Plant Communications
Review Article

CellPress
PARTNER JOURNAL

The Past, Present, and Future of Maize Improvement: Domestication, Genomics, and Functional Genomic Routes toward Crop Enhancement

Jie Liu[1,*], Alisdair R. Fernie[2] and Jianbing Yan[1,*]

[1]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

[2]Department of Molecular Physiology, Max-Planck-Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

***Correspondence: Jie Liu (jieliu@mail.hzau.edu.cn), Jianbing Yan (yjianbing@mail.hzau.edu.cn)**

https://doi.org/10.1016/j.xplc.2019.100010

ABSTRACT

After being domesticated from teosinte, cultivated maize (*Zea mays* ssp. *mays*) spread worldwide and now is one of the most important staple crops. Due to its tremendous phenotypic and genotypic diversity, maize also becomes to be one of the most widely used model plant species for fundamental research, with many important discoveries reported by maize researchers. Here, we provide an overview of the history of maize domestication and key genes controlling major domestication-related traits, review the currently available resources for functional genomics studies in maize, and discuss the functions of most of the maize genes that have been positionally cloned and can be used for crop improvement. Finally, we provide some perspectives on future directions regarding functional genomics research and the breeding of maize and other crops.

Key words: maize, domestication, genomics, functional genomics, improvement

Liu J., Fernie A.R., and Yan J. (2020). The Past, Present, and Future of Maize Improvement: Domestication, Genomics, and Functional Genomic Routes toward Crop Enhancement. Plant Comm. **1**, 100010.

INTRODUCTION

After Columbus arrived in the New World and brought maize to Europe, it subsequently spread throughout the world (Brandolini and Brandolini, 2009; Nunn and Nancy, 2010). Nowadays, its production exceeds that of wheat (*Triticum aestivum*) and rice (*Oryza sativa*) (http://www.fao.org). Indeed maize has become the most productive cereal crop and therefore has a massively important role, directly or indirectly, in feeding human beings. At the beginning of the last century, a controversy arose concerning the origin of maize. One hypothesis stated that maize is a species resulting from hybridization of a close relative of maize and *Tripsacum*; a second stated that the oldest wild maize is the ancestor of modern maize (Beadle, 1972). Subsequent genetic and archaeological evidence indicated that one subgroup of teosinte, *Zea mays* ssp. *parviglumis*, is the only ancestor of maize, a finding that is now widely accepted (Benz, 2001; Piperno and Flannery, 2001; Matsuoka et al., 2002; Vallebueno-Estrada et al., 2016). The divergence between *Z. mays* ssp. *parviglumis* and maize occurred about 9000 years ago around the Balsas River of southwest Mexico (Matsuoka et al., 2002).

Some key traits of teosinte are very different from those of maize, for example, its abundant branches and tillers, increased number of ears per plant, reduced number of kernels per ear (5–12 per ear for teosinte and several hundred for maize), and small kernels with a hardened fruitcase (reviewed in Doebley, 2004). How many loci were involved in the domestication of maize? More than 40 years ago, Beadle (1972) reported that he had planted a teosinte–maize F_2 population consisting of 50 000 individuals and found that the frequency of parental types was ~1 in 500, then estimated that there were four or five major loci involved in maize domestication. Later, Doebley and Stec (1993) mapped five major quantitative trait loci (QTLs) plus some minor-effect QTLs for key traits in which teosinte and maize differ. This result was consistent with Beadle's estimation and indicated that a small number of loci were responsible for the teosinte–maize morphological difference. Wright et al. (2005) investigated 774 genes and estimated that 2%–4% of maize genes were selected during maize domestication and subsequent improvement. According to recently released gene annotations of high-quality maize

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, CAS.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Gene	Phenotype	Functional annotation	Selection type	References
tb1	Plant architecture	TCP transcription factor	Increased expression	Doebley et al., 1997; Studer et al., 2011
gt1	Plant architecture	Homeodomain leucine zipper	Increased expression	Whipple et al., 2011; Wills et al., 2013
tru1	Plant architecture	BTB/POZ ankyrin repeat protein	Increased expression	Dong et al., 2017
tga1	Hardened fruitcase	SBP-domain transcription factor	Protein function	Wang et al., 2005
ZmSWEET4c	Seed filling	Hexose transporter	Increased expression	Sosso et al., 2015
UB3	Kernel row number	SBP-box transcription factor	Altered expression	
ids1/Ts6	Kernel row number	AP2 transcription factor	Increased expression	Wang et al., 2019
ZmSh1-1 *ZmSh1-5.1+ZmSh1-5.2*	Shattering	YABBY transcription factor	Protein function	Lin et al., 2012
ra1	Inflorescence architecture	Transcription factor	Altered expression	Sigmon and Vollbrecht, 2010

Table 1. Key Domestication Genes Cloned in Maize.

genomes, modern maize contains ∼39 000–42 000 protein-coding genes (Jiao et al., 2017; Springer et al., 2018; Sun et al., 2018; Yang et al., 2019a), indicating that ∼800–1700 protein-coding genes (39 000 × 2% ≈ 800; 42 000 × 4% ≈ 1700) underwent selection during the process of domestication. Enabled by the recent massive decrease in the cost of next-generation sequencing (NGS) technologies, population-scale genome resequencing has provided considerably more details about these selected regions. By resequencing improved maize varieties, landraces, and wild relatives, then scanning for selective sweeps, Hufford et al. (2012) identified 484 domestication loci, 107 of which were further selected during improvement, and 695 improvement loci covering ∼7.6% of the maize genome. These findings suggest that only a limited number of genes and a small proportion of the genome were selected during the maize domestication and improvement processes.

Among the candidate genes or loci identified to have undergone selection were some important domestication genes that had already been cloned (see Table 1). *Teosinte branched 1* (*tb1*), which encodes a *TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL NUCLEAR ANTIGEN FACTOR* (TCP) family transcription factor, was mapped with respect to key traits (i.e., inflorescence development and apical dominance) that distinguish teosinte from maize (Doebley et al., 1995). *tb1* was then cloned by transposon tagging and shown to inhibit the outgrowth of axillary buds and enable the formation of female inflorescences (Doebley et al., 1997; reviewed in Doebley, 2004 and Doebley et al., 2006). The causal variant (a transposable element [TE]) of *tb1* is located ∼60 kb upstream (according to the B73 V4 reference) of this gene (Figure 1A), which suggests that long-range chromatin interactions may have had an important role in maize domestication (Li et al., 2019b). This TE could enhance the expression of *tb1* and then lead to repression of axillary growth and increased apical dominance through regulation of the expression of important development-related genes (such as cell-cycle genes *pcna2* and *mcm2*) and direct/indirect gene-regulatory networks (Doebley et al., 2006; Studer et al., 2017). The role of *tb1* in affecting apical dominance is not restricted to maize but has been conserved in other plant species. The rice homolog of *tb1*, *OsTB1*, also functions as a negative regulator of branching in rice (Takeda et al., 2003); RNAi of *BRANCHED1* (*BRC1*) in *Arabidopsis thaliana* leads to a higher number of branches (Aguilar-Martínez et al., 2007); barley plants with the mutant allele of *HvTB1* have a significantly higher number of tillers at the juvenile stage (Ramsay et al., 2011); and *TaTB1* interacts with *FLOWERING LOCUS T1* and regulates inflorescence architecture in bread wheat (*Triticum aestivum*) (Dixon et al., 2018). The conserved function of *tb1* suggests its great potential for domesticating existing wild plants.

Another well-characterized maize domestication gene is *tga1* (*teosinte glume architecture 1*), which encodes a squamosa-promoter binding protein (SBP) transcription factor. *tga1* is responsible for the shift from kernels encased in a hardened fruitcase in teosinte to naked kernels exposed on the ear of modern maize. This QTL was mapped to a single Mendelian locus on chromosome 4 (Dorweiler et al., 1993) and was subsequently cloned (Wang et al., 2005). The causal variant of *tga1* is a single-nucleotide polymorphism (SNP) in the first exon that leads to a single amino acid change (Lys to Asn) at the sixth amino acid of TGA1, a change that affects the binding of TGA1 to its targets (Wang et al., 2005). Interestingly, *tga1* is regulated by *tb1* through direct binding of TB1 to two GGNCCC motifs in the promoter region of *tga1* (Studer et al., 2017).

In addition to *tb1* and *tga1*, dozes of genes have been isolated that control key traits for the transition from teosinte to maize (Table 1). Most of these domestication genes are transcription factors, and their expression levels were upregulated during domestication (Figure 1B–1D). Thus, their altered expression likely affected the expression of large numbers of downstream genes, including additional transcription factors, as illustrated by the regulatory network in which *tb1* and *tga1* are involved (Studer et al., 2017). This suggests that selection at the level of a regulatory network might be much more efficient in driving phenotypic changes as compared with variations that affect the function of an individual protein (Liu et al., 2015c).

Recently, two studies using ChIA-PET (chromatin interaction analysis by paired-end tag sequencing) technology mapped

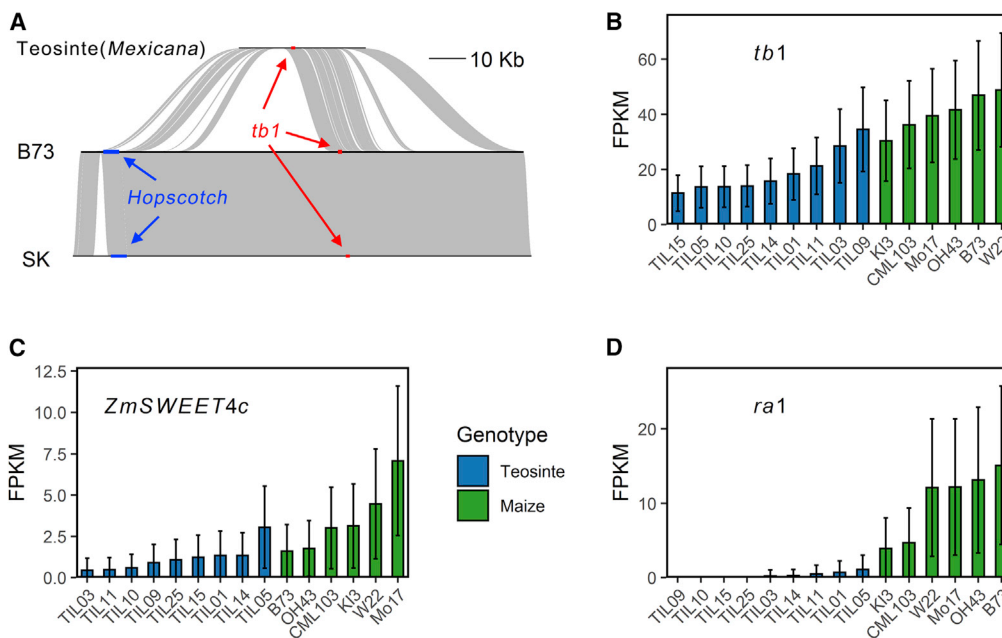

Figure 1. Genomic Sequence Variants in the *tb1* Regions of Teosinte and Tropical and Temperate Maize Lines.

(A) The red rectangles indicate the position of *tb1* and the blue rectangles indicate the position of the *Hopscotch* TE. This TE is the functional variant of *tb1* and is absent in teosinte. This *Hopscotch* TE is located ~60 kb upstream of *tb1* in temperate line B73 (RefGen_v4, Jiao et al., 2017) and tropical line SK (Yang et al., 2019a).

(B–D) The increased expression levels of representative selected genes (*tb1* in **B**, *ZmSWEET4c* in **C**, *ra1* in **D**) in modern elite maize lines compared with teosinte, the ancestor of maize. The expression profile was obtained by analyzing RNA-seq data generated by Lemmon et al. (2014).

genome-wide chromatin interactions and revealed their connections to gene-expression regulation (Li et al., 2019b; Peng et al., 2019), including the chromatin interactions in which *TB1*, *UB3*, *ZmCCT9*, *Vgt1* involved. Likely, population-scale identification of chromatin interactions would allow the detection of many more important regulatory elements, providing a number of useful selection targets for future maize improvement.

MAIZE FUNCTIONAL GENOMICS RESEARCH RESOURCES

Reference Genome Sequences

The first maize reference genome sequence (B73 RefGen_v1) was released in 2009 (Schnable et al., 2009) and relied on the sequencing of bacterial artificial chromosomes and fosmids; a subsequent release in 2017 was improved by the use of single-molecule technologies (Jiao et al., 2017). Since the release of the B73 reference genome sequence, it has been extensively used for maize functional genomics research. By comparing other inbred lines with B73, millions of SNPs and insertions/deletions (InDels; e.g., those of *tb1* as illustrated in Figure 1) were identified, along with many structural variations (SVs) (examples of which are shown in Figure 2), presence/absence variation (PAV), expression presence/absence variation (ePAV), and copy-number variation (Springer et al., 2009; Lai et al., 2010; Jiao et al., 2012; Fu et al., 2013; Hirsch et al., 2014; Lu et al., 2015; Jin et al., 2016; Bukowski et al., 2018; Sun et al., 2018; Yang et al., 2019a).

The use of high-density SNPs, which were identified by resequencing maize lines and populations and mapping those sequences to the B73 reference genome, has greatly expanded our knowledge concerning maize domestication and development (Hufford et al., 2012) and the genetic basis of important traits such as flowering time (Buckler et al., 2009), disease resistance (Kump et al., 2011), oil and vitamin metabolism (Li et al., 2013; Wang et al., 2018a), plant height (Pan et al., 2017), and yield-related traits (Xiao et al., 2016; Liu et al., 2017c). Maize SNP arrays designed for molecular breeding by integrating SNPs associated with known genes and traits have been reported (Yan et al., 2010b; Xu et al., 2017). However, most resequencing research in maize and other crops has focused on identifying SNPs through mapping to a single reference genome, which limits the ability to make full use of genome data, detect SVs, and fully reveal the genetic diversity (Yang et al., 2019b). Recently, human graph-reference genomes, into which existing variations could be incorporated, have been used for alignments and genotyping (Garrison et al., 2018; Kim et al., 2019; Rakocevic et al., 2019). This strategy can take advantage of multiple genome sequences and reported variations and can increase the variant-calling accuracy, especially for SVs. We speculate that this graph-reference genome approach or other newly developed methods will soon be adopted by researchers studying maize and other crop genomes in order to integrate the increasing amounts of variation data and genome sequences that are becoming available.

A further disadvantage of resequencing studies is that most of them fail to detect associations between SVs and traits, despite

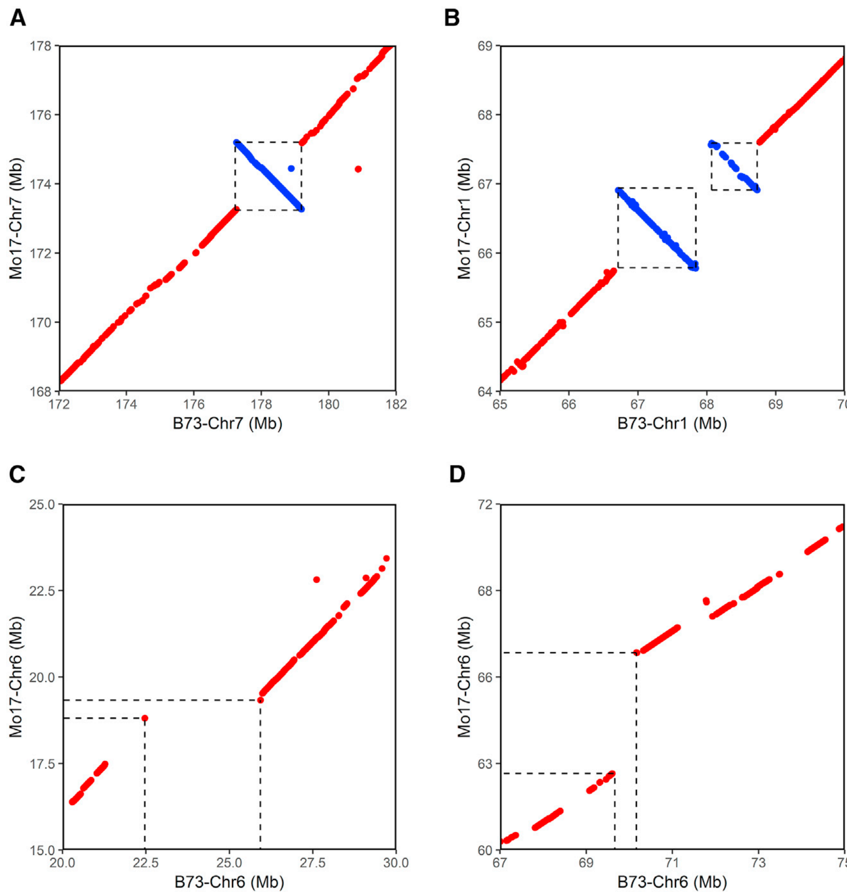

Figure 2. Examples of Large Structural Variations between Two Maize Lines, B73 and Mo17.

(A) An ~2-Mb inversion on chromosome 7, as indicated by blue dots and dashed box.

(B) Two adjacent inversions covering ~2 Mb on chromosome 1, as indicated by blue dots and dashed boxes.

(C) An ~3-Mb deletion in Mo17, as compared with B73, on chromosome 6.

(D) An ~3-Mb insertion in Mo17, as compared with B73, on chromosome 6.

The alignment was performed using MUMMER 3.23 (Kurtz et al., 2004) with a minimum match length of 1 kb.

growing evidence that SVs affect important agronomic traits, e.g., flowering time, plant height, kernel weight, disease resistance, and metabolic traits (Hirsch et al., 2014; Lu et al., 2015; Zuo et al., 2015; Jin et al., 2016). At a genome-wide level, it was demonstrated that 22% of SVs could not be detected using SNPs (Yang et al., 2019a). This indicates that more reference genome sequences are required to fully characterize the maize functional genome.

With the development of new sequencing technologies and the decreased costs of sequencing, several additional maize genome sequences have been reported since 2016. These include the genomes of PH207 (Hirsch et al., 2016), *mexicana* (Yang et al., 2017a), Mo17 (Yang et al., 2017a; Sun et al., 2018), W22 (Springer et al., 2018), HZS (Li et al., 2019a), and SK (Yang et al., 2019a). Among these released genomes, the B73, Mo17, W22, and SK genomes have high-quality assemblies (Table 2) and thus could be representative genomes for maize genomics research. The high-quality assemblies enabled by single-molecular sequencing, which are currently available for B73 Ref_V4, Mo17, and SK, have greater advantages in annotation of genome features and identification of important elements (such as promoters and TEs) and variations (especially structure variations) compared with assemblies based on short-read sequencing (Jiao et al., 2017; Sun et al., 2018; Yang et al., 2019a). Without the high-quality assembly of SK achieved by long-read sequencing, it would have taken tremendous effort to identify the large structure variations located upstream of *ZmBAM1d*, the gene underlying a major QTL affecting maize kernel size and weight (Yang et al., 2019a). There is no doubt that these genome sequences will greatly facilitate the revelation of maize genetic diversity, the connection between genetic variation and phenotypic variation, and the maize improvement process.

Mutant Libraries

Mutants are important resources for plant functional genomics research. Unlike *Agrobacterium* T-DNA insertion mutant libraries in *Arabidopsis* and rice (Alonso et al., 2003; Wu et al., 2003; Zhang et al., 2006), maize researchers mainly use transposon mutagenesis approaches to tag genes for cloning. Many genes for important agronomic traits have been cloned by transposon tagging. The most well known is *tb1*, a major domestication locus affecting tiller number and inflorescence development (Doebley et al., 1997). Several research programs have tagged maize genes with active *Mu*, including the Trait Utility System for Corn, Maize Targeted Mutagenesis database, *Mu* array, *RescueMu*, Photosynthetic Mutant Screen, and UniformMu (reviewed in Brutnell, 2002). Among these programs, UniformMu is currently the most widely used (McCarty et al., 2005). However, only ~30% of maize genes have UniformMu insertions. Apart from transposon mutagenesis, targeting induced local lesions in genomes (TILLING) has also been used to construct maize mutant libraries (Till et al., 2004; Lu et al., 2018). In this approach, point mutations, caused by exposure to ethyl methanesulfonate, are detected with high-throughput methods or NGS. The mutant library constructed by Lu et al. (2018) covers ~80% of maize genes (32 069 of 39 591), with each line containing on average 4500 mutations.

Nurtured and Natural Populations

As molecular markers were applied to construct linkage maps and to map QTLs for quantitative traits, many different kinds of mapping populations were generated, not only in maize but also in other crops, including classic F_2, backcross, chromosome segment substitution line, recombinant inbred line (RIL), and multi-parent advanced generation intercross populations. In

Genome	Assembly size (Mb)	No. of scaffolds	Scaffold N50	Contig N50	No. of gene models	References
B73	2106	625	9.56 Mb	1.18 Mb	39 324	Jiao et al., 2017
Mo17 (CAU)	2183	2560	10.2 Mb	1.48 Mb	38 620	Sun et al., 2018
W22	2134	306	35.52 Mb	72.4 kb	40 789	Springer et al., 2018
SK	2161	708	73.24 Mb	15.78 Mb	43 271	Yang et al., 2019a
PH207	2102	127 488	654 kb	5.3 kb	37 613	Hirsch et al., 2016
Mo17 (Yan)	2042	48 268	3.00 Mb	60.5 kb	40 003	Yang et al., 2017a
mexicana	1204	107 418	108 kb	26.6 kb	31 387	Yang et al., 2017a
HZS	2209	12	223.93 Mb	78.2 kb	40 893	Li et al., 2019a

Table 2. Summary of Released Maize (Teosinte) Genomes.

addition to these commonly used populations, maize researchers have constructed other types of high-resolution mapping populations that have been used extensively for genetic mapping. One such population is the intermated B73-Mo17 (IBM) population, a type of intermated RIL (IRIL) population. Unlike regular RIL populations, the F_2 progenies of crosses between B73 and Mo17 were intermated for four generations before repeated selfing by single-seed descent to generate RILs. Because of the extra generations of intermating, the genetic map of this population has a 2.7-fold greater recombination fraction distance and a 3.86-fold longer map length (Lee et al., 2002), thus significantly improving the QTL mapping resolution. Another such population is a nested association mapping (NAM) population consisting of 25 RIL populations with a common parent line (B73) with a sequenced genome (Yu et al., 2008). This population captured ~136 000 recombination events with three recombinations per gene on average. This genetic design took advantage of both linkage mapping and association mapping (McMullen et al., 2009). Both IBM and NAM populations have been used to dissect the genetic basis of complex traits, such as flowering time and disease resistance (Balint-Kurti et al., 2007; Buckler et al., 2009; Kump et al., 2011), and the population resources, including parental lines, are now publicly available (https://maizegdb.org/stock_catalog). Recently, Xiao et al. (2016) developed a random-open-parent association mapping population. This type of population has several advantages, the most impressive of which is that newly developed populations can be integrated into existing populations (reviewed in Xiao et al., 2017). It has been successfully used to dissect the genetic basis of cob traits (Xiao et al., 2016), kernel traits (Liu et al., 2017c), and plant architecture (Pan et al., 2017).

An association test between *dwarf8* and flowering time was the first instance of association analysis being applied in plant genetics. This analysis tested associations between 123 polymorphic sites covering *dwarf8* and flowering time in an association mapping population consisting of 92 maize inbred lines, and used 141 simple sequence repeat loci to control for population structure (Thornsberry et al., 2001). Thereafter, many maize natural populations were used for candidate-gene-association analyses (reviewed in Yan et al., 2011; Xiao et al., 2017; Liu and Yan, 2019), and association tests were subsequently extended to genome-wide association studies (GWASs) in maize and other plants including, but not limited to, *Arabidopsis* (Atwell et al., 2010), barley (Cockram et al., 2010), rice (Huang et al., 2011), and tomato (Sauvage et al., 2014). In maize, some natural populations have been widely used for GWAS research. One such population is the "Goodman" maize diversity panel, which contains 302 inbred lines and captures the diversity present in public breeding programs (Flint-Garcia et al., 2005). This panel has been genotyped using the Illumina MaizeSNP50 Array (Olukolu et al., 2013) and genotyping-by-sequencing (GBS; Romay et al., 2013), and is now covered by maize HapMap 3 (Bukowski et al., 2018). The "Goodman" panel has been used to dissect the genetic basis of ear rot resistance (Zila et al., 2013), iron homeostasis (Benke et al., 2015), aflatoxin resistance and important agronomic traits (Farfan et al., 2015), resistance to *Fusarium verticillioides* infection (Stagnati et al., 2019), and metabolite production (Zhou et al., 2019b). Another panel-based resource is the MaizeGo panel (http://www.maizego.org/Resources.html), which consists of 540 well-chosen lines and represents the largest association mapping population panel assembled for maize (Yang et al., 2011; Liu et al., 2017b). Many efforts to dissect the genetic architecture of complex traits have been carried out with this population or with a subset of this population, including analyses of expression traits (Fu et al., 2013; Liu et al., 2015c, 2017b), metabolite levels (Li et al., 2012, 2013;; Wen et al., 2014, 2018; Deng et al., 2017; Wang et al., 2018a), agronomic traits (Yang et al., 2013, 2014), disease resistance (Chen et al., 2015; Ding et al., 2015; Li et al., 2019d), and drought tolerance (Liu et al., 2013; Mao et al., 2015; Wang et al., 2016; Xiang et al., 2017).

Databases and Datasets

The most valuable database for the maize research community is the Maize Genetics and Genomics Database (MaizeGDB; https://www.maizegdb.org/) (Portwood et al., 2019). This database was first released in the early 1990s and has since evolved into a "Swiss army knife" of datasets and tools for maize functional genomics research. First, it supports genome sequences of several maize inbred lines with a user-friendly genome browser. It also contains, among other features, large numbers of datasets covering stocks, polymorphisms, mutants and phenotypes, genetic maps, QTLs, gene annotations, and transcriptomic and proteomic data. Furthermore, it provides many tools for querying and visualizing datasets, such as SNPversity, which allows the visualization of the SNP diversity of a selected genomic region, and qTeller (https://qteller.maizegdb.org/), which allows the exploration of the expression of genes of interest. Finally, it provides additional services for the maize community such as the

coordination of the annual maize genetics conference, recommendation of publications by the MaizeGDB Editorial Board, and handling of stock requests through the Maize Genetics Cooperation Stock Center (https://www.maizegdb.org/data_center/stock).

The maize community has generated large numbers of datasets, but in this review we will focus on population-scale genotypic variation and gene-expression datasets given their important proven connections with phenotypic variation. Accompanying the release of the maize B73 reference genome, the first-generation haplotype map of maize was also released. This HapMap 1 was constructed with 3.3 million SNPs and InDels identified within the founder lines of the NAM population (Gore et al., 2009). This HapMap subsequently evolved into HapMap 2 following the deep sequencing of 103 inbred lines covering 19 teosinte accessions, 23 landraces, and 60 improved lines, including the NAM founders of HapMap 1. HapMap 2 consists of 55 million SNPs, 21% of which are located in genic regions (Chia et al., 2012), and has been used to identify genomic regions that underwent selection during maize domestication and improvement (Hufford et al., 2012). More recently, maize HapMap 3 was built from NGS data from ~1200 lines covering those used for HapMap 2, International Maize and Wheat Improvement Center (CIMMYT) lines, China Agricultural University (CAU) lines, and the Goodman panel. HapMap 3 contains ~83 million polymorphisms (Bukowski et al., 2018). In addition to the HapMap populations, the MaizeGo panel has been genotyped with SNP arrays, RNA sequencing (RNA-seq), and GBS and also has a very high SNP density (2.65 million SNPs) (Yang et al., 2011; Fu et al., 2013; Liu et al., 2017b). All information about MaizeGo panel was integrated into a database called MODEM (http://modem.hzau.edu.cn/) (Liu et al., 2016).

Temporal and spatial gene-expression profiles are critical both for the functional analysis of genes and for understanding developmental processes. Most gene-expression profiling studies in maize have focused on maize seeds or seed parts at different stages of development in order to identify networks important for kernel and endosperm development and, thus, crop yield. Using deep RNA-seq, Fu et al. (2013) profiled the expression levels of 28 769 genes and 42 211 transcripts in kernels at 15 days after pollination (DAP) from 368 inbred lines, which were a subset of the MaizeGo panel. Similarly, Li et al. (2014) reported the temporal patterns of gene expression in kernels and endosperm during eight stages within 0–12 DAP; Chen et al. (2014) and Yi et al. (2019) quantified gene-expression levels in the embryo, endosperm, and whole seed at 0–38 DAP and 0–144 h after pollination, respectively. Recently, Zhan et al. (2015) used laser-capture microdissection technology to isolate five different endosperm compartments from B73 kernels at 8 DAP and then quantified the transcription level of genes within these compartments using RNA-seq. In addition to gene-expression quantification in kernel-related tissues, there have been other reports of gene-expression profiling in multiple tissues across the entire maize life cycle. These include 79 tissues from B73 (Stelpflug et al., 2016), 23 tissues from B73 (Walley et al., 2016), and 23 tissues from B73, Mo17, and their F_1 hybrid (Zhou et al., 2019a). These diverse gene-expression profiles have already contributed much toward identifying networks for development and could also provide the first hints of the functions of genes of interest.

GENES UNDERLYING IMPORTANT QUANTITATIVE TRAITS

There are thousands of QTLs, and hundreds of genes have been identified that are important for maize agronomic traits. Among the resulting cloned genes, 464 "classic" genes were cloned based on visible morphological mutant phenotypes (Schnable and Freeling, 2011). Although these cloned mutant genes are functionally important for maize development, most of them cannot be used directly in maize-breeding programs due to their deleterious effects. In this review, we will pay much more attention to genes (Table 3) identified by positional cloning whose genetic polymorphisms are responsible for natural variation in important quantitative traits and are additionally promising candidates for marker-assisted selection (MAS).

Flowering Time-Related Genes

In the late 1990s, Vlăduţu et al. (1999) identified two tightly linked QTLs for flowering time and named them *Vegetative to generative transition 1* and *2* (*Vgt1* and *Vgt2*). *Vgt1* was then mapped to a 1.3-cM interval (Salvi et al., 2002) and was finally cloned (Salvi et al., 2007). *Vgt1* consists of a non-coding regulatory element located ~70 kb upstream of an *AP2*-like transcription factor, *ZmRap2.7*. The role of *ZmRap2.7* as a negative regulator of flowering time was confirmed in both overexpression and downregulation studies carried out by means of transgenesis (Salvi et al., 2007). Three polymorphisms (*G/A/indel324*, *MITE*, and *ATindel434*) in *Vgt1* were found to be significantly associated with flowering time in two independent association mapping studies (Salvi et al., 2007; Ducrocq et al., 2008). The miniature inverted-repeat transposable element (MITE) insertion in *Vgt1* was demonstrated to be highly methylated, and its flanking region was significantly more methylated compared with alleles without this MITE insertion (Castelletti et al., 2014). Allele-specific expression analysis of *ZmRap2.7* in F_1 plants showed that the two allele-specific transcripts exhibited significantly different expression levels (Castelletti et al., 2014).

ZmCCT is another major flowering-time QTL and was first mapped to a ~170-kb region containing five putative genes in one parental line and eight in the other (Ducrocq et al., 2009). A CCT gene, *ZmCCT10*, which is homologous to rice *Ghd7* (a large-effect locus for grain number, plant height, and heading date) (Xue et al., 2008), was proposed to be the candidate for this QTL (Ducrocq et al., 2009). This QTL was independently mapped to a ~202-kb region that includes part of *ZmCCT10*, the only gene present in the QTL interval (Hung et al., 2012). The effect of *ZmCCT10* on flowering time was confirmed by association analyses in two independent diverse panels (Hung et al., 2012; Yang et al., 2013) and also by transgenic analysis (Yang et al., 2013). The causal variation of *ZmCCT* is a 5122-bp *CACTA*-like TE located ~2.5 kb upstream of *ZmCCT10*. This TE could lead to a change in the methylation level of the *ZmCCT10* promoter that would affect the expression level of the gene (Yang et al., 2013). Another CCT gene, *ZmCCT9*, is the gene underlying *qDTA9*, a QTL for days to anthesis (DTA). This QTL was narrowed down to a ~2.4-kb region located ~57 kb

Gene	Phenotype	Functional annotation	Causal variation	References
ZmRap2.7	Flowering time	AP2-like transcription factor	*MITE* TE ~70 kb upstream of this gene	Salvi et al., 2007; Castelletti et al., 2014
ZmCCT10	Flowering time Disease resistance	CCT transcription factor	*CACTA*-like TE ~2.5 kb upstream of this gene	Hung et al., 2012; Yang et al., 2013; Wang et al., 2017
ZmCCT9	Flowering time	CCT transcription factor	*Harbinger*-like TE ~57 kb upstream of this gene	Huang et al., 2018
ZmCLA4	Leaf angle	Ortholog of *LAZY1* in rice	–	Zhang et al., 2014
ZmRAVL1	Leaf angle	B3 domain–containing protein	2-bp InDel ~9.5 kb upstream of this gene	Tian et al., 2019
brd1	Leaf angle	Brassinosteroid C-6 oxidase	–	Tian et al., 2019
ZmGA3ox2	Plant height	GA3 β-hydroxylase	Two small InDels in the promoter	Teng et al., 2013
qph1	Plant height	ABC transporter	A mis-sense SNP	Xing et al., 2015
UB3	Kernel row number	SBP-box transcription factor	1.2-kb transposon-containing insertion located ~60 kb downstream of this gene	Liu et al., 2015b
ids1/Ts6	Kernel row number	AP2 transcription factor	–	Wang et al., 2019
ZmBAM1d	Kernel size and weight	CLV1/BAM-related receptor kinase-like protein	–	Yang et al., 2019a
DGAT1-2	Oil content and composition	Diacylglycerol acyltransferase	Extra phenylalanine insertion	Zheng et al., 2008
Zmfatb	Palmitic acid	Acyl-ACP thioesterase	11-bp InDel in the sixth exon	Li et al., 2011
ZmPORB2	Tocopherol concentration	Protochlorophyllide oxidoreductase	A small InDel in 5' UTR	Zhan et al., 2019
ZmAuxRP1	Disease resistance	Domain of unknown function 966	–	Ye et al., 2019
ZmPLA1/MTL/NLD	Haploid induction rate	Phospholipase A	A 4-bp insertion led to a frame shift	Gilles et al., 2017; Kelliher et al., 2017; Liu et al., 2017a
ZmDMP	Haploid induction rate	DUF679 domain membrane protein	One single-nucleotide change	Zhong et al., 2019

Table 3. Positionally Cloned Maize Genes Related to Important Quantitative Traits.
"–" indicates that a causal variant has not yet been confirmed.

upstream of *ZmCCT9*. A Harbinger-like TE was the most significant polymorphism associated with DTA based on an association analysis. In addition, knocking out *ZmCCT9* with CRISPR/Cas9 resulted in earlier flowering under long-day conditions (Huang et al., 2018).

Interestingly, all three causal TEs underwent selection during maize adaptation to temperate regions. The frequency of the MITE allele of *Vgt1* varies from 0.3 in tropical lines to 0.87 in the European and Northern elite lines, and there is a significant correlation between latitude and MITE frequency (Ducrocq et al., 2008). The frequency of the TE alleles of *ZmCCT10* and *ZmCCT9* increased from 51.4% and 10.3% in tropical lines to 85.4% and 79% in temperate lines, respectively (Huang et al., 2018). *ZmCCT10*, *ZmCCT9*, and CCT genes in other plants (such as rice *Hd1* and *Ghd7*) have important roles in regulating flowering time, which indicates a conserved function among CCT genes. In maize there are 53 CCT genes, among which 15 are significantly associated with flowering time based on candidate-gene association analysis (Jin et al., 2018). These results underscore the substantial role of CCT genes and their related pathways in flowering-time manipulation.

Genes Regulating Plant Architecture

The use of dwarf and semi-dwarf cultivars in breeding programs was the key component of the first "Green Revolution." This suggests the importance of plant architecture in crop yield improvement. Although many mutant genes for plant architecture have been cloned, only a few QTLs for leaf angle or plant height have been positionally cloned in maize.

ZmCLA4, an ortholog of rice *OsLAZY1*, is the functional gene responsible for a leaf-angle QTL, *qLA4-1*, with a semi-dominant effect of ~15° (Zhang et al., 2014). *ZmCLA4* was differentially expressed between plants with large and small leaf angles, with

plants with larger leaf angles showing lower expression, indicating the negative role of *ZmCLA4* in controlling leaf angle. This negative role was confirmed by RNAi of *ZmCLA4* in maize and overexpression of *ZmCLA4* in rice (Zhang et al., 2014). Rice plants that overexpressed *ZmCLA4* had much smaller leaf angles than did controls, suggesting that *ZmCLA4* has a conserved role in regulating leaf angle among different crops. This is supported by the finding that both mutant and RNAi lines of *OsLAZY1* have larger tiller angles compared with their respective control lines (Li et al., 2007).

Recently, Tian et al. (2019) cloned two QTLs (*Upright Plant Architecture1* and *2* [*UPA1* and *UPA2*]) for upright plant architecture. The underlying genes for these two QTLs are *brassinosteroid C-6 oxidase1* (*brd1*) and *ZmRAVL1*, which is a homolog of rice *RAVL1*. The functional variant in *UPA2* is a 2-bp InDel located 9.5 kb upstream, which alters the binding affinity of the protein DRL encoded by *drooping leaf1*, another maize leaf-angle gene (Strable et al., 2017). While LG1, encoded by *liguleless1*, which is essential for leaf ligule and auricle development (Moreno et al., 1997), activates the expression of *ZmRAVL1*, the physical interaction between DRL1 and LG1 represses this activation (Tian et al., 2019). Interestingly, *ZmRAVL1* regulates the expression level of *brd1* and thus alters the endogenous brassinosteroid level and leaf angle (Tian et al., 2019).

qPH3.1 is a major QTL for plant height with a 10.0-cm additive effect and 3.7-cm dominant effect. This QTL was narrowed down to a 12.6-kb region, within which *ZmGA3ox2* was shown to be responsible for both *qPH3.1* and the *dwarf-1* mutant (Teng et al., 2013). This gene, which encodes a gibberellin (GA) 3-oxidase, is expressed constitutively, and higher expression results in greater endogenous GA_1 (a bioactive gibberellin) levels and taller plants. Association mapping results subsequently indicated that two 3-nucleotide repeats within this gene might be the causative polymorphism for plant-height variation (Teng et al., 2013).

qph1, another major QTL for plant height with a 17.7-cm additive effect and a 7.8-cm dominant effect, was mapped to a 1.59-kb region covered by the fifth exon of *Br2*. A mis-sense SNP leading to an amino acid change from Arg to Leu is the causal variant for this QTL (Xing et al., 2015). Mutants of maize *Br2*, which encodes an ATP-binding cassette transporter belonging to the multidrug-resistant class of P-glycoproteins, show shortening of the lower stalk internodes due to reduced auxin transport. In addition, an 882-bp duplication in exon 5 of the *Br2* homolog in sorghum is responsible for the *dw3* mutant phenotype and the instability of *dw3* caused by unequal recombination at this duplication site (Multani et al., 2003).

Yield-Related Genes

Ear- and kernel-related traits are important components of maize yield. Two major QTLs (*KRN4* and *KRN1*) have been cloned for kernel row number (KRN) and one (*qHKW1*) has been cloned for kernel size and weight.

KRN4 was mapped to a 3-kb intergenic region located ~60 kb downstream of an SBP-box gene, *UB3* (*unbranched3*), which negatively affects KRN (Chuck et al., 2014). A 1.2-kb PAV with a *harbinger* TE in this 3-kb region and a SNP (S35) in the third exon of *UB3* additively contribute to KRN, and these two sites both underwent selection during maize improvement (Liu et al., 2015a). In both rice and maize, the UB3 protein binds to genes encoding inflorescence-related transcription factors that are involved in the cytokinin pathway, and UB3 additionally targets genes involved in the auxin pathway and *CLV-WUS* pathway in maize (Du et al., 2017).

krn1 was identified in an $F_{2:3}$ population derived from crosses between the tropical line CML220 and landrace 1091AF, and had a phenotypic variation explained (PVE) value of 22%. This QTL was then fine-mapped to a 6.6-kb region, and the causal gene, which was confirmed by association analysis and two other independent mutation events, was shown to be a previously reported mutant gene, *ids1*/*Ts6* (*indeterminate spikelet 1*/*Tassel seed 6*) (Wang et al., 2019). This gene is homologous to the wheat gene *Q* (Simons et al., 2006). Mutants of these two genes have similar phenotypes including altered spike length, spike density, grain size, and plant height, and both genes also underwent selection during the domestication of these two crop species (Debernardi et al., 2017; Xie et al., 2018; Wang et al., 2019).

qHKW1 accounts for 18.4% of the variation in hundred-kernel weight (HKW) in an RIL population derived from the modern elite maize inbred line ZHENG58 and the landrace SK (Raihan et al., 2016; Liu et al., 2017c). This QTL was fine-mapped to a ~177-kb region containing only one gene, *ZmBAM1d*, which encodes a *CLAVATA1* (*CLV1*)/*BARELY ANY MERISTEM* (*BAM*)-related receptor kinase-like protein. The increased expression of *ZmBAM1d*, which is caused both by disrupted chromatin interactions resulting from distal large InDels in the promoter and altered methylation levels in the promoter region, leads to enhanced HKW. A similar effect was noted in transgenic lines that overexpress *ZmBAM1d*, which show a ~1.9-g increase in HKW (Yang et al., 2019a).

Disease Resistance-Related Genes

Disease is one of the largest contributors to maize yield loss. Strategies for cloning quantitative disease-resistance genes and the characterization of some of these cloned genes have been extensively reviewed in Yang et al. (2017b). Therefore, we focus here solely on two recently cloned disease-resistance genes, *ZmCCT10* and *ZmAuxRp1*.

Two major QTLs (*qRfg1* and *qRfg2*) conferring resistance to *Gibberella* stalk rot were identified, and *qRfg1* was mapped to a ~500-kb region (Yang et al., 2010a). *qRfg1* was then narrowed down to a ~170-kb region (Wang et al., 2017). Interestingly, *ZmCCT10* was the gene underlying this QTL, and the *CACTA*-like TE, which had been shown to be the causal polymorphism for flowering time (Yang et al., 2013), was also the causal variant for stalk rot resistance. This TE results in altered histone modification and DNA methylation in the *cis*-regulatory region of *ZmCCT10*, which leads to changes in its expression and, thus, enhanced disease resistance (Wang et al., 2017).

qRfg2 was first mapped to a ~300-kb region (Zhang et al., 2012) and then to a ~2.6-kb region containing only one gene,

ZmAuxRP1, which encodes a protein with a domain of unknown function (DUF966). RNAi of *ZmAuxRP1* results in plants with greater resistance, whereas plants that overexpress this gene are more susceptible to *Gibberella* stalk rot (Ye et al., 2019). It turns out that this gene regulates resistance to *Gibberella* stalk rot by affecting the biosynthesis of both indole-3-acetic acid (IAA) and benzoxazinoids (Ye et al., 2019).

Nutrition-Related Genes

The 21st century has seen an increase in concerns about nutrition, as indicated by a PubMed search with the keywords "nutrition" or "metabolome," which reveals a rapid increase in published studies since 2010. Searches with the phrase "maize nutrition" or "maize metabolome" produce similar results. Biofortification, the process of increasing nutrition in crops through plant breeding (Nestel et al., 2006), has been successfully applied in maize nutrition improvement through MAS for, for instance, protein content and provitamin A (Muthusamy et al., 2014; Zunjare et al., 2018). Selection of the recessive allele of *Opaque2* increases the lysine content of kernels (Mertz et al., 1964) while favorable alleles of *lcyE* and *crtRB1* enhance provitamin A levels in kernels (Harjes et al., 2008; Yan et al., 2010a).

A QTL (*qHO6*) that accounts for 11% and 9.5% of the variation in seed oil concentration and embryo oil concentration, respectively, was mapped to a 4.8-kb region. *DGAT1-2*, which encodes a type I acyl-coenzyme A:diacylglycerol acyltransferase, is the gene underlying this QTL (Zheng et al., 2008). Ectopic expression of the high-oil allele of this gene under the control of an embryo-specific promoter leads to a 27.9%, 26.1%, and 84.5% increase in seed oil, embryo oil, and oleic acid concentration, respectively, and to a 28.1% decrease in linoleic acid concentration. The causal polymorphism is the insertion of an extra phenylalanine at position 469 of DGAT1-2, which leads to the enhancement of its enzymatic activity (Zheng et al., 2008).

Palmitic acid (C16:0) is an important class of saturated fatty acids. Eight QTLs mapped using an RIL population derived from B73 (normal-oil line) and By804 (high-oil line) explain 65.3% of the total phenotypic variance in palmitic acid concentration, with *pal9* having the largest effect (42.0% PVE) (Yang et al., 2010b). *pal9* was then fine-mapped to a 90-kb region covering only one gene, *Zmfatb*, which encodes an acyl-ACP thioesterase. An 11-bp InDel in the last exon of *Zmfatb* is the causal variant detected by candidate-gene-association analysis, and this InDel introduces a premature stop codon that affects enzyme activity (Li et al., 2011).

ZmPORB2, encoding a protochlorophyllide oxidoreductase, was identified as the causal gene underlying a major QTL (*qVE5*) for tocopherol content in maize kernels mapped in three RIL populations (Wang et al., 2018a; Zhan et al., 2019). An InDel, called Indel058, located 59 bp upstream of the ATG of *ZmPORB2*, which allowed separation of diverse inbred lines into three groups (one group with no insertion, another two groups with 5- or 8-bp insertions), was the most significant variation detected by candidate-gene-association analysis and is likely the causal variant. Lines with no insertion had a much higher *ZmPORB2* expression level and tocopherol concentration. Overexpression of this gene led to increased tocopherol content in both the leaf and kernel (Zhan et al., 2019), which suggests great potential in vitamin E biofortification.

These are by no means the only metabolic QTLs that are important for nutrition in maize, with several recent QTL studies based on maize introgression or RILs, or alternatively on association panels, profiling such nutritionally important compounds as essential amino acids, lipids, vitamins and flavonoids, and other antioxidants (Riedelsheimer et al., 2012; Wen et al., 2015, 2016). In a similar vein, although such studies have not yet been reported on a large scale in maize, a huge number of QTLs for taste have been found in studies in many species, most notably in tomato (Tieman et al., 2017; Gao et al., 2019; Zhao et al., 2019) and cucumber (Shang et al., 2014; Zhu et al., 2019). Encompassing both traits, the first papers studying the domestication of the plant metabolome have recently been published for wheat, tomato, and maize (Beleggia et al., 2016; Zhu et al., 2018; Li et al., 2019c; Xu et al., 2019). The first of these studies carried out in wheat demonstrated that a reduction in unsaturated fatty acids was associated with selection during domestication of emmer while changes in the amino acid content due to selection mark the domestication of durum wheat (Beleggia et al., 2016). Similarly, tomato domestication was characterized by a shift in the ratios of nutritional and antinutritional steroidal glycoalkaloids as well as alterations in phenylpropanoid content driven by the Asian preference for pink tomatoes and considerable changes in metabolism that were due to genes hitchhiking on those genes selected to improve tomato yield (Zhu et al., 2018). Two recent studies have carried out a detailed comparison of metabolism in large populations derived from maize-by-teosinte crosses, revealing large-scale changes in both primary (particularly in amino acids and lipids [Li et al., 2019c; Xu et al., 2019]) and secondary (particularly in alkaloids and terpenoids [Xu et al., 2019]) metabolism that occurred during domestication. A current complication in using insights derived from the metabolomics data described above for crop improvement is that often the metabolite function is not well characterized (Alseekh and Fernie, 2018). With increasing use of wild germplasm, such studies identifying the changes that occurred during domestication to make our food more palatable (Zhu et al., 2018) will, however, become essential in order to prevent the reintroduction of antinutritional compounds.

Lessons Learned from Characterization of Positionally Cloned Genes

One lesson we have learned from these positionally cloned genes is that SVs and TEs are of great importance in shaping plants and enhancing their fitness. SVs and TEs are causal variations not only for important QTLs (*tb1*, *Vgt1*, *ZmCCT*, *ZmCCT9*, *KRN4*, *qHKW1*) as discussed above, but also for other genes identified by GWASs (such as *ZmNAC111* and *ZmVPP1* for drought tolerance; Mao et al., 2015; Wang et al., 2016). TEs can lead to phenotypic plasticity in plants and humans through regulating gene expression (Wei and Cao, 2016). The second lesson is that both natural variants and mutant alleles of some cloned genes (*ZmGA3ox2*, *Br2*, *UB3*, *ids1/Ts6*) are responsible for variation in important quantitative traits. Other studies have had similar findings. For example, a weak allele of *fea2* increases the KRN and the number of kernels per ear without changing ear length

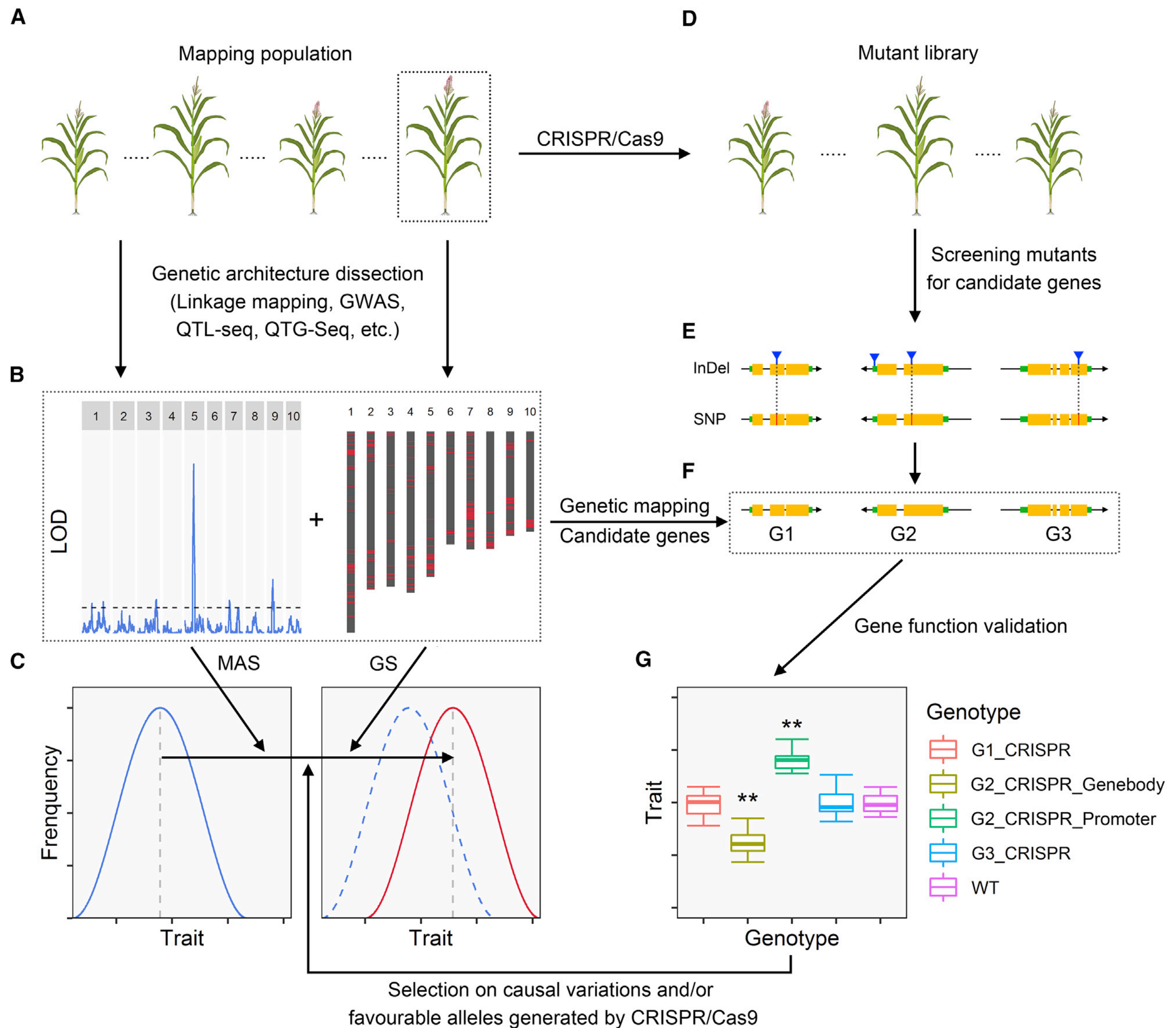

Figure 3. An Overview of Strategies for Rapid Gene Cloning and Maize Improvement.
(A–C) Genetic architecture of a target trait is dissected with, for example, linkage mapping, GWAS, QTG-seq, and QTG-seq. Different selection methods are then used based on the complexity of the genetic architecture of the target trait. For example, MAS is suitable for selection of GLS resistance, whereas GS is much more efficient for selection of hundred-kernel weight.
(D–G) After QTLs are mapped, candidate genes are proposed and the mutants for candidate genes are screened from a mutant library generated with CRISPR/Cas9. The causal gene is validated with these mutants. With causal variants and/or favorable alleles generated with CRISPR/Cas9 confirmed, they then can be applied directly in breeding programs. GWAS, genome-wide association study; MAS, marker-assisted selection; GS, genomic selection.

(Bommert et al., 2013). Liu et al. (2017b) found that some kernel-related mutant genes are enriched in QTLs for kernel traits and that SNPs in some of these genes are significantly associated with kernel traits, with the most significant SNPs being either located in the regulatory region or being synonymous substitutions. These findings together highlight the possibility of improving maize via a combination of knowledge concerning mutant genes and CRISPR/Cas9 technology, which, as has already been demonstrated (Eshed and Lippman, 2019; Fernie and Yan, 2019), could be used to generate considerably more favorable alleles.

The genetic basis of different traits varies substantially in maize. For flowering time and oil concentration, dozens of loci, each with a small effect, have been identified (Buckler et al., 2009; Li et al., 2013). Yield-related traits, such as plant architecture and ear and kernel traits, have a much more complex genetic architecture as indicated by the detection of several hundreds of loci (Xiao et al., 2016; Liu et al., 2017c; Pan et al., 2017). In contrast, there are fewer mapped QTLs for disease resistance. Taking gray leaf spot (GLS) as an example, only 16 loci were mapped in the NAM population, among which three had a large effect (Benson et al., 2015). This situation is not the same across crops. Unlike maize, rice heading and seed size and weight are regulated by only a handful of genes, each with a large effect (Huang et al., 2011). These differences suggest that the road to maize

improvement should differ from that of rice, and researchers should select an appropriate method based on the genetic complexity of the desired traits. For traits mainly controlled by major QTLs, i.e., disease resistance, MAS would be a more efficient strategy, whereas genomic selection would be more suitable for traits with a much more complex genetic basis, such as flowering-time and yield-related traits (Figure 3A–3C).

PERSPECTIVES

Based on the tremendous genotypic diversity revealed by resequencing studies in both maize and other plants (especially rice and *Arabidopsis*), we believe that more high-quality genome assemblies, including assemblies for teosinte, landraces, and tropical lines, are needed to comprehensively reveal maize variations. Another challenge is to rapidly clone and validate genes responsible for important agronomic traits. The step forward in methodology has greatly accelerated this process, especially the progress in transformation, CRISPR/Cas9 technology, statistical methods for speeding up QTL cloning (such as QTL sequencing [QTL-seq] and quantitative trait gene sequencing [QTG-seq] [Takagi et al., 2013; Zhang et al., 2019]), and haploid induction technologies. Using a combination of these methods to quickly map and validate candidate genes (with CRISPR/Cas9) could provide a revolutionary way to rapidly clone genes and generate favorable alleles that could be directly used in future breeding programs (Figure 3). In addition, many important QTLs, including *tb1*, *UPA1*, and *UPA2*, were mapped within populations derived from teosinte and cultivated maize, which indicates that teosinte is full of treasures waiting to be found. What studies with teosinte can provide is not only insights into domestication but also directions for future improvement and *de novo* domestication new crops (Fernie and Yan, 2019).

The Need for Additional Whole Genome Sequences from Diverse Maize Lines

Population-scale genome projects (the human 1000 Genomes Project, the *Arabidopsis* 1001 Genomes Project, and the rice 3000 Genomes Project) have shown that a handful of reference genome sequences is not enough to reveal the genetic variation within a species, especially low-frequency and rare variants that are also important components of the genetic basis of human diseases or plant traits. For example, according to the pilot phase results of the human 1000 Genomes Project, 55% of the ~15 million SNPs identified are novel (1000 Genomes Project Consortium, 2010); According to pan-genome analysis results from the rice 3000 Genome Project, each rice accession. has at least 26.5% distributed gene families (Wang et al., 2018b).

The genetic diversity of maize is much higher than that of humans (more than 10-fold higher; Tenaillon et al., 2001). However, most of the maize genome sequences released to date are from temperate lines. Genome sequences from tropical lines, landraces, and teosintes have seldom been reported, although many known variations exist that are specific to teosinte or tropical lines (examples are shown in Figure 1). With the advancement of sequencing technologies, it will become much easier and also cheaper to assemble many more maize genome sequences to fully reveal genetic variations, including widespread genome-structure variations (representative inversions and InDels >1 Mb are illustrated in Figure 2), and their relationship to phenotypes.

Technologies

Although some maize genome sequences and abundant variations are already available, the functions of most genes and also the genes underlying many traits and biological processes remain unclear. In the past decades, genotype limitations and the low efficiency of maize genetic transformation made maize genetic analysis much slower than that of *Arabidopsis* and rice. However, the greatly increased transformation efficiency that results from ectopic expression of *Baby Boom* (*Bbm*) and *Wuschel 2* (*Wus2*) and the use of promoters from *Zm-PLTP* and *Zm-Axig1* (which encode a phospholipid transferase protein and an Aux/IAA transcription factor, respectively) opened the gates for efficient and genotype-independent maize transformation (Lowe et al., 2016, 2018).

Rooted in these improved transformation approaches, CRISPR/Cas9 technology can be used to generate both multiple mutated alleles for a single gene and mutations of multiple genes in a single maize plant by integrating multiple guide RNAs into one vector and using just a single guide RNA targeting a conserved region of homologous genes, respectively (Doll et al., 2019). This would be a very convenient method for studying gene families with functional redundancy, as demonstrated in rice (Ma et al., 2015). Furthermore, the CRISPR/Cas9 system has been successfully used to construct genome-wide mutant libraries in rice (Lu et al., 2017; Meng et al., 2017). Lines in mutant libraries derived from TILLING or *Mu* tagging usually have many randomly mutated loci, and mutant loci caused by *Mu* insertions are sometimes not stable (Brutnell, 2002; Lu et al., 2018), which could exacerbate the difficulty in phenotype investigation and validation of gene function. In contrast, the CRISPR/Cas9 system introduces stable mutations precisely at specific sites as dictated by the single guide RNA with a much cleaner genetic background. There are also increasing reports about the application of CRISPR/Cas9 technology in maize. For example, Svitashev et al. (2016) delivered a Cas9-guide RNA ribonucleoprotein assembly into maize embryos and successfully obtained edited plants and achieved DNA-free genome editing; Feng et al. (2018) reported highly efficient genome editing using the *dmc1* promoter combined with the *U3* promoter to drive the CRISPR/Cas9 system. Thus, the technique shows great promise in both functional research and breeding of maize (Figure 3D–3G).

Traditional QTL mapping in populations derived from multiple generations of selfing, backcrossing, or intercrossing and positional cloning are time-consuming methods. There are two approaches, referred to as QTL-seq (Takagi et al., 2013) and QTG-seq (Zhang et al., 2019), that take advantage of both bulked segregant analysis and NGS to rapidly identify QTLs (Figure 3B). Using QTG-seq, researchers even mapped a plant-height QTL to a ~300-kb region (Zhang et al., 2019). There is also another way to quickly construct populations and, hence, map QTLs, which depends on the application of doubled haploids. A haploid, which contains only one set of chromosomes, could become a diploid inbred line in only one generation after a

chromosome-doubling process. One method widely used in maize to produce haploids is to cross plants with haploid inducer lines. A number of QTLs for haploid induction rate (HIR) have been reported, with *qhir1* and *qhir8* explaining most of the variation (66% and 20%, respectively) in this trait (Prigge et al., 2012; Hu et al., 2016). The cloning of these two loci and the dissection of the mechanisms underlying this process were reported recently (Gilles et al., 2017; Kelliher et al., 2017; Li et al., 2017; Liu et al., 2017a; Zhong et al., 2019). A 4-bp insertion and a nucleotide change in *ZmPLA1* (also named *MTL*, *NLD*) and *ZmDMP*, respectively, led to increased HIR (Gilles et al., 2017; Kelliher et al., 2017; Liu et al., 2017a; Zhong et al., 2019). Li et al. (2017) found severe chromosome fragmentation in pollen of an inducer line using single nucleus sequencing technology, which might be the reason for haploid induction. These findings have great promise for accelerating the breeding process, and there are already reports about the genetic manipulation of *ZmPLA1* in other crops to promote haploid induction, for example in rice and wheat (Yao et al., 2018; Liu et al., 2019).

Diversified Breeding Goals

Since the process of domestication of maize from its wild ancestors began several thousand years ago, yield has always been the central goal in breeding programs. In addition, there are now many new kinds of maize developed for specific purposes, the so-called ‘‘specialty corns,’’ such as silage corn, waxy corn, high-amylose corn, high-oil corn, high-quality-protein corn, sweet corn, and popcorn. The planting area of specialty corn has substantially increased since the early 20th century. For example, waxy corn was discovered in the 1900s in China, and its planting area now has reached ~800 000 hm^2 in China. Meanwhile, it has been reported that nearly 40% of maize produced in the United States is used to make ethanol for fuel (Ranum et al., 2014). These two observations reflect the increasingly diverse demands of maize breeding.

Based on the genomic changes during maize domestication and variations in positionally cloned maize genes that underlie QTLs for diverse traits, many traits are affected mainly by a limited number of genes. Taking the waxy endosperm trait as an example, mutation of only one gene (*Wx*) leads to an amylopectin content of nearly 100% in the endosperm (Sprague et al., 1943). Thus, it is possible that we could improve specific maize traits by manipulating only a handful of genes to meet modern demands for diverse maize types.

Reuse of Wild Relatives

It is estimated that maize retains 83% of the genetic diversity from its progenitor, *Zea mays* ssp. *parviglumis* (Hufford et al., 2012). This is actually a considerably higher proportion than found in analogous comparisons between other major crops and their wild species progenitors (Fernie et al., 2006), largely due to the fact that maize is an outcrossing species. Nevertheless, the teosintes represent important resources for pinpointing the genes/variations that are important for traits but have been lost during maize artificial selection (Li et al., 2019c; Tian et al., 2019). There are several ways to discover such genes/variations. First, traditional QTL mapping and positional cloning have been successfully used to clone genes within populations derived from crosses between maize and teosinte. For example, *ZmCCT10* and *ZmCCT9* were mapped and cloned using this method in a maize-teosinte BC_2S_3 population (Hung et al., 2012; Huang et al., 2018). Association mapping in a teosinte population has also been demonstrated to be a powerful approach by which to detect significant associations between genes and traits (Weber et al., 2007, 2008). However, these two studies performed candidate-gene-association analyses using <150 markers. It is expected that more loci, including some novel loci, would be detected by a GWAS with a much higher density of molecular markers. Recently, a *k*-mer-based association mapping method called association genetics with resistance gene-enrichment sequencing (AgRenSeq) was used to successfully discover and clone disease-resistance genes from crop wild relatives (Arora et al., 2019). Multi-omics data are also very powerful for identifying genes for traits (Liu et al., 2016; de Abreu E Lima et al., 2018; McLoughlin et al., 2018) The teosinte allele of cloned genes/variations could subsequently be directly applied in maize-breeding programs through introgression into modern elite maize lines.

An alternative possibility to rewilding maize would be to *de novo* domesticate teosinte, as was recently proposed by Fernie and Yan (2019), via either long-term backcrossing or short-term genome editing. The latter would require a successful transformation method for teosinte. As noted above, only a few hundred genes underwent selection during maize domestication. If we want to change only a few key traits to meet the aforementioned diverse breeding goals for maize, there may be as few as <100 genes that need to be edited. The resulting costs of this process—in terms of both time and money spent—will be greatly reduced relative to the previous process of human selection, which has already lasted for thousands of years.

FUNDING

This work was supported by the National Natural Science Foundation of China (31961133002, 31901549, and 31525017), and the Ministry of Agriculture of China (2018ZX0801004B).

AUTHOR CONTRIBUTIONS

J.L., A.F., and J.Y. discussed and wrote the manuscript.

ACKNOWLEDGMENTS

No conflict of interest declared.

Received: September 27, 2019
Revised: November 7, 2019
Accepted: November 22, 2019
Published: January 13, 2020

REFERENCES

1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. Nature **467**:1061–1073.

de Abreu E Lima, F., Li, K., Wen, W., Yan, J., Nikoloski, Z., Willmitzer, L., and Brotman, Y. (2018). Unraveling lipid metabolism in maize with time-resolved multi-omics data. Plant J. **93**:1102–1115.

Aguilar-Martínez, J.A., Poza-Carrión, C., and Cubas, P. (2007). *Arabidopsis BRANCHED1* acts as an integrator of branching signals within axillary buds. Plant Cell **19**:458–472.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science **301**:653–657.

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

Arora, S., Steuernagel, B., Gaurav, K., Chandramohan, S., Long, Y., Matny, O., Johnson, R., Enk, J., Periyannan, S., Singh, N., et al. (2019). Resistance gene cloning from a wild crop relative by sequence capture and association genetics. Nat. Biotechnol. **37**:139–143.

Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A.M., Hu, T.T., et al. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature **465**:627–631.

Balint-Kurti, P.J., Zwonitzer, J.C., Wisser, R.J., Carson, M.L., Oropeza-Rosas, M.A., Holland, J.B., and Szalma, S.J. (2007). Precise mapping of quantitative trait loci for resistance to southern leaf blight, caused by *Cochliobolus heterostrophus* race O, and flowering time using advanced intercross maize lines. Genetics **176**:645–657.

Beadle, G.W. (1972). The mystery of maize. Field Museum Nat. Hist. Bull. **43**:2–11.

Beleggia, R., Rau, D., Laidò, G., Platani, C., Nigro, F., Fragasso, M., De Vita, P., Scossa, F., Fernie, A.R., Nikoloski, Z., et al. (2016). Evolutionary metabolomics reveals domestication-associated changes in tetraploid wheat kernels. Mol. Biol. Evol. **33**:1740–1753.

Benke, A., Urbany, C., and Stich, B. (2015). Genome-wide association mapping of iron homeostasis in the maize association population. BMC Genet. **16**:1.

Benson, J.M., Poland, J.A., Benson, B.M., Stromberg, E.L., and Nelson, R.J. (2015). Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. PLoS Genet. **11**:e1005045.

Benz, B.F. (2001). Archaeological evidence of teosinte domestication from Guilá Naquitz, Oaxaca. Proc. Natl. Acad. Sci. U S A **98**:2104–2106.

Bommert, P., Nagasawa, N.S., and Jackson, D. (2013). Quantitative variation in maize kernel row number is controlled by the *FASCIATED EAR2* locus. Nat. Genet. **45**:334–337.

Brandolini, A., and Brandolini, A. (2009). Maize introduction, evolution and diffusion in Italy. Maydica **54**:233–242.

Brutnell, T.P. (2002). Transposon tagging in maize. Funct. Integr. Genom. **2**:4–12.

Buckler, E.S., Holland, J.B., Bradbury, P.J., Acharya, C.B., Brown, P.J., Browne, C., Ersoz, E., Flint-Garcia, S., Garcia, A., Glaubitz, J.C., et al. (2009). The genetic architecture of maize flowering time. Science **325**:714–718.

Bukowski, R., Guo, X., Lu, Y., Zou, C., He, B., Rong, Z., Wang, B., Xu, D., Yang, B., Xie, C., et al. (2018). Construction of the third-generation *Zea mays* haplotype map. Gigascience **7**:1–12.

Castelletti, S., Tuberosa, R., Pindo, M., and Salvi, S. (2014). A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL *Vgt1*. G3 (Bethesda) **4**:805–812.

Chen, J., Zeng, B., Zhang, M., Xie, S., Wang, G., Hauck, A., and Lai, J. (2014). Dynamic transcriptome landscape of maize embryo and endosperm development. Plant Physiol. **166**:252–264.

Chen, G., Wang, X., Hao, J., Yan, J., and Ding, J. (2015). Genome-wide association implicates candidate genes conferring resistance to maize rough dwarf disease in maize. PLoS One **10**:e0142001.

Chia, J.M., Song, C., Bradbury, P.J., Costich, D., de Leon, N., Doebley, J., Elshire, R.J., Gaut, B., Geller, L., Glaubitz, J.C., et al. (2012). Maize HapMap2 identifies extant variation from a genome in flux. Nat. Genet. **44**:803–807.

Chuck, G.S., Brown, P.J., Meeley, R., and Hake, S. (2014). Maize SBP-box transcription factors *unbranched2* and *unbranched3* affect yield traits by regulating the rate of lateral primordia initiation. Proc. Natl. Acad. Sci. U S A **111**:18775–18780.

Cockram, J., White, J., Zuluaga, D.L., Smith, D., Comadran, J., Macaulay, M., Luo, Z., Kearsey, M.J., Werner, P., Harrap, D., et al. (2010). Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. Proc. Natl. Acad. Sci. U S A **107**:21611–21616.

Debernardi, J.M., Lin, H., Chuck, G., Faris, J.D., and Dubcovsky, J. (2017). *microRNA172* plays a crucial role in wheat spike morphogenesis and grain threshability. Development **144**:1966–1975.

Deng, M., Li, D., Luo, J., Xiao, Y., Liu, H., Pan, Q., Zhang, X., Jin, M., Zhao, M., and Yan, J. (2017). The genetic architecture of amino acids dissection by association and linkage analysis in maize. Plant Biotechnol. J. **15**:1250–1263.

Ding, J., Ali, F., Chen, G., Li, H., Mahuku, G., Yang, N., Narro, L., Magorokosho, C., Makumbi, D., and Yan, J. (2015). Genome-wide association mapping reveals novel sources of resistance to northern corn leaf blight in maize. BMC Plant Biol. **15**:206.

Dixon, L.E., Greenwood, J.R., Bencivenga, S., Zhang, P., Cockram, J., Mellers, G., Ramm, K., Cavanagh, C., Swain, S.M., and Boden, S.A. (2018). *TEOSINTE BRANCHED1* regulates inflorescence architecture and development in bread wheat (*Triticum aestivum*). Plant Cell **30**:563–581.

Doebley, J. (2004). The genetics of maize evolution. Annu. Rev. Genet. **38**:37–59.

Doebley, J., and Stec, A. (1993). Inheritance of the morphological differences between maize and teosinte: comparison of results for two F_2 populations. Genetics **134**:559–570.

Doebley, J., Stec, A., and Gustus, C. (1995). *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics **141**:333–346.

Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize. Nature **386**:485–488.

Doebley, J., Gaut, B., and Smith, B. (2006). The molecular genetics of crop domestication. Cell **127**:1309–1321.

Doll, N.M., Gilles, L.M., Gérentes, M.F., Richard, C., Just, J., Fierlej, Y., Borrelli, V.M.G., Gendrot, G., Ingram, G.C., Rogowsky, P.M., et al. (2019). Single and multiple gene knockouts by CRISPR-Cas9 in maize. Plant Cell Rep. **38**:487–501.

Dong, Z., Li, W., Unger-Wallace, E., Yang, J., Vollbrecht, E., and Chuck, G. (2017). Ideal crop plant architecture is mediated by *tassels replace upper ears1*, a BTB/POZ ankyrin repeat gene directly targeted by TEOSINTE BRANCHED1. Proc. Natl. Acad. Sci. U S A **114**:E8656–E8664.

Dorweiler, J., Stec, A., Kermicle, J., and Doebley, J. (1993). *Teosinte glume architecture* 1: a genetic locus controlling a key step in maize evolution. Science **262**:233–235.

Du, Y., Liu, L., Li, M., Fang, S., Shen, X., Chu, J., and Zhang, Z. (2017). *UNBRANCHED3* regulates branching by modulating cytokinin biosynthesis and signaling in maize and rice. New Phytol. **214**:721–733.

Ducrocq, S., Madur, D., Veyrieras, J.B., Camus-Kulandaivelu, L., Kloiber-Maitz, M., Presterl, T., Ouzunova, M., Manicacci, D., and Charcosset, A. (2008). Key impact of *Vgt1* on flowering time adaptation in maize: evidence from association mapping and ecogeographical information. Genetics **178**:2433–2437.

Ducrocq, S., Giauffret, C., Madur, D., Combes, V., Dumas, F., Jouanne, S., Coubriche, D., Jamin, P., Moreau, L., and Charcosset, A. (2009). Fine mapping and haplotype structure analysis of a major flowering time quantitative trait locus on maize chromosome 10. Genetics **183**:1555–1563.

Eshed, Y., and Lippman, Z.B. (2019). Revolutions in agriculture chart a course for targeted breeding of old and new crops. Science https://doi.org/10.1126/science.aax0025.

Farfan, I.D., De La Fuente, G.N., Murray, S.C., Isakeit, T., Huang, P.C., Warburton, M., Williams, P., Windham, G.L., and Kolomiets, M. (2015). Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics. PLoS One **10**:e0117737.

Feng, C., Su, H., Bai, H., Wang, R., Liu, Y., Guo, X., Liu, C., Zhang, J., Yuan, J., Birchler, J.A., et al. (2018). High-efficiency genome editing using a *dmc1* promoter-controlled CRISPR/Cas9 system in maize. Plant Biotechnol. J. **16**:1848–1857.

Fernie, A.R., and Yan, J. (2019). *De Novo* domestication: an alternative route toward new crops for the future. Mol. Plant **12**:615–631.

Fernie, A.R., Tadmor, Y., and Zamir, D. (2006). Natural genetic variation for improving crop quality. Curr. Opin. Plant Biol. **9**:196–202.

Flint-Garcia, S.A., Thuillet, A.C., Yu, J., Pressoir, G., Romero, S.M., Mitchell, S.E., Doebley, J., Kresovich, S., Goodman, M.M., and Buckler, E.S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J. **44**:1054–1064.

Fu, J., Cheng, Y., Linghu, J., Yang, X., Kang, L., Zhang, Z., Zhang, J., He, C., Du, X., Peng, Z., et al. (2013). RNA sequencing reveals the complex regulatory network in the maize kernel. Nat. Commun. **4**:2832.

Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D.M., Burzynski-Chang, E.A., Fish, T.L., Stromberg, K.A., Sacks, G.L., et al. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. Nat. Genet. **51**:1044–1051.

Garrison, E., Sirén, J., Novak, A.M., Hickey, G., Eizenga, J.M., Dawson, E.T., Jones, W., Garg, S., Markello, C., Lin, M.F., et al. (2018). Variation graph toolkit improves read mapping by representing genetic variation in the reference. Nat. Biotechnol. **36**:875–879.

Gilles, L.M., Khaled, A., Laffaire, J.B., Chaignon, S., Gendrot, G., Laplaige, J., Bergès, H., Beydon, G., Bayle, V., Barret, P., et al. (2017). Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. EMBO J. **36**:707–717.

Gore, M.A., Chia, J.M., Elshire, R.J., Sun, Q., Ersoz, E.S., Hurwitz, B.L., Peiffer, J.A., McMullen, M.D., Grills, G.S., Ross-Ibarra, J., et al. (2009). A first-generation haplotype map of maize. Science **326**:1115–1117.

Harjes, C.E., Rocheford, T.R., Bai, L., Brutnell, T.P., Kandianis, C.B., Sowinski, S.G., Stapleton, A.E., Vallabhaneni, R., Williams, M., Wurtzel, E.T., et al. (2008). Natural genetic variation in *lycopene epsilon cyclase* tapped for maize biofortification. Science **319**:330–333.

Hirsch, C.N., Foerster, J.M., Johnson, J.M., Sekhon, R.S., Muttoni, G., Vaillancourt, B., Peñagaricano, F., Lindquist, E., Pedraza, M.A., Barry, K., et al. (2014). Insights into the maize pan-genome and pan-transcriptome. Plant Cell **26**:121–135.

Hirsch, C.N., Hirsch, C.D., Brohammer, A.B., Bowman, M.J., Soifer, I., Barad, O., Shem-Tov, D., Baruch, K., Lu, F., Hernandez, A.G., et al. (2016). Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. Plant Cell **28**:2700–2714.

Hu, H., Schrag, T.A., Peis, R., Unterseer, S., Schipprack, W., Chen, S., Lai, J., Yan, J., Prasanna, B.M., Nair, S.K., et al. (2016). The genetic basis of haploid induction in maize identified with a novel genome-wide association method. Genetics **202**:1267–1276.

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., Li, W., Guo, Y., Deng, L., Zhu, C., et al. (2011). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat. Genet. **44**:32–39.

Huang, C., Sun, H., Xu, D., Chen, Q., Liang, Y., Wang, X., Xu, G., Tian, J., Wang, C., Li, D., et al. (2018). *ZmCCT9* enhances maize adaptation to higher latitudes. Proc. Natl. Acad. Sci. U S A **115**:E334–E341.

Hufford, M.B., Xu, X., van Heerwaarden, J., Pyhäjärvi, T., Chia, J.M., Cartwright, R.A., Elshire, R.J., Glaubitz, J.C., Guill, K.E., Kaeppler, S.M., et al. (2012). Comparative population genomics of maize domestication and improvement. Nat. Genet. **44**:808–811.

Hung, H.Y., Shannon, L.M., Tian, F., Bradbury, P.J., Chen, C., Flint-Garcia, S.A., McMullen, M.D., Ware, D., Buckler, E.S., Doebley, J.F., et al. (2012). *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc. Natl. Acad. Sci. U S A **109**:E1913–E1921.

Jiao, Y., Zhao, H., Ren, L., Song, W., Zeng, B., Guo, J., Wang, B., Liu, Z., Chen, J., Li, W., et al. (2012). Genome-wide genetic changes during modern breeding of maize. Nat. Genet. **44**:812–815.

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M.C., Wang, B., Campbell, M.S., Stein, J.C., Wei, X., Chin, C.S., et al. (2017). Improved maize reference genome with single-molecule technologies. Nature **546**:524–527.

Jin, M., Liu, H., He, C., Fu, J., Xiao, Y., Wang, Y., Xie, W., Wang, G., and Yan, J. (2016). Maize pan-transcriptome provides novel insights into genome complexity and quantitative trait variation. Sci. Rep. **6**:18936.

Jin, M., Liu, X., Jia, W., Liu, H., Li, W., Peng, Y., Du, Y., Wang, Y., Yin, Y., Zhang, X., et al. (2018). ZmCOL3, a CCT gene represses flowering in maize by interfering with the circadian clock and activating expression of ZmCCT. J. Integr. Plant Biol. **60**:465–480.

Kelliher, T., Starr, D., Richbourg, L., Chintamanani, S., Delzer, B., Nuccio, M.L., Green, J., Chen, Z., McCuiston, J., Wang, W., et al. (2017). MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature **542**:105–109.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. **37**:907–915.

Kump, K.L., Bradbury, P.J., Wisser, R.J., Buckler, E.S., Belcher, A.R., Oropeza-Rosas, M.A., Zwonitzer, J.C., Kresovich, S., McMullen, M.D., Ware, D., 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.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol. **5**:R12.

Lai, J., Li, R., Xu, X., Jin, W., Xu, M., Zhao, H., Xiang, Z., Song, W., Ying, K., Zhang, M., et al. (2010). Genome-wide patterns of genetic variation among elite maize inbred lines. Nat. Genet. **42**:1027–1030.

Lee, M., Sharopova, N., Beavis, W.D., Grant, D., Katt, M., Blair, D., and Hallauer, A. (2002). Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. Plant Mol. Biol. **48**:453–461.

Lemmon, Z.H., Bukowski, R., Sun, Q., and Doebley, J.F. (2014). The role of cis regulatory evolution in maize domestication. PLoS Genet. **10**:e1004745.

Li, P., Wang, Y., Qian, Q., Fu, Z., Wang, M., Zeng, D., Li, B., Wang, X., and Li, J. (2007). *LAZY1* controls rice shoot gravitropism through regulating polar auxin transport. Cell Res. **17**:402–410.

Li, L., Li, H., Li, Q., Yang, X., Zheng, D., Warburton, M., Chai, Y., Zhang, P., Guo, Y., Yan, J., 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.

Li, Q., Yang, X., Xu, S., Cai, Y., Zhang, D., Han, Y., Li, L., Zhang, Z., Gao, S., Li, J., et al. (2012). Genome-wide association studies identified three independent polymorphisms associated with α-tocopherol content in maize kernels. PLoS One **7**:e36807.

Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., Han, Y., Chai, Y., Guo, T., Yang, N., et al. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat. Genet. **45**:43–50.

Li, G., Wang, D., Yang, R., Logan, K., Chen, H., Zhang, S., Skaggs, M.I., Lloyd, A., Burnett, W.J., Laurie, J.D., et al. (2014). Temporal patterns of gene expression in developing maize endosperm identified through transcriptome sequencing. Proc. Natl. Acad. Sci. U S A **111**:7582–7587.

Li, X., Meng, D., Chen, S., Luo, H., Zhang, Q., Jin, W., and Yan, J. (2017). Single nucleus sequencing reveals spermatid chromosome fragmentation as a possible cause of maize haploid induction. Nat. Commun. **8**:991.

Li, C., Song, W., Luo, Y., Gao, S., Zhang, R., Shi, Z., Wang, X., Wang, R., Wang, F., Wang, J., et al. (2019a). The HuangZaoSi maize genome provides insights into genomic variation and improvement history of maize. Mol. Plant **12**:402–409.

Li, E., Liu, H., Huang, L., Zhang, X., Dong, X., Song, W., Zhao, H., and Lai, J. (2019b). Long-range interactions between proximal and distal regulatory regions in maize. Nat. Commun. **10**:2633.

Li, K., Wen, W., Alseekh, S., Yang, X., Guo, H., Li, W., Wang, L., Pan, Q., Zhan, W., Liu, J., et al. (2019c). Large-scale metabolite quantitative trait locus analysis provides new insights for high-quality maize improvement. Plant J. **99**:216–230.

Li, N., Lin, B., Wang, H., Li, X., Yang, F., Ding, X., Yan, J., and Chu, Z. (2019d). Natural variation in *ZmFBL41* confers banded leaf and sheath blight resistance in maize. Nat. Genet. **51**:1540–1548.

Lin, Z., Li, X., Shannon, L.M., Yeh, C.T., Wang, M.L., Bai, G., Peng, Z., Li, J., Trick, H.N., Clemente, T.E., et al. (2012). Parallel domestication of the *Shattering1* genes in cereals. Nat. Genet. **44**:720–724.

Liu, H., and Yan, J. (2019). Crop genome-wide association study: a harvest of biological relevance. Plant J. **97**:8–18.

Liu, S., Wang, X., Wang, H., Xin, H., Yang, X., Yan, J., Li, J., Tran, L.S., Shinozaki, K., Yamaguchi-Shinozaki, K., et al. (2013). Genome-wide analysis of *ZmDREB* genes and their association with natural variation in drought tolerance at seedling stage of *Zea mays* L. PLoS Genet. **9**:e1003790.

Liu, L., Du, Y., Shen, X., Li, M., Sun, W., Huang, J., Liu, Z., Tao, Y., Zheng, Y., Yan, J., et al. (2015a). *KRN4* controls quantitative variation in maize kernel row number. PLoS Genet. **11**:e1005670.

Liu, L., Jeffers, D., Zhang, Y., Ding, M., Chen, W., Kang, M.S., and Fan, X. (2015b). Introgression of the *crtRB1* gene into quality protein maize inbred lines using molecular markers. Mol. Breed. **35**:154.

Liu, H., Wang, X., Warburton, M.L., Wen, W., Jin, M., Deng, M., Liu, J., Tong, H., Pan, Q., Yang, X., et al. (2015c). Genomic, transcriptomic, and phenomic variation reveals the complex adaptation of modern maize breeding. Mol. Plant **8**:871–884.

Liu, H., Wang, F., Xiao, Y., Tian, Z., Wen, W., Zhang, X., Chen, X., Liu, N., Li, W., Liu, L., et al. (2016). MODEM: multi-omics data envelopment and mining in maize. Database (Oxford) https://doi.org/10.1093/database/baw117.

Liu, C., Li, X., Meng, D., Zhong, Y., Chen, C., Dong, X., Xu, X., Chen, B., Li, W., Li, L., et al. (2017a). A 4-bp insertion at *ZmPLA1* encoding a putative phospholipase A generates haploid induction in maize. Mol. Plant **10**:520–522.

Liu, H., Luo, X., Niu, L., Xiao, Y., Chen, L., Liu, J., Wang, X., Jin, M., Li, W., Zhang, Q., et al. (2017b). Distant eQTL and non-coding sequences play critical roles in regulating gene expression and quantitative trait variation in maize. Mol. Plant **10**:414–426.

Liu, J., Huang, J., Guo, H., Lan, L., Wang, H., Xu, Y., Yang, X., Li, W., Tong, H., Xiao, Y., et al. (2017c). The conserved and unique genetic architecture of kernel size and weight in maize and rice. Plant Physiol. **175**:774–785.

Liu, C., Zhong, Y., Qi, X., Chen, M., Liu, Z., Chen, C., Tian, X., Li, J., Jiao, Y., Wang, D., et al. (2019). Extension of the *in vivo* haploid induction system from diploid maize to hexaploid wheat. Plant Biotechnol. J. https://doi.org/10.1111/pbi.13218.

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.J., Scelonge, C., Lenderts, B., Chamberlin, M., Cushatt, J., et al. (2016). Morphogenic regulators *Baby boom* and *Wuschel* improve monocot transformation. Plant Cell **28**:1998–2015.

Lowe, K., La Rota, M., Hoerster, G., Hastings, C., Wang, N., Chamberlin, M., Wu, E., Jones, T., and Gordon-Kamm, W. (2018). Rapid genotype "independent" *Zea mays* L. (maize) transformation via direct somatic embryogenesis. In Vitro Cell Dev. Biol. Plant **54**:240–252.

Lu, F., Romay, M.C., Glaubitz, J.C., Bradbury, P.J., Elshire, R.J., Wang, T., Li, Y., Li, Y., Semagn, K., Zhang, X., et al. (2015). High-resolution genetic mapping of maize pan-genome sequence anchors. Nat. Commun. **6**:6914.

Lu, Y., Ye, X., Guo, R., Huang, J., Wang, W., Tang, J., Tan, L., Zhu, J.K., Chu, C., and Qian, Y. (2017). Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system. Mol. Plant **10**:1242–1245.

Lu, X., Liu, J., Ren, W., Yang, Q., Chai, Z., Chen, R., Wang, L., Zhao, J., Lang, Z., Wang, H., et al. (2018). Gene-indexed mutations in maize. Mol. Plant **11**:496–504.

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., Yang, Z., Li, H., Lin, Y., et al. (2015). A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Mol. Plant **8**:1274–1284.

Mao, H., Wang, H., Liu, S., Li, Z., Yang, X., Yan, J., Li, J., Tran, L.S., and Qin, F. (2015). A transposable element in a *NAC* gene is associated with drought tolerance in maize seedlings. Nat. Commun. **6**:8326.

Matsuoka, Y., Vigouroux, Y., Goodman, M.M., Sanchez, G.J., Buckler, E., and Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. Proc. Natl. Acad. Sci. U S A **99**:6080–6084.

McCarty, D.R., Settles, A.M., Suzuki, M., Tan, B.C., Latshaw, S., Porch, T., Robin, K., Baier, J., Avigne, W., Lai, J., et al. (2005). Steady-state transposon mutagenesis in inbred maize. Plant J. **44**:52–61.

McLoughlin, F., Augustine, R.C., Marshall, R.S., Li, F., Kirkpatrick, L.D., Otegui, M.S., and Vierstra, R.D. (2018). Maize multi-omics reveal roles for autophagic recycling in proteome remodelling and lipid turnover. Nat. Plants **4**:1056–1070.

McMullen, M.D., Kresovich, S., Villeda, H.S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., et al. (2009). Genetic properties of the maize nested association mapping population. Science **325**:737–740.

Meng, X., Yu, H., Zhang, Y., Zhuang, F., Song, X., Gao, S., Gao, C., and Li, J. (2017). Construction of a genome-wide mutant library in rice using CRISPR/Cas9. Mol. Plant **10**:1238–1241.

Mertz, E.T., Bates, L.S., and Nelson, O.E. (1964). Mutant gene that changes protein composition and increases lysine content of maize endosperm. Science **145**:279–280.

Moreno, M.A., Harper, L.C., Krueger, R.W., Dellaporta, S.L., and Freeling, M. (1997). liguleless1 encodes a nuclear-localized protein required for induction of ligules and auricles during maize leaf organogenesis. Genes Dev. **11**:616–628.

Multani, D.S., Briggs, S.P., Chamberlin, M.A., Blakeslee, J.J., Murphy, A.S., and Johal, G.S. (2003). Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. Science **302**:81–84.

Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Choudhary, M., Saha, S., Bhat, J.S., Prasanna, B.M., and Gupta, H.S. (2014). Development of β-carotene rich maize hybrids through marker-assisted introgression of *β-carotene hydroxylase* allele. PLoS One **9**:e113583.

Nestel, P., Bouis, H.E., Meenakshi, J.V., and Pfeiffer, W. (2006). Biofortification of staple food crops. J. Nutr. **136**:1064–1067.

Nunn, N., and Nancy, Q. (2010). The Columbian exchange: a history of disease, food, and ideas. J. Econ. Perspect. **24**:163–188.

Olukolu, B.A., Negeri, A., Dhawan, R., Venkata, B.P., Sharma, P., Garg, A., Gachomo, E., Marla, S., Chu, K., Hasan, A., et al. (2013). A connected set of genes associated with programmed cell death implicated in controlling the hypersensitive response in maize. Genetics **193**:609–620.

Pan, Q., Xu, Y., Li, K., Peng, Y., Zhan, W., Li, W., Li, L., and Yan, J. (2017). The genetic basis of plant architecture in 10 maize recombinant inbred line populations. Plant Physiol. **175**:858–873.

Peng, Y., Xiong, D., Zhao, L., Ouyang, W., Wang, S., Sun, J., Zhang, Q., Guan, P., Xie, L., Li, W., et al. (2019). Chromatin interaction maps reveal genetic regulation for quantitative traits in maize. Nat. Commun. **10**:2632.

Piperno, D.R., and Flannery, K.V. (2001). The earliest archaeological maize (*Zea mays L.*) from highland Mexico: new accelerator mass spectrometry dates and their implications. Proc. Natl. Acad. Sci. U S A **98**:2101–2103.

Portwood, J.L., Woodhouse, M.R., Cannon, E.K., Gardiner, J.M., Harper, L.C., Schaeffer, M.L., Walsh, J.R., Sen, T.Z., Cho, K.T., Schott, D.A., et al. (2019). MaizeGDB 2018: the maize multi-genome genetics and genomics database. Nucleic Acids Res. **47**:D1146–D1154.

Prigge, V., Xu, X., Li, L., Babu, R., Chen, S., Atlin, G.N., and Melchinger, A.E. (2012). New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. Genetics **190**:781–793.

Raihan, M.S., Liu, J., Huang, J., Guo, H., Pan, Q., and Yan, J. (2016). Multi-environment QTL analysis of grain morphology traits and fine mapping of a kernel-width QTL in Zheng58 × SK maize population. Theor. Appl. Genet. **129**:1465–1477.

Rakocevic, G., Semenyuk, V., Lee, W.P., Spencer, J., Browning, J., Johnson, I.J., Arsenijevic, V., Nadj, J., Ghose, K., Suciu, M.C., et al. (2019). Fast and accurate genomic analyses using genome graphs. Nat. Genet. **51**:354–362.

Ramsay, L., Comadran, J., Druka, A., Marshall, D.F., Thomas, W.T., Macaulay, M., MacKenzie, K., Simpson, C., Fuller, J., Bonar, N., et al. (2011). INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. Nat. Genet. **43**:169–172.

Ranum, P., Peña-Rosas, J.P., and Garcia-Casal, M.N. (2014). Global maize production, utilization, and consumption. Ann. N Y Acad. Sci. **1312**:105–112.

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.

Romay, M.C., Millard, M.J., Glaubitz, J.C., Peiffer, J.A., Swarts, K.L., Casstevens, T.M., Elshire, R.J., Acharya, C.B., Mitchell, S.E., Flint-Garcia, S.A., et al. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol. **14**:R55.

Salvi, S., Tuberosa, R., Chiapparino, E., Maccaferri, M., Veillet, S., van Beuningen, L., Isaac, P., Edwards, K., and Phillips, R.L. (2002). Toward positional cloning of *Vgt1*, a QTL controlling the transition from the vegetative to the reproductive phase in maize. Plant Mol. Biol. **48**:601–613.

Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., Meeley, R., Ananiev, E.V., Svitashev, S., Bruggemann, E., et al. (2007). Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc. Natl. Acad. Sci. U S A **104**:11376–11381.

Sauvage, C., Segura, V., Bauchet, G., Stevens, R., Do, P.T., Nikoloski, Z., Fernie, A.R., and Causse, M. (2014). Genome-wide association in tomato reveals 44 candidate loci for fruit metabolic traits. Plant Physiol. **165**:1120–1132.

Schnable, J.C., and Freeling, M. (2011). Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize. PLoS One **6**:e17855.

Schnable, S.P., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. Science **326**:1112–1115.

Shang, Y., Ma, Y., Zhou, Y., Zhang, H., Duan, L., Chen, H., Zeng, J., Zhou, Q., Wang, S., Gu, W., et al. (2014). Biosynthesis, regulation, and domestication of bitterness in cucumber. Science **346**:1084–1088.

Sigmon, B., and Vollbrecht, E. (2010). Evidence of selection at the *ramosa1* locus during maize domestication. Mol. Ecol. **19**:1296–1311.

Simons, K.J., Fellers, J.P., Trick, H.N., Zhang, Z., Tai, Y.S., Gill, B.S., and Faris, J.D. (2006). Molecular characterization of the major wheat domestication gene Q. Genetics **172**:547–555.

Sosso, D., Luo, D., Li, Q.B., Sasse, J., Yang, J., Gendrot, G., Suzuki, M., Koch, K.E., McCarty, D.R., Chourey, P.S., et al. (2015). Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport. Nat. Genet. **47**:1489–1493.

Sprague, G.F., Brimhall, B., and Hixon, R.M. (1943). Some effects of the *waxy* gene in corn on properties of the endosperm starch. J. Am. Soc. Agron. **35**:817–822.

Springer, N.M., Ying, K., Fu, Y., Ji, T., Yeh, C.T., Jia, Y., Wu, W., Richmond, T., Kitzman, J., Rosenbaum, H., et al. (2009). Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PLoS Genet. **5**:e1000734.

Springer, N.M., Anderson, S.N., Andorf, C.M., Ahern, K.R., Bai, F., Barad, O., Barbazuk, W.B., Bass, H.W., Baruch, K., Ben-Zvi, G., et al. (2018). The maize W22 genome provides a foundation for functional genomics and transposon biology. Nat. Genet. **50**:1282–1288.

Stagnati, L., Lanubile, A., Samayoa, L.F., Bragalanti, M., Giorni, P., Busconi, M., Holland, J.B., and Marocco, A. (2019). A genome wide association study reveals markers and genes associated with resistance to *fusarium verticillioides* infection of seedlings in a maize diversity panel. G3 (Bethesda) **9**:571–579.

Stelpflug, S.C., Sekhon, R.S., Vaillancourt, B., Hirsch, C.N., Buell, C.R., de Leon, N., and Kaeppler, S.M. (2016). An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. Plant Genome **9**. https://doi.org/10.3835/plantgenome2015.04.0025.

Strable, J., Wallace, J.G., Unger-Wallace, E., Briggs, S., Bradbury, P.J., Buckler, E.S., and Vollbrecht, E. (2017). Maize YABBY genes *drooping leaf1* and *drooping leaf2* regulate plant architecture. Plant Cell **29**:1622–1641.

Studer, A., Zhao, Q., Ross-Ibarra, J., and Doebley, J. (2011). Identification of a functional transposon insertion in the maize domestication gene tb1. Nat. Genet. **43**:1160–1163.

Studer, A.J., Wang, H., and Doebley, J.F. (2017). Selection during maize domestication targeted a gene network controlling plant and inflorescence architecture. Genetics **207**:755–765.

Sun, S., Zhou, Y., Chen, J., Shi, J., Zhao, H., Zhao, H., Song, W., Zhang, M., Cui, Y., Dong, X., et al. (2018). Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes. Nat. Genet. **50**:1289–1295.

Svitashev, S., Schwartz, C., Lenderts, B., Young, J.K., and Mark Cigan, A. (2016). Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. Nat. Commun. **7**:13274.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., et al. (2013). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J. **74**:174–183.

Takeda, T., Suwa, Y., Suzuki, M., Kitano, H., Ueguchi-Tanaka, M., Ashikari, M., Matsuoka, M., and Ueguchi, C. (2003). The *OsTB1* gene negatively regulates lateral branching in rice. Plant J. **33**:513–520.

Tenaillon, M.I., Sawkins, M.C., Long, A.D., Gaut, R.L., Doebley, J.F., and Gaut, B.S. (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.

Teng, F., Zhai, L., Liu, R., Bai, W., Wang, L., Huo, D., Tao, Y., Zheng, Y., and Zhang, Z. (2013). *ZmGA3ox2*, a candidate gene for a major QTL, *qPH3.1*, for plant height in maize. Plant J. **73**:405–416.

Thornsberry, J.M., Goodman, M.M., Doebley, J., Kresovich, S., Nielsen, D., and Buckler, E.S. (2001). *Dwarf8* polymorphisms associate with variation in flowering time. Nat. Genet. **28**:286–289.

Tian, J., Wang, C., Xia, J., Wu, L., Xu, G., Wu, W., Li, D., Qin, W., Han, X., Chen, Q., et al. (2019). Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. Science **365**:658–664.

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

Till, B.J., Reynolds, S.H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Codomo, C.A., Enns, L.C., Odden, A.R., et al. (2004). Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biol. **4**:12.

Vallebueno-Estrada, M., Rodríguez-Arévalo, I., Rougon-Cardoso, A., Martínez González, J., García Cook, A., Montiel, R., and Vielle-Calzada, J.P. (2016). The earliest maize from San Marcos Tehuacán is a partial domesticate with genomic evidence of inbreeding. Proc. Natl. Acad. Sci. U S A **113**:14151–14156.

Vlăduţu, C., McLaughlin, J., and Phillips, R.L. (1999). Fine mapping and characterization of linked quantitative trait loci involved in the transition of the maize apical meristem from vegetative to generative structures. Genetics **153**:993–1007.

Walley, J.W., Sartor, R.C., Shen, Z., Schmitz, R.J., Wu, K.J., Urich, M.A., Nery, J.R., Smith, L.G., Schnable, J.C., Ecker, J.R., et al. (2016). Integration of omic networks in a developmental atlas of maize. Science **353**:814–818.

Wang, H., Nussbaum-Wagler, T., Li, B., Zhao, Q., Vigouroux, Y., Faller, M., Bomblies, K., Lukens, L., and Doebley, J.F. (2005). The origin of the naked grains of maize. Nature **436**:714–719.

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

Wang, C., Yang, Q., Wang, W., Li, Y., Guo, Y., Zhang, D., Ma, X., Song, W., Zhao, J., and Xu, M. (2017). A transposon-directed epigenetic change in *ZmCCT* underlies quantitative resistance to Gibberella stalk rot in maize. New Phytol. **215**:1503–1515.

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

Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., Li, M., Zheng, T., Fuentes, R.R., Zhang, F., et al. (2018b). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature **557**:43–49.

Wang, J., Lin, Z., Zhang, X., Liu, H., Zhou, L., Zhong, S., Li, Y., Zhu, C., and Lin, Z. (2019). *krn1*, a major quantitative trait locus for kernel row number in maize. New Phytol. **223**:1634–1646.

Weber, A., Clark, R.M., Vaughn, L., Sánchez-Gonzalez Jde, J., Yu, J., Yandell, B.S., Bradbury, P., and Doebley, J. (2007). Major regulatory genes in maize contribute to standing variation in teosinte (*Zea mays* ssp. *parviglumis*). Genetics **177**:2349–2359.

Weber, A.L., Briggs, W.H., Rucker, J., Baltazar, B.M., de Jesús Sánchez-Gonzalez, J., Feng, P., Buckler, E.S., and Doebley, J. (2008). The genetic architecture of complex traits in teosinte (*Zea mays* ssp. *parviglumis*): new evidence from association mapping. Genetics **180**:1221–1232.

Wei, L., and Cao, X. (2016). The effect of transposable elements on phenotypic variation: insights from plants to humans. Sci. China Life Sci. **59**:24–37.

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

Wen, W., Li, K., Alseekh, S., Omranian, N., Zhao, L., Zhou, Y., Xiao, Y., Jin, M., Yang, N., Liu, H., 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.

Wen, W., Brotman, Y., Willmitzer, L., Yan, J., and Fernie, A.R. (2016). Broadening our portfolio in the genetic improvement of maize chemical composition. Trends Genet. **32**:459–469.

Wen, W., Jin, M., Li, K., Liu, H., Xiao, Y., Zhao, M., Alseekh, S., Li, W., de Abreu E Lima, F., Brotman, Y., et al. (2018). An integrated multi-layered analysis of the metabolic networks of different tissues uncovers key genetic components of primary metabolism in maize. Plant J. **93**:1116–1128.

Whipple, C.J., Kebrom, T.H., Weber, A.L., Yang, F., Hall, D., Meeley, R., Schmidt, R., Doebley, J., Brutnell, T.P., and Jackson, D.P. (2011). *Grassy tillers1* promotes apical dominance in maize and responds to shade signals in the grasses. Proc. Natl. Acad. Sci. U S A **108**:E506–E512.

Wills, D.M., Whipple, C.J., Takuno, S., Kursel, L.E., Shannon, L.M., Ross-Ibarra, J., and Doebley, J.F. (2013). From many, one: genetic control of prolificacy during maize domestication. PLoS Genet. **9**:e1003604.

Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D., and Gaut, B.S. (2005). The effects of artificial selection on the maize genome. Science **308**:1310–1314.

Wu, C., Li, X., Yuan, W., Chen, G., Kilian, A., Li, J., Xu, C., Li, X., Zhou, D.X., Wang, S., et al. (2003). Development of enhancer trap lines for functional analysis of the rice genome. Plant J. **35**:418–427.

Xiang, Y., Sun, X., Gao, S., Qin, F., and Dai, M. (2017). Deletion of an endoplasmic reticulum stress response element in a *ZmPP2C-A* gene facilitates drought tolerance of maize seedlings. Mol. Plant **10**:456–469.

Xiao, Y., Tong, H., Yang, X., Xu, S., Pan, Q., Qiao, F., Raihan, M.S., Luo, Y., Liu, H., Zhang, X., et al. (2016). Genome-wide dissection of the maize ear genetic architecture using multiple populations. New Phytol. **210**:1095–1106.

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

Xie, Q., Li, N., Yang, Y., Lv, Y., Yao, H., Wei, R., Sparkes, D.L., and Ma, Z. (2018). Pleiotropic effects of the wheat domestication gene *Q* on yield and grain morphology. Planta **247**:1089–1098.

Xing, A., Gao, Y., Ye, L., Zhang, W., Cai, L., Ching, A., Llaca, V., Johnson, B., Liu, L., Yang, X., et al. (2015). A rare SNP mutation in *Brachytic2* moderately reduces plant height and increases yield potential in maize. J. Exp. Bot. **66**:3791–3802.

Xu, C., Ren, Y., Jian, Y., Guo, Z., Zhang, Y., Xie, C., Fu, J., Wang, H., Wang, G., Xu, Y., et al. (2017). Development of a maize 55 K SNP array with improved genome coverage for molecular breeding. Mol. Breed. **37**:20.

Xu, G., Cao, J., Wang, X., Chen, Q., Jin, W., Li, Z., and Tian, F. (2019). Evolutionary metabolomics identifies substantial metabolic divergence between maize and its wild ancestor, teosinte. Plant Cell **31**:1990–2009.

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., Zhou, H., Yu, S., Xu, C., Li, X., et al. (2008). Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. Nat. Genet. **40**:761–767.

Yan, J., Kandianis, C.B., Harjes, C.E., Bai, L., Kim, E.H., Yang, X., Skinner, D.J., Fu, Z., Mitchell, S., Li, Q., et al. (2010a). Rare genetic variation at *Zea mays crtRB1* increases beta-carotene in maize grain. Nat. Genet. **42**:322–327.

Yan, J., Yang, X., Shah, T., Sánchez-Villeda, H., Li, J., Warburton, M., Zhou, Y., Crouch, J., and Xu, Y. (2010b). High-throughput SNP genotyping with the GoldenGate assay in maize. Mol. Breed. **25**:441–451.

Yan, J., Warburton, M., and Crouch, J. (2011). Association mapping for enhancing maize (*Zea mays* L.) genetic improvement. Crop Sci. **51**:433–449.

Yang, Q., Yin, G., Guo, Y., Zhang, D., Chen, S., and Xu, M. (2010a). A major QTL for resistance to *Gibberella* stalk rot in maize. Theor. Appl. Genet. **121**:673–687.

Yang, X., Guo, Y., Yan, J., Zhang, J., Song, T., Rocheford, T., and Li, J.S. (2010b). Major and minor QTL and epistasis contribute to fatty acid compositions and oil concentration in high-oil maize. Theor. Appl. Genet. **120**:665–678.

Yang, X., Gao, S., Xu, S., Zhang, Z., Prasanna, B., Li, L., Li, J., and Yan, J. (2011). Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. Mol. Breed. **28**:511–526.

Yang, Q., Li, Z., Li, W., Ku, L., Wang, C., Ye, J., Li, K., Yang, N., Li, Y., Zhong, T., et al. (2013). *CACTA-like* transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. Proc. Natl. Acad. Sci. U S A **110**:16969–16974.

Yang, N., Lu, Y., Yang, X., Huang, J., Zhou, Y., Ali, F., Wen, W., Liu, J., Li, J., and Yan, J. (2014). Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17 agronomic traits in an enlarged maize association panel. PLoS Genet. **10**:e1004573.

Yang, N., Xu, X.W., Wang, R.R., Peng, W.L., Cai, L., Song, J.M., Li, W., Luo, X., Niu, L., Wang, Y., et al. (2017a). Contributions of *Zea mays* subspecies *mexicana* haplotypes to modern maize. Nat. Commun. **8**:1874.

Yang, Q., Balint-Kurti, P., and Xu, M. (2017b). Quantitative disease resistance: dissection and adoption in maize. Mol. Plant **10**:402–413.

Yang, N., Liu, J., Gao, Q., Gui, S., Chen, L., Yang, L., Huang, J., Deng, T., Luo, J., He, L., et al. (2019a). Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement. Nat. Genet. **51**:1052–1059.

Yang, N., Wu, S., and Yan, J. (2019b). Structural variation in complex genome: detection, integration and function. Sci. China Life Sci. **62**:1098–1100.

Yao, L., Zhang, Y., Liu, C., Liu, Y., Wang, Y., Liang, D., Liu, J., Sahoo, G., and Kelliher, T. (2018). *OsMATL* mutation induces haploid seed formation in indica rice. Nat. Plants **4**:530–533.

Ye, J., Zhong, T., Zhang, D., Ma, C., Wang, L., Yao, L., Zhang, Q., Zhu, M., and Xu, M. (2019). The auxin-regulated protein ZmAuxRP1 coordinates the balance between root growth and stalk rot disease resistance in maize. Mol. Plant **12**:360–373.

Yi, F., Gu, W., Chen, J., Song, N., Gao, X., Zhang, X., Zhou, Y., Ma, X., Song, W., Zhao, H., et al. (2019). High-temporal-resolution transcriptome landscape of early maize seed development. Plant Cell **31**:974–992.

Yu, J., Holland, J.B., McMullen, M.D., and Buckler, E.S. (2008). Genetic design and statistical power of nested association mapping in maize. Genetics **178**:539–551.

Zhan, J., Thakare, D., Ma, C., Lloyd, A., Nixon, N.M., Arakaki, A.M., Burnett, W.J., Logan, K.O., Wang, D., Wang, X., et al. (2015). RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. Plant Cell **27**:513–531.

Zhan, W., Liu, J., Pan, Q., Wang, H., Yan, S., Li, K., Deng, M., Li, W., Liu, N., Kong, Q., et al. (2019). An allele of *ZmPORB2* encoding a protochlorophyllide oxidoreductase promotes tocopherol accumulation in both leaves and kernels of maize. Plant J. **100**:114–127.

Zhang, J., Li, C., Wu, C., Xiong, L., Chen, G., Zhang, Q., and Wang, S. (2006). RMD: a rice mutant database for functional analysis of the rice genome. Nucleic Acids Res. **34**:D745–D748.

Zhang, D., Liu, Y., Guo, Y., Yang, Q., Ye, J., Chen, S., and Xu, M. (2012). Fine-mapping of *qRfg2*, a QTL for resistance to *Gibberella* stalk rot in maize. Theor. Appl. Genet. **124**:585–596.

Zhang, J., Ku, L.X., Han, Z.P., Guo, S.L., Liu, H.J., Zhang, Z.Z., Cao, L.R., Cui, X.J., and Chen, Y.H. (2014). The *ZmCLA4* gene in the *qLA4-1* QTL controls leaf angle in maize (*Zea mays* L.). J. Exp. Bot. **65**:5063–5076.

Zhang, H., Wang, X., Pan, Q., Li, P., Liu, Y., Lu, X., Zhong, W., Li, M., Han, L., Li, J., et al. (2019). QTG-Seq accelerates QTL fine mapping through QTL partitioning and whole-genome sequencing of bulked segregant samples. Mol. Plant **12**:426–437.

Zhao, J., Sauvage, C., Zhao, J., Bitton, F., Bauchet, G., Liu, D., Huang, S., Tieman, D.M., Klee, H.J., and Causse, M. (2019). Meta-analysis of genome-wide association studies provides insights into genetic control of tomato flavor. Nat. Commun. **10**:1534.

Zheng, P., Allen, W.B., Roesler, K., Williams, M.E., Zhang, S., Li, J., Glassman, K., Ranch, J., Nubel, D., Solawetz, W., et al. (2008). A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat. Genet. **40**:367–372.

Zhong, Y., Liu, C., Qi, X., Jiao, Y., Wang, D., Wang, Y., Liu, Z., Chen, C., Chen, B., Tian, X., et al. (2019). Mutation of *ZmDMP* enhances haploid induction in maize. Nat. Plants **5**:575–580.

Zhou, P., Hirsch, C.N., Briggs, S.P., and Springer, N.M. (2019a). Dynamic patterns of gene expression additivity and regulatory variation throughout maize development. Mol. Plant **12**:410–425.

Zhou, S., Kremling, K., Bandillo, N., Richter, A., Zhang, Y.K., Ahern, K.R., Artyukhin, A.B., Hui, J.X., Younkin, G.C., Schroeder, F.C., et al. (2019b). Metabolome-scale genome-wide association studies reveal chemical diversity and genetic control of maize specialized metabolites. Plant Cell **31**:937–955.

Zhu, G., Gou, J., Klee, H., and Huang, S. (2019). Next-gen approaches to flavor-related metabolism. Annu. Rev. Plant Biol. **70**:187–212.

Zhu, G., Wang, S., Huang, Z., et al. (2018). Rewiring of the fruit metabolome in tomato breeding. Cell **172**:249–261.

Zila, C.T., Samayoa, L.F., Santiago, R., Butrón, A., and Holland, J.B. (2013). A genome-wide association study reveals genes associated with *fusarium* ear rot resistance in a maize core diversity panel. G3 (Bethesda) **3**:2095–2104.

Zunjare, R.U., Hossain, F., Muthusamy, V., Baveja, A., Chauhan, H.S., Bhat, J.S., Thirunavukkarasu, N., Saha, S., and Gupta, H.S. (2018). Development of biofortified maize hybrids through marker-assisted stacking of *β-carotene hydroxylase*, *Lycopene-ε-cyclase* and *Opaque2* genes. Front Plant Sci. **9**:178.

Zuo, W., Chao, Q., Zhang, N., Ye, J., Tan, G., Li, B., Xing, Y., Zhang, B., Liu, H., Fengler, K.A., et al. (2015). A maize wall-associated kinase confers quantitative resistance to head smut. Nat. Genet. **47**:151–157.