., Imaishi, H., Endo, T.R., Ohkawa, H., and Iwamura, H. (2002). Molecular characterization and chromosomal localization of cytochrome P450 genes involved in the biosynthesis of cyclic hydroxamic acids in hexaploid wheat. Mol. Genet. Genomics. 267:210–217.

Nomura, T., Ishihara, A., Imaishi, H., Okhawa, H., and Iwamura, H. (2003). Rearrangement of the genes for the biosynthesis of benzoxazinones in the evolution of Triticaceae species. Planta 217:776–782.

Oliver, S. (1998). Systematic functional analysis of the yeast genome. Trends. Biotechnol. 16:373–378.

Paine, J.A., Shippton, C.A., Chaggar, S., Howells, R.M., Kennedy, M.J., Vernon, G., Wright, S.Y., Hincliffe, E., Adams, J.L., Silverstone, A.L., et al. (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. Nat. Biotechnol. 23:482–487.

Peng, M., Shahzad, R., Gul, A., Subthain, H., Shen, S., Loi, L., Zheng, Z., Zhou, J., Lu, D., Wang, S., et al. (2017). Differentially evolved glucosyltransferases determine natural variation of rice flavone accumulation and UV-tolerance. Nat. Commun. 8:1975.

Qiu, W., Su, W., Cai, Z., Dong, L., Li, C., Xin, M., Fang, W., Liu, Y., Wang, X., Huang, Z., et al. (2020). Combined analysis of transcriptome and metabolome reveals the potential mechanism of coloration and fruit quality in yellow and purple Passiflora edulis Sims. J. Agr. Food Chem. 68:12096–12106.
Plant Communications

Quadranra, L., Almeida, J., Asis, R., Duffy, T., Dominguez, P.G., Bermúdez, L., Conti, G., Correa da Silva, J.V., Peralta, I.E., Colot, V., et al. (2014). Natural occurring epialleles determine vitamin E accumulation in tomato fruits. Nat. Commun. 5:3027.

Rasheed, A., and Xia, X. (2019). From markers to genome-based breeding in wheat. Theor. Appl. Genet. 132:767–784.

Reichardt, S., Budahn, H., Lamprecht, D., Rieupe, D., Ulrich, D., Dunemann, F., and Kopertekh, L. (2020). The carrot monoterpene synthase gene cluster on chromosome 4 harbours genes encoding flavour-associated sabinene synthases. Hortic. Res. 7:190.

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

Rieupe, L., Rey, M., Gerin, F., Wisniewski-Dye, F., Prigent-Combaret, C., and Comte, G. (2021). A cross-metabolomic approach shows that wheat interferes with fluorescent pseudomonas physiology through its root metabolites. Metabolites 11:84.

Roepenack-Lahaye, E. von, Degenkolb, T., Zerjeski, M., Franz, M., Roth, U., Wessjohann, L., Schmidt, J., Scheel, D., and Clemens, S. (2004). Profiling of Arabidopsis secondary metabolites by capillary liquid chromatography coupled to electrospray ionization quadrupole time-of-flight mass spectrometry. Plant Physiol. 134:548–559.

Rottmann, T., Zierer, W., Subert, C., Sauer, N., and Stadler, R. (2016). STP10 encodes a high-affinity monosaccharide transporter and is induced under low-glucose conditions in pollen tubes of Arabidopsis. J. Exp. Bot. 67:2387–2399.

Rowe, H.C., Hansen, B.G., Halkier, B.A., and Kliebenstein, D.J. (2021). A cross-metabolomic approach shows that wheat interferes with fluorescent pseudomonas physiology through its root metabolites. Metabolites 11:84.

Ruef, M., Technow, F., Sulpice, R., Altmann, T., Stitt, M., Willmitzer, L., and Melchinger, A.E. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. Nat. Genet. 44:217–220.

Saia, S., Fragasso, M., Vita, P. de, and Beleggia, R. (2019). Metabolomics provides valuable insight for the study of durum wheat: a review. J. Agr. Food Chem. 67:3069–3085.

Saito, K., and Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. Annu. Rev. Plant Biol. 61:463–489.

Sayre, R., Beeching, J.R., Cahoon, E.B., Egesi, C., Faquett, C., Fellman, J., Fregene, M., Grissem, W., Mallowa, S., Manary, M., et al. (2011). The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. Annu. Rev. Plant Biol. 62:251–272.

Schafer, N., Semel, Y., Roessner, U., Gurr, A., Balbo, I., Carrari, F., Pleban, T., Perez-Melis, A., Brudigam, C., Kopka, J., et al. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. Nat. Biotechnol. 24:447–454.

Schuy, C., Groth, J., Ammon, A., Eydam, J., Baier, S., Schweizer, G., Hanemann, A., Herz, M., Voll, L.M., and Sonnewald, U. (2019). Deciphering the genetic basis for vitamin E accumulation in leaves and grains of different barley accessions. Sci. Rep. 9:9470.

Shi, C., Zheng, Y., Geng, J., Liu, C., Pei, H., Ren, Y., Dong, Z., Zhao, L., Zhang, N., and Chen, F. (2020a). Identification of herbcide resistance loci using a genome-wide association study and linkage mapping in Chinese common wheat. Crop J. 8:666–675.

Shi, Q., Du, J., Zhu, D., Li, X., and Li, X. (2020b). Metabolomic and transcriptomic analyses of anthocyanin biosynthesis mechanisms in the color mutant Ziziphus jujuba cv. Tailihong. J. Agr. Food Chem. 68:15186–15198.

Shi, T., Zhu, A., Jia, J., Hu, X., Chen, J., Liu, W., Ren, X., Sun, D., Fernie, A.R., Cui, F., et al. (2020c). Metabolomics analysis and metabolite-agronomic trait associations using kernels of wheat (Triticum aestivum) recombinant inbred lines. Plant J. 103:279–292.

Song, Y., Zhang, N., Shi, S., Li, J., Zhang, Q., Zhao, Y., Jiang, Y., and Tu, P. (2015). Large-scale qualitative and quantitative characterization of components in Shenfu injection by integrating hydrophilic interaction chromatography, reversed phase liquid chromatography, and tandem mass spectrometry. J. Chromatogr. A 1407:106–118.

Souza, L.P. de, Borghi, M., and Fernie, A. (2020). Plant single-cell metabolomics-challenges and perspectives. Int. J. Mol. Sci. 21:8987.

Sue, M., Nakamura, C., and Nomura, T. (2011). Dispersed benzoxazinone gene cluster: molecular characterization and chromosomal localization of glucosytransferase and glucosidase genes in wheat and rye. Plant Physiol. 157:985–997.

Sue, M., Yamazaki, K., Yajima, S., Nomura, T., Matsuoka, T., Iwamura, H., and Miyamoto, T. (2006). Molecular and structural characterization of hexameric beta-D-glucosidases in wheat and rye. Plant Physiol. 141:1237–1247.

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

Tiemann, D., Zhu, G., Resende, M.F.R., Lin, T., Nguyen, C., Bies, D., Rambla, J.L., Beltran, K.S.O., Taylor, M., Zhang, B., et al. (2017). A chemical genetic roadmap to improved tomato flavor. Science 355:391–394.

Tohe, T., Wendenburg, R., Ishihara, H., Nakabayashi, R., Watanabe, M., Sulpice, R., Hoeferg, R., Takayama, H., Saito, K., Stitt, M., et al. (2016). Characterization of a recently evolved flavonol-phenylacetylene synthase gene provides signatures of natural light selection in Brassicaceae. Nat. Commun. 7:12399.

Trela, A., and Szymanska, R. (2019). Less widespread plant oils as a good source of vitamin E. Food Chem. 296:160–166.

Tsugawa, H., Nakabayashi, R., Mori, T., Yamada, Y., Takahashi, M., Rai, A., Sugiyama, R., Yamamoto, H., Nakaya, T., Yamazaki, M., et al. (2019). A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms. Nat. Methods 16:295–298.

Vergara-Diaz, O., Vatter, T., Kefauver, S.C., Obata, T., Fernie, A.R., and Araus, J.L. (2020a). Assessing durum wheat ear and leaf metabolomes in the field through hyperspectral data. Plant J. 102:615–630.

Vergara-Diaz, O., Vatter, T., Vicente, R., Obata, T., Nieto-Taladriz, M.T., Aparicio, N., Carlisle Kefauver, S., Fernie, A., and Araus, J.L. (2020b). Metabolome profiling supports the key role of the spike in wheat yield performance. Cells 9:1023.

Vos, R.C.H. de, Moco, S., Lommen, A., Keurentjes, J.J.B., Bino, R.J., and Hall, R.D. (2007). Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat. Protoc. 2:778–791.

Wang, H., Xu, S., Fan, Y., Liu, N., Zhan, W., Liu, H., Xiao, Y., Li, K., Pan, Q., Li, W., et al. (2018). Beyond pathways: genetic dissection of tocopherol content in maize kernels by combining linkage and association analyses. Plant Biotechnol. J. 16:1464–1475.

Wang, S., Alseekh, S., Fernie, A.R., and Luo, J. (2019). The structure and function of major plant metabolite modifications. Mol. Plant 12:899–919.

Wang, S.C., Wong, D.B., Forrest, K., Allen, A., Chao, S.M., Huang, B.E., Maccarelli, M., Salvi, S., Milner, S.G., Cattivelli, L., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnol. J. 12:797–796.
Exploring genic resources underlying metabolites

Wang, X., Zhang, X., Hou, H., Ma, X., Sun, S., Wang, H., and Kong, L. (2020). Metabolomics and gene expression analysis reveal the accumulation patterns of phenylpropanoids and flavonoids in different colored-grain wheats (Triticum aestivum L.). Food Res. Int. 138:109711.

Wang, Z., Cui, Y., Vainstein, A., Chen, S., and Ma, H. (2017). Regulation of fig (Ficus carica L.) fruit color: metabolomic and transcriptomic analyses of the flavonoid biosynthetic pathway. Front. Plant Sci. 8:1990.

Winfield, M.O., Allen, A.M., Burridge, A.J., Barker, G.L.A., Benbow, W.D., Xiong, L., and Yan, J. (2020). Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives. Mol. Plant 13:187–214.

Ye, X., Al-Babili, S., Kloti, A., Zhang, J., Lucca, P., Beyer, P., and Potrykus, I. (2000). Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 287:303–305.

Yin, R., Han, K., Heller, W., Albert, A., Dobreva, P.I., Zazimalová, E., and Schaffner, A.R. (2013). Kaempferol 3-O-rhamnoside-7-O-rhamnoside is an endogenous flavonol inhibitor of polar auxin transport in Arabidopsis shoots. New Phytol. 201:466–475.

Yu, Y., Jian, L., Zhao, J., Yan, Z., Wang, G., and Guo, D. (2019). Molecular cloning and characterization of a grapevine (Vitis vinifera L.) serotonin N-acetyltransferase (VvSNAT2) gene involved in plant defense. BMC. Genomics. 20:580.

Zhan, X., Qiu, J., Zhou, B., and Mao, B. (2020). Metabolomic and transcriptomic analyses reveal the regulation of pigmentation in the purple variety of Dendrobium officinale. Sci. Rep. 10:17700.

Zhang, W., Zhang, Y., Qiu, H., Guo, Y., Wang, H., Zhang, X., Scossa, F., Alseekh, S., Zhang, Q., Wang, P., and et al. (2020). Genome-wide dissection of co-selected UV-B responsive pathways in the UV-B adaptation of qingke. Mol. Plant 13:112–127.

Zhang, X., Li, P., Yun, P., Wang, L., Gao, G., Zhang, Q., Luo, L., Zhang, C., and He, Y. (2019). Genetic dissection of vitamin E content in rice (Oryza sativa) grains using recombinant inbred lines derived from a cross between ‘Zhenshan97B’ and ‘Nanyangzhan’. Plant Breeding 138:820–829.

Zhou, Y., Li, Z., Liu, G., Jiang, Y., Maurer, H.P., Wurschum, T., Mock, H.-P., Matros, A., Ebmeyer, E., Schachschneider, R., and et al. (2015). Genome-based establishment of a high-yielding heterotic pattern for hybrid wheat breeding. Proc. Natl. Acad. Sci. U S A 112:15624–15629.

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., Yang, C., and et al. (2018). Rewiring of the fruit metabolome in tomato breeding. Cell 172:249–261.