Annual Review of Plant Biology

Natural Variation in Crops: Realized Understanding, Continuing Promise

Yameng Liang,[1,*] Hai-Jun Liu,[2,*] Jianbing Yan,[3] and Feng Tian[1]

[1]State Key Laboratory of Plant Physiology and Biochemistry, National Maize Improvement Center, Key Laboratory of Biology and Genetic Improvement of Maize (MOA), Beijing Key Laboratory of Crop Genetic Improvement, China Agricultural University, Beijing 100193, China; email: yml1992@cau.edu.cn, ft55@cau.edu.cn

[2]Gregor Mendel Institute, Austrian Academy of Sciences, Vienna BioCenter, 1030 Vienna, Austria; email: haijun.liu@gmi.oeaw.ac.at

[3]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China; email: yjianbing@mail.hzau.edu.cn

Annu. Rev. Plant Biol. 2021. 72:7.1–7.29

The *Annual Review of Plant Biology* is online at plant.annualreviews.org

https://doi.org/10.1146/annurev-arplant-080720-090632

*These authors contributed equally to this article

Keywords

natural alleles, quantitative trait locus, quantitative trait nucleotide, regulatory and evolutionary mechanisms, genome editing, breeding by design

Abstract

Crops feed the world's population and shape human civilization. The improvement of crop productivity has been ongoing for almost 10,000 years and has evolved from an experience-based to a knowledge-driven practice over the past three decades. Natural alleles and their reshuffling are long-standing genetic changes that affect how crops respond to various environmental conditions and agricultural practices. Decoding the genetic basis of natural variation is central to understanding crop evolution and, in turn, improving crop breeding. Here, we review current advances in the approaches used to map the causal alleles of natural variation, provide refined insights into the genetics and evolution of natural variation, and outline how this knowledge promises to drive the development of sustainable agriculture under the dome of emerging technologies.

Contents

1. INTRODUCTION: THE GREAT VALUE OF NATURAL VARIATION 7.2
2. DISSECTING THE GENETIC BASIS OF NATURAL VARIATION 7.3
2.1. Linkage Analysis 7.3
2.2. Genome-Wide Association Study 7.3
2.3. Joint Linkage Association Mapping 7.5
2.4. From Quantitative Trait Loci to Quantitative Trait Nucleotides 7.6
3. QUANTITATIVE TRAIT LOCUS CLONING PROGRESS IN CROPS 7.7
3.1. Critical Quantitative Trait Locus Alleles for Crop Domestication 7.7
3.2. Critical Quantitative Trait Locus Alleles for Crop Adaptation 7.10
3.3. Critical Quantitative Trait Locus Alleles for Agronomic Traits 7.10
3.4. Critical Quantitative Trait Locus Alleles for Stress Responses 7.11
3.5. Critical Quantitative Trait Locus Alleles for Fertility 7.12
4. THE REGULATORY AND EVOLUTIONARY MECHANISMS OF CROP QUANTITATIVE TRAIT LOCI 7.12
4.1. The Molecular Functions of Identified Causal Genes 7.12
4.2. The Molecular Regulatory Mechanisms of Quantitative Trait Loci 7.13
4.3. The Evolutionary Mechanisms of Quantitative Trait Loci 7.14
4.4. The Mechanisms of Pleiotropic Quantitative Trait Loci 7.16
5. IDENTIFYING AND CREATING ALLELES FOR CROP BREEDING 7.16
5.1. Employing Favorable Natural Alleles 7.16
5.2. Creating Novel Alleles Beyond Nature 7.18
5.3. Engineering New Traits 7.19
6. CONCLUSIONS AND PERSPECTIVE 7.19

1. INTRODUCTION: THE GREAT VALUE OF NATURAL VARIATION

Feeding an ever-increasing global population poses a grand challenge in light of the declining availability of cropland under changing climate conditions. One of the best solutions to this problem is to continuously and sustainably improve crop productivity. The genetic improvement of crops involves selecting and combining different favorable traits, which are usually controlled by many quantitative trait loci (QTLs). The accumulated knowledge of underlying genetic causes of agronomically important traits increases the predictability of customized breeding and enables the de novo design of new traits and crops. This knowledge-guided approach to crop improvement is thus superior in both effectiveness and actionability compared to conventional breeding based solely on phenotypic variation.

In the foreseeable future, mapping the QTLs and (ultimately) the causal quantitative trait nucleotides (QTNs) underlying complex traits will be a cornerstone for breeding design. Natural populations provide unparalleled experimental designs built up over thousands of years, as these populations simultaneously contain segregating variants at millions of loci and hundreds of allelic replications at each locus. These natural changes are ideal not only for discovering genetic causation but also for improving mechanistic insights into genotype-phenotype connections. During the past decade, the collection and development of diverse crop populations, dramatic innovations in genomic technologies, emerging quantitative genetics methods, and increasing availability of molecular biology tools have enabled the genetic basis of natural variation to be decoded in an unprecedented manner.

In this review, we summarize the progress, problems, and potential improvements in QTL-to-QTN mapping in crops, using rice, maize, wheat, tomato, soybean, and barley as examples. The first crop QTL—*Pto*, which confers disease resistance in tomato—was identified in 1993 (108). We surveyed the literature and found that 364 QTLs have been cloned (that is, the causal genes have been identified and validated) since then in the six crops, as of May 2020 (**Figure 1*a***; **Supplemental Table 1**). We use this representative data set to provide a general picture of the regulation and evolution of QTLs in crops as well as their similarities and differences within and between species. We end with a discussion of how this knowledge can be exploited for traditional and novel approaches to crop improvement.

Pseudo-overdominance: seeming overdominance caused by two dominant loci that are closely linked in repulsion

Repulsion phase: for two genes in repulsion phase, each homologous chromosome contains one dominant and one recessive allele (Ab and aB)

2. DISSECTING THE GENETIC BASIS OF NATURAL VARIATION

2.1. Linkage Analysis

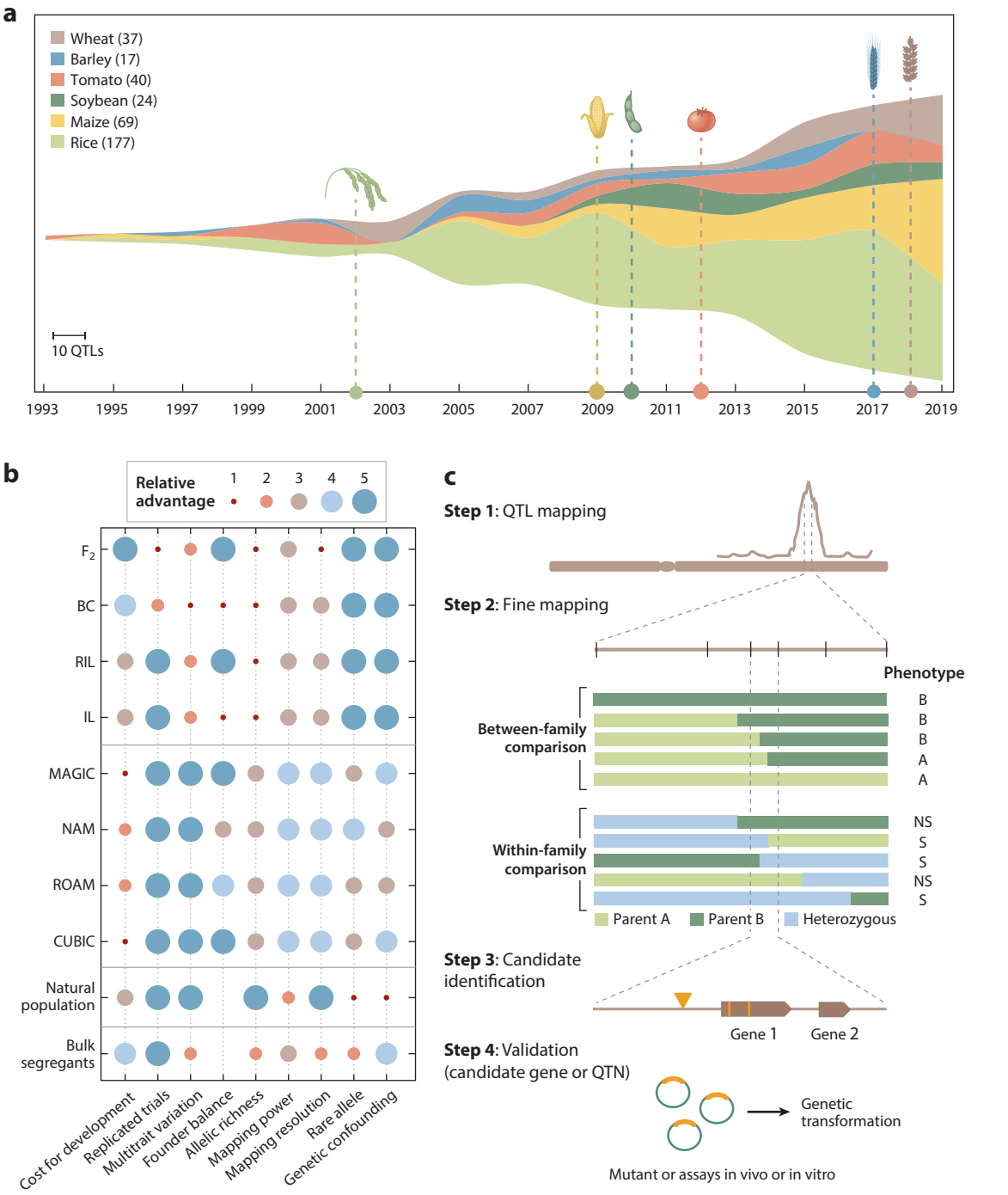
Linkage mapping using artificially created segregating populations has been the most successful method used to dissect the genetic basis of crop traits. Unlike humans and animals, crops can be studied using diverse genetic designs with distinct properties (143, 188) (**Figure 1*b***), and these genetic populations have been exploited to identify thousands of QTLs for hundreds of agronomic traits. Among the types of linkage populations, recombinant inbred lines are the most popular because of their simple development, balanced parental mixture, repeated phenotyping, and relatively high mapping power. Introgression lines, together with other interspecies advanced backcross populations, are also widely used to study domestication-related traits and to quickly initiate QTL fine mapping. In a pioneering study, an introgression line population was successfully used to map the first yield-related QTL in tomato: *fw2.2* (28). A large maize-teosinte BC_2S_3 population was recently developed and used to successfully clone QTLs for a variety of traits important for domestication and adaptation (42, 48, 49, 56, 88, 160, 184).

Although linkage mapping is a straightforward process, the results obtained using this approach should be carefully interpreted. While it is well known that linkage analysis underestimates QTL number, and that QTL effects are biased in different backgrounds and/or environments, the effect size of a QTL can also be overestimated in small (<500) populations (the Beavis effect) (192). In addition, the results could be completely misleading due to complicated modes of inheritance (e.g., pseudo-overdominance resulting from two closely linked loci in repulsion phase) (86).

Given that linkage mapping is based on the occurrence of recombination events between genetic loci, existing populations could be reused to maximize efficiency and value. For instance, the widespread residual heterozygosity among advanced inbred lines can provide a quick starting point for fine-mapping target QTLs (163). Recent advances could partially remedy the limited mapping resolution of linkage analysis, such as (*a*) genotyping larger populations for recombinant screening, (*b*) speed breeding techniques (57, 182) to promote generations of recombination at the same period, and (*c*) increasing the recombination frequency by using optimized alternative parents (104) or manipulating potential targets (such as *RecQ-7*) (37). However, the low allelic diversity of biparental populations remains an impediment for identifying many other causal genes.

2.2. Genome-Wide Association Study

Genome-wide association study (GWAS) using natural populations (preferably collections of homozygous lines) has become a powerful and routine approach for dissecting trait variation in crops. Compared with linkage populations, natural populations harbor numerous mutations and abundant historical recombinations and are cost-effective for population development, genotyping, and repeated phenotyping (116) (**Figure 1*b***). The unique properties of natural populations provide



(Caption appears on following page)

Figure 1 (*Figure appears on preceding page*)

A timeline of the progress of QTL cloning, comparison of mapping populations, and schematic diagram of QTN identification in crops. (*a*) The number of QTLs cloned from 1993 to May 2020 in rice, maize, wheat, tomato, soybean, and barley. The total number of cloned QTLs are indicated in parentheses after each species. The years in which the crop reference genomes were released are indicated by dashed lines. (*b*) A comparison of trade-offs when developing and implementing different genetic designs. The larger circles represent a higher relative advantage, not the extent, of a given criterion. For example, the biparental populations are generally less confounded, which is more favorable for genetic mapping. Natural population here represents the inbreds for a collection of natural, unrelated individuals. (*c*) A simplified procedure for QTL-to-QTN mapping. A and B represent that the phenotype of recombinant progeny is similar to corresponding parents. Abbreviations: CUBIC, complete-diallel design plus unbalanced breeding-like inter-cross; MAGIC, multiparent advanced generation inter-cross; NAM, nested association mapping; NS, nonsegregating; QTL, quantitative trait locus; QTN, quantitative trait nucleotide; ROAM, random-open-parent association mapping; S, segregating.

GWASs with higher mapping resolution, which allows multiple functional alleles at a given locus to be surveyed. With dramatic increases in the power of sequencing technologies and statistical methods, GWAS has been revolutionized over the past decade. Its applications have shifted from providing validation to identifying new loci with common alleles or even small effects (100, 106, 177, 188), from exploring traditional agronomic traits to omics-based molecular phenotypes (51, 188), from trait-targeted GWASs to genome–phenome-wide association study (89), and so on.

Mixed models are commonly utilized in GWASs of crops. These models incorporate both population structure and familial relatedness to control genetic confounding from stratified and cryptic genetic backgrounds (165). These corrections allow valid associations to be distinguished from spurious ones; however, true marker-phenotype associations could be missed during the analysis, especially for traits correlated with subpopulation differentiation or local adaptation (e.g., flowering time). Another major source of false negatives is the lack of detection power of substantially lower-frequency and rare functional variants, although several statistical strategies have been proposed to address this issue (5, 138).

Additionally, the lack of high-quality nonreference genomes and the presence of large structural variation may account for a great portion of causality—the hidden genetic landscapes of traits. For example, de novo assembly of a tropical maize inbred line led to the identification of *ZmBAM1d*, a gene responsible for kernel weight and whose expression is regulated by large upstream insertions affecting chromatin interactions and methylation (199). The large number of transposable elements and the widespread presence of duplications in crops can lead to greater regulatory complexity as well as missing phenotypic contributions, yet these effects are largely underestimated. Fragmented pan-genomes assembled from short resequencing reads (62, 158) and k-mer-based GWAS without assembly (4, 166) can be used to explore hidden genetic causality genome wide. Furthermore, the upcoming era of rapid reference-quality assemblies at the population level combined with ever-evolving long-read sequencing technologies (2, 100) would bring more progress.

Despite the improved mapping resolution, identifying the causal genes for association signals can be challenging in some cases. Particularly, the associations might be located far from functional genes because of complex linkage disequilibrium architecture caused by allelic heterogeneity and the presence of distal noncoding regulatory elements (11, 96). Several comprehensive methods have been proposed to pinpoint causality, especially by incorporating information from genomic functional elements and transcriptomic and/or proteomic variation (139–141). Importantly, recently developed high-throughput genome-editing techniques provide a rapid, large-scale method for identifying and validating gene function (93).

Allelic heterogeneity: different mutations at a given locus independently resulting in the same (or very similar) phenotype(s)

2.3. Joint Linkage Association Mapping

The presence of rare alleles and genetic confounding are two major limitations of GWASs in natural populations, but both problems can be circumvented to some extent using controlled

segregating populations. Using a multiple-founder-derived genetic design, which integrates the complementary strengths of association panels and traditional biparental populations, is a promising approach for enhancing both allelic richness and mapping power in plants (**Figure 1*b***). Various multiparent genetic designs have been proposed, including nested association mapping (NAM) (110, 202), MAGIC multiparent advanced generation inter-cross (MAGIC) (47), random-open-parent association mapping (ROAM) (124, 189), and the recently described complete-diallel design plus unbalanced breeding-like inter-cross (CUBIC) (95).

These designs have greatly facilitated the genetic dissection of complex traits. For instance, flowering time is a highly complex trait that is strongly correlated with population structure (76). Therefore, it is often difficult to identify statistically significant associations with flowering time in a natural population when including population structure as a covariate in statistical model. Indeed, in a GWAS using a diverse set of maize inbred lines, only one locus (*ZmCCT10*) was identified that was associated with photoperiod response with genome-wide significance (201), and several candidate loci validated by other pieces of evidence only displayed moderate associations with this trait (61). By contrast, many small-effect flowering-time QTLs were identified in NAM (10, 56) and CUBIC populations (95). Beyond identifying many rare alleles and small-effect loci (10, 159), multiparent designs are also powerful for discovering QTLs with multiple alleles (95) or allelic series with opposite effects (189).

In theory, joint linkage association mapping using these multiparent designs offers a way to increase mapping resolution, given the simultaneous employment of both historical and recent recombination. For example, in the maize ROAM population consisting of 10 recombinant inbred line populations derived from 14 founders, there were ∼14,600 genetic bins in the linkage map and over 185,000 segments in the linkage disequilibrium map (189). However, due to the limited number of founder lines, outcrossing events, and progeny sample size, the mapping resolution of these multiparent populations is still not sufficient in practice. Given the low sequencing costs, a larger number of parents (e.g., more than 100) could be recruited to achieve not only a higher mapping resolution but also allelic richness, providing greater representation of rare alleles.

So far, it seems as though no specific design is the best, but in the near future, a pyramidal assembly could be established for most crops. Each pyramid would include one large collection of thousands of individuals consisting of wild, landrace, and cultivated lines, several diverse large multiparent designs, and many specific biparental populations. Community-wide efforts are required to realize this goal, with data sets of all types of variation processed under the same guidelines and made publicly available. Such community-assembled data sets would set the standard for mapping directly to individual genes and validating the associated effects.

2.4. From Quantitative Trait Loci to Quantitative Trait Nucleotides

Various mapping populations not only provide great power to detect QTLs but also lay the foundation for identifying QTNs. Positional cloning (or map-based cloning) is the most widely used method for QTL cloning. Among the 364 cloned QTLs we gleaned from literature, 297 (81.6%) were cloned by positional cloning alone or in combination with other methods (**Supplemental Table 1**). Mendelizing the QTL of interest from background loci is required for QTL cloning, as most QTL effects are moderate or even minor (106). Two major strategies are generally used to develop QTL near-isogenic lines. One strategy involves consecutive backcrossing with a marker-assisted selection of the target and background region (157). The other strategy involves using a heterogeneous inbred family (163) that is heterozygous only at the target QTL but fixed in background loci. Another key aspect of QTL cloning is to precisely phenotype the progeny of recombinants, which can be conducted within or between recombinant families (3, 88, 149, 150)

(**Figure 1*c***). Which strategy is more appropriate depends on the effect size of the target QTL. In general, the between-family comparison is suitable for moderate- to large-effect QTLs (3, 150), whereas within-family comparison is especially helpful for minor QTLs, as it can effectively control the interference from the genetic background and/or the environment (49, 88).

When a QTL is delimited to a sufficiently small region by fine-mapping or narrowed down to an acceptable resolution by GWASs, candidate gene-based association could be used to quickly identify causative variation (42, 49, 135, 178, 201). Nevertheless, like GWASs, candidate gene association analysis also has limited power to detect rare alleles. Bioinformatics analysis of potential functional elements in the causal region together with various in vivo or in vitro assays of critical sequences help identify the causative polymorphism underlying the target QTL (6, 55, 149, 160).

With the rapid development and continuously decreasing cost of sequencing technology, whole-genome sequencing pools of individuals with extreme phenotypes together with mapping-by-sequencing approaches, including GradedPool-Seq (168), QTL-seq (153), QTG-seq (204), and BSR-Seq (98), provide a quick and efficient way to identify genetic causes controlling trait differences. However, these bulked segregate analyses still have limited power to dissect minor-effect QTLs, and their mapping resolutions depend largely on pooling strategy and the size of each pool.

The last critical step in QTL cloning is to validate the causality of both the candidates and alleles (**Figure 1*c***). Common approaches used to verify the functions of candidate genes include mutant screening, genetic complementation testing, overexpression, and gene knockdown or knockout by RNA interference and other various genome-editing technologies. Validating the function of a QTN in vivo remains challenging if the QTN resides in a noncoding region, especially if it is distant from its functional gene. Nevertheless, CRISPR/Cas9 provides a promising way to validate a QTN in noncoding regions (29).

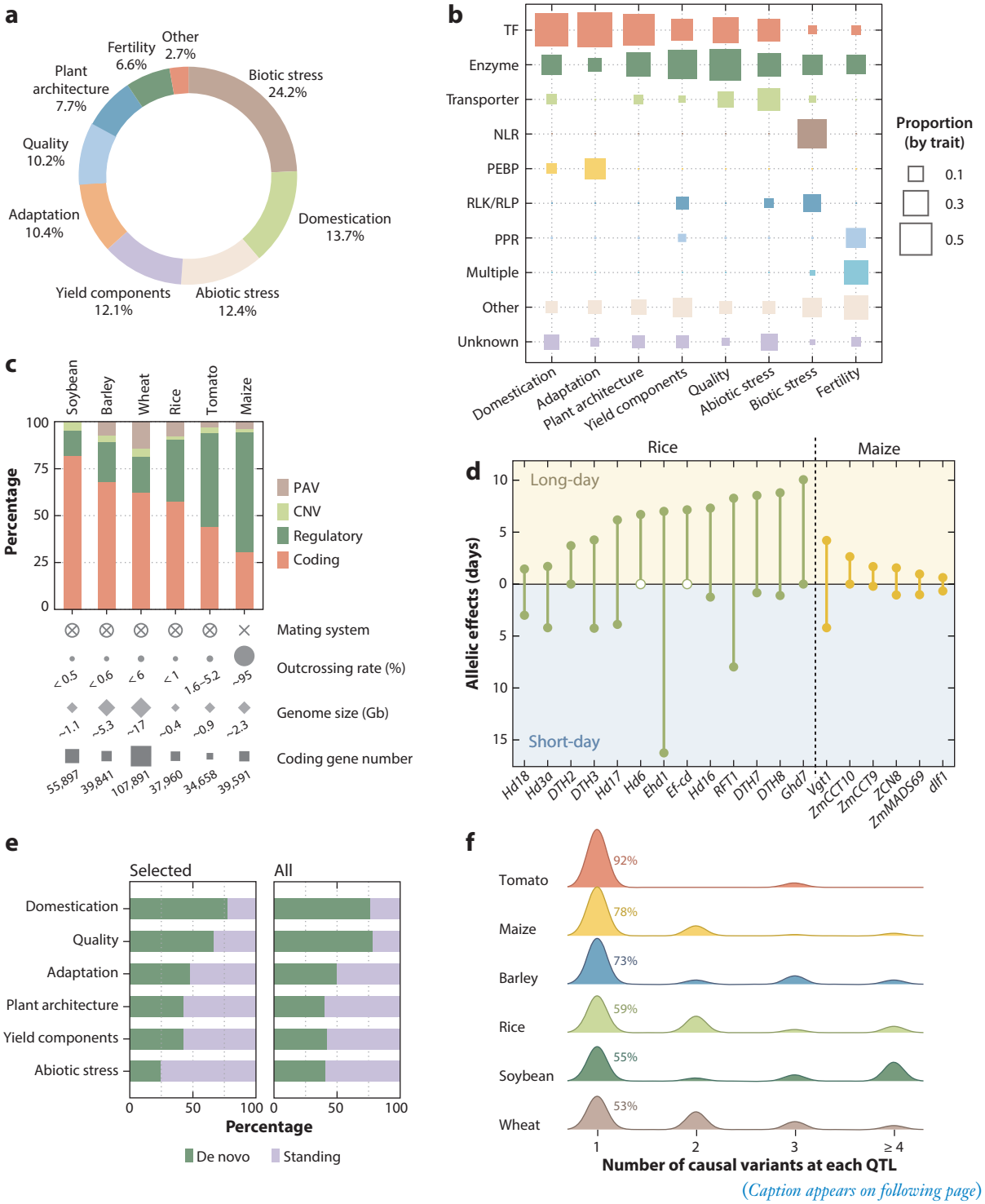
New strategies promise to revolutionize the fine-mapping process by making it more effective or even less costly, including assembling founder genomes using long-read sequencing technologies and genotyping segregating descendants via short-read sequencing, combined with additional omics measurements. The validated QTN not only provides an entry point for uncovering the regulatory mechanism but also serves as an ideal functional marker for breeding.

3. QUANTITATIVE TRAIT LOCUS CLONING PROGRESS IN CROPS

Natural variation contributes greatly to heritable trait variation that responds to long-term natural and artificial selection. Therefore, identifying causal genotype-phenotype links will increase our understanding of plant growth and development as well as crop domestication and adaptation. A total of 364 QTLs have been cloned in the six crops from 1993 to May 2020, with most—nearly half of them—in rice, followed by maize (**Figure 1*a***). The release of crop reference genome sequences and technological advances in genotyping and genetic transformation have significantly accelerated the cloning of QTLs. Characterizing this set of cloned QTLs provides us with an unprecedented opportunity to understand the genetic bases and molecular mechanisms of natural variation and their roles in crop evolution.

3.1. Critical Quantitative Trait Locus Alleles for Crop Domestication

Within the collection of cloned QTLs, 50 control domestication traits in the six crops (**Figure 2*a***). During domestication, crops undergo a common suite of trait changes that distinguish them from their progenitors. This suite of changes is known as the domestication syndrome and includes decreased seed shattering, loss of seed dormancy, larger seed size or increased fruit



(Caption appears on following page)

Figure 2 (*Figure appears on preceding page*)

Insights obtained from QTL cloning in six crops (maize, tomato, rice, wheat, barley, and soybean). (*a*) Currently cloned QTL across trait categories. (*b*) Comparison of functional types of causal genes classified based on trait categories. Multiple indicates that more than one gene is functioning coordinately in the given locus, and unknown indicates that it is without clear annotation. (*c*) Proportional differences of underlying polymorphisms among crops. The six crops were sorted based on decreasing (increasing) ratios of coding (regulatory) variants. Four characteristics associated with the six crops are indicated below the graph. × symbols indicate outcrossing, while circled × symbols indicate selfing. The estimated genome size and gene number statistics are from Ensembl Plants (**http://plants.ensembl.org/index.html**), and the outcrossing rate data of domesticated crops are from References 1, 17, 26, 78, 111, and 131. (*d*) Allelic effects (measured in days) of cloned flowering-time QTLs in rice and maize under long- and short-day conditions. The flowering-time effects of QTL NILs are obtained from published studies, and the mean QTL allelic effects across different studies are calculated. The white-filled circles represent information not available. (*e*) The relative importance of de novo mutation versus standing variation for different traits. (*Left*) Causal genes are detected with selection signals; (*right*) all causal genes are shown for cloned QTL. The categories of biotic stress and fertility are not presented because fewer than three QTLs are available in the selected set. Please note that the ratios of de novo mutation might be (largely) overrated due to the usually underexamined wild population in the original study. (*f*) The distribution of the number of causal variants identified at each QTL across crops. Abbreviations: CNV, copy number variation; NIL, near-isogenic line; NLR, nucleotide-binding leucine-rich repeat receptor; PAV, presence/absence variation; PEBP, phosphatidylethanolamine-binding protein; PPR, pentatricopeptide repeat; QTL, quantitative trait locus; RLK/RLP, receptor-like kinase or protein; TF, transcription factor.

weight, and more determinate growth or increased apical dominance (43, 45). It appears that crop domestication syndrome traits are controlled by both conserved and species-specific genes.

Nonshattering is believed to be the most critical event during the initial domestication of a crop. Seed shattering in cereals appears to depend largely on the development of an abscission layer, as the selected alleles of 9 of the 12 cloned shattering QTLs disrupt normal abscission layer development. Conserved genes controlling shattering include *Sh1* in sorghum, rice, and maize (91) and *Btr1* in barley and wheat (129, 208). Species-specific genes have also been identified, including *sh4* (80) and *qSH1* (69) in rice and *Btr2* (129) in barley. The genetic basis of soybean pod shattering tends to be different from that of cereals. *SHAT1-5* and *Pdh1* control soybean pod shattering by regulating the lignification of fiber cap cells and the torsion of dried pod walls, respectively (22, 33).

The loss of seed dormancy is another critical event in crop domestication. Eight QTLs controlling this trait have been cloned. *G* is a conserved gene controlling seed dormancy in soybean, rice, and tomato (175). The *MKK3* (*mitogen-activated protein kinase kinase 3*) underlying the *SD2* locus in barley (118) and the *Phs1* locus in wheat also play conserved roles in controlling seed dormancy (162). Rice *Sdr4* (150), barley *SD1* (137), and wheat *PM19* (7) and *TaMFT* (117) function in a species-specific manner. In addition, reduced seed dormancy in barley and wheat is often associated with preharvest sprouting (7, 117, 118, 137, 162). Therefore, identifying and pyramiding natural alleles with optimum effects represents an efficient way to balance seed dormancy and preharvest sprouting.

Another striking change during domestication is the increased apical dominance (43, 45). *tb1*, the first domestication QTL cloned in crops, controls apical dominance in maize (20, 149) and functions at the top of the hierarchical maize domestication gene network, where it directly regulates a suite of domestication genes, including *gt1*, *tga1*, *tru1*, and *Zagl1* (23, 148). *tb1* homologs in other crops also regulate tillering or branching (18, 130, 155). The major regulator controlling apical dominance in rice is *PROG1*, and its loss-of-function allele promotes the transition from prostrate to erect growth during rice domestication (60, 156).

fw2.2 (31), *fw3.2* (12), and *fw11.3* (115) are three QTLs regulating fruit weight during tomato domestication, and a recent study suggested that they are also involved in tomato improvement (90). *fw2.2* encodes a cell number regulator whose ectopic expression during early fruit development increased pericarp cell number (31). Its ortholog in rice, *OsCNR1*, also confers grain weight (133).

In addition to domestication syndrome traits, each crop has its specific domestication characteristics. For example, the transition from stony fruitcase–enveloped to naked kernel is a critical event during maize domestication but not for rice and wheat. Mining the underlying causal factors of common and specific domestication traits is important not only for understanding the past domestication history but also for de novo domestication of new crops.

3.2. Critical Quantitative Trait Locus Alleles for Crop Adaptation

After their initial domestication, crops spread from their centers of origin to diverse ecological and geographical areas. Flowering time is a major determinant of plant local adaptation. The set of cloned QTLs includes 38 for flowering time (**Figure 2*a***). Most of the favorable alleles are associated with attenuated photoperiod sensitivity.

Rice and maize originated in low latitudes (52, 109). Most of the variation contributing to their spread from low to high latitudes are loss-of-function or weakened alleles of long-day suppressors such as *Hd6* (154), *Ghd7* (193), *Ghd8/DTH8* (16, 183, 197), and *DTH7* (35, 99) in rice, as well as *ZmCCT10* (201) and *ZmCCT9* (49) in maize, or enhanced flowering activators under both long days and short days such as *Hd3a* (66), *RFT1* (120, 211), and *Ehd1* (21) in rice and *ZCN8* (42) and *ZmMADS69* (88) in maize.

Unlike maize and rice, soybean was domesticated from temperate regions in China between 32°N and 40°N (87). Therefore, soybean dispersion involved adaptation to both lower and higher latitudes. *J* is a soybean flowering activator under short days that downregulates *E1*, a legume-specific flowering suppressor (102). Loss-of-function alleles of *J* were selected to extend the vegetative phase and improve yields at lower latitudes (102). The adaptation of soybean to higher latitudes was helped by natural loss-of-function alleles of five other flowering suppressors, including *E1* (187), *E2* (181), *E3* (180), and the paralogs *Tof11/Gp11* and *Tof12/Gp12* (40, 82, 101).

Photoperiod and vernalization are important factors affecting flowering time for the long-day plants, wheat and barley. *VRN1* (196), *VRN2* (195), and *VRN3* (194) in wheat and *Ppd-1* (164), *EPS2/HvCEN* (14), and *VRN-H3* (194) in barley function in response to photoperiod and vernalization. Loss-of-function alleles of *VRN2* exhibit increased expression of *VRN1* and *VRN3*, thereby converting wheat from a winter to a spring growth habit (194, 195).

Although different alleles conferring prolonged or shortened flowering time were selected to help crops respond to diverse environments, the underlying genes are quite conserved, such as florigen and MADS-box family. Therefore, critical flowering-time genes identified in one crop are very likely to function in other species as well.

3.3. Critical Quantitative Trait Locus Alleles for Agronomic Traits

Improved yield and quality are two key traits that humans have long pursued, while plant architecture is another important trait affecting planting mode and yield potential. These agronomic traits have been studied extensively, accounting for 30% of the cloned QTLs (**Figure 2*a***). Loss-of-function of the Green Revolution gene *sd1* in rice and gain-of-function of *Rht1* in wheat led to the first breakthrough in increasing grain yields in the 1960s by reducing plant height and improving lodging resistance (128, 136). The proposal for new plant type breeding (64) lays the foundation for another advancement in rice yields. *IPA1* encodes an SBP-like transcription factor (TF). Mutations in the miR156 target site and tandem repeats in the upstream region of *IPA1* resulted in its higher expression, thereby conferring reduced tiller number, more grains per panicle, and thicker culms (59, 113, 205).

Different from other cereal crops, maize is an outcrossing species with separate male (tassel) and female (ear) inflorescences. Maize yield is mainly determined by kernel row number, kernel number per row, and kernel weight. *KNR6* controls kernel number per row (58), while *FEA2* (9), *krn1* (170), and *KRN4* (97) regulate kernel row number, and *qHKW1* (199) and *qKW9* (50) affect kernel weight. Increased planting densities play a critical role in boosting maize yield, and high-density planting requires an upright plant architecture. *UPA1* and *UPA2* control leaf angle; introgressing the wild allele of *UPA2* into modern hybrids produced narrower plants and enhanced maize grain yields under high planting densities (68, 160). Pyramiding the favorable alleles of yield components and plant architecture represents an important strategy for enhancing maize yield potential.

Beyond yield, grain quality has attracted increasing attention, and many studies have identified a suite of alleles that improve the appearance, cooking quality, and nutritional quality of crops (**Supplemental Table 1**). The quality traits of tomato are quite diverse, including fruit shape, fruit color, sugar-acid ratio, vitamin content, and the levels of important metabolites. One-third of the QTLs cloned in tomato are associated with quality traits. The major QTL *Brix9-2-5* encodes a flower- and fruit-specific invertase (32). Introgression of the wild allele of *Brix9-2-5* improved the soluble sugar content of cultivated tomato fruit (32), again confirming that alleles from wild germplasms are valuable for modern breeding.

3.4. Critical Quantitative Trait Locus Alleles for Stress Responses

Biotic and abiotic stresses severely reduce crop yields, and improving stress resistance could lead to more stable yields. More than one-third of the cloned QTLs (133 of 364) are related to stress responses, including 45 for abiotic and 88 for biotic stress responses (**Figure 2*a***). Different crops face similar abiotic stresses, including high or low temperatures, drought, salt, and nutrient limitation. Various alleles for resistance to these widespread stresses have been identified (112). For example, the submergence-induced *Sub1A-1* improves rice tolerance to submergence (191), and a 366-bp insertion in the *ZmVPP1* promoter confers drought-inducible expression and enhances drought resistance in maize (178).

By contrast, most biotic stresses, including diseases and pests, are crop specific. Bacterial blight, fungus blast, and brown planthopper are major biotic stresses in rice. Head smut, stalk root, rough dwarf, and (southern, northern, and gray) leaf blight pose grave threats to maize yields, while leaf rust, stripe rust, yellow rust, powdery mildew, and head blight are the most devastating diseases to wheat. Among the cloned QTLs, nucleotide-binding and leucine-rich repeat receptors (NLRs) and receptor-like kinases/receptor-like proteins (RLKs/RLPs) contribute greatly to biotic stress resistance in crops. The specific molecular mechanisms for biotic resistance have been summarized in previous reviews (24, 85, 206).

As pathogens often rapidly adapt to plant resistance genes, creating crops with broad, durable resistance alleles is therefore an important objective of crop breeding. *Bph6* encodes an uncharacterized protein whose amino acid changes provide rice with broad resistance to both the white-backed and brown planthopper without sacrificing yield (41). *ZmCCoAOMT2*, a caffeoyl-CoA *O*-methyltransferase, simultaneously confers resistance to southern leaf blight, northern leaf blight, and gray leaf spot in maize (200). Polymorphisms in the coding region of the putative ABC transporter gene *Lr34* provide wheat with durable resistance to leaf rust, stripe rust, and powdery mildew (71). The horizontal transfer of the glutathione *S*-transferase gene *Fhb7* from fungus to wheat results in broad resistance to head blight and crown rot without yield penalty (169).

Strong resistance often comes with yield costs. How yield and immunity can be balanced has been reviewed previously (119, 171). Plant hormones, TFs, and microRNAs play important roles

in regulating the trade-off between plant immunity and growth (119, 171). Some NLRs and cell wall-associated kinase (WAK) proteins could balance immunity and yields, making them excellent targets for breeding crop varieties with high yields and strong resistance to pathogens (119).

3.5. Critical Quantitative Trait Locus Alleles for Fertility

In addition to the Green Revolution, hybrid breeding represents another breakthrough in crop improvement. The two-line (thermo-sensitive or photoperiod-sensitive genic male sterility) and three-line [cytoplasmic male sterility (CMS)] hybrid breeding systems were developed in crops, especially in rice, to make better use of heterosis. The most widely used system, three-line hybrid breeding, includes three key elements: male sterility, fertility restoration, and hybrid compatibility. As most CMS male sterility genes of the CMS system are in the mitochondrial genome, QTL mapping is not appropriate for detecting these genes. The 24 fertility QTLs cloned to date are mainly involved in fertility restoration and hybrid compatibility.

Rf2, a maize gene encoding an aldehyde dehydrogenase, was the first restorer gene isolated (15). It can restore the fertility of CMS-T maize and has been used to produce approximately 85% of hybrid seeds in the United States prior to the southern corn leaf blight epidemic in 1970 (15). The selfish genetic element *qHSM7* controls rice hybrid compatibility, and two closely linked genes, *ORF2* and *ORF3*, underlie *qHSM7* (203). *ORF2* encodes a toxic genetic element, whereas *ORF3* encodes the antidote to protect pollen from being eliminated by *ORF2* (203). Additional cases and the mechanisms underlying fertility were reviewed previously (13, 122). In general, the functional genes of most fertility restoration QTLs encode pentatricopeptide repeat (PPR) proteins, and more than one closely linked gene usually underlies a hybrid compatibility QTL with segregation distortion.

4. THE REGULATORY AND EVOLUTIONARY MECHANISMS OF CROP QUANTITATIVE TRAIT LOCI

Analyzing the representative set of QTLs provides an excellent opportunity to refine our understanding of how natural variation is regulated and has evolved. It should be noted that the QTLs cloned to date might not reflect a complete polygenic architecture and are inherently biased towards large-effect QTLs and ascertainment biases could also be present among crops and trait categories. Therefore, the conclusions obtained from the current observations might be biased. However, the following overview of this comparative data set could be still valuable to provide a deeper understanding of the biological relevance of QTLs and fresh insights for future studies and breeding programs.

4.1. The Molecular Functions of Identified Causal Genes

According to the gene functions, the causal genes from the 364 cloned QTLs can be classified into ten types (**Figure 2*b***). The relative contributions of different gene types to each trait category vary greatly. Domestication, adaptation, and plant architecture traits are mainly regulated by TFs, while yield components and quality traits are primarily regulated by enzymes. This notion is consistent with previous views (19). The likely reason for this difference might be that TFs usually function at the top of the hierarchy of gene networks, and changes in TFs can have comprehensive effects.

The molecular basis of biotic stress resistance is different from that of abiotic stress resistance (24, 85, 112, 206). NLR genes (42%) predominately contribute to pathogen or insect resistance, followed by enzymes (19.3%) and RLK/RLPs (13.6%). However, abiotic stress is often regulated

by enzymes (26.7%), TFs (24.4%), and transporters (24.4%). Difference also exists among different abiotic stress responses. For instance, salt tolerance is typically regulated by transporters, while drought tolerance is mainly controlled by TFs and enzymes.

Interacting pairs or multiple closely linked genes contribute greatly to hybrid sterility, while fertility restoration is controlled predominantly by PPR genes (**Figure 2*b***). Noncoding RNAs appear to play an important role in photoperiod-sensitive male sterility (65). These general features of molecular functions of causal genes for different trait categories provide important guidance when identifying the underlying gene for new QTLs in future studies.

Coding variation: causal variants that occur in the coding sequence and result in amino acid changes

Regulatory variation: causal variants that regulate the expression of a functional gene

Presence/absence variation (PAV): causal variants that involve the presence or absence of a functional gene

Copy number variation (CNV): causal variants that involve the copy number of the functional gene

4.2. The Molecular Regulatory Mechanisms of Quantitative Trait Loci

The relative importance of protein-coding versus regulatory changes in phenotypic evolution is a long-standing debate. Among the 364 cloned QTLs, causative polymorphisms were identified in 291 (79.9%), including four epigenetic variants. The causal genetic polymorphisms are further classified into coding variation, regulatory variation, presence/absence variation (PAV), and copy number variation (CNV). Most causal variants identified to date are coding variation and regulatory variation, whereas CNV and PAV account for only a small percentage of total causal variation (**Figure 2*c***). The underrepresentation of CNV and PAV might primarily be due to technical limitations rather than biological significance, as increasing studies have revealed their roles in trait variation (2, 100, 105, 158).

Notably, the proportions of the two major variation types, coding versus regulatory, differ among the six crop species (**Figure 2*c***). In rice, 57.5% of QTLs are caused by coding variation, while 33% are caused by regulatory variation. Wheat, barley, and soybean exhibit a higher proportion of coding (>62.5%) than regulatory (<21.4%) variation. By contrast, of the 57 maize QTLs, 30.7% are caused by coding variation and 64% are caused by regulatory variation. The predominance of coding versus regulatory variation does not appear to be associated with genome size or predicted coding gene number (**Figure 2*c***). The mating systems of these crops might explain these differences. Soybean, rice, wheat, and barley are selfing plants (selfers), whereas maize is a typical outcrossing plant (outcrosser). Long-term self-fertilization decreases the effective recombination and population size (*Ne*), thus reducing the efficacy of selection (39). It has been shown that selfers tend to have a higher ratio of nonsynonymous to synonymous substitutions (39). This increased coding variation in selfers might reflect a release from selective constraint (39).

Unexpectedly, the selfer tomato exhibits an intermediate pattern of coding versus regulatory variation (43.9% versus 50%). Among the six crops, rice and tomato experienced an outcrossing-to-selfing transition during domestication (114, 132). We found that the mapping populations used in tomato involved a very high proportion of wild lines compared to rice, with 72.5% of tomato QTLs identified in wild × cultivated cross-populations compared with 22.6% in rice (**Supplemental Table 1**). Since regulatory variation tends to be enriched in outcrossing species, the high frequency of using wild lines might contribute to the detection of a larger proportion of regulatory variation in tomato.

Likely due to the greater prevalence of causal coding polymorphisms in selfers, QTL allele effects for comparable traits are usually larger in selfers than in outcrossers (using flowering time in rice and maize as an example) (**Figure 2*d***). The contrasting distributions of variation types suggest that the genetic architecture of complex traits in maize is usually controlled by many loci of minor effects, as frequently observed in genetic studies of various quantitative traits in maize (10, 127, 159).

Four epigenetic polymorphisms underlying trait variation have been reported, including *WFP* (113), *Gnp4/LAX2* (207), and *GLR1/GL-1* (81) in rice and *Cnr* (107) in tomato. The *WFP* locus,

controlling plant architecture in rice, was narrowed down to a 2.6-kb region upstream of *OsSPL14*. Surprisingly, no nucleotide difference in the causal region was detected between the two parental lines, indicating epigenetic causality (113). Similarly, no nucleotide sequence was detected in the causal region of the *Cnr* locus that regulates tomato fruit ripening (107). Differential DNA methylation in the *LeSPL-CNR* was found to be responsible for the phenotypic variation (107). These heritable epialleles might represent another important natural source of trait variation. With the rapid accumulation of epigenomic data, more epialleles will likely be identified, and the relative importance of epigenetic regulation in trait variation should be assessed fairly in the future.

Standing genetic variation: variants that preexist in wild ancestral populations

De novo mutation: mutations that do not exist in wild populations and only emerge during domestication or postdomestication

4.3. The Evolutionary Mechanisms of Quantitative Trait Loci

The relative importance of standing genetic variation versus de novo mutation in crop evolution has long been debated. The classification of standing variation versus de novo mutation here is based on the published studies (**Supplemental Table 1**). It is worth noting that de novo mutation can be overrated, as the number of wild lines examined in the published studies is relatively limited in general compared with the large number of cultivated lines investigated. However, the general insights obtained in the present review remain intriguing. Among the 167 cloned QTLs for which the origin of the causative variant has been identified, standing variation and de novo mutation contribute 51.8% and 47.3% across all traits, respectively. Among the 79 QTLs that were identified as selection targets during domestication or improvement and with clear information about variation origin, 53.8% result from de novo mutations and the rest, 45.6%, result from standing variation. For specific trait categories, de novo mutations are predominant for domestication and quality traits, whereas standing variation makes larger contributions to other traits, especially for abiotic stress (**Figure 2*e***). Notably, the selected alleles of three out of four maize domestication QTLs (*tb1*, *gt1*, and *ZmSh1-1*) (91, 149, 184, 198) involved standing variation, suggesting that standing variation might have played more important roles during maize domestication than it did in five other crop domestications. By contrast, at 12 out of 16 rice domestication QTLs, the selected alleles involved de novo mutation, indicating the predominant importance of de novo mutations in driving rice domestication. Another interesting observation is that loss-of-function alleles are more frequently associated with the domestication traits of selfing crops (67, 150, 175), while alleles selected during maize domestication more frequently involved gene expression regulation (149, 184).

By counting the number of causative variants within each cloned QTL, we found that rice, barley, and wheat have higher proportions of QTLs with allelic series (**Figure 2*f***). For example, nine causative variants were identified at the rice flowering-time locus *DTH7/Ghd7.1/qHD7.2* (35, 99). We propose three possible reasons for the differences in the prevalence of allelic series among crops. First, since the coding variation is inherently easier to detect and validate than regulatory variation, multiple causal alleles are more likely to be identified in crops with a majority of coding alleles (i.e., the three high-allelic series crops). Second, the origins of rice and wheat are more complex than those of maize, barley, and tomato. Wheat originated from an intercross between tetraploid *Triticum turgidum* and diploid *Aegilops tauschii*, and tetraploid *T. turgidum* was domesticated from wild emmer wheat that resulted from the hybridization between diploid *Triticum urartu* and *Aegilops speltoides* (30). Rice has two cultivated species, *Oryza sativa* and *Oryza glaberrima*, which were domesticated independently in Asia and Africa, respectively (176), and two subgroups of wild rice (*Oryza rufipogon*) are thought to be ancestors of Asian rice (*O. sativa*) (52, 177). Third, the selfing mating system limits the spread of favorable alleles, leading to a higher probability of maintaining independent functional variants in different groups or regions.

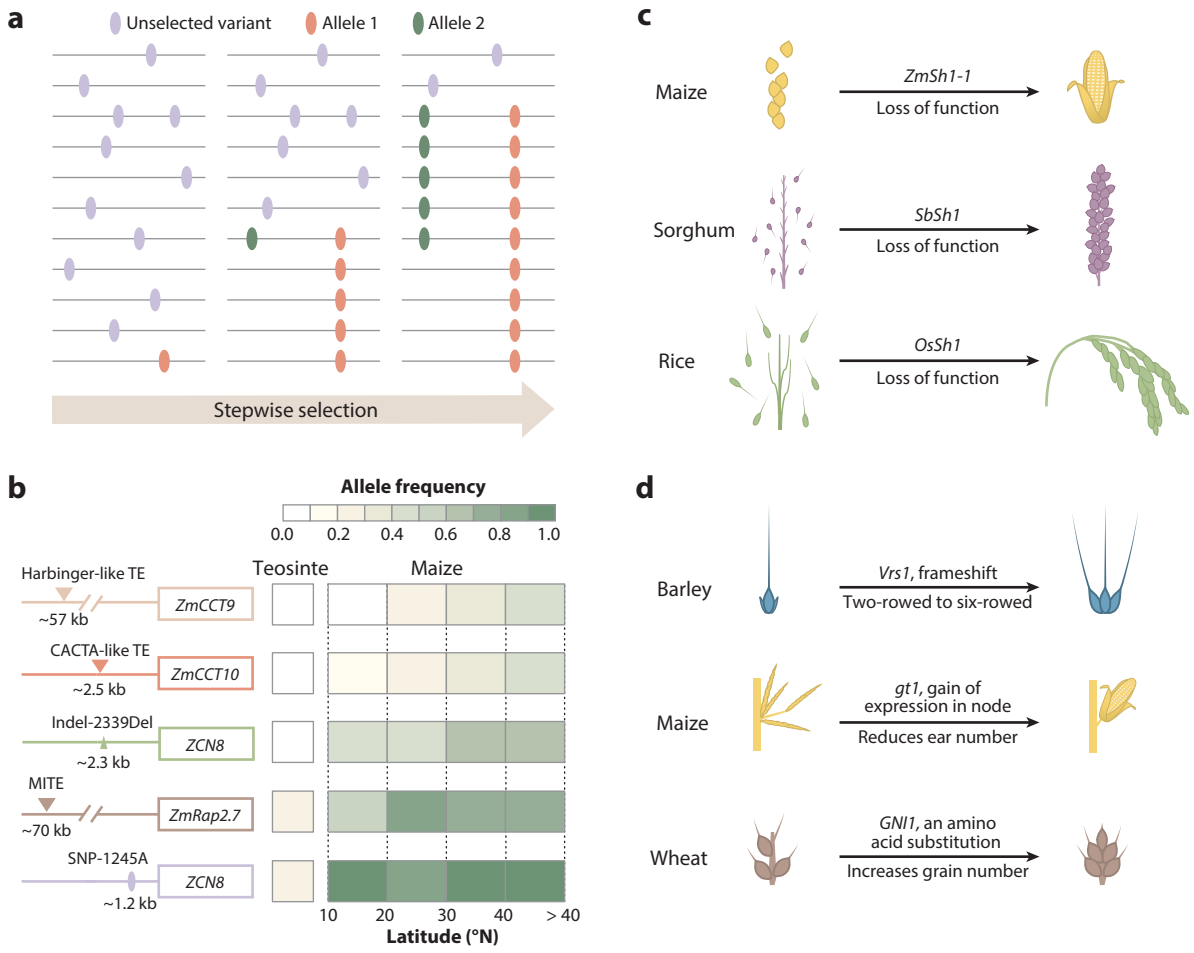


Figure 3

Evolutionary patterns. (*a*) Stepwise selection. Favorable allele 1 occurs first and is selected. Favorable allele 2 then occurs in the haplotype of allele 1 and is further selected in response to human demands or local conditions. (*b*) Sequential selection of polygenes using maize flowering time as an example. Five early flowering alleles of four maize flowering-time genes were sequentially selected as maize spread from its tropical origin to higher latitudes (42). (*c*) Parallel selection. Orthologs among crops are selected to control the same trait and evolve in the same direction. (*d*) Divergent selection. Orthologs across crops are selected to control different traits or evolve in distinct directions. Abbreviation: TE, transposable element.

Examining how those QTLs with allelic series have evolved might provide important insight into crop local adaptation. Researchers have frequently observed that different alleles within a locus were selected independently as a crop adapted to distinct ecological regions (35, 102, 183). Interestingly, recent studies showed that different causal variants in the same gene can be selected in a stepwise manner, such as *FZP* (6, 55) and *DTH2* (186) in rice and *ZCN8* (42) in maize (**Figure 3*a***). Two early flowering alleles, *SNP-1245_A* and *Indel-2339_Del*, were identified in the *ZCN8* promoter. The *SNP-1245_A* allele was initially selected during the early domestication stage of maize, and the *Indel-2339_Del* allele, which was introgressed into the *SNP-1245_A* haplotype, was subsequently selected as maize spread into the Mexican highlands (42). In addition to the stepwise selection of different alleles in the same gene, favorable alleles of different genes

controlling the same trait could also be selected sequentially to meet human demands or help plants adapt to different environments. For example, five early flowering alleles at four flowering-time genes (*ZmCCT9*, *ZmCCT10*, *ZmRap2.7*, and *ZCN8*) were sequentially selected to enhance the adaptation of maize to higher latitudes of temperate regions (42, 49) (**Figure 3*b***).

As discussed in Section 3.1, some conserved genes have been selected in a parallel manner and exhibit similar functions in different crops, such as the shattering genes *Sh1* (91) (**Figure 3*c***) and *Btr1* (129, 208) and the seed dormancy gene *G* (175). Interestingly, some orthologous genes that were targeted by selection have divergent functions in different crops. The gain-of-expression of the HD-ZIP TF *gt1* in the nodes of the upper branch of maize reduces ear number (184). Conversely, the loss-of-function alleles of its ortholog *Vrs1* in barley and *GNI1* in wheat were selected to increase spikelet and floret number, respectively (67, 70, 134) (**Figure 3*d***).

4.4. The Mechanisms of Pleiotropic Quantitative Trait Loci

Pleiotropy refers to the phenomenon in which a single locus affects two or more unrelated phenotypic traits (144), and it is prevalent among the cloned QTLs. Analyzing the QTL collection derived from a broad set of traits in diverse crops provides us with clues about the mechanisms underlying pleiotropy. Pleiotropic genes usually function as hubs in multiple temporally and/or spatially divergent regulatory networks. Rice *IPA1* provides a great example: (*a*) It suppresses rice tillering by directly promoting the expression of *OsTB1* in the shoot apex (103); (*b*) it increases grain number per panicle by directly activating *DEP1* in young panicles (173); and (*c*) upon pathogen attack, IPA1 becomes phosphorylated and activates *WRKY45* expression, leading to enhanced blast resistance (173).

Pleiotropy might be mediated by distinct *cis*-regulatory variants of the pleiotropic genes. Taking *GS2/GL2/qGRN2/OsGRF4* controlling both yield and nitrogen-use efficiency (46, 83) as an example, a mutation within the miR396 target site significantly affects grain size and yield (46), while *cis*-polymorphisms in its promoter respond exclusively to nitrogen sensitivity (83). These mechanism-based examples highlight the relevance and diverse origins of pleiotropy and provide valuable targets for further manipulating overall crop performance.

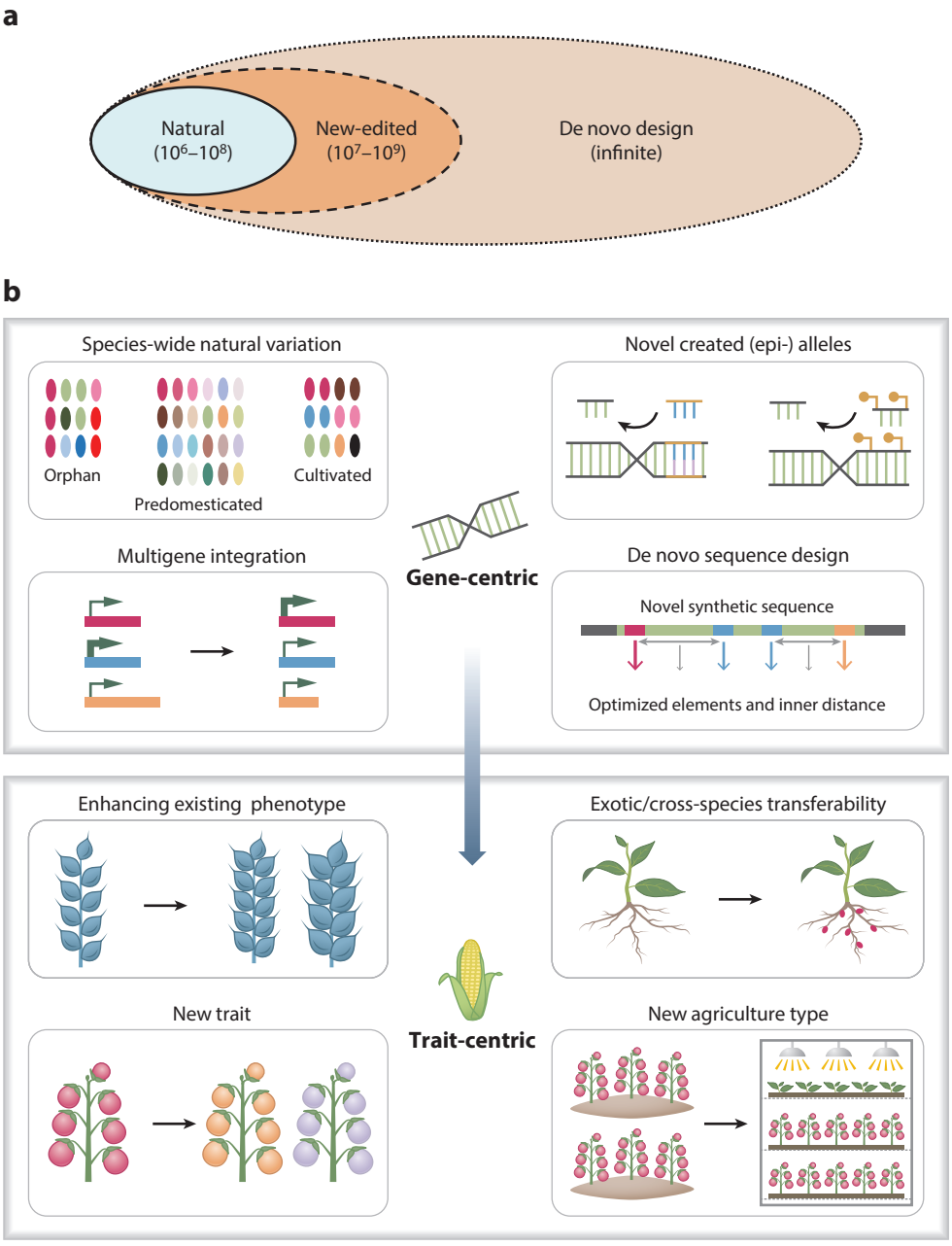
5. IDENTIFYING AND CREATING ALLELES FOR CROP BREEDING

The polymorphisms within diverse natural accessions provide an existing variation scope (**Figure 4*a***), independently or in combination, for various traits or physiological or biological processes. The identification of increasing numbers of functional genes, causal alleles, and the corresponding molecular mechanisms has caused breeding paradigms to evolve from modifying individual genes and traits to modifying entire gene sets and overall performance and even to generating novel crops to achieve new agricultural types (**Figure 4*b***). All these exciting advances make breeding much more targeted and more customizable, efficient, and sustainable.

5.1. Employing Favorable Natural Alleles

Cloned genes and beneficial natural alleles directly provide valuable materials for precisely targeted trait improvement. In addition to genes controlling specific individual traits, key hub genes that simultaneously affect multiple traits have great potential for improving overall crop performance. As described above, fine-tuning the tissue-specific abundance and transient turnover of the phosphorylated modification of IPA1 offers a way to optimize plant type to achieve the highest grain potential (172, 174, 205) and enhance resistance without yield penalties (173). In practice, the

ipa1-1D allele (a microRNA target site mutation leading to higher expression of *IPA1*) increased yields by at least 10% and 30% under normal and high blast disease conditions, respectively (59, 173, 205). *DEP1* has been shown to regulate both rice yield (53) and nitrogen responses (151), and its natural allele from the *japonica* rice variety Qianzhonglang2 coordinately increases nitrogen use efficiency and yield (151). Interestingly, *DEP1* is a direct downstream target of *IPA1* (173); thus,



(*Caption appears on following page*)

Figure 4 (*Figure appears on preceding page*)

The value and implementation of genetic engineering for crop improvement. (*a*) A rough estimate of the variation space from the natural population, genome editing, and de novo design for a crop. Here, the number of natural variations only represents a part that can be aligned and is estimated from current HapMap-like data sets plus the number of structure variation (SV) space identified using long-read sequencing technologies (2, 100). The scale of new-edited variation space is roughly estimated as about $3 \times 10^6 - 1 \times 10^8$ Cas9 single-guide RNAs that could be targeted to genic regions of the six crops (152), and generally ∼3–25 independent mutations can be generated for each (93). The edited alleles could be repeated with natural ones, but researchers would need to consider that other Cas proteins or other editing technologies (such as base editing) would add an additional severalfold (or even tenfold) to the number of potential mutations. For the infinite de novo design, if we assume to only design a 140-bp sequence, the possible sequences are $4^{140} \approx 2 \times 10^{84}$, which is significantly larger than the estimated number (10^{78}–10^{82}) of all atoms in the observable universe. (*b*) Sources of variation and current genetic engineering approaches used for breeding. Each subpanel represents a specific case described in the main text and is defined as follows: Species-wide natural variation: employs favorable natural alleles from the species-wide pool. Novel created (epi-) alleles: de novo creation of favorable (epi-) alleles compared with natural alleles. Multigene integration: creates better allelic combinations of multiple genes. De novo sequence design: designs a de novo sequence of a given gene. Enhancing existing phenotype: diverse and customized enhancement of existing phenotypes. Exotic/cross-species transferability: agricultural bioengineering through synthetic transkingdom signaling. New trait: develops new traits to better meet the present or novel breeding objectives. New agriculture type: expands the types and boundaries of traditional farming (here, urban vertical farming is used as an example).

manipulating these two genes could provide the opportunity to develop Green Super Rice (185) with less need for fertilization and pesticide treatment but higher yields.

Another prospect is to directly characterize heterosis-related genes to guide heterosis utilization. For example, the rice flowering-time gene *Hd3a* and tiller angle gene *TAC1* show single-gene overdominance (i.e., the heterozygous alleles perform better than both homozygous alleles); plants harboring these mutually heterozygous alleles exhibited optimized flowering time and yield as well as plant architecture adapted to dense planting (54). In addition, the yield of a heterozygote of the maize *KY4q19* locus containing *Ub3* (an *IPA1* ortholog) was over 13% higher than that of the homozygous genotype (92).

During crop domestication and improvement or local adaptation, a significant amount of genetic diversity was lost due to genetic bottlenecks and selection, which could have excluded favorable alleles that are valuable for modern breeding. Introgression of the teosinte *UPA2* allele, conferring upright leaf angle, enhanced maize yield under dense planting in both inbreds and hybrids (160). Similarly, a number of wild alleles could improve the flavor of tomato fruit (161, 209). Probing species-wide pan-genome or pan-family data would provide great opportunities to define the optimal landscape of genes and alleles (158), such as by identifying and engineering nonreference and wild-specific genes for improving agronomic traits (199), disease resistance (4, 36), and nutritional metabolites (36).

Base editing: a genome-editing tool used to directly and irreversibly convert one specific nucleotide to another without generating double-stranded breaks

5.2. Creating Novel Alleles Beyond Nature

The knowledge gained from uncovering genotype-phenotype links using natural variation could be utilized to greatly increase the variation space through genome editing (**Figure 4*b***). Importantly, recently developed base-editing-mediated technologies could facilitate the directed evolution of one or a few plant genes by introducing saturation mutations (72, 79). Targeting tomato *J2* with CRISPR/Cas9 produced desirable pedicel and branch traits in fresh-market tomato breeding lines (147). Furthermore, waxy corn was engineered to overcome the production declines that have accompanied the introgression of the waxy haplotype during traditional breeding (34). In addition to individual loci, novel allelic combinations are increasingly being produced by genome editing.

The MADS-box genes *J2* and *EJ2* in tomato have negative epistatic effects on inflorescence development and yield. New alleles from gene editing for these two genes and another MADS-box gene (*LIN*) allowed inflorescence architecture to be optimized, leading to improved yields (146). Targeting diverse *SWEET* genes simultaneously in rice provided a cost-effective way to achieve broad-spectrum pathogen resistance with no obvious growth penalties (27, 121).

Engineering stable gene transcriptional changes and inherited epialleles could also potentially accelerate crop improvement in the context of various stress conditions and global climate change. This technique has been successfully utilized to alter flowering time and improve drought tolerance in *Arabidopsis* (123, 126). Although the basic epigenomic analysis in crops is still in its infancy and few specific epitargets for editing are currently available, additional GWASs of gene expression and epigenetic modifications should provide promising target sites in the future (94, 190).

5.3. Engineering New Traits

The improvement of existing traits usually involves a gene-centric approach, in which cloned genes with natural or created alleles are employed as starting points. However, the emerging innovative modes of crop production and climate change, new traits, and even new species are becoming the breeding objectives in a trait-centric approach (**Figure 4*b***). This type of breeding will typically require the modification of the activities of new genes and/or series of genes and could be accomplished using various synthetic biology technologies.

The most straightforward approach is to create a crop ideotype from wild relatives through rapid de novo domestication. Under the guidance of the numerous underlying genes and their interactions, new modern crops could be engineered by molecular domestication to increase productivity (77), as has already been done, while simultaneously keeping desirable characteristics from wild relatives, including resistance to diverse stresses and high nutritional value (84, 212).

The types and boundaries of traditional farming are being expanded. For example, urban vertical farming, especially when it involves artificial intelligence–powered gardening, represents a new production mode to help cope with the growing urban population, continuing climate change, and diminishing resource availability (8, 25). Urban farming requires new traits that are different from those suited for field agriculture. In a pioneering study, one-step targeting of genes responsible for flowering time (*SP5G*), growth termination (*SP*), and stem length (*SlER*) successfully produced highly compact, rapid-flowering fruit crops suitable for indoor agriculture (74). Additionally, the use of genome editing to engineer apomixis for the fixation of hybrid rice through clonally propagated seeds (63, 167) represents a revolutionary approach for hybrid seed production in naturally selfing crops.

Beyond genome editing, other emerging synthetic biology techniques can enrich bioengineering applications in agriculture, allowing for the fine manipulation of crops to meet diverse demands. Examples include reconfiguring metabolic flux to enrich nutrients (44, 210), increasing carbon fixation such as by enhancing the efficiency of Rubisco (the CO_2-fixing enzyme in photosynthesis) (73), suppressing photorespiration by altering pathways (142, 145) or reducing transpiration using a synthetic light-gated ion channel (125), promoting rhizobial symbioses via synthetic transkingdom signaling (38), and employing molecular farming—the artificial introduction of genes into plants—to develop a COVID-19 vaccine (75).

Synthetic biology: the design or redesign of novel biological parts or systems for new purposes

Vertical farming: a new crop production practice in which crops are grown in vertically stacked layers in a controlled environment

Metabolic flux: the movement and turnover of metabolites through pathways regulated by enzymes and cofactors over time and across conditions

6. CONCLUSIONS AND PERSPECTIVE

Crops, perhaps the most exciting human inventions, in turn, shape human civilization. Since the Green Revolution, the pace of crop improvement has not been sufficient to meet the ever-increasing demands of the growing worldwide population, especially in light of the increasingly

changing climate and diminishing resources. Natural variation provides us with ideal experimental disturbances that could be used to characterize the genetic basis of trait variation, providing a deeper understanding of biological insights and promoting knowledge-guided crop improvement. As we are currently facing an era of rapidly developing technologies, our ability to map QTLs and pinpoint QTNs will become even more efficient and high-throughput.

Although natural variants in crops have been around for tens of millions of years, they represent only a tiny portion of the theoretical variation space. Furthermore, during most of evolution, these variants have led to the development of phenotypes that are not tailored for today's or tomorrow's agricultural requirements. Genome editing and de novo synthetic design (179) have greatly expanded the vast scope of the potential variation space (**Figure 4*a***) at rates many orders of magnitude faster than natural evolution. Many biologically inspired novel alleles and new traits have been created, holding great promise for helping us face problems raised by current and emerging agricultural paradigms.

The continuous harnessing of natural alleles is essential for discovering new targets and increasing the predictability of new-to-nature design. Further exploration and exploitation of these alleles would offer unparalleled power, from reading to engineering. It is reasonable to expect that the next generation of crops could be rationally (re)designed to enhance sustainable development for both humans and the planet.

SUMMARY POINTS

1. Natural alleles are ideal experimental disturbances for characterizing causal genotype-phenotype connections and the corresponding molecular mechanisms.
2. It is important to understand the trade-offs of different genetic designs. The increasing availability of independent data sets provides the opportunity for replicating and identifying novel associations and obtaining new biological insights.
3. During the last three decades, natural variants of many genes corresponding to various important traits have been cloned, providing fundamental resources for knowledge-guided crop improvement.
4. A general survey of quantitative trait loci cloned to date provides a picture of the molecular, regulatory, and evolutionary mechanisms of natural variation.
5. Both basic research and applied research are needed for studies ranging from genetic mapping of individual traits to analyzing overall crop performance. New agricultural paradigms will continue to emerge, and new phenotypes are required to meet growing demands.
6. Natural alleles are rich in scope, but emerging genome-editing techniques and other synthetic biology technologies could further expand the mutational space by quickly creating de novo alleles, pyramiding favorable alleles from multiple genes, designing novel sequences for enhancing traits, or even generating new traits.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Professor Xiaofeng Cao (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences), Professor John Doebley (University of Wisconsin–Madison), and Professor Jeffrey Ross-Ibarra (University of California, Davis) for their critical comments and discussions. We thank Xu Han, Wenchao Qin, and Shumin Wang for helping to collect QTL data. We sincerely apologize to the authors whose work was not cited due to space limitations. This work was mainly completed during the COVID-19 pandemic: We would like to express our deepest admiration and gratitude to all health professionals. This work was supported by the National Natural Science Foundation of China (91935305), the National Transgenic Major Project of China (2019ZX08010003-002-013), the National Key Research and Development Program of China (2020YFE0202300), the Recruitment Program of Global Experts, and the Fundamental Research Funds for the Central Universities.

LITERATURE CITED

1. Abdelghani AH, Parzies HK, Omary A, Geiger HH. 2004. Estimating the outcrossing rate of barley landraces and wild barley populations collected from ecologically different regions of Jordan. *Theor. Appl. Genet.* 109:588–95
2. Alonge M, Wang X, Benoit M, Soyk S, Pereira L, et al. 2020. Major impacts of widespread structural variation on gene expression and crop improvement in tomato. *Cell* 182:145–61. e23
3. Alpert KB, Tanksley SD. 1996. High-resolution mapping and isolation of a yeast artificial chromosome contig containing *fw2.2*: a major fruit weight quantitative trait locus in tomato. *PNAS* 93:15503–7
4. Arora S, Steuernagel B, Gaurav K, Chandramohan S, Long Y, et al. 2019. Resistance gene cloning from a wild crop relative by sequence capture and association genetics. *Nat. Biotechnol.* 37:139–43
5. Auer PL, Lettre G. 2015. Rare variant association studies: considerations, challenges and opportunities. *Genome Med.* 7:16
6. Bai X, Huang Y, Hu Y, Liu H, Zhang B, et al. 2017. Duplication of an upstream silencer of *FZP* increases grain yield in rice. *Nat. Plants* 3:885–93
7. Barrero JM, Cavanagh CR, Verbyla KL, Tibbits J, Verbyla AP, et al. 2015. Transcriptomic analysis of wheat near-isogenic lines identifies *PM19-A1* and *A2* as candidates for a major dormancy QTL. *Genome Biol.* 16:93
8. Benke KK, Tomkins B. 2017. Future food-production systems: vertical farming and controlled-environment agriculture. *Sustain. Sci. Pract. Policy* 13:13–26
9. Bommert P, Nagasawa NS, Jackson D. 2013. Quantitative variation in maize kernel row number is controlled by the *FASCIATED EAR2* locus. *Nat. Genet.* 45:334–37
10. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, et al. 2009. The genetic architecture of maize flowering time. *Science* 325:714–18
11. Cannon ME, Mohlke KL. 2018. Deciphering the emerging complexities of molecular mechanisms at GWAS loci. *Am. J. Hum. Genet.* 103:637–53
12. Chakrabarti M, Zhang N, Sauvage C, Munos S, Blanca J, et al. 2013. A cytochrome P450 regulates a domestication trait in cultivated tomato. *PNAS* 110:17125–30
13. **Chen L, Liu Y. 2014. Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* 65:579–606**
14. Comadran J, Kilian B, Russell J, Ramsay L, Stein N, et al. 2012. Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44:1388–92
15. Cui X, Wise RP, Schnable PS. 1996. The *rf2* nuclear restorer gene of male-sterile T-cytoplasm maize. *Science* 272:1334–36
16. Dai X, Ding Y, Tan L, Fu Y, Liu F, et al. 2012. *LHD1*, an allele of *DTH8/Ghd8*, controls late heading date in common wild rice (*Oryza rufipogon*). *J. Integr. Plant Biol.* 54:790–99

13. Reviews the mechanisms underlying male sterility, fertility restoration, and hybrid compatibility (see also 122).

17. De Wiel CCMV, Lotz LAP. 2006. Outcrossing and coexistence of genetically modified with (genetically) unmodified crops: a case study of the situation in the Netherlands. *NJAS Wagen. J. Life Sci.* 54:17–35
18. Dixon LE, Greenwood JR, Bencivenga S, Zhang P, Cockram J, et al. 2018. *TEOSINTE BRANCHED1* regulates inflorescence architecture and development in bread wheat (*Triticum aestivum*). *Plant Cell* 30:563–81
19. Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* 127:1309–21
20. Doebley JF, Stec A, Hubbard L. 1997. The evolution of apical dominance in maize. *Nature* 386:485–88
21. Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, et al. 2004. *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT-like* gene expression independently of *Hd1*. *Genes Dev.* 18:926–36
22. Dong Y, Yang X, Liu J, Wang B, Liu BL, Wang Y. 2014. Pod shattering resistance associated with domestication is mediated by a NAC gene in soybean. *Nat. Commun.* 5:3352
23. Dong Z, Xiao Y, Govindarajulu R, Feil R, Siddoway ML, et al. 2019. The regulatory landscape of a core maize domestication module controlling bud dormancy and growth repression. *Nat. Commun.* 10:3810
24. Douglas AE. 2018. Strategies for enhanced crop resistance to insect pests. *Annu. Rev. Plant Biol.* 69:637–60
25. Elichukwu NC. 2019. Applications of artificial intelligence in agriculture: a review. *Eng. Technol. Appl. Sci. Res.* 9:4377–83
26. Enjalbert J, David JL. 2000. Inferring recent outcrossing rates using multilocus individual heterozygosity: application to evolving wheat populations. *Genetics* 156:1973–82
27. Eom JS, Luo D, Atienza-Grande G, Yang J, Ji C, et al. 2019. Diagnostic kit for rice blight resistance. *Nat. Biotechnol.* 37:1372–79
28. Eshed Y, Zamir D. 1995. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* 141:1147–62
29. Fang J, Zhang F, Wang H, Wang W, Zhao F, et al. 2019. *Ef-cd* locus shortens rice maturity duration without yield penalty. *PNAS* 116:18717–22
30. Faris JD. 2014. Wheat domestication: key to agricultural revolutions past and future. In *Genomics of Plant Genetic Resources*, Vol. 1: *Managing, Sequencing and Mining Genetic Resources*, ed. R Tuberosa, A Graner, E Frison, pp. 439–64. Dordrecht, Neth.: Springer
31. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, et al. 2000. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* 289:85–88
32. Fridman E, Pleban T, Zamir D. 2000. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *PNAS* 97:4718–23
33. Funatsuki H, Suzuki M, Hirose A, Inaba H, Yamada T, et al. 2014. Molecular basis of a shattering resistance boosting global dissemination of soybean. *PNAS* 111:17797–802
34. Gao H, Gadlage MJ, Lafitte HR, Lenderts B, Yang M, et al. 2020. Superior field performance of waxy corn engineered using CRISPR-Cas9. *Nat. Biotechnol.* 38:579–81
35. Gao H, Jin M, Zheng XM, Chen J, Yuan D, et al. 2014. *Days to heading 7*, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. *PNAS* 111:16337–42
36. Gao L, Gonda I, Sun H, Ma Q, Bao K, et al. 2019. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* 51:1044–51
37. Gardiner LJ, Wingen LU, Bailey P, Joynson R, Brabbs T, et al. 2019. Analysis of the recombination landscape of hexaploid bread wheat reveals genes controlling recombination and gene conversion frequency. *Genome Biol.* 20:69
38. Geddes BA, Paramasivan P, Joffrin A, Thompson AL, Christensen K, et al. 2019. Engineering transkingdom signalling in plants to control gene expression in rhizosphere bacteria. *Nat. Commun.* 10:3430
39. Glémin S, François CM, Galtier N. 2019. Genome evolution in outcrossing versus selfing versus asexual species. In *Evolutionary Genomics: Statistical and Computational Methods*, ed. M Anisimova, pp. 331–69. New York: Springer
40. Gong Z. 2020. Flowering phenology as a core domestication trait in soybean. *J. Integr. Plant Biol.* 62:546–49
41. Guo J, Xu C, Wu D, Zhao Y, Qiu Y, et al. 2018. *Bph6* encodes an exocyst-localized protein and confers broad resistance to planthoppers in rice. *Nat. Genet.* 50:297–306

39. Illustrates the profound influences of mating system on population genetic processes and genome evolution.

42. Guo L, Wang X, Zhao M, Huang C, Li C, et al. 2018. Stepwise *cis*-regulatory changes in *ZCN8* contribute to maize flowering-time adaptation. *Curr. Biol.* 28:3005–15.e4

43. Hammer K. 1984. Das Domestikationsyndrom [The domestication syndrome]. *Kulturpflanze* 32:11–34
44. Han L, Usher S, Sandgrind S, Hassall K, Sayanova O, et al. 2020. High level accumulation of EPA and DHA in field-grown transgenic Camelina—a multi-territory evaluation of TAG accumulation and heterogeneity. *Plant Biotechnol. J.* 18:2280–91
45. Harlan JR. 1992. The dynamics of domestication. In *Crops and Man*, ed. JR Harlan, pp. 115–33. Madison, WI: Am. Soc. Agron. 2nd ed.
46. Hu J, Wang Y, Fang Y, Zeng L, Xu J, et al. 2015. A rare allele of *GS2* enhances grain size and grain yield in rice. *Mol. Plant* 8:1455–65
47. Huang BE, Verbyla KL, Verbyla AP, Raghavan C, Singh VK, et al. 2015. MAGIC populations in crops: current status and future prospects. *Theor. Appl. Genet.* 128:999–1017
48. Huang C, Chen Q, Xu G, Xu D, Tian J, Tian F. 2016. Identification and fine mapping of quantitative trait loci for the number of vascular bundle in maize stem. *J. Integr. Plant Biol.* 58:81–90
49. Huang C, Sun H, Xu D, Chen Q, Liang Y, et al. 2018. *ZmCCT9* enhances maize adaptation to higher latitudes. *PNAS* 115:E334–41
50. Huang J, Lu G, Liu L, Raihan MS, Xu J, et al. 2020. The kernel size-related quantitative trait locus *qKW9* encodes a pentatricopeptide repeat protein that affects photosynthesis and grain filling. *Plant Physiol.* 183:1696–709

51. Huang X, Han B. 2014. Natural variations and genome-wide association studies in crop plants. *Annu. Rev. Plant Biol.* 65:531–51

52. Huang X, Kurata N, Wei X, Wang ZX, Wang A, et al. 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490:497–501
53. Huang X, Qian Q, Liu Z, Sun H, He S, et al. 2009. Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat. Genet.* 41:494–97
54. Huang X, Yang S, Gong J, Zhao Q, Feng Q, et al. 2016. Genomic architecture of heterosis for yield traits in rice. *Nature* 537:629–33
55. Huang Y, Zhao S, Fu Y, Sun H, Ma X, et al. 2018. Variation in the regulatory region of *FZP* causes increases in secondary inflorescence branching and grain yield in rice domestication. *Plant J.* 96:716–33
56. Hung HY, Shannon LM, Tian F, Bradbury PJ, Chen C, et al. 2012. *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *PNAS* 109:E1913–21
57. Jähne F, Hahn V, Würschum T, Leiser WL. 2020. Speed breeding short-day crops by LED-controlled light schemes. *Theor. Appl. Genet.* 133:2335–42
58. Jia H, Li M, Li W, Liu L, Jian Y, et al. 2020. A serine/threonine protein kinase encoding gene *KERNEL NUMBER PER ROW6* regulates maize grain yield. *Nat. Commun.* 11:988

59. Jiao Y, Wang Y, Xue D, Wang J, Yan M, et al. 2010. Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* 42:541–44

60. Jin J, Huang W, Gao J, Yang J, Shi M, et al. 2008. Genetic control of rice plant architecture under domestication. *Nat. Genet.* 40:1365–69
61. Jin M, Liu X, Jia W, Liu H, Li W, 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.* 60:465–80
62. Khan AW, Garg V, Roorkiwal M, Golicz AA, Edwards D, Varshney RK. 2020. Super-pangenome by integrating the wild side of a species for accelerated crop improvement. *Trends Plant Sci.* 25:148–58
63. Khanday I, Skinner DJ, Yang B, Mercier R, Sundaresan V. 2019. A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature* 565:91–95
64. Khush GS. 1995. Breaking the yield frontier of rice. *GeoJournal* 35:329–32
65. Kim Y, Zhang D. 2018. Molecular control of male fertility for crop hybrid breeding. *Trends Plant Sci.* 23:53–65
66. Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, et al. 2002. *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol.* 43:1096–105
67. Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, et al. 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *PNAS* 104:1424–29

42. Provides an example of stepwise selection and wild introgression that help maize adapt temperate regions.

51. Presents an overview of sequencing-based genotyping, GWAS, and subsequent innovations to crop breeding.

59. Provides a great example of pleiotropy and an excellent breeding target (see also 103, 113, 172–174, and 205).

68. Kong D, Wang B, Wang H. 2020. *UPA2* and *ZmRAVL1*: promising targets of genetic improvement of maize plant architecture. *J. Integr. Plant Biol.* 62:394–97
69. Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, et al. 2006. An SNP caused loss of seed shattering during rice domestication. *Science* 312:1392–96
70. Koppolu R, Schnurbusch T. 2019. Developmental pathways for shaping spike inflorescence architecture in barley and wheat. *J. Integr. Plant Biol.* 61:278–95
71. Krattinger SG, Lagudah ES, Spielmeyer W, Singh RP, Huerta-Espino J, et al. 2009. A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323:1360–63
72. Kuang Y, Li S, Ren B, Yan F, Spetz C, et al. 2020. Base-editing-mediated artificial evolution of *OsALS1 in planta* to develop novel herbicide-tolerant rice germplasms. *Mol. Plant* 13:565–72
73. Kubis A, Bareven A. 2019. Synthetic biology approaches for improving photosynthesis. *J. Exp. Bot.* 70:1425–33
74. Kwon CT, Heo J, Lemmon ZH, Capua Y, Hutton SF, et al. 2020. Rapid customization of Solanaceae fruit crops for urban agriculture. *Nat. Biotechnol.* 38:182–88
75. Labios L. 2020. *Marrying molecular farming and advanced manufacturing to develop a COVID-19 vaccine*. UC San Diego News Center. Press Release, April 20. **https://ucsdnews.ucsd.edu/pressrelease/marrying-molecular-farming-and-advanced-manufacturing-to-develop-a-covid-19-vaccine**
76. Larsson SJ, Lipka AE, Buckler ES. 2013. Lessons from *Dwarf8* on the strengths and weaknesses of structured association mapping. *PLOS Genet.* 9:e1003246
77. Lemmon ZH, Reem NT, Dalrymple J, Soyk S, Swartwood KE, et al. 2018. Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat. Plants* 4:766–70
78. Lesley JW. 1948. Plant breeding methods and current problems in developing improved varieties of tomatoes. *Econ. Bot.* 2:100–10
79. Li C, Zhang R, Meng X, Chen S, Zong Y, et al. 2020. Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nat. Biotechnol.* 38:875–82
80. Li C, Zhou A, Sang T. 2006. Rice domestication by reducing shattering. *Science* 311:1936–39
81. Li J, Yuan Y, Lu Z, Yang L, Gao R, et al. 2012. *Glabrous Rice 1*, encoding a homeodomain protein, regulates trichome development in rice. *Rice* 5:32
82. Li MW, Liu W, Lam HM, Gendron JM. 2019. Characterization of two growth period QTLs reveals modification of *PRR3* genes during soybean domestication. *Plant Cell Physiol.* 60:407–20
83. Li S, Tian Y, Wu K, Ye Y, Yu J, et al. 2018. Modulating plant growth–metabolism coordination for sustainable agriculture. *Nature* 560:595–600
84. Li T, Yang X, Yu Y, Si X, Zhai X, et al. 2018. Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* 36:1160–63
85. Li W, Deng Y, Ning Y, He Z, Wang G. 2020. Exploiting broad-spectrum disease resistance in crops: from molecular dissection to breeding. *Annu. Rev. Plant Biol.* 71:575–603
86. Li X, Li X, Fridman E, Tesso TT, Yu J. 2015. Dissecting repulsion linkage in the dwarfing gene *Dw3* region for sorghum plant height provides insights into heterosis. *PNAS* 112:11823–28
87. Li Y, Guan R, Liu Z, Ma Y, Wang L, et al. 2008. Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. *Theor. Appl. Genet.* 117:857–71
88. Liang Y, Liu Q, Wang X, Huang C, Xu G, et al. 2019. *ZmMADS69* functions as a flowering activator through the *ZmRap2.7-ZCN8* regulatory module and contributes to maize flowering time adaptation. *New Phytol.* 221:2335–47
89. Liang Z, Qiu Y, Schnable JC. 2020. Genome–phenome wide association in maize and *Arabidopsis* identifies a common molecular and evolutionary signature. *Mol. Plant* 13:907–22
90. Lin T, Zhu G, Zhang J, Xu X, Yu Q, et al. 2014. Genomic analyses provide insights into the history of tomato breeding. *Nat. Genet.* 46:1220–26
91. Lin Z, Li X, Shannon LM, Yeh CT, Wang ML, et al. 2012. Parallel domestication of the *Shattering1* genes in cereals. *Nat. Genet.* 44:720–24
92. Liu H, Wang Q, Chen M, Ding Y, Yang X, et al. 2020. Genome-wide identification and analysis of heterotic loci in three maize hybrids. *Plant Biotechnol. J.* 18:185–94
93. Liu HJ, Jian L, Xu J, Zhang Q, Zhang M, et al. 2020. High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize. *Plant Cell* 32:1397–413

72. A pioneer study to accelerate crop breeding with directed evolution of target genes by functional screening saturation mutations induced by CRISPR-based base editors (see also 79).

77. Demonstrates the prospect of fast de novo domestication from wild or orphan crops with genome editing toward the creation of ideal crops (see also 84 and 212).

94. Liu HJ, Luo X, Niu L, Xiao Y, Chen L, et al. 2017. Distant eQTLs and non-coding sequences play critical roles in regulating gene expression and quantitative trait variation in maize. *Mol. Plant* 10:414–26
95. Liu HJ, Wang X, Xiao Y, Luo J, Qiao F, et al. 2020. CUBIC: an atlas of genetic architecture promises directed maize improvement. *Genome Biol*. 21:20
96. Liu HJ, Yan J. 2019. Crop genome-wide association study: a harvest of biological relevance. *Plant J*. 97:8–18
97. Liu L, Du Y, Shen X, Li M, Sun W, et al. 2015. *KRN4* controls quantitative variation in maize kernel row number. *PLOS Genet*. 11:e1005670
98. Liu S, Yeh CT, Tang HM, Nettleton D, Schnable PS. 2012. Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PLOS ONE* 7:e36406
99. Liu T, Liu H, Zhang H, Xing Y. 2013. Validation and characterization of *Ghd7.1*, a major quantitative trait locus with pleiotropic effects on spikelets per panicle, plant height, and heading date in rice (*Oryza sativa* L.). *J. Integr. Plant Biol.* 55:917–27
100. Liu Y, Du H, Li P, Shen Y, Peng H, et al. 2020. Pan-genome of wild and cultivated soybeans. *Cell* 182:162–76.e13
101. Lu S, Dong L, Fang C, Liu S, Kong L, et al. 2020. Stepwise selection on homeologous *PRR* genes controlling flowering and maturity during soybean domestication. *Nat. Genet.* 52:428–36
102. Lu S, Zhao X, Hu Y, Liu S, Nan H, et al. 2017. Natural variation at the soybean *J* locus improves adaptation to the tropics and enhances yield. *Nat. Genet.* 49:773–79
103. Lu Z, Yu H, Xiong G, Wang J, Jiao Y, et al. 2013. Genome-wide binding analysis of the transcription activator *Ideal Plant Architecture1* reveals a complex network regulating rice plant architecture. *Plant Cell* 25:3743–59
104. Luo C, Li X, Zhang Q, Yan J. 2019. Single gametophyte sequencing reveals that crossover events differ between sexes in maize. *Nat. Commun.* 10:785
105. Lye ZN, Purugganan MD. 2019. Copy number variation in domestication. *Trends Plant Sci*. 24:352–65
106. Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* 10:565–77
107. Manning K, Tor M, Poole M, Hong Y, Thompson AJ, et al. 2006. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* 38:948–52
108. Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, et al. 1993. Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262:1432–36
109. Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler ES, Doebley J. 2002. A single domestication for maize shown by multilocus microsatellite genotyping. *PNAS* 99:6080–84
110. McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, et al. 2009. Genetic properties of the maize nested association mapping population. *Science* 325:737–40
111. Messeguer J, Fogher C, Guiderdoni E, Marfa V, Catala MM, et al. 2001. Field assessments of gene flow from transgenic to cultivated rice (*Oryza sativa* L.) using a herbicide resistance gene as tracer marker. *Theor. Appl. Genet.* 103:1151–59
112. Mickelbart MV, Hasegawa PM, Baileyserres J. 2015. Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. *Nat. Rev. Genet.* 16:237–51
113. Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, et al. 2010. *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* 42:545–49
114. Morishima H, Barbier P. 1990. Mating system and genetic structure of natural populations in wild rice *Oryza rufipogon*. *Plant Species Biol*. 5:31–39
115. Mu Q, Huang Z, Chakrabarti M, Illaberenguer E, Liu X, et al. 2017. Fruit weight is controlled by *Cell Size Regulator* encoding a novel protein that is expressed in maturing tomato fruits. *PLOS Genet*. 13:e1006930
116. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, et al. 2009. Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21:2194–202
117. Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, et al. 2011. A wheat homolog of *MOTHER OF FT AND TFL1* acts in the regulation of germination. *Plant Cell* 23:3215–29

118. Nakamura S, Pourkheirandish M, Morishige H, Kubo Y, Nakamura M, et al. 2016. *Mitogen-Activated Protein Kinase Kinase 3* regulates seed dormancy in barley. *Curr. Biol.* 26:775–81
119. Ning Y, Liu W, Wang G. 2017. Balancing immunity and yield in crop plants. *Trends Plant Sci.* 22:1069–79
120. Ogisotanaka E, Matsubara K, Yamamoto S, Nonoue Y, Wu J, et al. 2013. Natural variation of the *RICE FLOWERING LOCUS T 1* contributes to flowering time divergence in rice. *PLOS ONE* 8:e75959
121. Oliva R, Ji C, Atienzagrande G, Huguettapia JC, Perezquintero AL, et al. 2019. Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* 37:1344–50
122. Ouyang Y, Zhang Q. 2013. Understanding reproductive isolation based on the rice model. *Annu. Rev. Plant Biol.* 64:111–35
123. Paixao JFR, Gillet F, Ribeiro TP, Bournaud C, Lourencotessutti IT, et al. 2019. Improved drought stress tolerance in *Arabidopsis* by CRISPR/dCas9 fusion with a histone acetyltransferase. *Sci. Rep.* 9:8080
124. Pan Q, Xu Y, Li K, Peng Y, Zhan W, et al. 2017. The genetic basis of plant architecture in 10 maize recombinant inbred line populations. *Plant Physiol.* 175:858–73
125. Papanatsiou M, Petersen J, Henderson L, Wang Y, Christie JM, Blatt MR. 2019. Optogenetic manipulation of stomatal kinetics improves carbon assimilation, water use, and growth. *Science* 363:1456–59
126. Papikian A, Liu W, Gallegobartolome J, Jacobsen SE. 2019. Site-specific manipulation of *Arabidopsis* loci using CRISPR-Cas9 SunTag systems. *Nat. Commun.* 10:729
127. Peiffer JA, Romay MC, Gore MA, Flintgarcia S, Zhang Z, et al. 2014. The genetic architecture of maize height. *Genetics* 196:1337–56
128. Peng J, Richards DE, Hartley NM, Murphy G, Devos KM, et al. 1999. 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400:256–61
129. Pourkheirandish M, Hensel G, Kilian B, Senthil N, Chen G, et al. 2015. Evolution of the grain dispersal system in barley. *Cell* 162:527–39
130. Ramsay L, Comadran J, Druka A, Marshall D, Thomas WTB, 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.* 43:169–72
131. Ray JD, Kilen TC, Abel CA, Paris RL. 2003. Soybean natural cross-pollination rates under field conditions. *Environ. Biosaf. Res.* 2:133–38
132. Rick CM, Holle M, Thorp RW. 1978. Rates of cross-pollination in *Lycopersicon pimpinellifolium*: impact of genetic variation in floral characters. *Plant Syst. Evol.* 129:31–44
133. Ruan B, Shang L, Zhang B, Hu J, Wang Y, et al. 2020. Natural variation in the promoter of *TGW2* determines grain width and weight in rice. *New Phytol.* 227:629–40
134. Sakuma S, Golan G, Guo Z, Ogawa T, Tagiri A, et al. 2019. Unleashing floret fertility in wheat through the mutation of a homeobox gene. *PNAS* 116:5182–87
135. Salvi S, Sponza G, Morgante M, Tomes DT, Niu X, et al. 2007. Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *PNAS* 104:11376–81
136. Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, et al. 2002. Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature* 416:701–2
137. Sato K, Yamane M, Yamaji N, Kanamori H, Tagiri A, et al. 2016. Alanine aminotransferase controls seed dormancy in barley. *Nat. Commun.* 7:11625
138. Sazonovs A, Barrett JC. 2018. Rare-variant studies to complement genome-wide association studies. *Annu. Rev. Genom. Hum. Genet.* 19:97–112
139. Schaefer RJ, Michno JM, Jeffers J, Hoekenga O, Dilkes B, et al. 2018. Integrating coexpression networks with GWAS to prioritize causal genes in maize. *Plant Cell* 30:2922–42
140. Schaid DJ, Chen W, Larson NB. 2018. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat. Rev. Genet.* 19:491–504
141. Sey NYA, Hu B, Mah W, Fauni H, McAfee JC, et al. 2020. A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. *Nat. Neurosci.* 23:583–93
142. Shen B, Wang L, Lin X, Yao Z, Xu H, et al. 2019. Engineering a new chloroplastic photorespiratory bypass to increase photosynthetic efficiency and productivity in rice. *Mol. Plant* 12:199–214
143. Singh BD, Singh AK. 2015. Mapping populations. In *Marker-Assisted Plant Breeding: Principles and Practices*, pp. 125–50. New Delhi: Springer India

140. Introduces the influencing factors and statistical approaches to fine-map functional genetic variants.

144. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. 2013. Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* 14:483–95
145. South PF, Cavanagh AP, Liu HW, Ort DR. 2019. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* 363:eaat9077
146. Soyk S, Lemmon ZH, Oved M, Fisher J, Liberatore KL, et al. 2017. Bypassing negative epistasis on yield in tomato imposed by a domestication gene. *Cell* 169:1142–55.e12
147. Soyk S, Lemmon ZH, Sedlazeck FJ, Jimenezgomez JM, Alonge M, et al. 2019. Duplication of a domestication locus neutralized a cryptic variant that caused a breeding barrier in tomato. *Nat. Plants* 5:471–79
148. Studer AJ, Wang H, Doebley JF. 2017. Selection during maize domestication targeted a gene network controlling plant and inflorescence architecture. *Genetics* 207:755–65
149. Studer AJ, Zhao Q, Ross-Ibarra J, Doebley J. 2011. Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat. Genet.* 43:1160–63
150. Sugimoto K, Takeuchi Y, Ebana K, Miyao A, Hirochika H, et al. 2010. Molecular cloning of *Sdr4*, a regulator involved in seed dormancy and domestication of rice. *PNAS* 107:5792–97
151. Sun H, Qian Q, Wu K, Luo J, Wang S, et al. 2014. Heterotrimeric G proteins regulate nitrogen-use efficiency in rice. *Nat. Genet.* 46:652–56
152. Sun J, Liu H, Liu J, Cheng S, Peng Y, et al. 2019. CRISPR-Local: a local single-guide RNA (sgRNA) design tool for non-reference plant genomes. *Bioinformatics* 35:2501–3
153. Takagi H, Abe A, Yoshida K, Kosugi S, Natsume 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.* 74:174–83
154. Takahashi Y, Shomura A, Sasaki T, Yano M. 2001. *Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the α subunit of protein kinase CK2. *PNAS* 98:7922–27
155. Takeda T, Suwa Y, Suzuki M, Kitano H, Ueguchitanaka M, et al. 2003. The *OsTB1* gene negatively regulates lateral branching in rice. *Plant J.* 33:513–20
156. Tan L, Li X, Liu F, Sun X, Li C, et al. 2008. Control of a key transition from prostrate to erect growth in rice domestication. *Nat. Genet.* 40:1360–64
157. Tanksley SD, Nelson JC. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92:191–203
158. Tao Y, Zhao X, Mace E, Henry R, Jordan D. 2019. Exploring and exploiting pan-genomics for crop improvement. *Mol. Plant* 12:156–69
159. Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, et al. 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* 43:159–62
160. Tian J, Wang C, Xia J, Wu L, Xu G, et al. 2019. Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. *Science* 365:658–64
161. Tieman D, Zhu G, Resende MF Jr., Lin T, Nguyen C, et al. 2017. A chemical genetic roadmap to improved tomato flavor. *Science* 355:391–94
162. Torada A, Koike M, Ogawa T, Takenouchi Y, Tadamura K, et al. 2016. A causal gene for seed dormancy on wheat chromosome 4A encodes a MAP kinase kinase. *Curr. Biol.* 26:782–87
163. Tuinstra MR, Ejeta G, Goldsbrough PB. 1997. Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. *Theor. Appl. Genet.* 95:1005–11
164. Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005. The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310:1031–34
165. Vilhjálmsson BJ, Nordborg M. 2013. The nature of confounding in genome-wide association studies. *Nat. Rev. Genet.* 14:1–2
166. Voichek Y, Weigel D. 2020. Identifying genetic variants underlying phenotypic variation in plants without complete genomes. *Nat. Genet.* 52:534–40
167. Wang C, Liu Q, Shen Y, Hua Y, Wang J, et al. 2019. Clonal seeds from hybrid rice by simultaneous genome engineering of meiosis and fertilization genes. *Nat. Biotechnol.* 37:283–86
168. Wang C, Tang S, Zhan Q, Hou Q, Zhao Y, et al. 2019. Dissecting a heterotic gene through GradedPool-Seq mapping informs a rice-improvement strategy. *Nat. Commun.* 10:2982
169. Wang H, Sun S, Ge W, Zhao L, Hou B, et al. 2020. Horizontal gene transfer of *Fhb7* from fungus underlies *Fusarium* head blight resistance in wheat. *Science* 368:eaba5435

160. Provides a classic example that a teosinte allele lost during domestication could significantly improve modern maize yields.

165. Explains the true underlying sources of, and how to effectively address, the confounding in GWAS.

170. Wang J, Lin Z, Zhang X, Liu H, Zhou L, et al. 2019. *krn1*, a major quantitative trait locus for kernel row number in maize. *New Phytol*. 223:1634–46
171. Wang J, Long X, Chern M, Chen X. 2020. Understanding the molecular mechanisms of trade-offs between plant growth and immunity. *Sci. China Life Sci*. In press. **https://doi.org/10.1007/s11427-020-1719-y**
172. Wang J, Yu H, Xiong G, Lu Z, Jiao Y, et al. 2017. Tissue-specific ubiquitination by IPA1 INTERACTING PROTEIN1 modulates IPA1 protein levels to regulate plant architecture in rice. *Plant Cell* 29:697–707
173. Wang J, Zhou L, Shi H, Chern M, Yu H, et al. 2018. A single transcription factor promotes both yield and immunity in rice. *Science* 361:1026–28
174. Wang L, Sun S, Jin J, Fu D, Yang X, et al. 2015. Coordinated regulation of vegetative and reproductive branching in rice. *PNAS* 112:15504–9
175. Wang M, Li W, Fang C, Xu F, Liu Y, et al. 2018. Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nat. Genet.* 50:1435–41
176. Wang M, Yu Y, Haberer G, Marri PR, Fan C, et al. 2014. The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nat. Genet.* 46:982–88
177. Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, et al. 2018. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557:43–49
178. Wang X, Wang H, Liu S, Ferjani A, Li J, et al. 2016. Genetic variation in *ZmVPP1* contributes to drought tolerance in maize seedlings. *Nat. Genet.* 48:1233–41
179. Wang Y, Wang H, Wei L, Li S, Liu L, Wang X. 2020. Synthetic promoter design in *Escherichia coli* based on a deep generative network. *Nucleic Acids Res*. 48:6403–12
180. Watanabe S, Hideshima R, Xia Z, Tsubokura Y, Sato S, et al. 2009. Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics* 182:1251–62
181. Watanabe S, Xia Z, Hideshima R, Tsubokura Y, Sato S, et al. 2011. A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics* 188:395–407
182. Watson A, Ghosh S, Williams MJ, Cuddy WS, Simmonds J, et al. 2018. Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 4:23–29
183. Wei X, Xu J, Guo H, Jiang L, Chen S, et al. 2010. *DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol*. 153:1747–58
184. Wills DM, Whipple CJ, Takuno S, Kursel LE, Shannon LM, et al. 2013. From many, one: genetic control of prolificacy during maize domestication. *PLOS Genet*. 9:e1003604
185. Wing RA, Purugganan MD, Zhang Q. 2018. The rice genome revolution: from an ancient grain to Green Super Rice. *Nat. Rev. Genet.* 19:505–17
186. Wu W, Zheng X, Lu G, Zhong Z, Gao H, et al. 2013. Association of functional nucleotide polymorphisms at *DTH2* with the northward expansion of rice cultivation in Asia. *PNAS* 110:2775–80
187. Xia Z, Watanabe S, Yamada T, Tsubokura Y, Nakashima H, et al. 2012. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *PNAS* 109:12852–53
188. Xiao Y, Liu H, Wu L, Warburton M, Yan J. 2017. Genome-wide association studies in maize: praise and stargaze. *Mol. Plant* 10:359–74
189. Xiao Y, Tong H, Yang X, Xu S, Pan Q, et al. 2016. Genome-wide dissection of the maize ear genetic architecture using multiple populations. *New Phytol*. 210:1095–106
190. Xu J, Chen G, Hermanson PJ, Xu Q, Sun C, et al. 2019. Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. *Genome Biol*. 20:243
191. Xu K, Xu X, Fukao T, Canlas PE, Maghirang-Rodriguez R, et al. 2006. *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* 442:705–8
192. Xu S. 2003. Theoretical basis of the Beavis effect. *Genetics* 165:2259–68
193. Xue W, Xing Y, Weng X, Zhao Y, Tang W, et al. 2008. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* 40:761–67

194. Yan L, Fu D, Li C, Blechl AE, Tranquilli G, et al. 2006. The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *PNAS* 103:19581–86
195. Yan L, Loukoianov A, Blechl AE, Tranquilli G, Ramakrishna W, et al. 2004. The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303:1640–44
196. Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J. 2003. Positional cloning of the wheat vernalization gene *VRN1*. *PNAS* 100:6263–68
197. Yan WH, Wang P, Chen HX, Zhou HJ, Li QP, et al. 2011. A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol. Plant* 4:319–30
198. Yang CJ, Kursel LE, Studer AJ, Bartlett ME, Whipple CJ, Doebley J. 2016. A gene for genetic background in *Zea Mays*: fine-mapping *enhancer of teosinte branched1.2* to a YABBY class transcription factor. *Genetics* 204:1573–85
199. Yang N, Liu J, Gao Q, Gui S, Chen L, et al. 2019. Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement. *Nat. Genet.* 51:1052–59
200. Yang Q, He Y, Kabahuma M, Chaya T, Kelly A, et al. 2017. A gene encoding maize caffeoyl-CoA *O*-methyltransferase confers quantitative resistance to multiple pathogens. *Nat. Genet.* 49:1364–72
201. Yang Q, Li Z, Li W, Ku L, Wang C, et al. 2013. CACTA-like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *PNAS* 110:16969–74
202. Yu J, Holland JB, McMullen MD, Buckler ES. 2008. Genetic design and statistical power of nested association mapping in maize. *Genetics* 178:539–51
203. Yu X, Zhao Z, Zheng X, Zhou J, Kong W, et al. 2018. A selfish genetic element confers non-Mendelian inheritance in rice. *Science* 360:1130–32
204. Zhang H, Wang X, Pan Q, Li P, Liu Y, et al. 2019. QTG-seq accelerates QTL fine mapping through QTL partitioning and whole-genome sequencing of bulked segregant samples. *Mol. Plant* 12:426–37
205. Zhang L, Yu H, Ma B, Liu G, Wang J, et al. 2017. A natural tandem array alleviates epigenetic repression of *IPA1* and leads to superior yielding rice. *Nat. Commun.* 8:14789
206. Zhang Y, Lubberstedt T, Xu M. 2013. The genetic and molecular basis of plant resistance to pathogens. *J. Genet. Genom.* 40:23–35
207. Zhang Z, Li J, Tang Z, Sun X, Zhang H, et al. 2018. Gnp4/LAX2, a RAWUL protein, interferes with the OsIAA3–OsARF25 interaction to regulate grain length via the auxin signaling pathway in rice. *J. Exp. Bot.* 69:4723–37
208. Zhao Y, Xie P, Guan P, Wang Y, Li Y, et al. 2019. *Btr1-A* induces grain shattering and affects spike morphology and yield-related traits in wheat. *Plant Cell Physiol.* 60:1342–53
209. Zhu G, Wang S, Huang Z, Zhang S, Liao Q, et al. 2018. Rewiring of the fruit metabolome in tomato breeding. *Cell* 172:249–61.e12
210. Zhu Q, Zeng D, Yu S, Cui C, Li J, et al. 2018. From golden rice to aSTARice: bioengineering astaxanthin biosynthesis in rice endosperm. *Mol. Plant* 11:1440–48
211. Zhu YJ, Fan YY, Wang K, Huang DR, Liu WZ, et al. 2017. Rice *Flowering Locus T 1* plays an important role in heading date influencing yield traits in rice. *Sci. Rep.* 7:4918
212. Zsogon A, Cermak T, Naves ER, Notini MM, Edel KH, et al. 2018. De novo domestication of wild tomato using genome editing. *Nat. Biotechnol.* 36:1211–16