Germplasm Resources for Mapping Quantitative Traits in Maize

10

Anna Glowinski and Sherry Flint-Garcia

Abstract

The expression of quantitative traits is complex, often the result of multiple genes acting in concert, and interacting with the environment. Determining the genetic control of quantitative traits can be accomplished using a number of methods to link genotype to phenotype, such as linkage-based quantitative trait locus (QTL) mapping, genome-wide association mapping (GWAS), and multi-parent mapping including nested association mapping (NAM) and multi-parent advanced generation intercrosses (MAGIC). A wide array of germplasm resources are available for mapping QTL in maize. The purpose of this chapter is to provide a brief overview of QTL mapping methods, to provide background about commonly used germplasm resources, and to discuss the strengths and weakness of each.

10.1 Introduction

Quantitative trait locus (QTL) mapping entails finding an association between a genetic marker and a measurable phenotype. Researchers work from the phenotype to the genotype, using statistical techniques to localize chromosomal regions that contain genes and/or non-coding sequences contributing to the phenotypic variation of a quantitative trait in a given population. Most traits of interest in plant breeding show quantitative inheritance, which complicates the selection process since phenotypic performances only partially reflect the genetic values of individuals. The genetic variation of a quantitative trait is controlled by the collective effects of QTL (epistasis), interactions between QTL, the environment, and QTL by environment interactions.

The goal of this chapter is to provide an overview of methods to detect genotype-to-phenotype associations for quantitative traits and to provide information about populations that are already available. This is not meant to be an exhaustive review of all traits or QTL associations as each trait group is covered in a different chapter of this book.

A. Glowinski · S. Flint-Garcia (✉)
Division of Plant Sciences, University of Missouri,
301 Curtis Hall, Columbia, MO 65211, USA
e-mail: Sherry.Flint-Garcia@ars.usda.gov

S. Flint-Garcia
USDA-ARS Plant Genetics Research Unit,
Columbia, MO 65211, USA

J. Bennetzen et al. (eds.), *The Maize Genome*, Compendium of Plant Genomes,
https://doi.org/10.1007/978-3-319-97427-9_10

10.1.1 Qualitative Versus Quantitative Traits

In genetics, we can divide traits into two categories based on their effects on phenotype: qualitative and quantitative. Qualitative traits have discontinuous phenotypic variation, meaning a qualitative trait can fit into discrete categories. These are traits that can be referred to simply as "yes or no" traits, where an individual either has the trait or it does not. Usually, a single gene or a small number of genes control qualitative traits.

Conversely, quantitative traits display a continuous range of variation. Examples of quantitative traits include plant height, flowering time, and yield. These traits do not fit into discrete categories and have a continuous distribution. Generally, a larger number of genes control quantitative traits. Due to the continuous distribution of phenotypic values, quantitative genetics must employ statistical methods to link phenotypes to genotypes.

10.1.2 What Are QTL?

A QTL is a region of DNA that is associated with a particular trait, which varies in degree and which can be attributed to polygenic effects (i.e., the product of two or more genes and the environment) (Members of the Complex Trait 2003). The number of QTL which explain variation in the phenotypic trait is indicative of the genetic architecture for that trait; the more QTL, the more complex the trait.

A QTL is not a gene; at least not in the initial stages of discovery. A QTL is a large region of the genome (usually many centiMorgans and Mbp of DNA) which is linked to or contains the gene(s) that control a trait. QTL mapping is often a first step in identifying the actual genes underlying the trait because it is used to identify candidate genes in the genomic region. For genes whose function is already known, candidates can be identified based on pathways and gene expression networks. Genes of unknown function in the region can be compared to other species to identify homology-based candidates.

10.1.3 Unknown Genetic Architectures of Traits

Not all aspects of the genetic architecture of a particular trait are known. There are a variety of sources that contribute to the heritability and the genetic architecture of the trait.

10.1.3.1 Heritability

The goal of the plant breeder is to improve phenotypic values in a population by identifying and selecting superior genotypes. Because environment also affects the phenotype, there is not a perfect correspondence between phenotypic and genotypic values. To predict the outcome of selection in a collection of genotypes, a breeder must know the level of correspondence between phenotypic and genotypic values; this is known as heritability. Specifically, heritability is the percentage of the phenotypic variance that is attributable to differences among individuals in genotypic value and ranges from 0 (completely environmental) to 1 (completely genetic).

Although the heritability of a trait depends on how it is measured, in what environment(s) it is measured, and which plants are measured, different traits of maize tend to have different values of heritability. Qualitative traits, such as cob color and pericarp color, often have a value of heritability close to 1. Heritability values for quantitative traits are typically less but can vary greatly. For example, heritability can be very high for flowering (0.94; Buckler et al. 2009) and kernel protein content (0.83; Cook et al. 2012). In contrast, grain yield often has a significantly lower heritability. As a rule, traits with greater heritability can be modified more easily by selection and breeding than traits with lower heritability.

10.1.3.2 Number of Causal Loci

The term causal locus is defined as a functional genetic locus that influences and helps to explain the trait of interest. The number of causal loci contributing to a phenotype varies for different traits. Some traits are governed by many loci with smaller effects as is the case with flowering time (at least 39 QTL; Buckler et al. 2009), while

others are governed by fewer loci with larger effects (average of 4 to 5 QTL for amino acid content in grain; Deng et al. 2017). Generally, the more causal loci, the smaller the effect of each locus.

10.1.3.3 Magnitude of the Effects of Loci

Loci with larger effect sizes are more easily detected, while loci with smaller effect sizes are harder to detect. As a result of this, a large fraction of the genetic architecture of many complex traits is not well understood. Small-effect QTLs are often physically linked in a cluster or linked to large-effect QTL and fractionate during fine mapping, and there are often extensive epistatic interactions between small- and large-effect QTLs (Studer and Doebley 2011). A more complete understanding of quantitative traits will require a better understanding of the numbers, effect sizes, and genetic interactions of small-effect QTL.

10.2 Linkage-Based QTL Mapping

In linkage-based QTL mapping, QTL are mapped by identifying molecular markers that correlate with an observed trait (Veldboom et al. 1994; Grimmer et al. 2007; Zhang et al. 2009). This type of mapping depends upon recent genetic recombination between two different plant lines (as a result of a genetic cross) to identify general regions of interest.

10.2.1 Population Structure

Linkage-based QTL mapping requires the development of a mapping population, usually by crossing parents differing for the trait(s) of interest; e.g., tall x short, resistant x susceptible, high x low. The most common population structures include (1) F2 populations, (2) F2:3 populations which are created by self-pollinating the F2s to allow for replicated phenotypic trials, (3) BC1 populations where the F1 is backcrossed to the parent with the low/susceptible phenotype, (4) recombinant inbred line (RIL) populations where the F2s are self-pollinated many generations to near homozygosity in order to stabilize the genetics within each family, and (5) intermated RIL populations which allow for additional recombination prior to inbreeding. Because there is limited opportunity for recombination during the development of these population types, the linkage blocks are large and require only moderate marker density to define the recombination events.

10.2.2 Statistical Analyses

Analysis of variance (ANOVA), also known as single-marker regression, is the simplest method of linkage-based QTL mapping and was commonly used in the 1990s. The ANOVA method involves a marker regression at the marker, and provides an F statistic and associated p value for each marker. When the markers are widely spaced, the QTL may be quite far from all markers, causing low power for QTL detection.

Interval mapping makes use of a genetic map of the markers to interpolate locations between markers and, like ANOVA, assumes the presence of single QTL (Lander and Botstein 1989). Each locus is considered at one time, and the logarithm of odds ratio (LOD score) can be calculated for the comparison of two hypotheses: the presence of a QTL at a given position versus a model with no QTL at that position. A significance level is calculated by performing permutation testing (Churchill and Doerge 1994).

Composite interval mapping (CIM) can determine the location and effect size of QTL more accurately than single-QTL approaches, especially in small mapping populations where the effect of correlation between genotypes in the mapping population may be problematic. CIM is performed by using a subset of marker loci, usually identified by single-marker ANOVA, as covariates. These markers serve as proxies for unlinked QTL to increase the resolution of

interval mapping, by accounting for linked QTL and reducing the residual variation (Lynch and Walsh 1998).

10.2.3 Advantages of Linkage-Based QTL Mapping

There are many methods for linkage-based QTL mapping, and each has its advantages and disadvantages. Single-marker analysis is generally a good choice when the goal is simple detection of a QTL linked to a marker, rather than estimation of its actual position. Interval mapping offers a further increase in power of detection and more precise estimates of QTL effects and position. CIM considers the intervals between markers plus a few other well-chosen markers in each analysis, attempting to reduce or remove bias that occurs when multiple QTLs are linked to the marker/interval being considered (Lynch and Walsh 1998). One advantage all methods have in common is that linkage-based QTL mapping ***requires few genetic markers*** to ensure genome-wide coverage. In addition, depending on the population structure, the allelic classes are more balanced as compared to association analysis; see below. For example, in an F2 population, 50% of the alleles are expected to originate from each parent leading to ***higher statistical power*** per allele.

10.2.4 Disadvantages of Linkage-Based QTL Mapping

Linkage-based QTL mapping is limited to the genetic diversity present in the parents of the segregating population, leading to ***low allele richness***. Both single-marker regression and interval mapping approaches are biased when multiple QTL are linked to the maker/interval being tested. When using CIM, the biggest concern is finding suitable marker loci to serve as covariates in the analysis to help remove or reduce bias. The primary disadvantage of linkage-based approaches is the ***low mapping resolution*** due to limited recombination during population development. This low resolution can be alleviated by greatly increasing population size and/or increasing recombination through the use of advanced intercrosses.

10.3 Association Mapping

Association mapping was originally designed for the analysis of human diseases, but is now extensively used in plant genetics research as either a candidate gene association by studying single-nucleotide polymorphisms within candidate genes or as a genome-wide association study (GWAS) using anonymous molecular markers distributed across the whole genome.

Association mapping, also known as linkage disequilibrium (LD) mapping, is a method of mapping QTL that takes advantage of historic linkage disequilibrium to link phenotypes to genotypes, uncovering genetic associations (Buckler and Thornsberry 2002). It is based on the idea that polymorphisms underlying the trait that have entered a population only recently will be linked to the surrounding genetic sequence of the original evolutionary ancestor, or in other words will more often be found within a given haplotype, than outside of it.

10.3.1 Population Structure and Linkage Disequilibrium

Association mapping is generally conducted in germplasm panels consisting of pre-existing unrelated materials; i.e., no population development is required. The more diverse the germplasm is, the more rapidly LD decays within the population and the better the mapping resolution, but the more markers are required. However, if the population is too diverse, there will be a high proportion of low-frequency alleles which are either filtered out (e.g., minor allele frequencies less than 0.05) or have low statistical power. Finally, if the population has genetic structure, then the uneven distribution of alleles among the

subpopulations could lead to false positives unless population structure is accounted for in the statistical model (Yu et al. 2006). The art of assembling an association panel lies in balancing these factors.

10.3.2 Types of Association Mapping

10.3.2.1 Candidate Gene Based

As mentioned above, association mapping can be candidate gene based in which single-nucleotide polymorphisms are studied within candidate genes (Castiblanco et al. 2017). Genes associated with a phenotype of interest are selected for association mapping, and polymorphisms in only these pre-selected genes are identified and tested for association with the trait. Candidate gene-based approaches remain the most effective way of dissecting complex traits for species where sufficiently dense marker assays are not yet developed (Thavamanikumar et al. 2011), a situation that is becoming increasingly rare with the advent of next-generation sequencing-based marker systems.

10.3.2.2 Genome-Wide Association Studies (GWAS)

GWAS studies investigate the entire genome, by rapidly scanning markers across a genome to identify SNPs associated with a particular phenotype. GWAS requires corrections for population structure using PCA and/or kinship matrices in order to prevent false positives (Yu et al. 2006). Depending on the germplasm and the extent of linkage disequilibrium, GWAS generally cannot identify which polymorphisms are causal but often identifies the likely candidate gene. To date, GWAS experiments have been performed for a variety of traits in maize; see below for examples.

GWAS experiments are performed by scanning the entire genome for significant associations between a panel of SNPs and a particular phenotype. Associations must then be independently verified in order to show that they either (a) contribute to the trait of interest directly, or (b) are linked to/in linkage disequilibrium with a QTL that contributes to the trait of interest.

10.3.3 Advantages of Association Mapping

Association mapping has several advantages over linkage mapping in traditional biparental populations: (1) Currently ***existing populations are used*** rather than generating a population via a biparental cross; (2) a potentially ***large number of alleles*** per locus—compared to only two—can be surveyed simultaneously; and (3) **dramatically *increased resolution*** can be achieved (Flint-Garcia et al. 2005). Given enough statistical power and marker coverage, the low LD in maize may allow for the identification of the causative polymorphism within a candidate gene.

10.3.4 Disadvantages of Association Mapping

Association mapping requires extensive knowledge of SNP relationships within the genome, particularly in maize where LD breaks down rapidly in diverse germplasm (Flint-Garcia et al. 2003), implying that ***tens of millions of SNPs may be required*** to characterize the haplotype structure. GWAS may have ***reduced statistical power*** for detecting rare alleles because the power for detecting a QTL is determined by the frequency of alleles (Myles et al. 2009). ***False positives*** can be seen due to population structure; however, there are ways to correct for population structure (Thornsberry et al. 2001). If population structure contributes to the variation in your trait, over-correcting for population structure may lead you to many false negatives.

10.4 Nested Association Mapping

Linkage analysis and association mapping are two commonly used approaches to dissect the genetic architecture of quantitative traits (Mackay 2001;

Lander and Schork 1994). Linkage analysis and association mapping are complementary approaches: Linkage analysis often identifies large chromosome regions of interest with relatively low marker coverage, while association mapping provides high resolution with very high marker coverage (Thornsberry et al. 2001; Hirschhorn and Daly 2005). Nested association mapping (NAM) aims to create an integrated mapping population specifically designed for a full genome scan with high mapping resolution and high power for QTL with different effect sizes. The NAM strategy addresses complex trait dissection at a fundamental level through generating a common mapping resource that enables researchers to efficiently exploit genetic, genomic, and systems biology tools (Yu et al. 2008). Currently, the NAM strategy has been employed for maize (McMullen et al. 2009) and other species (Fragoso et al. 2017; Song et al. 2017; Bajgain et al. 2016), with NAM populations under development for many other species.

10.4.1 Advantages of NAM

NAM takes advantage of both historic and recent recombination events in order to achieve ***low marker density requirements***, ***high allele richness***, ***high mapping resolution***, and ***high statistical power***, with none of the disadvantages of either linkage analysis or association mapping. This allows for the discovery of QTL with greater precision and accuracy. Parental alleles are shuffled over several generations through segregation and genetic recombination providing new combinations of alleles for study. NAM populations also have an added benefit, in that they can function as an archive for genetic diversity.

10.4.2 Disadvantages of NAM

Unless a NAM population already exists as in the case of maize, a new ***population must be generated*** which utilizes both time and resources. Challenges include ensuring that the pedigree of each cross is maintained while advancing to the next generation, that the founders are diverse enough to carry different alleles for important characteristics, and that near-complete homozygosity is reached in the final population. While NAM captures thousands of recombination events, recombination and segregation distortion vary among different families which can limit the precision of genetic dissection of quantitative traits (McMullen et al. 2009; Ladejobi et al. 2016).

10.5 Available QTL Mapping Populations and Germplasm Resources

10.5.1 Intermated B73/Mo17 (IBM)

Hundreds of linkage-based QTL populations have been created over the past 30 years; each is focused on a specific trait(s) but rarely made available to the public. However, the maize community in the public sector has championed the use of the intermated B73 × Mo17 (IBM) population as a central linkage mapping resource because of the historical value of the two parents and the value of merging the genetic map (IBM) to the maize genome (B73) (Coe et al. 2002).

The IBM population was the first widely used QTL population derived from additional generations of intermating prior to inbreeding (Lee et al. 2002). The IBM is comprised of approximately 300 RILs, 94 of which are referred to as "the core set." The RILs were derived from the single-cross hybrid of inbreds B73 (female) and Mo17 (Lee et al. 2002). A single F1 plant was self-pollinated to produce the F2 generation. In the F2, plants were used once, as male or female, in a cross with another plant so that 250 pairs of plants were mated. A single kernel was taken from each ear and bulked with the seed of the other ears to form the Syn1 generation. This procedure was repeated for four additional generations to produce the Syn5 generation. The increased opportunity for recombination in IBM has resulted in an almost four-fold increase in the

genetic map distance compared with conventional non-intermated RIL populations, allowing for more precise definition of QTLs. IBM has been widely used for developing genetic markers and anchoring them to the genetic map as well as the physical map (Coe et al. 2002) and for studying the genetic architecture of numerous traits (e.g., Eichten et al. 2011; Ordas et al. 2009; Rodríguez et al. 2008; Zhang et al. 2010a; Baxter et al. 2013; Balint-Kurti et al. 2007; Dubois et al. 2010; Hazen et al. 2003).

10.5.2 Association Panels

Describing the commonly used association panels is a somewhat difficult task. They do not require additional population development as they are typically collections of materials previously created by multiple groups. Because of this, association panels are extremely easy to modify by merging panels together, dropping various groups of germplasm from a panel based on phenology (e.g., adaptation to temperate or tropical environments) and/or germplasm availability (e.g., not all germplasm is publicly available), and customizing panels for specific traits by adding lines chosen for extremes in the trait. Studies often report phenotypes on multiple panels for the same trait(s), and the results are compared in the context of allele frequencies and population structure. The following is a short list of the most commonly used association panels.

10.5.2.1 Maize 282 Association Panel

The maize 282 association panel was one the earliest association panels in maize and consists of breeding lines assembled by Major Goodman (Flint-Garcia et al. 2005). The very first maize association panel by Thornsberry et al. (2001) consisted of 102 inbred lines, but it was quickly realized that this small of a panel had insufficient power to detect QTL; hence, it was increased to 302 inbred lines (Flint-Garcia et al. 2005) based on pedigree information (e.g., Gerdes et al. 1993) and prior to the availability of SNPs to characterize germplasm relationships. After genotyping, a number of isolines (highly related lines derived from backcrossing with B73, for example) were identified and removed from the panel yielding the current 282 association panel. The 282 panel, also known as the Goodman–Buckler panel, represents a sample of the diversity present in the public sector including current breeding lines (at the time of development) as well as historically important lines from both temperate and tropical programs. This association panel has been used for a variety of association studies since its creation (e.g., Krill et al. 2010; Hung et al. 2012; Cook et al. 2012; Hu et al. 2018; Diepenbrock et al. 2017; Hu et al. 2017; Olukolu et al. 2016; Zhang et al. 2010b; Butron et al. 2010; Harjes et al. 2008; Benke et al. 2015; Samayoa et al. 2015; Olukolu et al. 2013).

10.5.2.2 Ames Association Panel

The USDA North Central Regional Plant Introduction Station (NCRPIS) in Ames, Iowa, maintains over 3000 maize inbreds from around the world. The Ames panel was created by choosing over 2500 inbred lines from the NCRPIS inbred collection based only on sufficient seed availability and a minimum of five generations of self-pollination to ensure an inbred nature, and represents nearly a century of maize breeding efforts. The panel has been genotypically characterized by genotyping-by-sequencing (GBS; Elshire et al. 2011) in order to assist with curatorial management of germplasm collections and to evaluate diversity within breeding programs (Romay et al. 2013) and for use in association mapping (Lu et al. 2015; Xue et al. 2016; Peiffer et al. 2014; Zila et al. 2013). Because the population is so large and genetically diverse, subsets of lines from the Ames association panel have been used successfully to characterize many different traits (Pace et al. 2015).

10.5.2.3 Wisconsin Diversity Panel

A subset of the Ames panel with a reduced phenology for adaptation to the northern corn belt was chosen and is known as the Wisconsin Diversity Panel (WiDiv). This panel contains 627 lines selected based on flowering in the target

environment of Wisconsin, agronomic suitability, uniformity, and seed supply (Hansey et al. 2011).

Many of the WiDiv panel lines trace back to eight open-pollinated populations including Iowa Stiff Stalk, Minnesota No. 13, Reid Yellow Dent, Lancaster Surecrop, Golden Glow, Funk Yellow Dent, Pride of Saline, and Krug among others. Having multiple genotypes derived from the same open-pollinated population helps to maintain a balance of allele frequencies which results in increased statistical power. The WiDiv has been used in a number of high throughput image analysis projects to investigate stalk, tassel, ear, and kernel morphology traits (Miller et al. 2017; Gustin et al. 2013; Heckwolf et al. 2015; Muttoni et al. 2012), and for GWAS studies of juvenile-to-adult vegetative and vegetative-to-reproductive developmental transitions (Hirsch et al. 2014b) and tassel traits (Gage et al. 2018).

10.5.2.4 CIMMYT Association Panels

The International Center for Maize and Wheat Improvement (CIMMYT) has a global mandate for improving the productivity and sustainability of maize and wheat in developing countries (Hoisington et al. 1999). The CIMMYT maize germplasm bank contains over 28,000 seed samples, including inbred lines, breeding populations, landraces, and wild relatives. To leverage this germplasm resource, CIMMYT has developed a number of GWAS panels, primarily to study grain carotenoid content and drought tolerance. The carotenoid research was conducted on two panels: a set of 245 diverse maize inbred lines predominantly derived from tropical and subtropical adapted maize germplasm (Yan et al. 2010) and the carotenoid association mapping (CAM) panel consisting of 380 primarily diverse tropical and subtropical lines assembled by the HarvestPlus-funded program at CIMMYT (Suwarno et al. 2015).

The CIMMYT drought panel of 350 inbred lines was used to test candidates in abscisic acid (ABA) in response to drought (Setter et al. 2011) and to conduct GWAS for agronomic trait (Xue et al. 2013) and metabolic (Zhang et al. 2016) responses to drought as compared to well-watered conditions. In addition, the drought-tolerant maize for Africa (DTMA; ~250–300 lines) and improved maize for African soils (IMAS; ~400 lines) panels were combined to identify a major QTL for resistance to tar spot complex (Mahuku et al. 2016) and maize lethal necrosis disease (Gowda et al. 2015). Finally, another collection of 940 African lines was genotyped (Semagn et al. 2012) and evaluated disease resistance including Fusarium ear rot (Chen et al. 2016).

10.5.2.5 Chinese Association Panels

The first Chinese association panel was composed of 155 diverse temperate-adapted maize inbred lines from China (Yang et al. 2010) and was later referred to as the Chinese association mapping (CAM155) panel in subsequent publications by the lead authors (Li et al. 2011). This panel was used in a GWAS study of kernel carotenoids (Yan et al. 2010), before being merged with other germplasm to form additional panels such as the AM508 (see below).

A broader global diverse line panel of 527/513 inbred lines representative of tropical, subtropical, and temperate germplasm was collected to construct a larger association panel (Yang et al. 2011). This collection includes 527 lines from the GEM project, CIMMYT maize breeding programs, elite parents of commercial hybrids widely used in China, lines derived from Chinese landraces, and high-oil and high-provitamin A lines; 513 lines were genotyped with the Illumina MaizeSNP50 array and used for GWAS of kernel α-tocopherol content (Li et al. 2012), maize rough dwarf mosaic virus (Chen et al. 2015), and a large number of plant, ear, kernel, and yield-related traits (Yang et al. 2014), among many other traits.

More recently, the 527-/513-inbred line panel above was reduced to a 508-inbred line panel known as AM508 which was first described in an investigation of kernel oil content (Li et al. 2013). A commonly used subset of the AM508 is a set of 368 lines which was subjected to RNA-seq (Fu et al. 2013). This 368-line panel was used to investigate many traits and phenomena such as expression QTL (eQTL),

regulatory networks, non-coding sequences, and metabolites in the developing kernel (Fu et al. 2013; Wen et al. 2014; Liu et al. 2017).

10.5.2.6 European Association Panels

A widely used panel in Europe consisted of 375 inbred lines representative of American, European, and tropical maize (Camus-Kulandaivelu et al. 2006) which included the original 102-inbred subset of the 282 (Thornsberry et al. 2001) and a unique set of 153 inbreds derived from self-pollinating European landraces. This panel was used to study epistatic interactions in *Opaque2* for kernel traits (Manicacci et al. 2009), flowering time (Durand et al. 2012; Camus-Kulandaivelu et al. 2006), and phenology and plant architecture traits (Bouchet et al. 2016).

A set of 289 diverse dent inbred lines from the Americas, Europe, and China has been assembled to investigate genomic and metabolic prediction of heterosis (Riedelsheimer et al. 2012a) as well as GWAS for leaf metabolites and biomass-related traits (Riedelsheimer et al. 2012b).

Another recent association panel, of sorts, is comprised of two separate panels of 306 dent and 292 flint maize inbred lines based on collections of Spanish, French, and German breeders from the Cornfed Project. These are often evaluated as hybrids with the opposite heterotic group (i.e., flint panel crossed with dent tester and vice versa). This panel has been investigated for cold tolerance (Revilla et al. 2016).

10.5.2.7 Other Association Panels

Private industry has also used association analysis using their propriety germplasm, though few GWAS studies have been published by industry. Of those published, Belo et al. (2008) used 553 historically important and current elite maize inbred lines from Pioneer Hi-Bred to conduct GWAS for fatty acid content in kernels; 1,487 inbred lines from Limagrain representing elite European and North American germplasm were used to investigate northern corn leaf blight (Van Inghelandt et al. 2012); and Dow AgroSciences used 300 inbreds, including 215 Dow proprietary lines of North and South American origin, to validate QTL for gray leaf spot (Mammadov et al. 2015).

Additional trait- and geography-specific panels have been assembled. For example, the 300-inbred line panel of Warburton et al. (2013) was used to investigate resistance *Aspergillus flavus*, aflatoxin accumulation, and drought (Warburton et al. 2013; Farfan et al. 2015). A subset of 287 these 300 lines has been used to map resistance to corn earworm and associated metabolic pathways (Warburton et al. 2018). A Brazilian panel of 183 lines was assembled to conduct GWAS for Fusarium ear rot resistance (Coan et al. 2018), but is being expanded to 335 (M. Warburton, personal communication). A set of 240 Indian and CIMMYT lines were analyzed for associations with yield and yield component traits under drought conditions (Thirunavukkarasu et al. 2014).

10.5.3 US NAM

The US NAM consists of 5000 RILs derived by crossing 25 diverse maize lines to B73 (McMullen et al. 2009). The 25 diverse inbred lines were chosen as parents to maximize diversity encompassed in the 282 association panel (Flint-Garcia et al. 2005) and preserve historic linkage disequilibrium (Yu et al. 2008). Each parental line was crossed to B73, the inbred chosen for the reference genome. The F1 plants were then self-pollinated by single seed descent for six generations to create a total of 200 homozygous RILs per family, for a total of 5000 RILs which were originally genotyped with 1536 SNPs (McMullen et al. 2009) and subsequently by GBS. Benefits of using the US NAM population for QTL mapping include broader genetic diversity, higher mapping resolution than individual biparental populations, and an increase in statistical power because allele frequencies are balanced within each family. While the US NAM population only taps the diversity of 25 founder lines, it is large enough to address questions regarding magnitudes of QTL effects, heterosis, and the mapping of numerous genes controlling various traits (e.g., Buckler et al. 2009; Poland

et al. 2011; Cook et al. 2012; Hirsch et al. 2014a; Handrick et al. 2016; Kump et al. 2011; Tian et al. 2011; Benson et al. 2015; Brown et al. 2011).

10.5.4 European NAM

The European NAM population was created by creating two half-sib panels of 11 and 13 half-sib families, one for European Dent and one for European Flint maize, respectively (Bauer et al. 2013). Each of the two panels consists of a common parent crossed to founder lines that represent important and diverse breeding lines of the European maize germplasm. In the Dent panel, a central line (F353 from France) was crossed with ten Dent founder lines. In the Flint panel, the central line (UH007 from Germany) was crossed with 11 Flint founder lines. In addition, each of the common parents was crossed with B73, and the reciprocal populations F353xUH007 and UH007xF353 were generated. These additional populations were made to connect the two panels to each other and with the US NAM population. All progenies are homozygous doubled haploid lines obtained from F1 plants. The resulting 24 doubled haploid populations each consist of 35–129 lines, for a total of 2,267 doubled haploid lines, and have been genotyped with the Illumina MaizeSNP50 array (Ganal et al. 2011). The European NAM population has been used to study recombination rate (Bauer et al. 2013) and genomic prediction of yield (Lehermeier et al. 2014).

10.5.5 Chinese NAM

The Chinese NAM (CN-NAM) population was developed by crossing 11 diverse inbred lines representing the heterotic groups used in Chinese maize breeding with the common parent "HZS" which has wide adaptability and good combining ability. The F2s were self-pollinated to create 1971 RILs which were genotyped by GBS (Li et al. 2015). The CN-NAM population has been used to dissect drought tolerance (Li et al. 2016a), inflorescence size (Wu et al. 2016), and flowering time (Li et al. 2015, 2016b).

10.5.6 Multi-parent Advanced Generation Intercrosses (MAGIC) and Other Multi-parent Populations

Multi-parent advanced generation intercross (MAGIC) populations have now been developed for a variety of species including maize, rice, wheat, and *Arabidopsis* (Dell'Acqua et al. 2015; Bandillo et al. 2013; Huang et al. 2012; Kover et al. 2009). The maize MAGIC population contains 1636 RILs derived from eight genetically diverse founder lines that were crossed in a funnel breeding design (Dell'Acqua et al. 2015). RILs were produced by pooling two-way, four-way, and eight-way hybrids in 35 independent breeding funnels (subfamilies). Each funnel was advanced by single seed descent to the F6 generation. This MAGIC population has been used to investigate flowering time, plant and ear height and grain size (Dell'Acqua et al. 2015).

The mapping power and resolution of MAGIC maize are strengthened by high minor allele frequencies and a rapid decay of linkage disequilibrium. Similar to the US NAM population, MAGIC maize has broader genetic diversity, higher resolutions than biparental populations, and a reduction of problems associated with the frequency of rare alleles (Holland 2015). These benefits make MAGIC maize a useful population for QTL mapping in maize.

A more recent multi-parent mapping method is called random-open-parent association mapping (ROAM), where RIL populations are derived from a number of inbred lines crossed in combinations without an a priori requirement to interconnect across populations (Xiao et al. 2016; Pan et al. 2016). In concept, the ROAM method could be used when merging multiple NAM panels together, such as the US, CN, and European NAM populations, where different common hub parents are used and where a variable number of populations are derived from each hub. An advantage of ROAM is that additional

families can be developed and added, as there is no a priori design to the larger ROAM population. A possible disadvantage is that there may be lower power for some alleles due to unbalanced allele frequencies among the parents (Xiao et al. 2017).

10.6 Conclusions

The goal of this chapter was to provide an overview of methods for genotype-to-phenotype associations and introduce some of the mapping resources that are available for study. Detailed reviews of QTL and GWAS analyses for insect resistance, fungal diseases, cold tolerance, root system architecture traits, nitrogen use efficiency, and kernel oil content can be found in other chapters of this book.

References

Bajgain P, Rouse MN, Tsilo TJ, Macharia GK, Bhavani S, Jin Y, Anderson JA (2016) Nested association mapping of stem rust resistance in wheat using genotyping by sequencing. PLOS ONE 11(5):e0155760. https://doi.org/10.1371/journal.pone.0155760

Balint-Kurti PJ, Zwonitzer JC, Wisser RJ, Carson ML, Oropeza-Rosas MA, Holland JB, Szalma SJ (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

Bandillo N, Raghavan C, Muyco PA, Sevilla MAL, Lobina IT, Dilla-Ermita CJ, Tung C-W, McCouch S, Thomson M, Mauleon R, Singh RK, Gregorio G, Redoña E, Leung H (2013) Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice 6(1):11. https://doi.org/10.1186/1939-8433-6-11

Bauer E, Falque M, Walter H, Bauland C, Camisan C, Campo L, Meyer N, Ranc N, Rincent R, Schipprack W, Altmann T, Flament P, Melchinger AE, Menz M, Moreno-González J, Ouzunova M, Revilla P, Charcosset A, Martin OC, Schön C-C (2013) Intraspecific variation of recombination rate in maize. Genome Biol 14(9):R103. https://doi.org/10.1186/gb-2013-14-9-r103

Baxter I, Gustin JL, Settles M, Hoekenga OA (2013) Ionomic characterization of maize kernels in the intermated B73 × Mo17 population. Crop Sci 3:208–220

Belo A, Zheng P, Luck S, Shen B, Meyer DJ, Li B, Tingey S, Rafalski A (2008) Whole genome scan detects an allelic variant of fad2 associated with increased oleic acid levels in maize. Mol Genet Genomics 279(1):1–10

Benke A, Urbany C, Stich B (2015) Genome-wide association mapping of iron homeostasis in the maize association population. BMC Genet 16(1):1. https://doi.org/10.1186/s12863-014-0153-0

Benson JM, Poland JA, Benson BM, Stromberg EL, Nelson RJ (2015) Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. PLOS Genetics 11(3):e1005045. https://doi.org/10.1371/journal.pgen.1005045

Bouchet S, Bertin P, Presterl T, Jamin P, Coubriche D, Gouesnard B, Laborde J, Charcosset A (2016) Association mapping for phenology and plant architecture in maize shows higher power for developmental traits compared with growth influenced traits. Heredity 118:249. https://doi.org/10.1038/hdy.2016.88. https://www.nature.com/articles/hdy201688#supplementary-information

Brown PJ, Upadyayula N, Mahone GS, Tian F, Bradbury PJ, Myles S, Holland JB, Flint-Garcia S, McMullen MD, Buckler ES, Rocheford TR (2011) Distinct genetic architectures for male and female inflorescence traits of maize. PLoS Genet 7(11):e1002383. https://doi.org/10.1371/journal.pgen.1002383

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC (2009) The genetic architecture of maize flowering time. Science 325(5941):714–718

Buckler ES, Thornsberry JM (2002) Plant molecular diversity and applications to genomics. Curr Opin Plant Biol 5(2):107–111. https://doi.org/10.1016/S1369-5266(02)00238-8

Butron A, Chen YC, Rottinghaus GE, McMullen MD (2010) Genetic variation at bx1 controls DIMBOA content in maize. Theor Appl Genet 120:721–734

Camus-Kulandaivelu L, Veyrieras JB, Madur D, Combes V, Fourmann M, Barraud S, Dubreuil P, Gouesnard B, Manicacci D, Charcosset A (2006) Maize adaptation to temperate climate: relationship between population structure and polymorphism in the Dwarf8 gene. Genetics 172(4):2449–2463

Castiblanco V, Marulanda JJ, Würschum T, Miedaner T (2017) Candidate gene based association mapping in Fusarium culmorum for field quantitative pathogenicity and mycotoxin production in wheat. BMC Genet 18(1):49. https://doi.org/10.1186/s12863-017-0511-9

Chen G, Wang X, Hao J, Yan J, Ding J (2015) Genome-wide association implicates candidate genes conferring resistance to maize rough dwarf disease in maize. PLOS ONE 10(11):e0142001. https://doi.org/10.1371/journal.pone.0142001

Chen J, Shrestha R, Ding J, Zheng H, Mu C, Wu J, Mahuku G (2016) Genome-wide association study and QTL mapping reveal genomic loci associated with *Fusarium* ear rot resistance in tropical maize

germplasm. G3 Genes Genomes Genet 6(12):3803–3815. https://doi.org/10.1534/g3.116.034561

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138: 963–971

Coan MMD, Senhorinho HJC, Pinto RJB, Scapim CA, Tessmann DJ, Williams WP, Warburton ML (2018) Genome-wide association study of resistance to ear rot by Fusarium verticillioides in a tropical field maize and popcorn core collection. Crop Sci 58(2):564–578. https://doi.org/10.2135/cropsci2017.05.0322

Coe E, Cone K, McMullen M, Chen S-S, Davis G, Gardiner J, Liscum E, Polacco M, Paterson A, Sanchez-Villeda H, Soderlund C, Wing R (2002) Access to the maize genome: an integrated physical and genetic map. Plant Physiol 128(1):9–12. https://doi.org/10.1104/pp.010953

Cook JP, McMullen MD, Holland JB, Tian F, Bradbury P, Ross-Ibarra J, Buckler ES, Flint-Garcia SA (2012) Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. Plant Physiol 158(2):824–834. https://doi.org/10.1104/pp.111.185033

Dell'Acqua M, Gatti DM, Pea G, Cattonaro F, Coppens F, Magris G, Hlaing AL, Aung HH, Nelissen H, Baute J, Frascaroli E, Churchill GA, Inzé D, Morgante M, Pè ME (2015) Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in Zea mays. Genome Biol 16(1):167. https://doi.org/10.1186/s13059-015-0716-z

Deng M, Li D, Luo J, Xiao Y, Liu H, Pan Q, Zhang X, Jin M, Zhao M, Yan J (2017) The genetic architecture of amino acids dissection by association and linkage analysis in maize. Plant Biotechnol J 15(10):1250–1263. https://doi.org/10.1111/pbi.12712

Diepenbrock CH, Kandianis CB, Lipka AE, Magallanes-Lundback M, Vaillancourt B, Góngora-Castillo E, Wallace JG, Cepela J, Mesberg A, Bradbury PJ, Ilut DC, Mateos-Hernandez M, Hamilton J, Owens BF, Tiede T, Buckler ES, Rocheford T, Buell CR, Gore MA, DellaPenna D (2017) Novel loci underlie natural variation in vitamin E levels in maize grain. Plant Cell 29(10):2374–2392. https://doi.org/10.1105/tpc.17.00475

Dubois PG, Olsefski GT, Flint-Garcia S, Setter TL, Hoekenga OA, Brutnell TP (2010) Physiological and genetic characterization of end-of-day far-red light response in maize seedlings. Plant Physiol 154 (1):173–186. https://doi.org/10.1104/pp.110.159830

Durand E, Bouchet S, Bertin P, Ressayre A, Jamin P, Charcosset A, Dillmann C, Tenaillon MI (2012) Flowering time in maize: linkage and epistasis at a major effect locus. Genetics 190:1547–1562. https://doi.org/10.1534/genetics.111.136903

Eichten SR, Foerster JM, de Leon N, Kai Y, Yeh C-T, Liu S, Jeddeloh JA, Schnable PS, Kaeppler SM, Springer NM (2011) B73-Mo17 near-isogenic lines demonstrate dispersed structural variation in maize. Plant Physiol 156(4):1679–1690. https://doi.org/10.1104/pp.111.174748

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6(5):e19379. https://doi.org/10.1371/journal.pone.0019379

Farfan IDB, De La Fuente GN, Murray SC, Isakeit T, Huang P-C, Warburton M, Williams P, Windham GL, 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(2):e0117737. https://doi.org/10.1371/journal.pone.0117737

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374

Flint-Garcia SA, Thuillet A-C, Yu J, Pressoir G, Romero SM, Mitchell SE, Doebley J, Kresovich S, Goodman MM, Buckler ES (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J 44(6):1054–1064. https://doi.org/10.1111/j.1365-313X.2005.02591.x

Fragoso CA, Moreno M, Wang Z, Heffelfinger C, Arbelaez LJ, Aguirre JA, Franco N, Romero LE, Labadie K, Zhao H, Dellaporta SL, Lorieux M (2017) Genetic architecture of a rice nested association mapping population. G3 Genes Genomes Genet 7(6): 1913–1926. https://doi.org/10.1534/g3.117.041608

Fu J, Cheng Y, Linghu J, Yang X, Kang L, Zhang Z, Zhang J, He C, Du X, Peng Z, Wang B, Zhai L, Dai C, Xu J, Wang W, Li X, Zheng J, Chen L, Luo L, Liu J, Qian X, Yan J, Wang J, Wang G (2013) RNA sequencing reveals the complex regulatory network in the maize kernel. Nat Commun 4:2832. https://doi.org/10.1038/ncomms3832. https://www.nature.com/articles/ncomms3832#supplementary-information

Gage JL, White MR, Edwards J, Kaeppler S, de Leon N (2018) Selection signatures underlying dramatic male inflorescence transformation during modern hybrid maize breeding. bioRxiv. https://doi.org/10.1101/284109

Ganal MW, Durstewitz G, Polley A, Bérard A, Buckler ES, Charcosset A, Clarke JD, Graner EM, Hansen M, Joets J, Le Paslier MC, McMullen MD, Montalent P, Rose M, Schon CC, Sun Q, Walter H, Martin OC, Falque M (2011) A large maize (Zea mays L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS One 6. https://doi.org/10.1371/journal.pone.0028334

Gerdes JT, Behr CF, Coors JG, Tracy WF (1993) Compilation of North American maize breeding germplasm. CSSA, Madison, WI

Gowda M, Das B, Makumbi D, Babu R, Semagn K, Mahuku G, Olsen MS, Bright JM, Beyene Y, Prasanna BM (2015) Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. Theor Appl Genet 128(10):1957–1968. https://doi.org/10.1007/s00122-015-2559-0

Grimmer MK, Trybush S, Hanley S, Francis SA, Karp A, Asher MJC (2007) An anchored linkage map for sugar

beet based on AFLP, SNP and RAPD markers and QTL mapping of a new source of resistance to Beet necrotic yellow vein virus. Theor Appl Genet 114 (7):1151–1160. https://doi.org/10.1007/s00122-007-0507-3

Gustin JL, Jackson S, Williams C, Patel A, Armstrong P, Peter GF, Settles AM (2013) Analysis of maize (Zea mays) kernel density and volume using microcomputed tomography and single-kernel near-infrared spectroscopy. J Agric Food Chem 61(46):10872–10880. https://doi.org/10.1021/jf403790v

Handrick V, Robert CAM, Ahern KR, Zhou S, Machado RAR, Maag D, Glauser G, Fernandez-Penny FE, Chandran JN, Rodgers-Melnik E, Schneider B, Buckler ES, Boland W, Gershenzon J, Jander G, Erb M, Köllner TG (2016) Biosynthesis of 8-*O*-methylated benzoxazinoid defense compounds in maize. Plant Cell 28(7):1682–1700. https://doi.org/10.1105/tpc.16.00065

Hansey CN, Johnson JM, Sekhon RS, Kaeppler SM, De Leon N (2011) Genetic diversity of a maize association population with restricted phenology. Crop Sci 51. https://doi.org/10.2135/cropsci2010.03.0178

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science 319(5861): 330–333

Hazen SP, Hawley RM, Davis GL, Henrissat B, Walton JD (2003) Quantitative trait loci and comparative genomics of cereal cell wall composition. Plant Physiol 132:263–271

Heckwolf S, Heckwolf M, Kaeppler SM, de Leon N, Spalding EP (2015) Image analysis of anatomical traits in stalk transections of maize and other grasses. Plant Methods 11(1):26. https://doi.org/10.1186/s13007-015-0070-x

Hirsch CN, Flint-Garcia SA, Beissinger TM, Eichten SR, Deshpande S, Barry K, McMullen MD, Holland JB, Buckler ES, Springer N, Buell CR, de Leon N, Kaeppler SM (2014a) Insights into the effects of long-term artificial selection on seed size in maize. Genetics 198(1):409–421. https://doi.org/10.1534/genetics.114.167155

Hirsch CN, Foerster JM, Johnson JM, Sekhon RS, Muttoni G, Vaillancourt B, Peñagaricano F, Lindquist E, Pedraza MA, Barry K, de Leon N, Kaeppler SM, Buell CR (2014b) Insights into the maize pan-genome and pan-transcriptome. Plant Cell 26(1):121–135. https://doi.org/10.1105/tpc.113.119982

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95. https://doi.org/10.1038/nrg1521

Hoisington D, Khairallah M, Reeves T, Ribaut J-M, Skovmand B, Taba S, Warburton M (1999) Plant genetic resources: what can they contribute toward increased crop productivity? Proc Natl Acad Sci 96 (11):5937–5943. https://doi.org/10.1073/pnas.96.11.5937

Holland JB (2015) MAGIC maize: a new resource for plant genetics. Genome Biol 16(1):163. https://doi.org/10.1186/s13059-015-0713-2

Hu G, Li Z, Lu Y, Li C, Gong S, Yan S, Li G, Wang M, Ren H, Guan H, Zhang Z, Qin D, Chai M, Yu J, Li Y, Yang D, Wang T, Zhang Z (2017) Genome-wide association study Identified multiple genetic loci on chilling resistance during germination in maize. Sci Rep 7(1):10840. https://doi.org/10.1038/s41598-017-11318-6

Hu Y, Ren J, Peng Z, Umana AA, Le H, Danilova T, Fu J, Wang H, Robertson A, Hulbert SH, White FF, Liu S (2018) Analysis of extreme phenotype bulk copy number variation (XP-CNV) identified the association of rp1 with resistance to Goss's wilt of maize. Front Plant Sci 9(110). https://doi.org/10.3389/fpls.2018.00110

Huang BE, George AW, Forrest KL, Kilian A, Hayden MJ, Morell MK (2012) A multiparent advanced generation inter-cross population for genetic analysis in wheat. Plant Biotechnol J 10. https://doi.org/10.1111/j.1467-7652.2012.00702.x

Hung H-Y, Shannon LM, Tian F, Bradbury PJ, Chen C, Flint-Garcia SA, McMullen MD, Ware D, Buckler ES, Doebley JF, Holland JB (2012) ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proc Natl Acad Sci 109(28):E1913–E1921. https://doi.org/10.1073/pnas.1203189109

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R (2009) A multiparent advanced generation inter-cross to fine-map quantitative traits in Arabidopsis thaliana. PLOS Genet 5(7):e1000551. https://doi.org/10.1371/journal.pgen.1000551

Krill AM, Kirst M, Kochian LV, Buckler ES, Hoekenga OA (2010) Association and linkage analysis of aluminum tolerance genes in maize. PLOS ONE 5(4):e9958. https://doi.org/10.1371/journal.pone.0009958

Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Zwonitzer JC, Kresovich S, McMullen MD, Ware D (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nat Genet 43. https://doi.org/10.1038/ng.747

Ladejobi O, Elderfield J, Gardner KA, Gaynor RC, Hickey J, Hibberd JM, Mackay IJ, Bentley AR (2016) Maximizing the potential of multi-parental crop populations. Appl Transl Genomics 11(Supplement C):9–17. https://doi.org/10.1016/j.atg.2016.10.002

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265(5181):2037–2048

Lee M, Sharopova N, Beavis WD, Grant D, Katt M, Blair D, Hallauer A (2002) Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. Plant Mol Biol 48(5):453–461. https://doi.org/10.1023/a:1014893521186

Lehermeier C, Krämer N, Bauer E, Bauland C, Camisan C, Campo L, Flament P, Melchinger AE, Menz M, Meyer N, Moreau L, Moreno-González J, Ouzunova M, Pausch H, Ranc N, Schipprack W, Schönleben M, Walter H, Charcosset A, Schön C-C (2014) Usefulness of multiparental populations of maize (*Zea mays* L.) for genome-based prediction. Genetics 198(1):3–16. https://doi.org/10.1534/genetics.114.161943

Li C, Li Y, Bradbury PJ, Wu X, Shi Y, Song Y, Zhang D, Rodgers-Melnick E, Buckler ES, Zhang Z, Li Y, Wang T (2015) Construction of high-quality recombination maps with low-coverage genomic sequencing for joint linkage analysis in maize. BMC Biol 13 (1):78. https://doi.org/10.1186/s12915-015-0187-4

Li C, Sun B, Li Y, Liu C, Wu X, Zhang D, Shi Y, Song Y, Buckler ES, Zhang Z, Wang T, Li Y (2016a) Numerous genetic loci identified for drought tolerance in the maize nested association mapping populations. BMC Genomics 17(1):894. https://doi.org/10.1186/s12864-016-3170-8

Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N, Liu J, Warburton ML, Cheng Y, Hao X, Zhang P, Zhao J, Liu Y, Wang G, Li J, Yan J (2013) Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. Nat Genet 45(1):43–50. http://www.nature.com/ng/journal/v45/n1/abs/ng.2484.html#supplementary-information

Li L, Li H, Li Q, Yang X, Zheng D, Warburton M, Chai Y, Zhang P, Guo Y, Yan J, Li J (2011) An 11-bp Insertion in Zea mays fatb reduces the palmitic acid content of fatty acids in maize grain. PLoS ONE 6(9): e24699. https://doi.org/10.1371/journal.pone.0024699

Li Q, Yang X, Xu S, Cai Y, Zhang D, Han Y, Li L, Zhang Z, Gao S, Li J, Yan J (2012) Genome-wide association studies identified three independent polymorphisms associated with α-tocopherol content in maize kernels. PLOS ONE 7(5):e36807. https://doi.org/10.1371/journal.pone.0036807

Li YX, Li C, Bradbury PJ, Liu X, Lu F, Romay CM, Glaubitz JC, Wu X, Peng B, Shi Y, Song Y, Zhang D, Buckler ES, Zhang Z, Li Y, Wang T (2016b) Identification of genetic variants associated with maize flowering time using an extremely large multi-genetic background population. Plant J 86(5):391–402. https://doi.org/10.1111/tpj.13174

Liu H, Luo X, Niu L, Xiao Y, Chen L, Liu J, Wang X, Jin M, Li W, Zhang Q, Yan J (2017) Distant eQTLs and non-coding sequences play critical roles in regulating gene expression and quantitative trait variation in maize. Mol Plant 10(3):414–426. https://doi.org/10.1016/j.molp.2016.06.016

Lu F, Romay MC, Glaubitz JC, Bradbury PJ, Elshire RJ, Wang T, Li Y, Li Y, Semagn K, Zhang X, Hernandez AG, Mikel MA, Soifer I, Barad O, Buckler ES (2015) High-resolution genetic mapping of maize pan-genome sequence anchors. Nat Commun 6:6914. https://doi.org/10.1038/ncomms7914. https://www.nature.com/articles/ncomms7914#supplementary-information

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer

Mackay TFC (2001) The genetic architecture of quantitative traits. Annu Rev Genet 35(1):303–339. https://doi.org/10.1146/annurev.genet.35.102401.090633

Mahuku G, Chen J, Shrestha R, Narro LA, Guerrero KVO, Arcos AL, Xu Y (2016) Combined linkage and association mapping identifies a major QTL (qRtsc8-1), conferring tar spot complex resistance in maize. Theor Appl Genet 129(6):1217–1229. https://doi.org/10.1007/s00122-016-2698-y

Mammadov J, Sun X, Gao Y, Ochsenfeld C, Bakker E, Ren R, Flora J, Wang X, Kumpatla S, Meyer D, Thompson S (2015) Combining powers of linkage and association mapping for precise dissection of QTL controlling resistance to gray leaf spot disease in maize (Zea mays L.). BMC Genomics 16(1):916. https://doi.org/10.1186/s12864-015-2171-3

Manicacci D, Camus-Kulandaivelu L, Fourmann M, Arar C, Barrault S, Rousselet A, Feminias N, Consoli L, Francès L, Méchin V, Murigneux A, Prioul J-L, Charcosset A, Damerval C (2009) Epistatic Interactions between *Opaque2* transcriptional activator and its target gene *CyPPDK1* control kernel trait variation in maize. Plant Physiol 150(1):506–520. https://doi.org/10.1104/pp.108.131888

McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q (2009) Genetic properties of the maize nested association mapping population. Science 325. https://doi.org/10.1126/science.1174320

Members of the Complex Trait C (2003) The nature and identification of quantitative trait loci: a community's view. Nat Rev Genet 4:911. https://doi.org/10.1038/nrg1206. https://www.nature.com/articles/nrg1206#supplementary-information

Miller ND, Haase NJ, Lee J, Kaeppler SM, Leon N, Spalding EP (2017) A robust, high-throughput method for computing maize ear, cob, and kernel attributes automatically from images. Plant J 89(1):169–178. https://doi.org/10.1111/tpj.13320

Muttoni G, Johnson JM, Santoro N, Rhiner CJ, von Mogel KJH, Kaeppler SM, de Leon N (2012) A high-throughput core sampling device for the evaluation of maize stalk composition. Biotechnol Biofuels 5(1):27. https://doi.org/10.1186/1754-6834-5-27

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21(8):2194–2202. https://doi.org/10.1105/tpc.109.068437

Olukolu BA, Bian Y, De Vries B, Tracy WF, Wisser RJ, Holland JB, Balint-Kurti PJ (2016) The genetics of leaf flecking in maize and its relationship to plant

defense and disease resistance. Plant Physiol 172 (3):1787–1803. https://doi.org/10.1104/pp.15.01870

Olukolu BA, Negeri A, Dhawan R, Venkata BP, Sharma P, Garg A, Gachomo E, Marla S, Chu K, Hasan A, Ji J, Chintamanani S, Green J, Shyu C-R, Wisser R, Holland J, Johal G, Balint-Kurti P (2013) A connected set of genes associated with programmed cell death implicated in controlling the hypersensitive response in maize. Genetics 193(2):609–620. https://doi.org/10.1534/genetics.112.147595

Ordas B, Malvar RA, Santiago R, Sandoya G, Romay MC, Butron A (2009) Mapping of QTL for resistance to the Mediterranean corn borer attack using the intermated B73 × Mo17 (IBM) population of maize. Theor Appl Genet 119(8):1451–1459. https://doi.org/10.1007/s00122-009-1147-6

Pace J, Gardner C, Romay C, Ganapathysubramanian B, Lübberstedt T (2015) Genome-wide association analysis of seedling root development in maize (Zea mays L.). BMC Genomics 16(1):47. https://doi.org/10.1186/s12864-015-1226-9

Pan Q, Li L, Yang X, Tong H, Xu S, Li Z, Li W, Muehlbauer GJ, Li J, Yan J (2016) Genome-wide recombination dynamics are associated with phenotypic variation in maize. New Phytol 210(3):1083–1094. https://doi.org/10.1111/nph.13810

Peiffer JA, Romay MC, Gore MA, Flint-Garcia SA, Zhang Z, Millard MJ, Gardner CAC, McMullen MD, Holland JB, Bradbury PJ, Buckler ES (2014) The genetic architecture of maize height. Genetics. https://doi.org/10.1534/genetics.113.159152

Poland JA, Bradbury PJ, Buckler ES, Nelson RJ (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. Proc Natl Acad Sci 108(17):6893–6898. https://doi.org/10.1073/pnas.1010894108

Revilla P, Rodríguez VM, Ordás A, Rincent R, Charcosset A, Giauffret C, Melchinger AE, Schön C-C, Bauer E, Altmann T, Brunel D, Moreno-González J, Campo L, Ouzunova M, Álvarez Á, Ruíz de Galarreta JI, Laborde J, Malvar RA (2016) Association mapping for cold tolerance in two large maize inbred panels. BMC Plant Biol 16(1):127. https://doi.org/10.1186/s12870-016-0816-2

Riedelsheimer C, Czedik-Eysenberg A, Grieder C, Lisec J, Technow F, Sulpice R, Altmann T, Stitt M, Willmitzer L, Melchinger AE (2012a) Genomic and metabolic prediction of complex heterotic traits in hybrid maize. Nat Genet 44:217. https://doi.org/10.1038/ng.1033. https://www.nature.com/articles/ng.1033#supplementary-information

Riedelsheimer C, Lisec J, Czedik-Eysenberg A, Sulpice R, Flis A, Grieder C, Altmann T, Stitt M, Willmitzer L, Melchinger AE (2012b) Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. Proc Natl Acad Sci 109(23):8872–8877. https://doi.org/10.1073/pnas.1120813109

Rodríguez VM, Butrón A, Malvar RA, Ordás A, Revilla P (2008) Quantitative trait loci for cold tolerance in the maize IBM population. Int J Plant Sci 169(4):551–556. https://doi.org/10.1086/528755

Romay MC, Millard MJ, Glaubitz JC, Peiffer JA, Swarts KL, Casstevens TM (2013) Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol 14. https://doi.org/10.1186/gb-2013-14-6-r55

Samayoa LF, Malvar RA, Olukolu BA, Holland JB, Butrón A (2015) Genome-wide association study reveals a set of genes associated with resistance to the Mediterranean corn borer (Sesamia nonagrioides L.) in a maize diversity panel. BMC Plant Biol 15 (1):35. https://doi.org/10.1186/s12870-014-0403-3

Semagn K, Magorokosho C, Vivek BS, Makumbi D, Beyene Y, Mugo S, Prasanna BM, Warburton ML (2012) Molecular characterization of diverse CIMMYT maize inbred lines from eastern and southern Africa using single nucleotide polymorphic markers. BMC Genomics 13(1):113. https://doi.org/10.1186/1471-2164-13-113

Setter TL, Yan J, Warburton M, Ribaut J-M, Xu Y, Sawkins M, Buckler ES, Zhang Z, Gore MA (2011) Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. J Exp Bot 62(2):701–716. https://doi.org/10.1093/jxb/erq308

Song Q, Yan L, Quigley C, Jordan BD, Fickus E, Schroeder S, Song B-H, Charles An Y-Q, Hyten D, Nelson R, Rainey K, Beavis WD, Specht J, Diers B, Cregan P (2017) Genetic characterization of the soybean nested association mapping population. Plant Genome 10(2). https://doi.org/10.3835/plantgenome2016.10.0109

Studer AJ, Doebley JF (2011) Do large effect QTL fractionate? A case study at the maize domestication QTL teosinte branched1. Genetics 188(3):673–681. https://doi.org/10.1534/genetics.111.126508

Suwarno WB, Pixley KV, Palacios-Rojas N, Kaeppler SM, Babu R (2015) Genome-wide association analysis reveals new targets for carotenoid biofortification in maize. Theor Appl Genet 128(5):851–864. https://doi.org/10.1007/s00122-015-2475-3

Thavamanikumar S, Tibbits J, McManus L, Ades P, Stackpole D, Hadjigol S, Vaillancourt R, Zhu P, Bossinger G (2011) Candidate gene-based association mapping of growth and wood quality traits in Eucalyptus globulus Labill. BMC Proc 5(Suppl 7):O15–O15. https://doi.org/10.1186/1753-6561-5-S7-O15

Thirunavukkarasu N, Hossain F, Arora K, Sharma R, Shiriga K, Mittal S, Mohan S, Namratha PM, Dogga S, Rani TS, Katragadda S, Rathore A, Shah T, Mohapatra T, Gupta HS (2014) Functional mechanisms of drought tolerance in subtropical maize (Zea mays L.) identified using genome-wide association mapping. BMC Genomics 15(1):1182. https://doi.org/10.1186/1471-2164-15-1182

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler Iv ES (2001) Dwarf8 polymorphisms associate with variation in flowering time. Nat Genet 28:286. https://doi.org/10.1038/90135

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. Nat Genet 43. https://doi.org/10.1038/ng.746

Van Inghelandt D, Melchinger AE, Martinant J-P, Stich B (2012) Genome-wide association mapping of flowering time and northern corn leaf blight (Setosphaeria turcica) resistance in a vast commercial maize germplasm set. BMC Plant Biol 12(1):56. https://doi.org/10.1186/1471-2229-12-56

Veldboom LR, Lee M, Woodman WL (1994) Molecular marker-facilitated studies in an elite maize population: I. Linkage analysis and determination of QTL for morphological traits. Theor Appl Genet 88(1):7–16. https://doi.org/10.1007/bf00222387

Warburton ML, Williams WP, Windham GL, Murray SC, Xu W, Hawkins LK, Duran JF (2013) Phenotypic and genetic characterization of a maize association mapping panel developed for the identification of new sources of resistance to Aspergillus flavus and aflatoxin accumulation. Crop Sci 53:2374–2383. https://doi.org/10.2135/cropsci2012.10.0616

Warburton ML, Womack ED, Tang JD, Thrash A, Smith JS, Xu W, Murray SC, Williams WP (2018) Genome-wide association and metabolic pathway analysis of corn earworm resistance in maize. Plant Genome 11. https://doi.org/10.3835/plantgenome2017.08.0069

Wen W, Li D, Li X, Gao Y, Li W, Li H, Liu J, Liu H, Chen W, Luo J, Yan J (2014) Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. Nat Commun 5:3438. https://doi.org/10.1038/ncomms4438. https://www.nature.com/articles/ncomms4438#supplementary-information

Wu X, Li Y, Shi Y, Song Y, Zhang D, Li C, Buckler ES, Li Y, Zhang Z, Wang T (2016) Joint-linkage mapping and GWAS reveal extensive genetic loci that regulate male inflorescence size in maize. Plant Biotechnol J 14(7):1551–1562. https://doi.org/10.1111/pbi.12519

Xiao Y, Liu H, Wu L, Warburton M, Yan J (2017) Genome-wide association studies in maize: praise and stargaze. Mol Plant 10(3):359–374. https://doi.org/10.1016/j.molp.2016.12.008

Xiao Y, Tong H, Yang X, Xu S, Pan Q, Qiao F, Raihan MS, Luo Y, Liu H, Zhang X, Yang N, Wang X, Deng M, Jin M, Zhao L, Luo X, Zhou Y, Li X, Liu J, Zhan W, Liu N, Wang H, Chen G, Cai Y, Xu G, Wang W, Zheng D, Yan J (2016) Genome-wide dissection of the maize ear genetic architecture using multiple populations. New Phytol 210(3):1095–1106. https://doi.org/10.1111/nph.13814

Xue S, Bradbury PJ, Casstevens T, Holland JB (2016) Genetic architecture of domestication-related traits in maize. Genetics 204(1):99–113. https://doi.org/10.1534/genetics.116.191106

Xue Y, Warburton ML, Sawkins M, Zhang X, Setter T, Xu Y, Grudloyma P, Gethi J, Ribaut J-M, Li W, Zhang X, Zheng Y, Yan J (2013) Genome-wide association analysis for nine agronomic traits in maize under well-watered and water-stressed conditions. Theor Appl Genet 126(10):2587–2596. https://doi.org/10.1007/s00122-013-2158-x

Yan J, Kandianis CB, Harjes CE, Bai L, Kim E-H, Yang X, Skinner DJ, Fu Z, Mitchell S, Li Q, Fernandez MGS, Zaharieva M, Babu R, Fu Y, Palacios N, Li J, DellaPenna D, Brutnell T, Buckler ES, Warburton ML, Rocheford T (2010) Rare genetic variation at Zea mays crtRB1 increases [beta]-carotene in maize grain. Nat Genet 42(4):322–327. http://www.nature.com/ng/journal/v42/n4/suppinfo/ng.551_S1.html

Yang N, Lu Y, Yang X, Huang J, Zhou Y, Ali F, Wen W, Liu J, Li J, 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(9):e1004573. https://doi.org/10.1371/journal.pgen.1004573

Yang X, Gao S, Xu S, Zhang Z, Prasanna BM, Li L, Li J, Yan J (2011) Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. Mol Breed 28(4):511–526. https://doi.org/10.1007/s11032-010-9500-7

Yang X, Yan J, Shah T, Warburton ML, Li Q, Li L, Gao Y, Chai Y, Fu Z, Zhou Y, Xu S, Bai G, Meng Y, Zheng Y, Li J (2010) Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. Theor Appl Genet 121(3):417–431. https://doi.org/10.1007/s00122-010-1320-y

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178(1):539–551. https://doi.org/10.1534/genetics.107.074245

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Zhang N, Gibon Y, Gur A, Chen C, Lepak N, Höhne M, Zhang Z, Kroon D, Tschoep H, Stitt M, Buckler ES (2010a) Fine quantitative trait loci mapping of carbon and nitrogen metabolism enzyme activities and seedling biomass in the intermated maize IBM mapping population. Plant Physiol. https://doi.org/10.1104/pp.110.165787

Zhang N, Gur A, Gibon Y, Sulpice R, Flint-Garcia S, McMullen MD, Stitt M, Buckler ES (2010b) Genetic analysis of central carbon metabolism unveils an amino acid substitution that alters maize NAD-dependent isocitrate dehydrogenase activity. PLoS ONE 5(4):e9991. https://doi.org/10.1371/journal.pone.0009991

Zhang X, Warburton ML, Setter T, Liu H, Xue Y, Yang N, Yan J, Xiao Y (2016) Genome-wide association studies of drought-related metabolic changes in maize using an enlarged SNP panel. Theor

Appl Genet 129(8):1449–1463. https://doi.org/10.1007/s00122-016-2716-0

Zhang Z-S, Hu M-C, Zhang J, Liu D-J, Zheng J, Zhang K, Wang W, Wan Q (2009) Construction of a comprehensive PCR-based marker linkage map and QTL mapping for fiber quality traits in upland cotton (Gossypium hirsutum L.). Mol Breed 24(1):49–61. https://doi.org/10.1007/s11032-009-9271-1

Zila CT, Samayoa LF, Santiago R, Butrón A, Holland JB (2013) A Genome-Wide Association Study Reveals Genes Associated with Fusarium Ear Rot Resistance in a Maize Core Diversity Panel. G3 Genes Genomes Genet 3(11):2095–2104. https://doi.org/10.1534/g3.113.007328