New horizons for dissecting epistasis in crop quantitative trait variation

Sebastian Soyk[1,*], Matthias Benoit[2,3], and Zachary B. Lippman[2,3,*]

[1]Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland
[2]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, United States
[3]Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, United States

*Correspondence: sebastian.soyk@unil.ch, lippman@cshl.edu

ABSTRACT

A frontier in plant genetics is to uncover the genes, variants, and interactions underlying crop diversity. Phenotypic variation often does not reflect the cumulative effect of individual gene mutations. This deviation is due to epistasis, interactions between alleles that are not predictable and frequently quantitative in effect. Recent advances in genomics and genome editing technologies are elevating the study of epistasis in crops. Using traits and developmental pathways that were major targets in domestication and breeding, we highlight how epistasis is central in guiding the behavior of genetic variation that shapes quantitative trait variation. We outline new strategies that will illuminate the relationship of quantitative epistasis with modified gene dosage that defines background dependencies. Advancing our understanding of epistasis in crops can reveal new principles and approaches to engineer targeted improvements in agriculture.

1. INTRODUCTION

A major goal of contemporary plant genetics is to study how complex genotypes translate into quantitative phenotypes. Beginning with the first reported quantitative trait locus (QTL) study in plants (88), classical approaches in quantitative genetics have enabled identification of genes and mutations that underlie complex developmental and agricultural traits. More recently, advances in high-throughput sequencing technologies and computational methods have accelerated the dissection of the polygenic architectures of many traits that were selected and modified during domestication and breeding. Nevertheless, identifying specific causative genes and variants remains challenging, and thus a large proportion of quantitative variation is unexplained. This is due to the tedious identification and characterization of QTL, especially those having small effects. Adding to this is emerging evidence of a vast reservoir of potential 'cryptic' variants in plant genomes, whose phenotypic effects are only revealed in the presence of interacting loci in specific genetic backgrounds.

In this review we discuss how epistasis was critical to dissect genetic pathways in model plants, and how a broader quantitative view of epistasis is emerging to expose complex genetic phenomena in crop breeding. Using recent examples from tomato as a foundation, we explore the contributions of epistasis to major agricultural traits that drove crop domestication and improvement. We highlight how new approaches integrating genomics-enabled quantitative genetics with genome editing can reveal and characterize genetic interactions at a scale and resolution that has never been possible in crops before. We conclude by outlining outstanding questions and approaches that we propose should drive the next phase of discoveries on the role of epistasis in plant quantitative genetics.

2. A BRIEF HISTORY OF EPISTASIS FROM A PLANT GENETICS PERSPECTIVE

Coining the term 'epistasis'

In its most general meaning, epistasis is a principle in genetics describing patterns of inheritance that deviate from Mendelian segregation. The term 'epistasis' was introduced

at the beginning of the 20th century by the British geneticist William Bateson, inspired by his studies of flower color inheritance in sweet pea (2). In one of many textbook examples, Bateson crossed two genetically distinct accessions of white-flowered *Lathyrus odoratus* and obtained a homogeneous purple-flowered F1 hybrid population. Upon self-fertilization, purple and white flowers segregated at a 9:7 ratio in the F2 population. This surprising outcome deviated substantially from the expected Mendelian segregation of 9:3:3:1 – typical for two non-interacting genes that affect distinct traits where complete dominant-recessive relationships are intact between mutant and wild type alleles at each gene. Bateson explained this phenomenon as "reversion" and the consequence of combining "distinct though complementary allelomorphic pairs" (2). Today, we know his observation is a direct genetic and molecular consequence of two independent recessive mutations in anthocyanin biosynthesis genes, which cause a loss of purple flower color in both single and double mutants. Although not appreciated at the time, Bateson's work was a prelude to similar, though often more complex, genetic scenarios that can arise upon bringing together different mutations and genetic backgrounds in breeding.

Bateson's example illustrates the classical definition of epistasis in Mendelian genetics, which refers to the masking of the genotypic effect at one locus by the genotype at a different locus, manifesting in a deviation from the expected Mendelian segregation ratios in a biparental cross. A decade after Bateson, Ronald Fisher rightly expanded the definition of epistasis based on inheritance studies on human body height (23). His statistical approaches illuminated that quantitative traits are polygenic and result from the often unequal contribution of many individual genetic loci. Specifically, Fisher's statistical method for describing heritability of body height with many loci slightly deviated from a linear model that assumed cumulative (additive) effects from each locus. This statistical abnormality, which Fisher termed at the time 'epistacy', expanded epistasis to include deviation from the expected additive effect from all loci contributing to a quantitative trait (77). Fisher's epistasis is the foundation of our modern definition, in which multiple genes and variants interact to determine a quantitative range of phenotypic outcomes (17). Aided by new technologies and tools, it is becoming clear that Bateson and Fisher defined two extremes along a continuum of epistasis, from simple qualitative (digenic) to highly

complex and quantitative (polygenic) interactions. Throughout this review, we present examples of genetic interactions in plants that reconcile and link Bateson's and Fisher's epistasis. We suggest that a deeper understanding and appreciation of this epistatic continuum can help resolve mechanisms of quantitative trait variation in crop domestication and improvement, and beyond.

The contribution of epistasis to hybrid behavior

Innumerable examples of epistasis in plants for myriad traits have been documents over the last century. Among these are the genetic architectures of different hybrid behaviors, which nicely illustrate the intersection of qualitative (Bateson) and quantitative (Fisher) concepts of epistasis. Hybridization of genetically-distinct parents often yields F1 hybrid offspring that are superior to the homozygous parental genotypes – the phenomenon of hybrid vigor or 'heterosis'. However, hybridization of genetically distinct parents can also lead to inferior offspring with many facets of detrimental phenotypes, including embryo lethality and dwarfism due to 'hybrid incompatibility'. Because of the economic importance of hybrids, the genetic determinants of heterosis and hybrid incompatibility have been intensively investigated in both model and crop plants (9, 90).

In the early 1960s, the studies of John Hermsen on wheat provided a preliminary description of the genetic mechanisms underlying one form of postzygotic hybrid incompatibility known as hybrid necrosis. From a set of test crosses involving hundreds of wheat varieties, Hermsen scored the degree of spontaneous tissue necrosis, premature senescence, and death in hybrids and then analyzed the segregation ratios in later generations. Surprisingly, these experiments revealed that most cases of necrosis were due to negative epistatic interactions between only two loci. Similar simple genetic architectures for hybrid necrosis have since been described in many other crop species (9), and the molecular basis of hybrid necrosis has been dissected in Arabidopsis. A large screen for hybrid necrosis in intraspecific hybrids among hundreds of Arabidopsis accessions revealed predominantly two major QTLs whose effects were highly dependent on genotypic background (8). These loci were pinpointed to naturally occurring variants of nucleotide-binding domain and leucine-rich repeat (NLR) immune receptor genes (8).

A negative epistatic interaction between distinct NLR variants results in aberrant activation of the plant immune system, triggering an autoimmunity response. These findings provided a first molecular mechanistic explanation for the classical Bateson-Dobzhansky-Muller model of postzygotic reproductive isolation, which posited (first in Drosophila) that hybrid incompatibility results from deleterious interactions between incompatible parental alleles. Importantly, hybrid necrosis can also be quantitative, as different severities of autoimmunity were found in different Arabidopsis hybrids and their progeny, suggesting modifiers (8, 12). Notably, similar genetic interactions involving NLR or functionally-related genes have been found in cases of hybrid necrosis in other crops, including lettuce, rice, and tomato (13, 39, 46, 115). The epistasis underlying this type of hybrid behavior follows the qualitative concept of Bateson, but with notably aspects of Fisher's epistasis coming from genotype combinations displaying quantitative effects on fitness.

The other extreme of hybrid behavior is heterosis, first described by George Shull in 1908 in his seminal article 'The composition of a field of maize' (94). Shull showed that cross-hybridization of genetically distant parents results in uniform hybrid offspring with increased performance (95). An important observation was that heterotic effects varied between different parental combinations, suggesting heterosis is a polygenic trait and also depends on genetic background. Heterosis is still the subject of intense investigation, and different models, genes, alleles, and phenotypes are likely involved, depending on the crop (5, 90). Briefly, the two classical models are 'dominance' and 'overdominance'. Dominance proposes that inbred parents carry different sets of slightly deleterious alleles, which complement each other upon hybridization. In contrast, overdominance proposes that there may be something special about heterozygosity and synergistic interactions between different alleles at one or more loci that causes hybrids to outperform parental genotypes.

Given the quantitative and polygenic nature of heterosis, it is not surprising that epistasis is involved, which may fall under the more recent gene dosage hypothesis for heterosis (4). The contribution of epistasis to heterosis has been reported in multiple plants, with

genetic architectures of varying complexities depending on the trait assessed, and contributions from dominance and overdominance (27, 30, 41, 45, 68). The gene dosage hypothesis proposes that modified levels of gene products acting in multiprotein complexes could optimize molecular function, and thus vigor, in unpredictable ways due to the inter-connectedness and feedbacks between protein complexes, their targets, and expression levels. In the context of heterosis, changes in gene dosage from heterozygosity at one or more loci affects the quantity or functional output of gene products, which translates into quantitative phenotypic changes. One can then extrapolate how epistasis may be intimately connected to the gene dosage model, as genetic variation at one or more points in pathways and circuits affected by altered dosage would influence quantitative phenotypic outcomes. This could perhaps explain in part why magnitudes of heterosis vary in different parental combinations.

One finely dissected heterotic effect that incorporates elements of both Bateson and Fisher epistasis involves a quantitative, dose-dependent interaction between the genes encoding the main tomato florigen and antiflorigen protein hormones (59). Loss-of-function mutations in the flower promoting florigen hormone gene *SINGLE FLOWER TRUSS* (*SFT*) delay flowering in favor of vegetative growth, while mutations in the flower repressing hormone *SELF PRUNING* (*SP*) gene lead to earlier flowering on side shoots and the transformation of a continuously growing ‘indeterminate’ growth habit to a compact ‘determinate’ architecture. Congruent with the epistasis concept of Batson, *SFT* is completely epistatic over *SP*, with strong *sft* mutations masking the effects of *sp*. However, elements from Fisher’s concept become apparent in *sft* heterozygotes, which show a quantitative dose-dependent suppression in plant size and determinate growth that ultimately increases fruit yield compared to both homozygous parents (40, 45). This remarkable heterotic effect can be interpreted as true single-gene overdominance with respect to *sft* heterozygosity, but depends on backgrounds that are homozygous for *sp*, revealing a critical epistatic component (45). The florigen-antiflorigen relationship is highly dose-sensitive, with heterozygosity for weak mutations in *SFT* and also an additional interacting factor in the florigen complex allowing quantitative tuning and optimization of the heterotic effect (40, 75). This example highlights how epistasis involving only a few

genes and alleles can confer a continuum of quantitative trait variation. Even more, the heterotic effects from these combined mutations can be influenced by additional unknown modifiers in different backgrounds (Soyk, Benoit, and Lippman, unpublished), thereby adding another quantitative, and likely epistatic, layer to this particular case of heterosis. Interestingly, epistasis in the florigen pathway is also associated with heterosis for yield in rice, which is modified by allelic variation at multiple florigen pathway genes (30). Although the polygenic nature has hindered further molecular characterization, this example of heterosis illustrates the spectrum of digenic to polygenic interactions that can influence quantitative trait variation, and the fluent transition between the concepts of Bateson and Fisher.

3. GENETIC INTERACTIONS IN CROP DOMESTICATION AND BREEDING: RECENT INSIGHTS FROM TOMATO

There are several examples of epistasis in fundamental developmental and biochemical pathways that were selected to diversify traits in crop domestication and breeding. Domestication studies have been most successful at identifying mutations in genes with large phenotypic effects, reflecting relatively simple genetic architectures (22). However, most domestication traits also involve additional smaller effect loci, indicating a more complex polygenic foundation and that genetic interactions could have contributed to plant domestication more than originally thought (102, 103).

Below, we present examples of epistatic interactions affecting quantitative traits that were major targets during domestication in tomato (**Figure 1**). We focus primarily on tomato, because of its rich history in dissecting quantitative traits, and also because tomato has risen above other systems in leveraging genomics and genome editing technologies to dissect mechanisms underlying quantitative trait variation, particularly those involving epistasis. These examples also illustrate how exposing epistatic interactions in developmental pathways that were modified during crop domestication were and will continue to be a rich resource of gene targets for crop breeding.

Genetic interactions in fruit traits

Diversification in the color, shape and size of fleshy edible fruits has been a major goal during domestication of fruit crops and remains a major focus for contemporary breeding (31). The molecular basis of this tremendous variation has exposed epistatic interactions among genes and alleles that have influenced fruit phenotypes.

The characteristic red color of tomato fruits from the accumulation of carotenoids and other antioxidants has an aesthetic value but also provides health benefits (44). A genetic conundrum involving epistasis in the fruit coloring pathway was first described more than 100 years ago. The recessive *yellow-flesh* mutant causes yellow fruits from a deficiency in a rate-limiting enzyme in the carotenoid biosynthesis pathway (24). A second recessive mutant, *tangerine*, is defective in another enzyme in the same pathway, and results in orange fruits due to accumulation of an alternative carotenoid. The enzyme encoded by *tangerine* functions downstream of the enzyme encoded by *yellow flesh* (32). Surprisingly, *tangerine* mutations are epistatic to *yellow-flesh*, with *yellow-flesh tangerine* double mutants developing orange fruits (37, 38, 106). This paradox was resolved when *yellow-flesh* mutation was mapped to the *cis*-regulatory region upstream of its coding sequence that leads to a loss of expression. However, in *yellow-flesh tangerine* double mutants, despite the *cis*-regulatory mutation, expression of *yellow-flesh* is nearly 200-fold higher than in the *yellow-flesh* single mutants, which sustains carotenoid biosynthesis (42). Though the molecular mechanism remains elusive, transcriptional rescue of *yellow-flesh* is detected in different *tangerine* alleles independent of genetic background, suggesting a specific digenic interaction between *yellow-flesh* and *tangerine* in the maintenance of carotenoid biosynthesis. While this example presents as a clear case of Bateson's qualitative epistasis, the regulatory nature of one of the interacting alleles hints at the potential for dosage-dependency and quantitative regulation of fruit coloring within Fisher's model of quantitative epistasis.

The shape of tomato fruits has been under intense selection during domestication and improvement. Two major loci that promote fruit elongation are *sun* and *ovate* (87, 113). The *sun* locus is the result of a transposon-mediated gene duplication that results in

increased expression of a growth regulator. The *ovate* locus carries a premature stop codon in a transcriptional repressor that restricts cellular growth. Although *sun* and *ovate* are widely distributed in the modern tomato germplasm and cause elongated fruits in most genetic backgrounds, quantitative modulation of *sun* and *ovate* phenotypes can occur in different genotypic backgrounds, suggesting epistatic interactions with weaker effect modifier fruit shape QTLs (87)**.** For example, accessions that carry *ovate* display a quantitative range in fruit shapes, including elongated, pear-shaped to round fruits. These background dependencies rely on two major QTLs, *suppressor of ovate1* (*sov1*) and *sov2* (86). Perhaps not surprisingly, the *sov1* locus is a variant in *SlOFP20*, another member of the *OVATE FAMILY PROTEIN* (*OFP*) gene family (112). Ectopic expression of *SlOFP20* suppresses *ovate* fruit elongation in the manner of Bateson epistasis. Conversely, a natural 31-Kbp deletion upstream of *SlOFP20* is associated with a quantitative reduction in expression and enhanced fruit elongation in *ovate* backgrounds (112). Altogether this suggests a continuous, dosage-sensitive epistatic relationship between *OVATE* and *SlOFP20*. The absence of wider allelic diversity at both *OVATE* and *SlOFP20* have for now impeded further dissection of quantitative epistatic relationships between these genes, and highlights the challenges in the detection and dissection of Fisher's epistasis.

Yield enhancement through increases in fruit size and weight have been central to tomato domestication, and a major determinant of these traits is the size of meristems. Shoot apical meristems are groups of stem cells that give rise all aerial organs. Meristem size is controlled through a negative feedback circuit involving the small signaling peptide CLAVATA3 (CLV3), its receptor kinase CLAVATA1 (CLV1), and the stem-cell promoting homeodomain transcription factor WUSCHEL (WUS), which is repressed by the CLV module and promotes *CLV3* expression (96, 97). The CLV-WUS circuit is highly conserved in flowering plants (25). First unraveled in Arabidopsis, loss of *WUS* leads to depletion of stem cells and the premature termination of meristems. Disruption of *CLV3* or *CLV1* (and its redundant family members) causes stem cell over-proliferation that increases meristem size and organ number, including in fruits. *WUS* is epistatic over *CLV1* and *CLV3*, with all *wus clv* double mutants resembling *wus* single mutants (91).

Although these early findings follow Bateson's epistasis, recent studies on CLV-WUS peptide and receptor components in crops suggest that meristem maintenance is controlled through more complex genetic interactions and can be quantitatively tuned (96).

In tomato, mutations in CLV-WUS circuit genes contributed to fruit size increases during domestication. The *fasciated* (*fas*) mutation is a weak loss-of-function mutation of the tomato *CLV3* ortholog (*SlCLV3*) originating from a 294 Kbp inversion that partially disrupts the promoter and reduces expression (114). The *locule number* (*lc*) mutation is a weak gain-of-function allele of tomato *WUS* (*SlWUS*) caused by two SNPs in a putative *cis*-regulatory element located downstream that presumably increases expression (72, 108, 114). Both *fas* and *lc* cause the development of additional seed compartments (locules) that results in an increase in fruit size and higher yields. The effect of *fas* is stronger than *lc,* and their combination increases locule number and fruit size non-additively (15, 85, 114). This quantitative epistatic effect from *fas* and *lc* is consistent with classical epistasis within the Arabidopsis CLV-WUS circuit, and again demonstrates how studying interactions between different allelic strengths reveals the link between Bateson's and Fisher's epistasis.

These links became even more apparent upon further dissection of the tomato CLV-WUS circuit using genome editing (84, 114). CRISPR-generated *SlCLV3* null mutants (*slclv3*^{CR}) develop severely enlarged meristems and many more fruit locules compared to the weaker *fas* mutant. Remarkably, loss of *SlCLV3* function in *slclv3*^{CR} triggers transcriptional upregulation of its closest paralog, *CLV3/Embryo Surrounding Region* (*SlCLE9*), suggesting an active compensatory mechanism through upregulation of a functionally related gene (19). The compensating function of *SlCLE9* becomes apparent in *slclv3*^{CR} *slcle9*^{CR} double mutants. While *slcle9*^{CR} mutants have no phenotype, combining with *slclv3*^{CR} results in exceptionally large meristems and fruits with twice as many locules compared to *slclv3*^{CR}. This epistatic mechanism, involving genetic compensation and dosage effects from *SlCLV3* and *SlCLE9*, has thus quantitatively modulated fruit size increases from *fas* during tomato domestication. Together with

reports for background dependence of *fas* and *lc* (87) and parallels with *CLV* homologs in maize (36, 84), the CLV-WUS circuit and surrounding epistatic compensators illustrate the contribution of Fisher's polygenic, dose-dependent epistasis to agriculturally important quantitative trait variation in crop plants.

Genetic interactions in shoot architecture

Diversity in shoot architecture is a major agricultural trait that is largely determined by the transition from vegetative to reproductive growth (flowering). Plants transition to flowering when shoot meristems cease production of vegetative organs and transition to producing flowers, fruits, and seeds. This process of meristem maturation is central to balancing vegetative and reproductive growth, and thus its modification, either genetically or environmentally, influences plant architecture (50, 74). This explains the prominent role for adaptations of flowering transitions during crop domestication and improvement, which allowed cultivation in wider geographical regions compared to wild crop relatives. Consequently, genes that encode components of flowering pathways were recurrent targets of selection in many crop species (6, 16, 33, 48).

An illustrative example is the adaptation of the universal flowering hormone (florigen) pathway in the domestication and breeding of modern cultivated tomato (*S. lycopersicum*). The wild ancestor of tomato (*S. pimpinellifolium*) and other closely related wild species only flower readily when light periods mimic the short days of their native habitats near the equator (100, 117). This response to day length was strongly mitigated during domestication, resulting in cultivated varieties that are nearly day-neutral and facilitating production in long days at northern latitudes. This loss of day-length sensitive flowering is largely based on two interacting QTLs that harbor two antagonistic florigen genes (100). The major QTL is a *cis*-regulatory mutation in a florigen homolog *SELF PRUNING 5G* (*SP5G*) that represses flowering in long days (100, 117). The gene underlying the second QTL has not been dissected, but maps to the florigen gene *SFT*, and interacts synergistically (i.e. epistatically) with *SP5G* to delay flowering in long days (100).

After the flowering transition, tomato growth continues by cycling between reproductive and vegetative growth, leading to indeterminate shoots that rapidly occupy the surrounding habitat and compete for resources with neighboring plants. Breeding yielded a spontaneous mutation in the antiflorigen gene *SP*, a homolog of the Arabidopsis antiflorigen T*ERMINATING FLOWER1* (*TFL1*) (79). The *sp* mutation transformed tomato from indeterminate vines into a determinate row crop suitable for field cultivation. In addition, plant maturation is accelerated in *sp* mutants, which causes faster flowering on side shoots and near synchronous fruit set and ripening, which is advantageous for mechanical harvesting in large-scale production. As discussed in the context of heterosis, epistatic interactions between *SP* with other florigen pathway genes allowed modification of *sp* determinacy. Again, heterozygosity for mutations in the main florigen *SFT* quantitatively suppresses *sp* determinate growth, allowing additional inflorescences and fruits to develop (40, 45). Similar but weaker effects result from mutations in *SSP*, encoding a transcription factor that interacts with florigen in a multimeric complex to regulate the expression of flowering transition genes (75, 105). Together, these dose-sensitive epistatic interactions among several florigen pathway genes can be harnessed to quantitatively manipulate and fine-tune tomato shoot architecture and yield (75, 100). Importantly, such manipulations for agricultural benefits could not have been realized without the background *cis*-regulatory domestication mutation in the antiflorigen gene *SP5G,* adding another layer of quantitative Fisher epistatic complexity to this system. Not surprisingly, florigen pathway genes and epistatic interactions among them were the foundation for recent steps towards the *de novo* domestication of wild *Solanaceae* species and the development of urban agriculture tomatoes by genome editing (49, 55, 57, 119).

Notably, revisiting Arabidopsis florigen (*FT*) and *TFL1* revealed a conserved dose-dependent epistatic relationship between these two opposing flowering hormone genes (35, 40), suggesting opportunities to fine-tune flowering by exploiting epistasis in the florigen systems of other crops, such as rice and soybean (30, 78)**.** Indeed, genomic regions associated with heterosis in rice harbor heterozygous mutations in the florigen gene homolog *Hd3a* (30). Interestingly, this heterosis effect is highly background

dependent, perhaps due to interactions with alleles of other florigen pathway genes that are present in different rice accessions (30). In soybean (*Glycine max*), mutations in florigen pathway genes altered growth habit during domestication and improvement. The indeterminate wild ancestor (*Glycine soya*) was naturally transformed into a determinate form, and early studies of soybean stem termination using hybrid populations between indeterminate and determinate lines led to the identification of two major alleles, the recessive *dt1* and dominant *Dt2* (3). A greater effect comes from *dt1*, which is mutated in the antiflorigen homolog *GmTFL1b* (64). Interestingly, *dt1* mutations are dose-sensitive, with *dt1* heterozygotes producing semi-determinate architectures (64). *Dt2* encodes a MADS-box transcription factor gene, and expression studies of different *dt2* genotypes suggest that the dominant *Dt2* allele downregulates *Dt1* (78). Beyond these genes are regulators of florigen expression, such as light receptor and circadian genes. Natural mutations in these genes were selected by breeders to quantitatively adjust flowering time and determinate growth to adapt soybean for different growing regions (110). Here again, the importance of Fisher's epistasis emerges in crop domestication and improvement.

Variation in florigen homologs and flowering pathways has been important in domestication and improvement of many crops (6, 16, 33, 48), and unexplored epistasis among these factors may offer new avenues to fine-tune and improve agricultural productivity.

Genetic interactions in inflorescence architecture

A major determinant of plant fitness in nature and agriculture is inflorescence architecture. One of the most spectacular inflorescence modifications that arose from domestication are the heads of broccoli (*B. oleracea ssp. italica*) and cauliflower (*B. oleracea ssp. botrytis*). Population and molecular genetics revealed an association of the cauliflower phenotype with a nonsense mutation in a MADS-box gene *BoCAL* (43, 80). However, the nonsense variant was also detected in wild cabbage (*B. oleracea ssp oleracea*) and kale (*B. oleracea ssp. acephala*), which develop regular inflorescences. This discrepancy indicates that *BoCAL* mutation is not sufficient for the cauliflower phenotype and that

additional interacting modifier loci are involved, which is supported by the identification of 67 QTL contributing to the cauliflower head (51).

An increase in inflorescence branching to improve flower and grain production has been a recurring target during domestication and improvement of several cereal crops including barley, maize, rice, and wheat (7, 21, 29, 81). However, for many fruit crops such as tomato and grape, inflorescences architecture remained largely unchanged from their wild ancestors (71, 76). Domesticated tomato and the wild ancestor *S. pimpinellifolium* develop multi-flowered inflorescences with several flowers arranged along a single branch. Several wild tomato relatives with weakly branched inflorescences exist but have been underexplored due to genetic incompatibilities and the polygenic nature of this trait (54, 66). There are accessions of domesticated tomato that develop highly branched inflorescences with hundreds of flowers due to natural mutations in the homeobox gene *COMPOUND INFLORESCENCE* (*S*) (63). However, such accessions with excessively branched inflorescences set fruit poorly, likely due to imbalances in source-sink relationships (101), and thus have been largely avoided by breeders.

Historical reports about breeding with mutations that improved tomato harvestability from a loss of the fruit abscission zone (the 'joint') alluded to branched inflorescences (82, 83). Introducing the *jointless-2* (*j2*) mutation into specific genotypes was problematic due to interacting modifier loci that caused undesirable branched inflorescences with reduces fruit yields (82, 83). More than 50 years later, both the *j2* mutation and the natural modifier mutation, termed *enhancer of j2* (*ej2*), were found to be mutated in two closely-related MADS-box genes (98). An intronic insertion in *J2* causes a complete loss of function, while an intronic insertion in *EJ2* leads to partial mis-splicing of the gene and a quantitative reduction of functional *EJ2* transcript. This results in a weak allele ($ej2^{W}$) that causes a quantitative elongation of sepals, the leafy organs on the flowers. However, when *j2* and $ej2^{W}$ are combined, inflorescences become excessively branching and fruit set is reduced. Analyses of $ej2^{W}$ allele frequencies indicated that the missplicing mutation arose early in domestication and became widespread in the domesticated germplasm, where it

collided with *j2* in a negative epistatic interaction during modern breeding and made accessions prone to undesirable inflorescence branching (98).

Genome editing has allowed a finer dissection of these interactions, including a role for a third MADS-box gene that helped expose one of the most informative cases linking Bateson and Fisher epistasis. CRISPR-engineered $j2^{CR}$ null mutations recapitulated the natural *j2* null and caused jointless fruits. Importantly, engineered $ej2^{CR}$ null mutations resulted in extremely elongated sepals compared to the natural $ej2^{W}$ allele, revealing a dosage effect from the natural $ej2^{W}$ missplicing allele. Combining both $j2^{CR}$ and $ej2^{CR}$ null mutations caused excessive inflorescence branching above the natural *j2* $ej2^{W}$ mutant, showing that dosage of *EJ2* also quantitatively modulates inflorescence branching. Engineering mutations in a closely related MADS-box gene (*LIN, LONG INFLORESCENCE*) result in elongated inflorescences with additional flowers and weak branching. Remarkably, lin^{CR} mutations further enhanced $j2^{CR}ej2^{CR}$ double mutants, with $j2^{CR}ej2^{CR}lin^{CR}$ triple mutants developing massively overproliferated inflorescence meristems without flowers. Thus, three doses of MADS box genes contribute to normal inflorescence development, and the serial loss of each gene results in progressively more severe branching. An even finer dosage relationship was revealed upon creating homozygous by heterozygous combinations among *j2* and *ej2* alleles, which produced a continuum of inflorescence complexity. Notably, *j2* $ej2^{W}$/+ hybrids developed weakly branched inflorescences that developed additional flowers but maintained high fertility, resulting in a heterotic effect for fruit yield (98).

Interestingly, negative epistasis between *j2* and $ej2^{W}$ is suppressed in specific breeding lines by the two *suppressor of branching1* (*sb1*) and (*sb3*) QTLs (99). The major effect QTL *sb3* contains a tandem duplication of the weak $ej2^{W}$ missplicing allele, which leads to an increase in *EJ2* expression to exceed a critical threshold of functional *EJ2* transcript for suppressing inflorescence branching in breeding lines (99). Remarkably, the *sb1* locus also contains a copy number variant of a MADS-box gene, a homolog of the Arabidopsis flowering and meristem identity regulator *SUPPRESSOR OF CONSTANS1* (*SOC*) (53). In breeding lines, lower copy number of this gene leads to reduced expression and a

quantitative suppression of inflorescence branching (Alonge et al., 2020, in press). Interestingly, the *sb1* and *sb3* copy number variants were present as cryptic variants in the tomato germplasm well before negative epistasis from *j2* and *ej2*^{W} emerged in modern breeding. Thus, breeders used this standing variation to stabilize inflorescence architecture, by taking advantage of the quantitative epistatic interactions among these MADS-box genes alleles.

4. NEW APPROACHES TO STUDY EPISTASIS IN CROPS

The examples of epistasis above reveal the continuum between Bateson's and Fisher's epistasis in crop domestication and breeding. They also raise new and exciting questions that highlight the challenges in elevating and expanding epistasis studies for both fundamental and applied value. For example, can epistatic loci underlying background effects across diverse germplasm resources be identified to facilitate predictable breeding? Is it possible to dissect, at scale, gene dosage relationships between two or more interacting genes that quantitatively modulate phenotypic variation? Can redundancy and compensation be untangled to refine fundamental understandings of genetic network dynamics relevant in crop breeding? These and related questions are the topic of our final section, where for simplicity and focus, we draw again from examples in tomato to illustrate how advances in genome editing have opened new opportunities to reveal epistasis that has shaped crop trait diversity and could be harnessed for crop engineering. We propose approaches that leverage genomics and genome editing to study epistasis imposed by genetic backgrounds, particularly to reveal cryptic variants whose functional relevance are only exposed when combined with other specific mutations. We further discuss the value in generating vast allelic variation for interacting genes, to dissect gene dosage dependencies that could be much more important than previously realized in shaping quantitative variation in crops and beyond.

Lots of Genetic Diversity, Lots of Potential Epistasis

Considering the breadth of genomic diversity within a crop and between their wild ancestors, there is certainly more epistasis to uncover (58, 62, 109, 116). In rice, for example, short-read re-sequencing more than 3000 *Oryza sativa* genomes led to the

identification of 12,465 novel genes absent from the reference genome (109). Similarly, thousands of genes show present-absent variation across more than 700 diverse tomato genomes (26). Natural variants often involve structural changes to the genome, but such structural variants (SVs: insertions, deletions, duplications, inversions, translocations) are difficult to resolve with short-read sequencing data (92). Several domestication and breeding traits are based on SVs (70, 98, 113). For example, the fruit shape *sun* locus in tomato arose when a retrotransposon carrying the *SUN* gene inserted into a new genomic region, which elevated *SUN* expression in flowers and developing fruits (113). In the context of epistasis, long-read sequencing was key to resolving the tandem duplications responsible for the two QTLs that suppress inflorescence branching caused by *j2 ej2w* epistasis in tomato (99) (Alonge et al., in press). More reference genomes and advanced computational tools for robust identification of all forms of genetic variation are needed to advance epistasis research in crops.

Epistasis and background dependencies

A primary message in this review is that the epistatic modifier loci that comprise background dependencies could be widespread. Conventional approaches for revealing natural modifier loci in plants involve crossing known mutants into many genetically distinct accessions or wild ancestors, and then phenotyping for transgressive variation in segregating populations. However, unless only a single recessive modifier mutation is involved, segregating ratios of individuals with phenotypic modifications from multiple modifiers are low. Prohibitively large populations are therefore needed to reliably dissect the genetic architectures of background dependencies. Even when population size is not limiting, the mode-of-inheritance and dosage effects from only a few modifiers may confound reliable phenotypic assessments. Systematic studies of background effects have been conducted in other multicellular organisms where achieving the needed population size and phenotyping at scale is more feasible, such as *C. elegans*. For example, knockdown of two conserved regulators of endoderm development in 96 unique *C. elegans* wild strains identified extensive cryptic variation within a regulatory network for developmental plasticity (107).

Though scale may still be limited, systematic approaches to reveal and study background effects in model and crop plants are now possible, enabled by recent advances in CRISPR-based genome editing systems (14). Genome editing can allow the introduction of mutations into a diverse accessions of a single crop species (**Figure 2**). For example, 'query mutations' that cause a known phenotype can be introduced into multiple genotypes to test for epistatic interactions across a population. Taking a known query-modifier pair from tomato as example, introducing stem abscission *j2* mutations into a large set of accessions would result in highly branched inflorescences in some accessions that carry *ej2*^{W} modifier. This approach could also reveal quantitative modulation of the excessive branching cause by *j2 ej2*^{W}, as some accessions carrying both mutations might be more or less severely branched compared to a reference background. Indeed, this would expose the *sb1* and *sb3* suppressor QTL, one of which is based on epistasis from another MADS-box gene distantly related from *J2* and *EJ2*. Similarly, introducing mutations for florigen pathway genes, which regulate flowering time and shoot architecture, could reveal weak background effects and epistasis in the florigen system. This approach can be applied to any gene and trait, and could be facilitated by using genomic data to focus on genotypes that capture the highest genetic diversity within a population. Conceptually, this approach may be supported by an underappreciated aspect of the domestication process. It seems likely that desirable large effect mutations that arose during domestication sensitized ancestral genomes, revealing standing (cryptic) genetic variation that could have accelerated domestication (52).

Systematic dissections of background dependencies and modifier alleles from natural populations will be especially useful for known domestication and breeding genes. During domestication, humans selected a limited number of large-effect beneficial traits on a small set of wild populations, a process that resulted in domestication syndromes (22). Continued selection of beneficial alleles leads to genomic islands of low genetic diversity. This reduced genetic diversity in crops compared to wild ancestors is universal (69). Reduction of genetic diversity after the domestication bottleneck also implies fixation of epistatic relationships between genes and alleles. To unleash such cases of cryptic epistasis, genome editing can now be used to introduce targeted mutations in

domestication and improvement genes into the wild ancestors and early domesticated forms (landraces). This could uncover background effects from epistasis that were fixed during domestication. For example, in two separate studies, mutations were engineered in the *SlCLV3* gene of both domesticated tomato and its closest wild ancestor (114, 119). Notably, these null mutations showed quantitative differences in locule number in the two genotypes, likely reflecting natural modifiers, one of which is almost certainly *lc* (87).

Using genome editing to reveal modifier alleles that were selected and fixed during domestication and breeding requires certain considerations. Domestication often favored weak mutations, frequently in *cis*-regulatory regions that modulate phenotypes in a dosage-dependent manner (104). This may represent a challenge for identifying epistatic interaction upon engineering domestication and breeding mutations in different backgrounds. Taking the florigen-antiflorigen (*SFT*-*SP*) interaction as example, *sft* null mutations are completely epistatic over *sp* and lead to highly vegetative plants regardless of whether *SP* is functional (45, 60). Only a quantitative reduction in *SFT* function, from *sft* heterozygosity and weak *sft* alleles (75), leads to a florigen-based dosage-dependent modulation of shoot architecture that only manifests in *sp* mutant backgrounds. This suggests that releasing natural modifiers of domestication and breeding traits that emerged from weak mutations may require recreating the specific natural alleles, or engineering alleles with similar allelic strength. Recent technological advances may allow for the recreation of domestication alleles that are caused by SNPs, transposable elements, and other SVs in any genetic background. SNPs, for example, can be recreated using base editing (28, 56, 93) and prime editing (1, 61). SVs such as deletions can in many cases be recreated using multiple CRISPR gRNAs, but insertions and duplications will require technological advancements (89).

The intersection of dosage and epistasis

Epistatic interactions selected during domestication and breeding may rely on changes in dosage from weak alleles or heterozygosity. Examples of quantitative, dosage-sensitive epistatic interactions were introduced in the previous sections. The *sp*-dependent *sft* heterotic effect on yield is triggered by reduction in *SFT* gene dosage, either through *sft*

heterozygosity or weak alleles (40, 45, 75). Similarly, heterozygosity from the weak natural *ej2*^{W} allele in a *j2* background provides a quantitative dosage-dependent benefit on inflorescence branching and yield (98). Similarly, fruit size variation is quantitatively modulated by individual gene dosage contributions in the *SlWUS*-*SlCLV3*-*SlCLE9* circuit, mediated by redundancy and compensation between *SLCLV3* and *SlCLE9* (15, 84, 85, 114). Such examples, while informative on relationships between gene dosage, epistasis and quantitative variation, rely on one or a few allelic variants that define isolated points on a possible epistasis continuum, and therefore may fail to capture the granularity of genetic interactions. Allelic series representing a range of gene dosage would address if and to what extent allelic strength affects interactions and phenotypic outputs. Here again, genome editing technologies offer an unprecedented opportunity to generate the tools needed to fill this gap. We discuss below methods that can deliver systematic manipulation of gene dosage to further our understanding of quantitative epistasis.

Understanding the dynamics in genetic interactions that arise from dosage effects relies on our ability to expand allelic diversity. Targeting protein coding regions using CRISPR-Cas systems in soybean was applied to generate a small, but informative, allelic series of in-frame mutations that modify dosage mostly through weak loss-of-function mutations (10). Recent refinements of CRISPR editing tools allow more delicate and systematic analyses. Base editing delivers precise nucleotide transition mutations and has been successfully applied to several crop models, including rice and wheat (56, 61, 118). Prime editing, a CRISPR system that relies on a reverse transcriptase, allows targeted insertions and deletions, and all transition and transversion mutations (1). This higher flexibility comes at the expense of lower efficiency, but prime editing has been applied in rice and wheat (1, 61). Saturated mutagenesis of coding sequences offers the possibility to interrogate every residue of a target peptide sequence and its contribution to an epistatic effect (56). In rice, dual-base editors were utilized to engineer herbicide resistance through saturation mutagenesis of the *OsACC* enzyme (61). Such semi-random mutagenesis of protein coding regions *in planta* could be applied to coding sequence allelic series in developmental genes (**Figure 2**). Since many single nucleotide edits will result in silent mutations, these approaches may be most suitable for targeting

genes encoding short proteins or peptides. Base editors and prime editors can also be used to target essential residues within known functional protein domains.

Another promising route to induce gene dosage variation is CRISPR-Cas9 mutagenesis of regulatory sequences. *Cis*-regulatory elements (CREs) in promoters regulate the timing, pattern and level of gene expression, and therefore contribute to phenotypic diversity (111). CRISPR-Cas9 editing of promoters has been proven efficient in tomato, providing new regulatory alleles for genes controlling fruit size, whole plant and inflorescence architecture (85). By generating even larger repertoires of expression alleles, promoter allelic series can reveal under which conditions two alleles between two interacting genes depart from additivity and trigger epistasis, and whether that epistasis is quantitative and scales linearly to a qualitative phenotypic output, or is threshold-based (**Figure 2**). Taking *SlCLV3*-*SlCLE9* epistasis as an example, a repertoire of weak regulatory alleles for both genes and systematically generating weak-by-null, null-by-weak and weak-by-weak combinations among them (or from a coding repertoire as discussed above) would reveal the extent that epistasis regulates the circuit, and particularly what are the initiators triggering *SlCLE9* compensation. The same approach can be envisioned for dissecting any other epistatic network, and is particularly suitable for dosage-sensitive systems such as the florigen pathway (75) and MADS-box genes (98). Also, editing regulatory sequences that underlie spatiotemporal-specific expression could allow to dissect the function of pleiotropic genes and the role of epistatic interactions in a particular tissue or developmental context. Moreover, mutating individual CREs of a specific gene separately and in combination could allow to dissect CRE specific function and reveal interactions between individual CREs within a single gene.

Overcoming limitations for dissecting epistasis in crops at scale

Epistatic interactions have been studied at large scale in simple model organisms, using high-throughput approaches to produce combinations of mutations from engineered mutant libraries. For example, in yeast, 410,399 digenic and 195,666 trigenic combinations were tested for fitness defects (47). Large mutant collections in multiple yeast strains can also be generated, allowing for background effects to be studied at

scale. The major bottleneck for large-scale approaches in plant genome editing, which is needed to expand genetic variation, is the absence of fast and efficient delivery of CRISPR modules into plant cells. Currently, this process is time and labor intensive. However, advances are emerging. In tomato, pooled CRISPR libraries have been used to mutate up to 15 genes in a single transformation experiment, promising higher throughput in generating single and higher-order mutant collections (34). Another promising avenue is a new technology for *de novo* meristem induction. This method, which is based on induction of transcription factors that promote stem cell production, has been applied in crops such as tomato, grape and tobacco, and could allow fast, inexpensive, genotype-independent genome editing (67). With this and other advanced plant cell delivery technologies (11, 18), large-scale CRISPR screens, which are already standard for epistasis studies in yeast and animals using cell-based systems (20, 73), may become a reality in crops at the whole plant level. A recent maize study combined CRISPR-gRNA libraries with deep sequencing was used to target 743 gene candidates connected to agronomic and nutritional traits. The approach yielded 412 edited alleles among 118 genes (65).

Development of *trans*-acting Cas9-gRNA cassettes offer the possibility of producing a collection of mutant alleles for a single gene or for inducing mutation in a single gene across different accessions from a limited number of transgenic plants. This was the foundation to rapidly generate promoter allelic series in tomato (85). Briefly, plants carrying an active Cas9-gRNA transgene are backcrossed to wild-type plants. Hybrids that carry the *trans*-acting Cas9-gRNA cassette are selected, in which the active Cas9-gRNA transgene can target the wild-type alleles in *trans* and thereby induce novel alleles. An important benefit of this approach is that hybrids also inherit a single strong allele from the transgenic mutant parent, which sensitizes the F1 plants to more easily reveal phenotypic effects from newly induced weak loss-of-function alleles.

Trans-acting Cas9-gRNA editing cassettes can also be used to dissect epistatic background effects from modifiers (**Figure 3**). Transgenics that are homozygous for the query mutation and carry the Cas9-gRNA transgene can be crossed with a diversity

panel. In hybrids, which are heterozygous for the query mutation and potential modifiers, Cas9-gRNA cassettes will target the remaining wild-type allele at the query locus *in trans*. This approach will allow the generation of F2 progeny that are all fixed (homozygous null loss-of-function) for query mutations, but segregate potential weak modifier loci whose effects are now enhanced and readily revealed. Thus, smaller F2 mapping populations may be sufficient for gene mapping, which becomes especially important when more than one modifier locus is segregating.

5. CONCLUSIONS AND OUTLOOK

Over the last decade, the dissection of quantitative trait variation in crops has revealed that human selection during domestication modified complex genetic networks. Multiple studies identified individual loci and their epistatic interactions as drivers of rapid trait evolution and phenotypic diversification. However, the full breadth and significance of epistasis in crop domestication and breeding has often been overlooked due to limitations in genetic and genomic tools and resources. Recent advances in long-read sequencing technologies and the availability of crop pan-genomes, combined with rapidly evolving genome editing tools are opening new horizons for a systematic dissection of epistasis in crops at unprecedented resolution. Based on principles defined from known epistatic interactions in tomato and other crops, we proposed new strategies that integrate genomics-enabled quantitative genetics with genome editing to reveal, resolve and harness epistasis in crops at scale. These novel approaches have the potential to expand our understanding of the molecular principles and the evolutionary scope of epistasis, by capturing dynamic and quantitative aspects of epistatic interactions arising from changes in gene dosage and background dependencies at the population level. We expect new insights of both fundamental and applied value, from identifying new layers of complexity in genetic networks, to harnessing principles of epistasis for predictable crop breeding. These advances may also illuminate the contribution of epistasis to missing heritability.

ACKNOWLEDGEMENTS

We thank members of the Soyk and Lippman labs and Y. Eshed for discussions. This work was supported by the Howard Hughes Medical Institute and funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC Starting Grant "EPICROP" Grant No. 802008) to S.S., a Swiss National Science Foundation (SNSF) Eccellenza Professorial Fellowship (Grant No. PCEFP3_181238) to S.S, the USDA National Institute of Food and Agriculture (AFRI Grant No. 2016-67013-24452) to Z.B.L., the BARD (United States-Israel Binational Agricultural Research and Development Fund, Grant No. IS-5120-18C) to Z.B.L., and the National Science Foundation Plant Genome Research Program (Grant Nos. IOS-1546837 and IOS-1732253) to Z.B.L.

REFERENCES

1. Anzalone A V., Randolph PB, Davis JR, Sousa AA, Koblan LW, et al. 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 576(7785):149–57
2. Bateson W, Saunders ER, Punnett RC. 1906. Further Experiments on Inheritance in Sweet Peas and Stocks: Preliminary Account. *Proc. R. Soc. B Biol. Sci.* 77(517):236–38
3. Bernard RL. 1972. Two Genes Affecting Stem Termination in Soybeans. *Crop Sci.* 12(2):235–39
4. Birchler JA, Veitia RA. 2012. Gene balance hypothesis: Connecting issues of dosage sensitivity across biological disciplines. *Proc. Natl. Acad. Sci.* 109(37):14746–53
5. Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA. 2010. Heterosis. *Plant Cell*. 22(7):2105–12
6. Blackman BK, Strasburg JL, Raduski AR, Michaels SD, Rieseberg LH. 2010. The Role of Recently Derived FT Paralogs in Sunflower Domestication. *Curr. Biol.* 20(7):629–35
7. Boden SA, Cavanagh C, Cullis BR, Ramm K, Greenwood J, et al. 2015. Ppd-1 is a key regulator of inflorescence architecture and paired spikelet development in wheat. *Nat. Plants*. 1(2):14016
8. Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, et al. 2007. Autoimmune Response as a Mechanism for a Dobzhansky-Muller-Type Incompatibility Syndrome in Plants. *PLoS Biol.* 5(9):e236
9. Bomblies K, Weigel D. 2007. Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nat. Rev. Genet.* 8(5):382–93
10. Campbell BW, Hoyle JW, Bucciarelli B, Stec AO, Samac DA, et al. 2019. Functional analysis and development of a CRISPR/Cas9 allelic series for a CPR5 ortholog necessary for proper growth of soybean trichomes. *Sci. Rep.* 9(1):1–11
11. Cao Y, Lim E, Xu M, Weng J, Marelli B. 2020. Precision Delivery of Multiscale Payloads to Tissue-Specific Targets in Plants. *Adv. Sci.* 1903551:1903551
12. Chae E, Bomblies K, Kim S-T, Karelina D, Zaidem M, et al. 2014. Species-wide Genetic Incompatibility Analysis Identifies Immune Genes as Hot Spots of Deleterious Epistasis. *Cell*. 159(6):1341–51
13. Chen C, Chen H, Lin Y-S, Shen J-B, Shan J-X, et al. 2014. A two-locus interaction causes interspecific hybrid weakness in rice. *Nat. Commun.* 5(1):3357
14. Chen K, Wang Y, Zhang R, Zhang H, Gao C. 2019. CRISPR/Cas Genome Editing and

Precision Plant Breeding in Agriculture. *Annu. Rev. Plant Biol.* 70(1):667–97
15. Chu YH, Jang JC, Huang Z, van der Knaap E. 2019. Tomato locule number and fruit size controlled by natural alleles of lc and fas. *Plant Direct*. 3(7):1–20
16. 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(12):1388–92
17. Costanzo M, Kuzmin E, van Leeuwen J, Mair B, Moffat J, et al. 2019. Global Genetic Networks and the Genotype-to-Phenotype Relationship. *Cell*. 177(1):85–100
18. Cunningham FJ, Goh NS, Demirer GS, Matos JL, Landry MP. 2018. Nanoparticle-Mediated Delivery towards Advancing Plant Genetic Engineering. *Trends Biotechnol.* 36(9):882–97
19. Diss G, Ascencio D, DeLuna A, Landry CR. 2014. Molecular mechanisms of paralogous compensation and the robustness of cellular networks. *J. Exp. Zool. Part B Mol. Dev. Evol.* 322(7):488–99
20. Dixit A, Parnas O, Li B, Chen J, Fulco CP, et al. 2016. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell*. 167(7):1853-1866.e17
21. Doebley J, Stec A, Hubbard L. 1997. The evolution of apical dominance in maize. *Nature*. 386(6624):485–88
22. Doebley JF, Gaut BS, Smith BD. 2006. The Molecular Genetics of Crop Domestication. *Cell*. 127(7):1309–21
23. Fisher RA. 1919. The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans. R. Soc. Edinburgh*. 52(2):399–433
24. Fray RG, Grierson D. 1993. Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing, complementation and co-suppression. *Plant Mol. Biol.* 22(4):589–602
25. Galli M, Gallavotti A. 2016. Expanding the Regulatory Network for Meristem Size in Plants. *Trends Genet.* 32(6):372–83
26. 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(6):1044–51
27. Garcia AAF, Wang S, Melchinger AE, Zeng Z-B. 2008. Quantitative Trait Loci Mapping and The Genetic Basis of Heterosis in Maize and Rice. *Genetics*. 180(3):1707–24
28. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, et al. 2017. Programmable base editing of T to G C in genomic DNA without DNA cleavage. *Nature*. 551(7681):464–

71
29. 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(4):494–97
30. Huang X, Yang S, Gong J, Zhao Q, Feng Q, et al. 2016. Genomic architecture of heterosis for yield traits in rice. *Nature*. 537(7622):629–33
31. Hufford MB, Berny Mier y Teran JC, Gepts P. 2019. Crop Biodiversity: An Unfinished Magnum Opus of Nature. *Annu. Rev. Plant Biol.* 70(1):727–51
32. Isaacson T, Ronen G, Zamir D, Hirschberg J. 2002. Cloning of tangerine from Tomato Reveals a Carotenoid Isomerase Essential for the Production of β-Carotene and Xanthophylls in Plants. *Plant Cell*. 14(2):333–42
33. Iwata H, Gaston A, Remay A, Thouroude T, Jeauffre J, et al. 2012. The TFL1 homologue KSN is a regulator of continuous flowering in rose and strawberry. *Plant J.* 69(1):116–25
34. Jacobs TB, Zhang N, Patel D, Martin GB. 2017. Generation of a collection of mutant tomato lines using pooled CRISPR libraries. *Plant Physiol.* 174(4):2023–37
35. Jaeger KE, Pullen N, Lamzin S, Morris RJ, Wigge PA. 2013. Interlocking Feedback Loops Govern the Dynamic Behavior of the Floral Transition in Arabidopsis. *Plant Cell*. 25(3):820–33
36. Je B Il, Gruel J, Lee YK, Bommert P, Arevalo ED, et al. 2016. Signaling from maize organ primordia via FASCIATED EAR3 regulates stem cell proliferation and yield traits. *Nat. Genet.* 48(7):785–91
37. Jenkins JA, Mackinney G. 1953. Inheritance of Carotenoid Differences in the Tomato Hybrid Yellow x Tangerine. *Genetics*. 38(2):107–10716
38. Jenkins JA, Mackinney G. 1955. Carotenoids of the Apricot Tomato and Its Hybrids with Yellow and Tangerine. *Genetics*. 40(5):715–20
39. Jeuken MJW, Zhang NW, McHale LK, Pelgrom K, den Boer E, et al. 2009. Rin4 Causes Hybrid Necrosis and Race-Specific Resistance in an Interspecific Lettuce Hybrid. *Plant Cell*. 21(10):3368–78
40. Jiang K, Liberatore KL, Park SJ, Alvarez JP, Lippman ZB. 2013. Tomato Yield Heterosis Is Triggered by a Dosage Sensitivity of the Florigen Pathway That Fine-Tunes Shoot Architecture. *PLoS Genet.* 9(12):
41. Jiang Y, Schmidt RH, Zhao Y, Reif JC. 2017. A quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. *Nat. Genet.* 49(12):1741–46
42. Kachanovsky DE, Filler S, Isaacson T, Hirschberg J. 2012. Epistasis in tomato color

mutations involves regulation of phytoene synthase 1 expression by cis-carotenoids. *Proc. Natl. Acad. Sci. U. S. A.* 109(46):19021–26

43. Kempin S, Savidge B, Yanofsky M. 1995. Molecular basis of the cauliflower phenotype in Arabidopsis. *Science (80-.).* 267(5197):522–25
44. Klee HJ, Giovannoni JJ. 2011. Genetics and Control of Tomato Fruit Ripening and Quality Attributes. *Annu. Rev. Genet.* 45(1):41–59
45. Krieger U, Lippman ZB, Zamir D. 2010. The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat. Genet.* 42(5):459–63
46. Krüger J, Thomas CM, Golstein C, Dixon MS, Smoker M, et al. 2002. A tomato cysteine protease required for Cf-2-dependent disease resistance and suppression of autonecrosis. *Science (80-.).* 296(5568):744–47
47. Kuzmin E, VanderSluis B, Wang W, Tan G, Deshpande R, et al. 2018. Systematic analysis of complex genetic interactions. *Science (80-.).* 360(6386):eaao1729
48. Kwak M, Velasco D, Gepts P. 2008. Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris). *J. Hered.* 99(3):283–91
49. 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(2):182–88
50. Kyozuka J, Tokunaga H, Yoshida A. 2014. Control of grass inflorescence form by the fine-tuning of meristem phase change. *Curr. Opin. Plant Biol.* 17(1):110–15
51. Lan TH, Paterson AH. 2000. Comparative mapping of quantitative trait loci sculpting the curd of Brassica oleracea. *Genetics*. 155(4):1927–54
52. Lauter N, Doebley J. 2002. Genetic variation for phenotypically invariant traits detected in teosinte: Implications for the evolution of novel forms. *Genetics*. 160(1):333–42
53. Lee J, Lee I. 2010. Regulation and function of SOC1, a flowering pathway integrator. *J. Exp. Bot.* 61(9):2247–54
54. Lemmon ZH, Park SJ, Jiang K, Van Eck J, Schatz MC, Lippman ZB. 2016. The evolution of inflorescence diversity in the nightshades and heterochrony during meristem maturation. *Genome Res.* 26(12):1676–86
55. 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(10):766–70
56. 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.*

57. 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(12):1160–63
58. Li Y, Zhou G, Ma J, Jiang W, Jin L, et al. 2014. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat. Biotechnol.* 32(10):1045–52
59. Lifschitz E, Ayre BG, Eshed Y. 2014. Florigen and anti-florigen - a systemic mechanism for coordinating growth and termination in flowering plants. *Front. Plant Sci.* 5(SEP):1–14
60. Lifschitz E, Eshed Y. 2006. Universal florigenic signals triggered by FT homologues regulate growth and flowering cycles in perennial day-neutral tomato. *J. Exp. Bot.* 57(13):3405–14
61. Lin Q, Zong Y, Xue C, Wang S, Jin S, et al. 2020. Prime genome editing in rice and wheat. *Nat. Biotechnol.*
62. 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(11):1220–26
63. Lippman ZB, Cohen O, Alvarez JP, Abu-Abied M, Pekker I, et al. 2008. The Making of a Compound Inflorescence in Tomato and Related Nightshades. *PLoS Biol.* 6(11):e288
64. Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, et al. 2010. The Soybean Stem Growth Habit Gene Dt1 Is an Ortholog of Arabidopsis TERMINAL FLOWER1. *Plant Physiol.* 153(1):198–210
65. Liu H, Jian L, Xu J, Zhang Q, Zhang M, et al. 2020. High-Throughput CRISPR/Cas9 Mutagenesis Streamlines Trait Gene Identification in Maize. *Plant Cell*. tpc.00934.2019
66. Macarthur JW, Chiasson LP. 1947. Cytogenetic Notes on Tomato Species and Hybrids. *Genetics*. 32(2):165–77
67. Maher MF, Nasti RA, Vollbrecht M, Starker CG, Clark MD, Voytas DF. 2020. Plant gene editing through de novo induction of meristems. *Nat. Biotechnol.* 38(1):84–89
68. Melchinger AE, Piepho H-P, Utz HF, Muminović J, Wegenast T, et al. 2007. Genetic Basis of Heterosis for Growth-Related Traits in Arabidopsis Investigated by Testcross Progenies of Near-Isogenic Lines Reveals a Significant Role of Epistasis. *Genetics*. 177(3):1827–37
69. Moyers BT, Morrell PL, McKay JK. 2018. Genetic Costs of Domestication and Improvement. *J. Hered.* 109(2):103–16
70. Müller NA, Zhang L, Koornneef M, Jiménez-Gómez JM. 2018. Mutations in EID1 and LNK2 caused light-conditional clock deceleration during tomato domestication. *Proc. Natl. Acad. Sci. U. S. A.* 115(27):7135–40

71. Mullins MG, Bouquet A, Williams LE. 1992. *Biology of the grapevine*. Cambridge University Press
72. Muños S, Ranc N, Botton E, Bérard A, Rolland S, et al. 2011. Increase in Tomato Locule Number Is Controlled by Two Single-Nucleotide Polymorphisms Located Near WUSCHEL. *Plant Physiol.* 156(4):2244–54
73. Norman TM, Horlbeck MA, Replogle JM, Ge AY, Xu A, et al. 2019. Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. *Science (80-.).* 365(6455):786–93
74. Park SJ, Eshed Y, Lippman ZB. 2014. Meristem maturation and inflorescence architecture-lessons from the Solanaceae. *Curr. Opin. Plant Biol.* 17(1):70–77
75. Park SJ, Jiang K, Tal L, Yichie Y, Gar O, et al. 2014. Optimization of crop productivity in tomato using induced mutations in the florigen pathway. *Nat. Genet.* 46(12):1337–42
76. Peralta IE, Spooner DM. 2005. Morphological characterization and relationships of wild tomatoes (Solanum L. Sect. Lycopersicon). *A Festschrift William G. D'arcy*. 104:227–57
77. Phillips PC. 2008. Epistasis — the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9(11):855–67
78. Ping J, Liu Y, Sun L, Zhao M, Li Y, et al. 2014. Dt2 is a gain-of-function MADS-domain factor gene that specifies semideterminacy in soybean. *Plant Cell*. 26(7):2831–42
79. Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, et al. 1998. The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development*. 125(11):1979–89
80. Purugganan MD, Boyles AL, Suddith JI. 2000. Variation and selection at the CAULIFLOWER floral homeotic gene accompanying the evolution of domesticated Brassica oleracea. *Genetics*. 155(2):855–62
81. Ramsay L, Comadran J, Druka A, Marshall DF, 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(2):169–72
82. Reynard GB. 1961. New source of the j2 gene governing Jointless pedicel in tomato. *Science (80-.).* 134(3496):2102
83. Rick CM. 1956. A new jointless gene from the Galapagos L. pimpinellifolium. *TGC Rep.*, p. 23
84. Rodriguez-Leal D, Xu C, Kwon CT, Soyars C, Demesa-Arevalo E, et al. 2019. Evolution of buffering in a genetic circuit controlling plant stem cell proliferation. *Nat. Genet.* 51(5):786–92

85. Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB. 2017. Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell*. 171(2):470-480.e8
86. Rodríguez GR, Kim HJ, Van Der Knaap E. 2013. Mapping of two suppressors of OVATE (sov) loci in tomato. *Heredity (Edinb).* 111(3):256–64
87. Rodríguez GR, Muños S, Anderson C, Sim S-C, Michel A, et al. 2011. Distribution of SUN, OVATE, LC , and FAS in the Tomato Germplasm and the Relationship to Fruit Shape Diversity. *Plant Physiol.* 156(1):275–85
88. Sax K. 1923. The Association of Size Differences with Seed-Coat Pattern and Pigmentation in Phaseolus vulgaris. *Genetics*. 8(6):552–60
89. Schmidt C, Pacher M, Puchta H. 2019. Efficient induction of heritable inversions in plant genomes using the CRISPR/Cas system. *Plant J.* 98(4):577–89
90. Schnable PS, Springer NM. 2013. Progress Toward Understanding Heterosis in Crop Plants. *Annu. Rev. Plant Biol.* 64(1):71–88
91. Schoof H, Lenhard M, Haecker A, Mayer KFX, Jürgens G, Laux T. 2000. The Stem Cell Population of Arabidopsis Shoot Meristems Is Maintained by a Regulatory Loop between the CLAVATA and WUSCHEL Genes. *Cell*. 100(6):635–44
92. Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, et al. 2018. Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods*. 15(6):461–68
93. Shimatani Z, Kashojiya S, Takayama M, Terada R, Arazoe T, et al. 2017. Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35(5):441–43
94. Shull GH. 1908. The Composition of a Field of Maize. *J. Hered.* os-4(1):296–301
95. Shull GH. 1910. Hybridization Methods in Corn Breeding. *J. Hered.* 1(2):98–107
96. Somssich M, Je B Il, Simon R, Jackson D. 2016. CLAVATA-WUSCHEL signaling in the shoot meristem. *Dev.* 143(18):3238–48
97. Soyars CL, James SR, Nimchuk ZL. 2016. Ready, aim, shoot: stem cell regulation of the shoot apical meristem. *Curr. Opin. Plant Biol.* 29:163–68
98. 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(6):1142-1155.e12
99. Soyk S, Lemmon ZH, Sedlazeck FJ, Jiménez-Gómez 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(5):471–79

100. Soyk S, Müller NA, Park SJ, Schmalenbach I, Jiang K, et al. 2017. Variation in the flowering gene SELF PRUNING 5G promotes day-neutrality and early yield in tomato. *Nat. Genet.* 49(1):162–68
101. Stephenson AG. 1981. Flower and Fruit Abortion: Proximate Causes and Ultimate Functions. *Annu. Rev. Ecol. Syst.* 12(1):253–79
102. Stetter MG, Gates DJ, Mei W, Ross-Ibarra J. 2017. How to make a domesticate. *Curr. Biol.* 27(17):R896–900
103. Stitzer MC, Ross-Ibarra J. 2018. Maize domestication and gene interaction. *New Phytol.* 220(2):395–408
104. Swinnen G, Goossens A, Pauwels L. 2016. Lessons from Domestication: Targeting Cis - Regulatory Elements for Crop Improvement. *Trends Plant Sci.* 21(6):506–15
105. Taoka K, Ohki I, Tsuji H, Furuita K, Hayashi K, et al. 2011. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature*. 476(7360):332–35
106. Tomes ML, Quackenbush FW, Nelson OE, North B. 1953. The Inheritance of Carotenoid Pigment Systems in the Tomato. *Genetics*. 38(2):117–11727
107. Torres Cleuren YN, Ewe CK, Chipman KC, Mears ER, Wood CG, et al. 2019. Extensive intraspecies cryptic variation in an ancient embryonic gene regulatory network. *Elife*. 8:1–31
108. van der Knaap E, Chakrabarti M, Chu YH, Clevenger JP, Illa-Berenguer E, et al. 2014. What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. *Front. Plant Sci.* 5(MAY):227
109. 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(7703):43–49
110. Weller JL, Ortega R. 2015. Genetic control of flowering time in legumes. *Front. Plant Sci.* 6(APR):1–13
111. Wittkopp PJ, Kalay G. 2012. Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13(1):59–69
112. Wu S, Zhang B, Keyhaninejad N, Rodríguez GR, Kim HJ, et al. 2018. A common genetic mechanism underlies morphological diversity in fruits and other plant organs. *Nat. Commun.* 9(1):1–12
113. Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E. 2008. A Retrotransposon-Mediated Gene Duplication Underlies Morphological Variation of Tomato Fruit. *Science (80-.).* 319(5869):1527–30

114. Xu C, Liberatore KL, MacAlister CA, Huang Z, Chu Y-H, et al. 2015. A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nat. Genet.* 47(7):784–92
115. Yamamoto E, Takashi T, Morinaka Y, Lin S, Wu J, et al. 2010. Gain of deleterious function causes an autoimmune response and Bateson–Dobzhansky–Muller incompatibility in rice. *Mol. Genet. Genomics*. 283(4):305–15
116. Zhang C, Wang P, Tang D, Yang Z, Lu F, et al. 2019. The genetic basis of inbreeding depression in potato. *Nat. Genet.* 51(3):374–78
117. Zhang S, Jiao Z, Liu L, Wang K, Zhong D, et al. 2018. Enhancer-Promoter Interaction of SELF PRUNING 5G Shapes Photoperiod Adaptation. *Plant Physiol.* 178(4):1631–42
118. Zong Y, Song Q, Li C, Jin S, Zhang D, et al. 2018. Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat. Biotechnol.* 36(10):950–53
119. Zsögön A, Čermák T, Naves ER, Notini MM, Edel KH, et al. 2018. De novo domestication of wild tomato using genome editing. *Nat. Biotechnol.* 36(12):1211–16

FIGURE LEGENDS

Figure 1: Epistasis and gene dosage effects in three yield traits in tomato. Genetic interactions between (**a**) the meristem maintenance genes *SlCLV3* and *SlWUS* (*lc*) regulate locule number and fruit size increase, (**b**) the florigen and antflorigen genes *SFT* and *SP* affect flowering time and shoot architecture, and (**c**) the MADS-box transcription factor genes *J2* and *EJ2* affect inflorescence branching. All three systems display quantitative dosage-sensitive epistatic relationships between the interacting gene pairs. Dosage changes from heterozygous and weak loss-of-function mutations can result in a continuum of quantitative epistasis and variation for each trait (dashed yellow curves). A balance between positive and negative developmental changes results in a dosage-dependent optimum, depending on desired phenotypic outputs (dashed red curves). Yellow and red dashed curves were inferred from published data (dots). Fine-tuning epistasis through gene dosage allows to pinpoint the optimal epistatic effect that results in a yield optimum for a specific agronomic target. The shape and optima of the "trait value" curves can shift depending on the trait and specific agronomic goals.

Figure 2: New approaches to dissect epistasis in crops using genome editing. (**a**) Genome editing to reveal genetic background effects on epistatic interactions. A 'query' mutation (red horizontal band) leads to a delay in flowering time (indicated by a red arrow) in a reference accession. Genome editing is used to introduce query mutations in the same gene into a collection of genetically-diverse genotypes (indicated by different colors). Background-specific natural variants (modifiers: multicolored horizontal bands) may or may not interact with the query mutation (indicated by curved lines with arrows). A genetic interaction between the query mutation and a modifier in a specific background leads to a quantitative deviation from the mutant phenotype in the reference background (multicolored straight arrows). (**b**) Genome editing to study gene dosage effects of epistatic interactions. An allelic series for a gene with a known and quantifiable phenotypic output (gene *A*) is generated by targeting of *cis*-regulatory regions or by saturation mutagenesis of coding sequences. An allelic series that translates into a range of quantitative variation (e.g. fruit size) is recovered. The allelic series is then combined with

a null mutation in a known interacting gene (gene *B*), and dosage-sensitivity of the interaction is quantified at the phenotype level. The reverse can also be tested using an allelic series of gene *B*, or combining a gene *A* allelic series by a gene *B* allelic series. A simple linear dose-dependent epistatic relationship is possible. Three hypothetical non-linear scenarios are shown, depicting the influence of altered dosage of gene *A* on trait values in the context of a loss-of-function gene *B*. The dosage of gene *A* at which the epistatic effect is non-linear is underlaid in grey.

Figure 3: Revealing epistasis and background dependencies using *trans*-acting Cas9-gRNA cassettes. First, a panel of transgenic 'query' mutants is developed by introducing mutations (orange bands) with known phenotypic effects on quantitative traits (e.g. inflorescence and shoot architecture) using CRISPR-Cas9 genome editing. Second, plants that are homozygous query mutation (orange bands) and for the Cas9-gRNA editing cassette (Cas9-gRNA; red bands) are crossed to a panel of genetically diverse accessions to sensitize the genetic background and expose weak phenotypic effects from cryptic modifier alleles (blue bands). In the F1 generation, the inherited Cas9-gRNA cassette targets the remaining functional allele of the query gene *in trans*. Resulting F2 mapping populations are null (homozygous for biallelic) for the query mutation and segregate potential modifier alleles whose phenotypic effects will be readily revealed. Causative modifier loci will be identified using established genomics and mapping strategies.

Figure 1

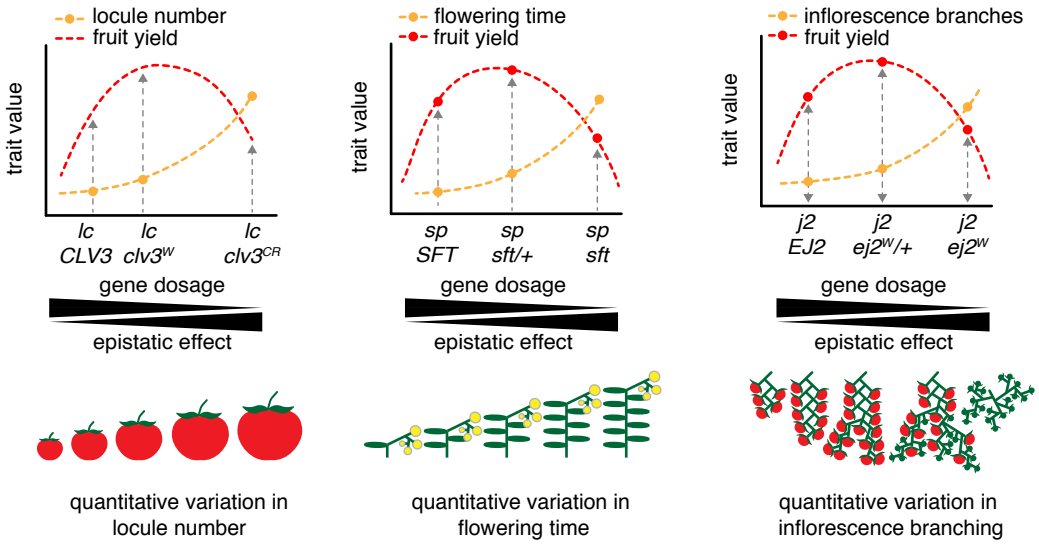

Figure 1: Epistasis and gene dosage effects in three yield traits in tomato. Genetic interactions between (**a**) the meristem maintenance genes *SlCLV3* and *SlWUS* (*lc*) regulate locule number and fruit size increase, (**b**) the florigen and antiflorigen genes *SFT* and *SP* affect flowering time and shoot architecture, and (**c**) the MADS-box transcription factor genes *J2* and *EJ2* affect inflorescence branching. All three systems display quantitative dosage-sensitive epistatic relationships between the interacting gene pairs. Dosage changes from heterozygous and weak loss-of-function mutations can result in a continuum of quantitative epistasis and variation for each trait (dashed yellow curves). A balance between positive and negative developmental changes results in a dosage-dependent optimum, depending on desired phenotypic outputs (dashed red curves). Yellow and red dashed curves were inferred from published data (dots). Fine-tuning epistasis through gene dosage allows to pinpoint the optimal epistatic effect that results in a yield optimum for a specific agronomic target. The shape and optima of the “trait value” curves can shift depending on the trait and specific agronomic goals.

Figure 2

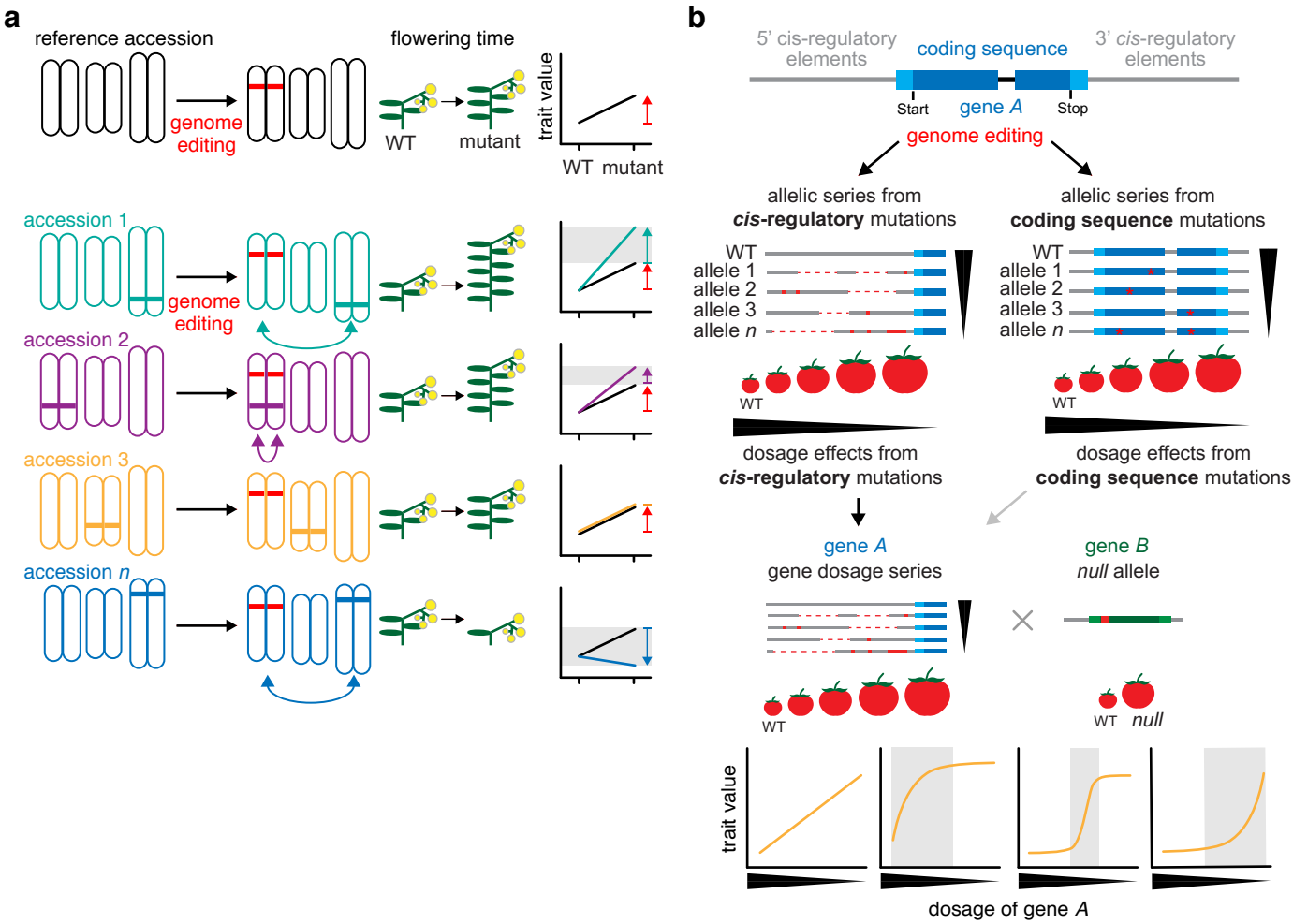

Figure 2: New approaches to dissect epistasis in crops using genome editing. (**a**) Genome editing to reveal genetic background effects on epistatic interactions. A 'query' mutation (red horizontal band) leads to a delay in flowering time (indicated by a red arrow) in a reference accession. Genome editing is used to introduce query mutations in the same gene into a collection of genetically-diverse genotypes (indicated by different colors). Background-specific natural variants (modifiers: multicolored horizontal bands) may or may not interact with the query mutation (indicated by curved lines with arrows). A genetic interaction between the query mutation and a modifier in a specific background leads to a quantitative deviation from the mutant phenotype in the reference background (multicolored straight arrows). (**b**) Genome editing to study gene dosage effects of epistatic interactions. An allelic series for a gene with a known and quantifiable phenotypic output (gene *A*) is generated by targeting of *cis*-regulatory regions or by saturation mutagenesis of coding sequences. An allelic series that translates into a range of quantitative variation (e.g. fruit size) is recovered. The allelic series is then combined with a null mutation in a known interacting gene (gene *B*), and dosage-sensitivity of the interaction is quantified at the phenotype level. The reverse can also be tested using an allelic series of gene *B*, or combining a gene *A* allelic series by a gene *B* allelic series. A simple linear dose-dependent epistatic relationship is possible. Three hypothetical non-linear scenarios are shown, depicting the influence of altered dosage of gene *A* on trait values in the context of a loss-of-function gene *B*. The dosage of gene *A* at which the epistatic effect is non-linear is underlaid in grey.

Figure 3

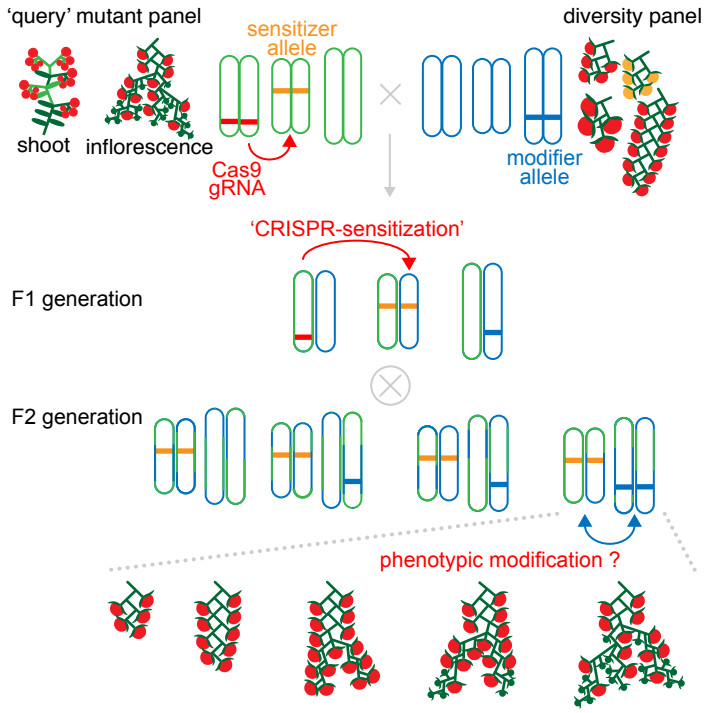

Figure 3: Revealing epistasis and background dependencies using trans-acting Cas9-gRNA cassettes. First, a panel of transgenic 'query' mutants is developed by introducing mutations (orange bands) with known phenotypic effects on quantitative traits (e.g. inflorescence and shoot architecture) using CRISPR-Cas9 genome editing. Second, plants that are homozygous query mutation (orange bands) and for the Cas9-gRNA editing cassette (Cas9-gRNA; red bands) are crossed to a panel of genetically diverse accessions to sensitize the genetic background and expose weak phenotypic effects from cryptic modifier alleles (blue bands). In the F1 generation, the inherited Cas9-gRNA cassette targets the remaining functional allele of the query gene in trans. Resulting F2 mapping populations are null (homozygous for biallelic) for the query mutation and segregate potential modifier alleles whose phenotypic effects will be readily revealed. Causative modifier loci will be identified using established genomics and mapping strategies.