Trends in Genetics

CellPress

Review

Broadening Our Portfolio in the Genetic Improvement of Maize Chemical Composition

Weiwei Wen,[1,2] Yariv Brotman,[1,3] Lothar Willmitzer,[1] Jianbing Yan,[2] and Alisdair R. Fernie[1,*]

The adoption of recombinant inbred line and introgression line populations, as well as the study of association mapping panels, has greatly accelerated our ability to identify the genes underlying plant phenotypic variance. In tandem, the development of metabolomics approaches has greatly enhanced our ability to comprehensively define cellular chemical composition. As a consequence, breeding for chemical composition is being extended beyond our traditional targets of oil and protein to include components such as essential amino acids, vitamins, and antioxidant secondary metabolites with considerable purported consequences for human health. Here, we review the above-mentioned developments paying particular attention to the genetic architecture of metabolic traits as well as updating the perspective for utilizing metabolomics in maize improvement.

Trends

Extensive breeding resources are currently available for maize.

Concurrently, our capacity for evaluating chemical composition of plants has greatly increased given the advent of metabolomics.

Taken together, these advances allow us to extend our portfolio for compositional quality oriented breeding.

The Long History of Plant Breeding

Plant breeding encompasses the creation, selection, and fixation of superior plant phenotypes in the development of improved cultivars suited to the needs of humans and has been a human pursuit since the advent of cultivation. The primary goals of crop plant breeding has focused on improved yield with considerable success being obtained by the development of hybrid maize (*Zea mays L.* ssp. *Mays*; [1]), introduction of the wheat (*Triticum aestivum*) and rice (*Oryza sativa*) varieties that enabled the Green Revolution [2], as well as the molecular marker assisted **introgression** (see Glossary) of defined genes or genomic regions from wild species and landraces [3]. However, the roots of plant breeding vastly predate these modern approaches. For example, prehistoric selection for visible phenotypes, which facilitated harvesting and increasing productivity, led to the domestication of the first crop varieties [4]. Additionally, revolutionary insights made both by Darwin and Mendel over 100 years ago paved the way towards the scientific approach to plant breeding [5]. However, these insights were only adopted in earnest once a better understanding of quantitative genetics was able to reconcile Mendelian principals with continuous trait variation [5,6]. Since this was achieved, successive iterations have adopted molecular biology, modern breeding technology such as **marker-assisted selection** (MAS), and most recently genomics [7–11] to further broaden the scientific basis of plant breeding. Here, we will focus on the application of these approaches to improving the chemical composition of maize focusing in parallel on the recent advances in crop genetics and methods for chemical analysis, and their application to better understand tolerance and resistance mechanisms, as well as improving human nutrition.

Maize Genetic Diversity

Maize is not only of global importance as a food and source of diverse industrially important products but is also a model system with tremendous genetic diversity. Maize was domesticated

[1]Max Planck Institute of Molecular Plant Physiol., 14476 Potsdam Golm, Germany
[2]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China
[3]Department of Life Sciences, Ben Gurion University of the Negev, Beersheva, Israel

*Correspondence: fernie@mpimp-golm.mpg.de (A.R. Fernie).

from the wild progenitor teosinte (*Zea mays* ssp. *Parviglumis*) approximately 10 000 years ago in Mexico [12]. Subsequent to domestication, maize landraces had been subject to intensive improvement efforts, culminating in the development of hybrid maize lines that are highly adapted to modern agricultural practices [13]. Teosinte is extremely diverse and maize retains much of the diversity of its wild ancestor, with any two maize varieties differing from one another by 1.4% at the DNA level [14]. It was documented that the level of nucleotide diversity found in maize is two to five times higher than that of other domesticated grass crops and is 14 times higher than that of humans [15]. Within the diverse germplasm, maize inbred lines represent a fundamental resource for studies in genetics and breeding.

The genome sequencing of a large number of maize lines and wild relative lines in recent years has provided significant insights to our understanding of the maize genome and evolution [16–18]. The B73 reference genome is 2.3 Gb in total size and is composed of approximately 85% repetitive sequences. In addition to single nucleotide polymorphisms (SNPs) and small insertion and deletions (indels), there are a large number of structural variants in the maize genomes, including causal polymorphisms for phenotypic variations. These advances in maize genomics further revolutionized our understanding of genetic diversity and provided an important foundation for designing strategies for maize improvement.

Most agriculturally and economically important traits have complex genetic underpinnings (i.e., determined by multiple **quantitative trait loci**, **QTLs**). Precisely locating and characterizing these functional loci facilitates crop improvement via MAS or biotechnology aided breeding. To dissect complex traits, **linkage analysis** and **association mapping** are commonly used. On the basis of the diverse maize germplasm, a variety of populations for the above-mentioned genetic analysis have been created or assembled. Here, we provide a short overview of different types of mapping populations with a discussion of their utilizations in maize in the following sections. For linkage analysis, recombinant inbred line (RIL) populations derived from two parental maize lines are frequently used. A noteworthy biparental population for linkage mapping is the **intermated B73 × Mo17 (IBM) population**. The Maize Mapping Project (MMP) constructed a genetic map (IBM2) that contains 2026 markers. The genetic map, the open access to seeds, and the available web resources have led to the wide use of IBM by the maize genetics community [19]. The adoption of linkage analysis provided great power and allowed the identification of epistatic interacting loci and loci exhibiting only minor effects [20–22]. However, this approach exhibits a relative paucity of alleles and unless huge mapping populations are used the mapping resolution is generally low. By contrast, association mapping possesses advantages over linkage analysis in terms of mapping resolution: allele richness and time investment. However, the power to detect minor effect loci and epistatic interactions using association mapping is often limited due to the complex genetic background of the population. Hence, the **germplasm collection** that encompasses sufficient genetic diversity covering most variations for the traits of interest and with rapid **linkage disequilibrium** (**LD**) decay is vital for a promising outcome of association mapping [23]. Linkage and association analysis can be complementary to each other and in some cases both have been used for cross-validation and causal genetic variant identification [22,24]. For the sake of combining the advantages and eliminating the disadvantages of linkage and association analysis, the **nested association mapping** (**NAM**) population that combines 25 RIL populations with 200 lines per family [25] and the **multiple parent advanced generation inter-cross** (**MAGIC**) population were developed [26]. These multiple biparental family sets and multiple parent populations were constructed to enable higher mapping resolution and power by introducing more recombination events and eliminating the confounding effects of population structure [25,26]. Compared with NAM, the MAGIC design does not use a common reference parent for all of the crosses and avoids the confounding effect of family structure on QTLs inheritance, which makes it more statistically efficient and easier to detect QTLs that contribute to differences among biparental families.

Glossary

Association mapping: also known as linkage disequilibrium mapping or association analysis; is a method that identifies the link between genomic variants and phenotypes, which takes advantage of historic linkage disequilibrium to detect and locate QTLs.

Genome-wide association study (GWAS): is an examination of associations between common genetic variants (usually genome-wide SNPs) and traits such as human diseases.

Genomic prediction: also known as genomic selection where genetic markers covering the whole genome of the training and breeding populations and the phenotypic data of the training population are used and integrated in a model to predict the performance of the breeding population. The selection decision will be made on the breeding population based on the breeding values.

Germplasm collection: collection of living genetic resources such as seeds or tissue that are maintained for the purpose of animal and plant breeding, preservation, and other research uses.

Illinois Long-Term Selection Experiment: an experiment initiated and conducted by researchers at the University of Illinois from 1896 for selecting maize lines with the highest or lowest concentrations of grain protein or oil. This study has created 12 populations that vary significantly in their grain protein and oil composition through 110 cycles of recurrent selection over a century.

Intermated B73 × Mo17 (IBM) population: the two inbred lines B73 and Mo17 were crossed to make the F1 hybrid and was then self-pollinated. F2 progenies were then intermated for four generations, followed by repeated selfing to generate recombinant inbred lines.

Introgression: the incorporation of genes from one species into the gene pool of a second species by hybridization and backcrossing. Introgression lines of a certain crop contain a genetic component artificially derived from a wild relative population through repeated backcrossing, which are used for gene or QTL mapping and breeding new varieties.

Linkage analysis: a tool for genetic mapping where the coinheritance of

Analysis of Crop Chemical Composition: Genetic Analysis

Crop chemical composition has been determined for decades with early studies in maize, rice, wheat, sunflower, soybean, rapeseed, pea, and oat concentrating on oil and protein content [3,27,28], while the genetic basis of both starch accumulation and structure–which is of high interest for industrial uses of the biopolymer–have been the subject of extensive investigation in potato and cereal species [29]. Similarly, QTLs for sugar content have been identified in a wide range of species including potato, tomato, melon, and sugarcane [30], while cell wall sugars in maize pericarp have also been identified relatively recently [31]. Vitamins, pigments, and antioxidants have also received a wealth of research attention, particularly in highly colored, genetically tractable crops such as tomato, peach, and melon [32]. The best place to start any discussion of maize chemical content is the **Illinois Long-Term Selection Experiment**, which examined both protein and oil content. This study began in 1896 and is the longest continuous genetic experiment in higher plants to date [27,28]. The Illinois Selection Strains span the known phenotypic extremes for maize kernel composition (8–12% protein and 4–6% oil), demonstrating the power of long-term selection and the variation they contain for altering the expression of complex traits. The progressive phenotypic responses for kernel composition and correlated traits observed in the Illinois Selection Strains have provided convincing evidence that these traits are controlled by many genes [28]. Indeed, to date, the Illinois long-term selection lines remain a source of favorable alleles that are associated with oil, protein, and starch accumulation. In addition, a series of studies using different types of genetic populations has been performed in the past decade to investigate the genetic basis and causal genomic variants controlling chemical compositions of maize (Table 1). Earlier studies mainly covered the quality or nutritional traits of maize such as starch, protein, and oil, and so on; the scope has been extended to a metabolomics level in very recent years (Table 1). Plant secondary metabolites have drawn more and more attention now owing to their nutritional and medical value as well as the important roles they play in plant defense to both biotic and abiotic stresses. Identification of genes associated with the level of maize secondary metabolites through genetic mapping enriched our knowledge of some secondary pathways, for example, the flavonoids and benzoxazinoid biosynthetic pathways (Table 1). Furthermore, some recently developed maize association mapping and multiparental populations have already been used for the genetic analysis of chemical compositions. For example, the NAM population was used to reveal the genetic architecture and identify candidate genes for kernel composition traits and carbon and nitrogen metabolic traits [33,34], as will be discussed in detail in the section on 'Attempts to improve maize nutritional composition' later.

Analysis of Crop Chemical Composition: Metabolomics Approaches

The advent of metabolomics in the late 1990s and its ongoing development has greatly enhanced our understanding concerning the metabolism of a great number of crop species but in particular tomato [35–37], maize [24,38,39], and rice [40–42]. While a wide assortment of techniques exist, the techniques tend to center around variants of mass spectrometry (MS) or nuclear magnetic resonance (NMR) [43], although the first applications of liquid chromatography (LC)–NMR–MS technologies are now being reported [44]. In the case of maize, the entire spectrum of methods has been applied and given that these have been comprehensively reviewed relatively recently [45], we will only discuss them briefly here. NMR has been used in multiple studies in maize frequently in the evaluation of transgenic maize (see [46] for an example), yet tends to be restricted to around 20–30 highly abundant metabolites. Similarly, capillary electrophoresis (CE)–MS-based approaches have also been used in maize [47], but these also provide information on a relatively small number of metabolites and are probably highly useful only in combinatorial approaches (see, for example, [48]). By contrast, the application of gas chromatography (GC)–MS to maize has been carried out to address a wide range of biological questions [38,49,50], typically some 100-odd polar analytes including 69 metabolites of known chemical structure can be measured comprising sugars and their

markers and traits is related to known genetic relationships between members of the same family or pedigree.

Linkage disequilibrium (LD): the nonrandom association between alleles at different loci within a population.

Marker-assisted backcrossing: a process of using molecular markers to assist in transferring a gene or genomic region of interest from a donor to a recipient (recurrent line) through at least five or six backcross generations. The recovery of the recurrent genotype can be accelerated with the use of molecular markers.

Marker-assisted selection: a process of using markers (mainly DNA-based markers nowadays) for selection of a genetic determinant(s) of a target trait to improve the efficiency and precision in plant or animal breeding.

Multiple parent advanced generation inter-cross (MAGIC): a creation of a large multiparent RIL population in plants. The MAGIC populations are created by inter-crossing *n* lines for *n*/2 generations until all founders are combined with equal proportions in the inter-crosses, which allows the use of both linkage and association methodologies without the difficulties of highly structured populations.

Nested association mapping (NAM): a technique designed by Buckler *et al.* [15] for dissecting the genetic architecture of complex traits in maize. The NAM population contains 25 families of 200 RILs per family by crossing 25 diverse inbred lines to the B73 reference line, enabling high power and high resolution through joint linkage association analysis.

Quantitative trait locus (QTL): a genomic site contributing to the genetic variability of a quantitative trait.

Table 1. Summary of Genetic Mapping Studies on Metabolic Traits and Nutritional Compositions of Maize[a]

Trait	Measurement	Gene	Mapping Population	Ref.
α-Carotene content	HPLC	*ZmcrtRB3*	RIL, association panel	[85]
α-Tocopherol content	HPLC	*ZmVTE4*	Association panel	[72]
Benzoxazinoid metabolites	LC–MS	*GlcMT*	RIL population	[86]
Carbon and nitrogen metabolites	Fluorescamine assay	Multiple genes	NAM population	[28]
Carotenoid content	HPLC	*PSY1*	RIL, association panel	[87]
Carotenoid content	HPLC	*LcyE*	RIL, association panel	[67]
Carotenoid content	HPLC	*crtRB1*	RIL, association panel	[88]
Carotenoid content	HPLC	QTL	$F_{2:3}$ population	[89]
Carotenoid content	HPLC	QTL	Association panel	[90]
Flavonoids	HPLC	*ZmF3′H1*	F_2 population	[91]
Lipids	LC–MS	QTL	Association panel	[51]
Maysin and chlorogenic acid accumulation	HPLC	QTL	RIL population	[92]
Multiple metabolic traits	LC–MS/MS	Multiple genes	RIL, association panel	[24]
Multiple metabolic traits	LC–MS/MS	Multiple genes	RIL, association panel	[76]
Oil and fatty acids content	NMR, GC	*DGAT*	BC_2 population	[58]
Oil and fatty acids content	Gas chromatogram	QTL	RIL population	[20]
Oil and fatty acids content	Gas chromatogram	Multiple genes	RIL, association panel	[57]
Palmitic acid content	Gas chromatogram	*Zmfatb*	NIL, association panel	[93]
Primary metabolites	GC–TOF–MS	Multiple genes	Association panel	[38]
Primary metabolites	GC–TOF–MS	Multiple genes	RIL population	[22]
Starch content	Near infrared	Multiple genes	Association panel	[66]
Starch content	Fermentable carbohydrate assay	QTL	RIL population	[64]
Starch, protein, oil	Near infrared	QTL	NAM, association panel	[27]
Tocopherol content	HPLC	QTL	Association panel	[73]

[a]GC–TOF–MS, gas chromatography–time of flight–mass spectrometry; HPLC, high-performance liquid chromatography; LC–MS, liquid chromatography–mass spectrometry; LC–MS/MS, liquid chromatography–tandem mass spectrometry; NMR, nuclear magnetic resonance; QTL, quantitative trait locus; *crtRB1*, β-carotene hydroxylase 1; *DGAT*, diacylglycerol-*O*-acyltransferase; *GlcMT*, 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA)-glucoside methyltransferase; *LcyE*, lycopene epsilon cyclase; *PSY1*, phytoene synthase 1; *ZmVTE4*, maize γ-tocopherol methyltransferase; *ZmcrtRB3*, carotenoid hydroxylase; *ZmF3′H1*, flavonoid 3′-hydroxylase; *Zmfatb*, acyl-ACP thioesterase; NAM, nested association mapping; NIL, near isogenic line; RIL, recombinant inbred line.

derivatives, organic and amino acids, and a few small secondary metabolites and vitamins. Additional analysis of the lipophilic components provides information concerning a further 40 peaks. However, coverage of the lipid species is far greater, information on over 500 species is available, using ultra performance liquid chromatography (UPLC)–Fourier transform (FT)–MS-based approaches [51]. Such approaches were also used to assess polar extracts of maize resulting in the identification of up to 700 compounds [24,48,52]. These examples thus demonstrate the power of metabolomics–in particular high-resolution LC–MS techniques–and set the basis for an increased portfolio for metabolic engineering strategies. The metabolite profiling can benefit from multiple metabolomics platforms if applicable.

Attempts to Improve Maize Nutritional Composition

The importance of cereal grains to the nutrition of millions of people globally is widely recognized. The high nutritional value of maize renders it a staple food across the world mainly due to its

starch, protein, and oil content. Arguably, the nutritional trait that has been most widely studied is the maize kernel oil content. Recently, the genetic architecture of maize kernel oil content was comprehensively dissected by using diverse populations including linkage population, association panel, and the NAM population, the complexity of which is demonstrated by the Illinois Long-Term Selection Experiment as well as the numerous identified oil QTLs [20,33,53–55] (Table 1). A total of 22 QTLs affecting oil levels have been identified by using the NAM population of 5000 lines and high-density markers. Owing to the limited number of parental lines in previous studies, the molecular basis of natural variation in oil biosynthesis has not been fully elucidated in maize despite a good understanding of the plant oil biosynthetic pathway and its relevant genes. A **genome-wide association study** (**GWAS**) [56] identified 74 loci significantly associated with kernel oil concentration and fatty acid composition. The 26 loci associated with oil concentration could explain up to 83% of phenotypic variation, and some of the identified oil-associated genes were validated by expression analysis and/or linkage analysis in biparental populations. This study confirmed the inheritance mode of oil concentration and composition in a mainly additive manner, implying that breeding using these identified genes would be more straightforward. For instance, it is feasible to improve the oil concentration of a target cultivar by introducing two or more genes with major or moderate effect via MAS. These results also provided evidence for the hypothesis that favorable allele accumulation is the genetic basis for oil concentration increase during the selection of high oil lines.

As a result of the negative correlation between kernel oil content and grain yield, for breeding high oil maize varieties a neat method should be adopted to increase the oil concentration without altering the performance of the grain yield and other agronomic traits. A phenylalanine in *DGAT1-2* (diacylglycerol *O*-acyltransferase) is a key determinant of oil content and composition in maize [57]. By resequencing the *DGAT1-2* region in a maize landrace set and in 155 inbred lines (35 high oil lines and 120 normal lines) and association analysis, the function of this amino acid (Phe) insertion for grain oil content was verified again [58]. On the basis of this information, PCR-based functional markers were further developed for improving oil content in maize kernels through MAS [58]. On top of these advances, another study used **marker-assisted backcrossing** to increase the oil content of a widely used hybrid 'Zhengdan958′ in China [59]. The oil content in improved Zhengdan958 reached 4.5%, with an increased relative content of 18%. According to the measurement across ten environments, the grain yield of the improved Zhengdan958 is similar to that of the original Zhengdan958, showing its market potential for maize production. Similarly, considerable parallel advances were made in improving protein content (reviewed in [60]). The deficiency in the essential amino acids (i.e., lysine and tryptophan) of zeins, which are main storage proteins of maize, has restricted the use of maize as the sole protein source for humans and monogastric livestock. High protein content maize containing considerably higher amounts of lysine and tryptophan were identified in the maize *Opaque 2* (*o2*) mutant in the early 1960s [61]; however, the original mutant also contained some undesirable properties and it took several years before CIMMYT (International Maize and Wheat Improvement Center) developed maize varieties, referred to as 'quality protein maize' (QPM), that retained the *o2* mutation and the quality protein trait but lacked the accompanying unfavorable agronomic characteristics [62,63]. Moreover, the genetic basis of maize kernel starch content was recently finely dissected based on an RIL population that was genotyped using high-density SNP markers [64]. In doing so, the authors were able to provide support for many genes previously putatively identified to be associated with starch content and starch compositional traits (e.g., those defined in [65]). Both starch and protein traits were also studied in the above-described NAM population, which provided a comprehensive study of traditional compositional targets in maize [27].

Other than these major chemical components, maize grain also exhibits considerable phenotypic variation for multiple micronutrients such as important B vitamins, folates, vitamin C, provitamin A, and minerals[i] [20,66]. Currently, the 'hidden hunger' brought about by the lack of

essential micronutrients is attracting increasing attention since close to 2 billion of the world's population suffers from malnutrition caused not by too few calories but by an insufficient intake of essential micronutrients from their daily diet [67]. Great efforts have also been devoted to the development of maize with improved levels of micronutrients, either through breeding or biotechnology. Traditional breeding methodologies use conventional phenotypic selection to accumulate favorable alleles. Relatively recently, the advances in genomics and genetics have generated numerous molecular markers and the number of QTL analysis and cloned genes have dramatically increased. The advent of MAS in this context provides an alternative for maize improvement that is more cost-effective and greatly accelerates the conventional breeding for nutritionally improved maize. Utilizing MAS, a few nutritional trait-associated genes or QTLs have been recently introgressed into elite maize lines for their quality improvement (Table 2). Comprehensive genetic studies are of fundamental importance for designing a feasible and effective breeding program. The genetic architecture of multiple nutritional traits in maize has been revealed by using different types of genetic populations (Table 1). A promising breeding strategy necessitates taking into account the dissected genetic architecture, which demonstrates the number of QTLs affecting the target traits and their effect size. Taking maize kernel carotenoid content as an example, studies using multiple populations (RIL, $F_{2:3}$, association panel) indicated that a small number of moderate to large effect loci (e.g., *PSY1*, *LcyE*, *CrtRB1*, and *CrtRB3*) can largely explain the phenotypic variance (Table 1). This implies that pyramiding favorable alleles of these major effect genes can result in provitamin A enhancement (i.e., enhancement of provitamin A level by selecting two or more than two genes at a time). For instance, introduction of a favorable *crtRB1* allele via MAS has led to rapidly increasing provitamin A content to >20 μg/g in maize [68]. In addition, it is vital to assess the effect and stability of target alleles or QTLs in different genetic backgrounds and environments before starting a breeding practice. After validating the effects of three polymorphisms (*LcyE*5′TE, *LcyE*3′Indel, and *CrtRB1*-3′TE) in 26 diverse tropical genetic backgrounds [69], it was recommended not to select for the favorable alleles of both *LcyE* and *CrtRB1* genes in breeding programs. The authors found that feedback inhibition may be reducing the total flux into the carotenoid pathway and that maximum total provitamin A concentrations were achieved in genotypes with homozygous unfavorable or heterozygous *LcyE*. To ensure good performance of the recurrent line or hybrid, it is also essential to evaluate if the target gene or QTL has pleiotropic effects on both target trait and other important traits such as agronomical and yield traits. A recent study examined the effect of crossing parental lines from two marker-based heterotic groups on carotenoid accumulation and agronomic performance in hybrids, and the results showed that several hybrids with high provitamin A content were competitive to a commercial hybrid in grain yield and other traits [70].

Table 2. Breeding Practices for Maize Nutritional Improvement in the Past Decade[a]

Improved Trait	Gene	Approach	Ref.
Carotenoid levels	Multiple	Genomic prediction	[90]
Lysine content	*opaque 2*, *opaque 16*	MAS (marker-assisted backcrossing)	[94]
Oil content	*DGAT1-2*	Development and evaluation of functional markers	[58]
Oil content	*DGAT1-2*	MAS (marker-assisted backcrossing)	[59]
Protein	*opaque 2*	MAS (marker-assisted backcrossing)	[95]
Provitamin A level	*lcyE*, *crtRB1*	Development and evaluation of functional markers	[69]
Provitamin A level	*PSY1*, *lcyE*, *crtRB1*	Development and evaluation of functional markers	[68]
Provitamin A level	*crtRB1*	MAS (marker-assisted backcrossing)	[96]

[a]*crtRB1*, β-carotene hydroxylase 1; *DGAT*, diacylglycerol-*O*-acyltransferase; *LcyE*, lycopene epsilon cyclase; *opaque 2*, bZIP transcription factor that regulates transcription of many genes (e.g., Zein genes) involved in a variety of pathways; *opaque 16*, associated with lysine content in the endosperm of the maize grain; *PSY1*, phytoene synthase 1.

This finding serves as the basis for developing and promoting hybrids with greater expression of heterosis (i.e., hybrid progeny have improved performance compared with both homozygous parents) in productivity and concentrations of provitamin A. While this is arguably the best example of vitamin biofortification in maize, GWASs have recently been reported for vitamin E [71,72], and there is increasing recent interest, at least at the biochemical level in the biosynthesis of B vitamins and folate, in maize [73].

Essential amino acids are the ones that humans cannot synthesize *de novo* but must acquire from their diet, namely His, Iso, Leu, Lys, Met, Phe, Val, Thr, and Trp [74]. Although enhancing the levels of the essential amino acids is partially covered in breeding for quality protein maize, considerable research has been carried out on elevating Lys content in maize [75]. Other compounds of maize that are of high potential nutritional value are the flavones. Detailed evaluation of the results from a recent metabolomics study of the maize kernel revealed that many of the QTLs for flavonoids were highly influenced by the *p1* locus [24,76]. *p1* has previously been identified to constitute a R2R3–MYB transcription factor, which regulates flavonoid biosynthesis in a manner that is apparently competitive with the formation of anthocyanins [77]. While potential targets for nutritional improvement were uncovered, it is important to note that understanding of flavonoid biosynthesis in maize lags considerably behind that of anthocyanin biosynthesis, which has been better studied given its importance in the change in kernel color on domestication.

Attempts to Metabolically Fortify Maize

In addition to improving the chemical composition of maize from the food and feed perspective, a certain amount of research has been carried out concerning the possibility to improve yield both under optimal growth conditions and under biotic (Box 1) and abiotic (Box 2) stresses. In the case of understanding the link between metabolism and growth under optimal conditions, two

Box 1. Metabolic Aspects of Biotic Tolerance

Maize plants are continually challenged by a wide range of pathogens and pests [97]. During domestication, selection for higher yield has resulted in many cases in reduced ability to generate an effective defense response [98]. This is, in great part, manifested by metabolite production, accumulation, and flux. Two main mechanisms are suggested for the reduction in defense capacity: first, energy and resources are allocated towards growth at the expense of secondary functions, such as defense; the resulting trade-off is a topic of intensive research [99]. Second, as selection criteria concentrated mainly on growth and grain quality, defense factors have been inadvertently eliminated. That said, the correlation between yield and defense capacity is not always straightforward [100].

Several studies exist with regard to the role that the chemical composition of the plant plays in biotic defense in maize. For example, the straightforward identification of defense metabolites, both at the site of infection and systemically [101,102], along with defense compounds and plant compounds that induced herbivore growth [101], ultimately leading to a definition of the differences between local and induced defense metabolites [101,102]. Recent efforts tend to be more towards a systems approach, whereby multilevel analyses are integrated to give a comprehensive picture of the pathways, molecular components, and regulation of defense [103]. In the cited article, for instance, benzoxazinoid biosynthesis genes were confirmed as aphid deterrents, both in the systems approach and empirically, using mutants. However, not all defense compounds directly fight or ward pathogens and herbivores off: indole was recently identified as a prominent volatile signaling molecule between maize plants, whereby a plant attacked by herbivore emits the signal to warn neighboring plants [104].

Examples also exist of pathogen- or herbivore-derived effectors interfering with defense measures of maize [105,106], and of herbivores that are able to neutralize plant defense compounds [86]. The latter study is also an example of utilizing natural variation to colocalize two defense-related traits, namely maize susceptibility to herbivory and abundance of a defense metabolite, to one genetic locus. Intriguingly, the identified encoded enzyme underlies a trade-off between the direct toxicity of the metabolite and its induction of other defense pathways [86]. Mapping disease resistance traits on a panel of maize inbred lines [107,108] resulted in the unraveling of numerous defense-related genes including several genes associated with metabolism. Thus, valuable information is being continuously accrued as a result of these multiple approaches, which will be potentially useful in breeding programs for elevated defense, and a better understanding of the chemotypes of resistant individuals will undoubtedly aid this process.

Box 2. Metabolic Aspects of Abiotic Resistance

Metabolites are key mediators of plant response to abiotic stress [43]. Unraveling the participating metabolites and the underlying mechanisms can help breeders develop better-adapted cultivars, which is an increasingly urgent need in light of climate change and environmental pollution. Drought is a major stress that causes considerable yield loss in maize, which is often also accompanied by heat stress under field conditions. The metabolic response of plant to multiple stresses has been studied several times. While having in some cases an additive nature, as shown for example in maize [109], examples in other species, using mostly transcriptomic data, show that prediction of response to stress combinations is different from the sum of the single stresses [110,111]. Metabolites have been demonstrated to be useful predictors of grain yield [109]; however, the connection between metabolites and plant performance cannot always be established [112].

Phytoalexins, secondary compounds that mainly function in defense against pathogens, have also been implicated in playing a role in maize drought tolerance [113]. In a field trial under drought, maize showed increased level of 5-hydroxynorvaline; this novel metabolite was then also shown to reduce aphid reproduction level [114]. Salinity, a stress to which maize is known to be exceptionally sensitive, was shown to compromise emission of herbivore-deterring volatile compounds of maize [111]. The enzyme phosphatidylinositol synthase in maize can induce tolerance to drought through modification of the lipid composition of the cell membrane, according to a study using a transgenic approach [115]. Aluminum, a pollutant present in huge areas of otherwise arable land, is tolerated by some plants by using organic anions, such as malate and citrate, to chelate it. When the gene mediating anion efflux is present in multiple copies in the maize genome, it is more highly expressed, resulting in higher aluminum tolerance; and the lines that have more copies originate from regions of aluminum-rich soil [116].

Abiotic stress tolerance is a complex feature comprising a wide gamut of trait categories from the molecular to the macrostructure: genetic regulation, cell physiology, photosynthetic capacity, plant morphology, chemical composition, to name but a few–all of which can be utilized by breeders for the creation of more tolerant cultivars. The chemical composition is one of the most obvious candidates for manipulation since it underlies all of these qualities.

Outstanding Questions

Despite considerable technical advances, coverage of the metabolome remains at the percentile level in eukaryotes; therefore, it is likely that many important compounds, especially micronutrients, are not covered by current approaches.

Given that the expansion of the number of metabolic traits has been very rapid, it follows that the amount of information available for the 'novel' traits is currently vastly outstripped by that afforded by long-studied traits such as protein and oil content. Once research on these novel traits matures, it will likely be highly informative to compare the genetic architecture, the heritability, and the robustness both within the novel traits and between the novel and traditionally studied traits.

As indicated in the section 'Attempts to metabolically fortify maize', understanding of the interaction between metabolism and growth remains fragmentary and as a result increasing the levels of many metabolites often confers a yield penalty. Thus, a greater understanding of the trade-offs between metabolite accumulation and plant growth is needed.

In the case of nutritionally important compounds, a better understanding with regard to their uptake and mode of action against chronic disease is still required for many compounds.

different types of experiments have been carried out. In one approach, detailed kinetic studies in which both metabolic and growth and differentiation parameters are recorded have been performed [78,79]. In the other approach, the phenomenon of heterosis [80] has been investigated from a metabolic perspective [39,81,82]. Starting with the integrated genomics approaches two studies are of particular note, both of which studied leaf gradients in maize at multiple levels. The first of these combined metabolome, transcriptome, chlorophyll, and protein measurements alongside dry weight determination revealing a list of potential regulators of the source sink gradient in this tissue, as well as allowing a revision of the key metabolites involved in maize C4 photosynthesis [79]. This will undoubtedly be highly important in the context of international efforts at improving the efficiency of cereal photosynthesis. The second adopted a similar approach but with an additional comparative aspect that of contrasting leaf gradients of maize with the C3 plant rice [78]. In addition to corroborating many of the findings of the first study, this work provided several further putative candidate regulators that appear to be important for differentially optimizing photosynthesis in the studied species. These studies represent an important starting point into unraveling the links between metabolism and growth under optimal growth conditions in maize. And it is worthwhile to conduct similar experiments in the future, for instance, using kinetic metabolomics analyses to identify the determinant factors of particular metabolite(s) on developing maize kernels. The other approach, that of studying heterosis, has a far longer history, although evaluation of metabolic aspects underpinning heterosis is relatively recent and seemingly fairly variable between species [82,83]. That said, a case study in maize revealed that profiling a panel on the basis of 130 primary metabolites [82] or over 560 lipid species [51] of 85 diverse Dent inbred lines allowed highly accurate prediction of biomass and bioenergy related traits, highlighting the utility of metabolite profiling, even in the absence of mechanistic understanding, for attempts to enhance biomass production in maize.

Concluding Remarks and Future Prospects

Crop improvement has been an important human pursuit since the very advent of agriculture; however, the twin problems of environmental deterioration and a rapidly expanding population are placing massive strain on global production demands. In parallel, clinical studies are

confirming the potential of an increasing number of plant-derived compounds to aid in combatting chronic human diseases such as heart disease, many cancers, type 2 diabetes, and obesity [84]. Here, we argue that the combination of metabolomics with contemporary genetics represents a powerful tool to uncover the genetic architecture underpinning the accumulation of metabolites important for both ensuring plant yield under adverse conditions and for providing health-conferring properties when consumed as food or feed. Parallel advances in sequencing, metabolite profiling, alongside the compilation and creation of large populations encompassing a considerable degree of natural diversity have thus enabled us to dramatically enhance the spectrum of metabolites for which clear breeding targets exist. Thus, our breeding portfolio extends far beyond the traditional staples of oil, protein, and starch to include vitamins and free amino acids, as well as phenylpropanoids and alkaloids, thus providing unprecedented opportunities for the biofortification of maize and indeed other food crops. Furthermore, the availability of metabolomics platforms or services for more and more laboratories will boost the efforts towards the integration of metabolic markers into breeding programs. Knowledge of the genetic basis and the metabolic network, coupled with data of various forms of markers, can be rationally incorporated and will eventually lead to ideal crops (see Outstanding Questions).

Acknowledgments

J.Y. was supported by the National Natural Science Foundation of China (31525017). J.Y. and A.R.F. were additionally supported by a China–Germany DAAD (German Academic Exchange Service) fellowship (project number 57136777).

Resources

[i] www.fao.org/docrep/t0395e/t0395e03.htm

References

1. Duvick, D.N. (2001) Biotechnology in the 1930s: the development of hybrid maize. *Nat. Rev. Genet.* 2, 69–74
2. Evenson, R.E. and Gollin, D. (2003) Assessing the impact of the Green Revolution, 1960 to 2000. *Science* 300, 758–762
3. Fernie, A.R. *et al.* (2006) Natural genetic variation for improving crop quality. *Curr. Opin. Plant Biol.* 9, 196–202
4. Moose, S.P. and Mumm, R.H. (2008) Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* 147, 969–977
5. Shull, G.E. (1909) A pure-line method in corn breeding. *Am. Breed. Assoc. Rep.* 5, 51–59
6. Paul, D.B. and Kimmelman, B.A. (1988) *Mendel in America: Theory and Practice, 1900–1919,* University of Pennsylvania Press
7. Dubcovsky, J. (2004) Marker-assisted selection in public breeding programs: the wheat experience. *Crop Sci.* 44, 1895–1898
8. Morgante, M. and Salamini, F. (2003) From plant genomics to breeding practice. *Curr. Opin. Biotechnol.* 14, 214–219
9. Varshney, R.K. *et al.* (2006) Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol.* 24, 490–499
10. Scossa, F. *et al.* (2016) Genomics-based strategies for the use of natural variation in the improvement of crop metabolism. *Plant Sci.* 242, 47–64
11. Fernie, A.R. and Schauer, N. (2009) Metabolomics-assisted breeding: a viable option for crop improvement? *Trends Genet.* 25, 39–48
12. van Heerwaarden, J. *et al.* (2011) Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1088–1092
13. Hufford, M.B. *et al.* (2012) Comparative population genomics of maize domestication and improvement. *Nat. Genet.* 44, 808–811
14. Tenaillon, M.I. *et al.* (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp *mays L.*). *Proc. Natl. Acad. Sci. U.S.A.* 98, 9161–9166
15. Buckler, E. and Stevens, N.M. (2005) Maize origins, domestication, and selection. In *Darwin's Harvest* (Motley, T.J. *et al.*, eds), pp. 67–90, Colombia University Press
16. Schnable, P.S. *et al.* (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115
17. Huang, X. and Han, B. (2012) A crop of maize variants. *Nat. Genet.* 44, 734–735
18. Lu, F. *et al.* (2015) High-resolution genetic mapping of maize pan-genome sequence anchors. *Nat. Commun.* 6, 6914
19. Coe, E. *et al.* (2002) Access to the maize genome: an integrated physical and genetic map. *Plant Physiol.* 128, 9–12
20. Yang, X. *et al.* (2010) Major and minor QTL and epistasis contribute to fatty acid compositions and oil concentration in high-oil maize. *Theor. Appl. Genet.* 120, 665–678
21. Mackay, T.F.C. (2014) Epistasis and quantitative traits: using model organisms to study gene–gene interactions. *Nat. Rev. Genet.* 15, 22–33
22. Wen, W. *et al.* (2015) Genetic determinants of the network of primary metabolism and their relationships to plant performance in a maize recombinant inbred line population. *Plant Cell* 27, 1839–1856
23. Yan, J. *et al.* (2011) Association mapping for enhancing maize (*Zea Mays* L.) genetic improvement. *Crop Sci.* 51, 433–449
24. Wen, W. *et al.* (2014) Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat. Commun.* 5, 3438
25. McMullen, M.D. *et al.* (2009) Genetic properties of the maize nested association mapping population. *Science* 325, 737–740
26. Dell'Acqua, M. *et al.* (2015) Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays*. *Genome Biol.* 16, 167
27. Hopkins, C.G. (1899) Improvement in the chemical composition of the corn kernel. *Illinois Agric. Exp. Stat. Bull.* 55, 205–240
28. Moose, S.P. *et al.* (2004) Maize selection passes the century mark: a unique resource for 21st century genomics. *Trends Plant Sci.* 9, 358–364
29. Fernie, A.R. *et al.* (2004) Carbohydrate metabolism. In *Handbook of Plant Biotechnology* (Christou, P. and Klee, H.J., eds), pp. 1230–1262, Wiley

30. Salvi, S. and Tuberosa, R. (2005) To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.* 10, 297–304
31. Hazen, S.P. *et al.* (2003) Quantitative trait loci and comparative genomics of cereal cell wall composition. *Plant Physiol.* 132, 263–271
32. Lewinsohn, E. *et al.* (2005) Carotenoid pigmentation affects the volatile composition of tomato and watermelon fruits, as revealed by comparative genetic analyses. *J. Agric. Food Chem.* 53, 3142–3148
33. Cook, J.P. *et al.* (2012) Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol.* 158, 824–834
34. Zhang, N. *et al.* (2015) Genome-wide association of carbon and nitrogen metabolism in the maize nested association mapping population. *Plant Physiol.* 168, 575–583
35. Alseekh, S. *et al.* (2015) Identification and mode of inheritance of quantitative trait loci for secondary metabolite abundance in tomato. *Plant Cell* 27, 485–512
36. Perez-Fons, L. *et al.* (2014) A genome-wide metabolomic resource for tomato fruit from *Solanum pennellii*. *Sci. Rep.* 4, 3859
37. Schauer, N. *et al.* (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* 24, 447–454
38. Riedelsheimer, C. *et al.* (2012) Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8872–8877
39. Lisec, J. *et al.* (2011) Corn hybrids display lower metabolite variability and complex metabolite inheritance patterns. *Plant J.* 68, 326–336
40. Matsuda, F. *et al.* (2012) Dissection of genotype–phenotype associations in rice grains using metabolome quantitative trait loci analysis. *Plant J.* 70, 624–636
41. Oikawa, A. *et al.* (2008) Rice metabolomics. *Rice* 1, 63–71
42. Hu, C. *et al.* (2014) Metabolic variation between *japonica* and *indica* rice cultivars as revealed by non-targeted metabolomics. *Sci. Rep.* 4, 5067
43. Obata, T. and Fernie, A.R. (2012) The use of metabolomics to dissect plant responses to abiotic stresses. *Cell. Mol. Life Sci.* 69, 3225–3243
44. Burton, K.I. *et al.* (1997) On-line liquid chromatography coupled with high field NMR and mass spectrometry (LC-NMR-MS): a new technique for drug metabolite structure elucidation. *J. Pharm. Biomed. Anal.* 15, 1903–1912
45. Khakimov, B. *et al.* (2014) High-throughput cereal metabolomics: current analytical technologies, challenges and perspectives. *J. Cereal Sci.* 59, 393–418
46. Manetti, C. *et al.* (2006) A metabonomic study of transgenic maize (*Zea mays*) seeds revealed variations in osmolytes and branched amino acids. *J. Exp. Bot.* 57, 2613–2625
47. Levandi, T. *et al.* (2008) Capillary electrophoresis time-of-flight mass spectrometry for comparative metabolomics of transgenic versus conventional maize. *Anal. Chem.* 80, 6329–6335
48. Leon, C. *et al.* (2009) Metabolomics of transgenic maize combining Fourier transform-ion cyclotron resonance-mass spectrometry, capillary electrophoresis-mass spectrometry and pressurized liquid extraction. *J. Chromatogr. A* 1216, 7314–7323
49. Frank, T. *et al.* (2012) Metabolite profiling of maize kernels–genetic modification versus environmental influence. *J. Agric. Food Chem.* 60, 3005–3012
50. Skogerson, K. *et al.* (2010) Impact of genetics and environment on the metabolite composition of maize grain. *J. Agric. Food Chem.* 58, 3600–3610
51. Riedelsheimer, C. *et al.* (2013) The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits. *Sci. Rep.* 3, 2479
52. Bueschl, C. *et al.* (2014) A novel stable isotope labelling assisted workflow for improved untargeted LC-HRMS based metabolomics research. *Metabolomics* 10, 754–769
53. Laurie, C.C. *et al.* (2004) The genetic architecture of response to long-term artificial selection for oil concentration in the maize kernel. *Genetics* 168, 2141–2155
54. Dudley, J.W. (2008) Epistatic interactions in crosses of Illinois high oil × Illinois low oil and of Illinois high protein × Illinois low protein corn strains. *Crop Sci.* 48, 59–68
55. Wassom, J.J. *et al.* (2008) QTL associated with maize kernel oil, protein, and starch concentrations; Kernel mass; and grain yield in Illinois high oil × B73 backcross-derived lines. *Crop Sci.* 48, 243–252
56. Li, H. *et al.* (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* 45, 43–50
57. Zheng, P. *et al.* (2008) A phenylalanine in DGAT is a key determinant of oil content and composition in maize. *Nat. Genet.* 40, 367–372
58. Chai, Y. *et al.* (2012) Validation of DGAT1-2 polymorphisms associated with oil content and development of functional markers for molecular breeding of high-oil maize. *Mol. Breed.* 29, 939–949
59. Hao, X. *et al.* (2014) Transferring a major QTL for oil content using marker-assisted backcrossing into an elite hybrid to increase the oil content in maize. *Mol. Breed.* 34, 739–748
60. Gibbon, B.C. and Larkins, B.A. (2005) Molecular genetic approaches to developing quality protein maize. *Trends Genet.* 21, 227–233
61. Mertz, E.T. *et al.* (1964) Mutant gene that changes protein composition and increases lysine content of maize endosperm. *Science* 145, 279–280
62. Bressani, R. (1991) Protein quality of high-lysine maize for humans. *Cereal Foods World* 36, 806–811
63. Food and Agriculture Organization (1992) Maize in human nutrition. In *FAO Food and Nutrition Series (no. 25)*. Food and Agriculture Organization (FAO) of the United Nations
64. Wang, T. *et al.* (2015) Genetic basis of maize kernel starch content revealed by high-density single nucleotide polymorphism markers in a recombinant inbred line population. *BMC Plant Biol.* 15, 288
65. Wilson, L.M. *et al.* (2004) Dissection of maize kernel composition and starch production by candidate gene association. *Plant Cell* 16, 2719–2733
66. Harjes, C.E. *et al.* (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319, 330–333
67. Von Grebmer, K. *et al.* (2014) *Global Hunger Index,* International Food Policy Research Institute
68. Azmach, G. *et al.* (2013) Marker-trait association analysis of functional gene markers for provitamin A levels across diverse tropical yellow maize inbred lines. *BMC Plant Biol.* 13, 227
69. Babu, R. *et al.* (2013) Validation of the effects of molecular marker polymorphisms in LcyE and CrtRB1 on provitamin A concentrations for 26 tropical maize populations. *Theor. Appl. Genet.* 126, 389–399
70. Menkir, A. *et al.* (2014) Carotenoid accumulation and agronomic performance of maize hybrids involving parental combinations from different marker-based groups. *Food Chem.* 148, 131–137
71. Li, Q. *et al.* (2012) Genome-wide association studies identified three independent polymorphisms associated with alpha-tocopherol content in maize kernels. *PLoS ONE* 7, e36807
72. Lipka, A.E. *et al.* (2013) Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. *G3 Genes Genomes Genet.* 3, 1287–1299
73. Fitzpatrick, T.B. *et al.* (2012) Vitamin deficiencies in humans: can plant science help? *Plant Cell* 24, 395–414
74. Galili, G. *et al.* (2016) The regulation of essential amino acid synthesis and accumulation in plants. *Annu. Rev. Plant Biol.* 67, 153–178
75. Ufaz, S. and Galili, G. (2008) Improving the content of essential amino acids in crop plants: goals and opportunities. *Plant Physiol.* 147, 954–961

76. Wen, W. *et al.* (2016) Combining quantitative genetics approaches with regulatory network analysis to dissect the complex metabolism of the maize kernel. *Plant Physiol.* 170, 136–146
77. Grotewold, E. *et al.* (1994) The myb-homologous p-gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell* 76, 543–553
78. Wang, L. *et al.* (2014) Comparative analyses of C-4 and C-3 photosynthesis in developing leaves of maize and rice. *Nat. Biotechnol.* 32, 1158–1165
79. Pick, T.R. *et al.* (2011) Systems analysis of a maize leaf developmental gradient redefines the current C-4 model and provides candidates for regulation. *Plant Cell* 23, 4208–4220
80. Shull, G.E. (1908) The composition of a field of maize. *Rep. Am. Breed. Assoc.* 4, 296–301
81. Chen, Z.J. (2013) Genomic and epigenetic insights into the molecular bases of heterosis. *Nat. Rev. Genet.* 14, 471–482
82. Riedelsheimer, C. *et al.* (2012) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat. Genet.* 44, 217–220
83. Schauer, N. *et al.* (2008) Mode of inheritance of primary metabolic traits in tomato. *Plant Cell* 20, 509–523
84. Martin, C. *et al.* (2013) Plants, diet, and health. *Annu. Rev. Plant Biol.* 64, 19–46
85. Zhou, Y. *et al.* (2012) *ZmcrtRB3* encodes a carotenoid hydroxylase that affects the accumulation of α-carotene in maize kernel. *J. Integr. Plant Biol.* 54, 260–269
86. Meihls, L.N. *et al.* (2013) Natural variation in maize aphid resistance is associated with 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside methyltransferase activity. *Plant Cell* 25, 2341–2355
87. Fu, Z. *et al.* (2013) Natural variation in the sequence of *PSY1* and frequency of favorable polymorphisms among tropical and temperate maize germplasm. *Theor. Appl. Genet.* 126, 923–935
88. Yan, J. *et al.* (2010) Rare genetic variation at *Zea mays* crtRB1 increases beta-carotene in maize grain. *Nat. Genet.* 42, 322–327
89. Kandianis, C.B. *et al.* (2013) Genetic architecture controlling variation in grain carotenoid composition and concentrations in two maize populations. *Theor. Appl. Genet.* 126, 2879–2895
90. Owens, B.F. *et al.* (2014) A foundation for provitamin A biofortification of maize: genome-wide association and genomic prediction models of carotenoid levels. *Genetics* 198, 1699–1709
91. Sharma, M. *et al.* (2011) Identification of the *Pr1* gene product completes the anthocyanin biosynthesis pathway of maize. *Genetics* 188, 69–79
92. Szalma, S.J. *et al.* (2005) Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. *Theor. Appl. Genet.* 110, 1324–1333
93. Li, L. *et al.* (2011) An 11-bp insertion in *Zea mays fatb* reduces the palmitic acid content of fatty acids in maize grain. *PLoS ONE* 6, e24699
94. Zhang, W. *et al.* (2013) Increasing lysine content of waxy maize through introgression of *Opaque-2* and *Opaque-16* genes using molecular assisted and biochemical development. *PLoS ONE* 8, e56227
95. Babu, R. *et al.* (2005) Two generation marker aided backcrossing for rapid conversion of normal maize lines to quality protein maize (QPM). *Theor. Appl. Genet.* 111, 888–897
96. Liu, L. *et al.* (2015) Introgression of the *crtRB1* gene into quality protein maize inbred lines using molecular markers. *Mol. Breed.* 35, 154
97. de Lange, E.S. *et al.* (2014) Insect and pathogen attack and resistance in maize and its wild ancestors, the teosintes. *New Phytol.* 204, 329–341
98. Davila-Flores, A.M. *et al.* (2013) Facilitated by nature and agriculture: performance of a specialist herbivore improves with host–plant life history evolution, domestication, and breeding. *Oecologia* 173, 1425–1437
99. Kempel, A. *et al.* (2011) Tradeoffs associated with constitutive and induced plant resistance against herbivory. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5685–5689
100. Maag, D. *et al.* (2015) Metabolomics in plant–herbivore interactions: challenges and applications. *Entomol. Exp. Appl.* 157, 18–29
101. Marti, G. *et al.* (2013) Metabolomics reveals herbivore-induced metabolites of resistance and susceptibility in maize leaves and roots. *Plant Cell Environ.* 36, 621–639
102. Balmer, D. *et al.* (2013) Induced resistance in maize is based on organ-specific defence responses. *Plant J.* 74, 213–225
103. Tzin, V. *et al.* (2015) Dynamic maize responses to aphid feeding are revealed by a time series of transcriptomic and metabolomic assays. *Plant Physiol.* 169, 1727–1743
104. Erb, M. *et al.* (2015) Indole is an essential herbivore-induced volatile priming signal in maize. *Nat. Commun.* 6, 6273
105. Djamei, A. *et al.* (2011) Metabolic priming by a secreted fungal effector. *Nature* 478, 395
106. Asselin, J.A.E. *et al.* (2015) Perturbation of maize phenylpropanoid metabolism by an AvrE family type III effector from *Pantoea stewartii*. *Plant Physiol.* 167, 1117–1135
107. Kump, K.L. *et al.* (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* 43, 163–168
108. Poland, J.A. *et al.* (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6893–6898
109. Obata, T. *et al.* (2015) Metabolite profiles of maize leaves in drought, heat, and combined stress field trials reveal the relationship between metabolism and grain yield. *Plant Physiol.* 169, 2665–2683
110. Suzuki, N. *et al.* (2014) Abiotic and biotic stress combinations. *New Phytol.* 203, 32–43
111. Forieri, I. *et al.* (2016) Salinity stress effects on direct and indirect defence metabolites in maize. *Environ. Exp. Bot.* 122, 68–77
112. Witt, S. *et al.* (2012) Metabolic and phenotypic responses of greenhouse-grown maize hybrids to experimentally controlled drought stress. *Mol. Plant* 5, 401–417
113. Vaughan, M.M. *et al.* (2015) Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. *Plant Cell Environ.* 38, 2195–2207
114. Yan, J. *et al.* (2015) Accumulation of 5-hydroxynorvaline in maize (*Zea mays*) leaves is induced by insect feeding and abiotic stress. *J. Exp. Bot.* 66, 593–602
115. Liu, X. *et al.* (2013) Overexpression of the phosphatidylinositol synthase gene (*ZmPIS*) conferring drought stress tolerance by altering membrane lipid composition and increasing ABA synthesis in maize. *Plant Cell Environ.* 36, 1037–1055
116. Maron, L.G. *et al.* (2013) Aluminum tolerance in maize is associated with higher MATE1 gene copy number. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5241–5246