A Bountiful Harvest: Genomic Insights into Crop Domestication Phenotypes

Kenneth M. Olsen[1] and Jonathan F. Wendel[2]

[1]Department of Biology, Washington University in St. Louis, St. Louis, Missouri 63130;
email: kolsen@wustl.edu

[2]Department of Ecology, Evolution, and Organismal Biology, Iowa State University,
Ames, Iowa 50011; email: jfw@iastate.edu

Keywords

Abstract

Human selection during crop domestication has resulted in remarkable transformations of plant phenotypes, providing a window into the genetic basis of morphological evolution. Recent progress in our understanding of the genetic architecture of novel plant traits has emerged from combining advanced molecular technologies with improved experimental designs, including nested association mapping, genome-wide association studies, population genetic screens for signatures of selection, and candidate gene approaches. These studies reveal a diversity of underlying causative mutations affecting phenotypes important in plant domestication and crop improvement, including coding sequence substitutions, presence/absence and copy number variation, transposon activation leading to novel gene structures and expression patterns, diversification following gene duplication, and polyploidy leading to altered combinatorial capabilities. The genomic regions unknowingly targeted by human selection include both structural and regulatory genes, often with results that propagate through the transcriptome as well as to other levels in the biosynthetic and morphogenetic networks.

Contents

INTRODUCTION: DOMESTICATION AS AN EVOLUTIONARY PROCESS

Modern civilization is dependent on only several dozen of the world's 300,000 plant species for its nourishment. These sustaining crop plants were derived, in most cases, by several thousand years or more of conscious as well as unintentional human selection, in the process transforming mostly unremarkable wild ancestors into high-yielding and otherwise useful domesticated descendants. In some cases

this domestication process entailed phenotypic changes that were sufficiently dramatic that the taxonomic origin of the domesticate was long obscure (e.g., maize from its wild ancestor, teosinte). Accordingly, evolutionary analysis and crop domestication have long been intertwined. Indeed, as Charles Darwin noted in the introduction to his most famous book (28), the strong directional, diversifying, and purifying selection practiced by aboriginal and modern domesticators provides a powerful lens for understanding the workings of the evolutionary process:

> At the commencement of my observations it seemed to me probable that a careful study of domesticated animals and of cultivated plants would offer the best chance of making out this obscure problem. Nor have I been disappointed; in this and in all other perplexing cases I have invariably found that our knowledge, imperfect though it be, of variation under domestication, afforded the best and safest clue. I may venture to express my conviction of the high value of such studies, although they have been very commonly neglected by naturalists.

Darwin's comments reflect the understanding that crop plants offer wonderful models for studying the evolutionary process, providing as they do a telescoped time frame in which both antecedent and descendant conditions remain extant and available for comparison. These comparisons have blossomed with the advent of the genomics era, leading to a number of novel insights into the enigmatic processes by which new phenotypes arise. Here we review some of these insights into the genetic and genomic basis of crop plant phenotypes.

Given their obvious economic significance and importance for humankind, it is perhaps not surprising that the majority of published plant genome sequences are for crop plants. Following the publication of the *Arabidopsis thaliana* reference genome in 2000 (2), the second completed angiosperm genome sequence was rice (63), which has since been followed by

the sequences of more than a dozen additional crop species (reviewed in 43; see also 29, 48, 97, 123). These and comparable efforts in many other species are greatly enabling, not just with respect to the myriad crop improvement applications in which a genome sequence may be leveraged, but also in terms of stimulating the development of a rich set of genomic and germplasm resources that have facilitated studies of the domestication process. For example, massively parallel (next-generation) sequencing and genotyping technologies in a number of genera are providing quantitative and qualitative insights into the structure of crop plant gene pools (e.g., 59, 139, 149), details on the geographic origin of domestication (e.g., 1, 80, 121), patterns of interspecific introgression or crop plant admixture (e.g., 53, 80), and exquisite detail on the shapes and sizes of genetic bottlenecks that accompanied the domestication process (e.g., 19, 60, 65, 73).

One of the most exciting areas of growth has been in characterizing the genetic basis of domestication-related phenotypes. As recently as 2006, the list of confirmed domestication-related genes numbered just over two dozen (32). Since that time, there has been an explosion in the numbers of crop species and traits examined, with emerging insights into the genetic architecture of domestication traits and processes of genome evolution in response to selection. Our focus here is on these recent advances in our understanding of the genetic architecture and molecular genetic basis of phenotypic changes favored during domestication and later crop improvement, aiming to provide an entry into this burgeoning literature.

Crop domestication imposes several microevolutionary forces on the plant genome. These fall into two categories based on how they are expected to reshape the genomic diversity of a crop in relation to its wild ancestor. Selectively neutral forces, which include genetic drift and gene flow, are expected to have genome-wide effects, with the former decreasing the genetic diversity of a crop compared with its wild relatives (e.g., 16, 135) and the latter maintaining or perhaps increasing

genetic diversity, particularly in the case of interspecific introgression (e.g., 27, 80). In contrast, selection is expected to differentially affect diversity, targeting as it does specific genomic regions that contain genes controlling the relevant phenotypes. Depending on the nature of the traits selected upon and their genetic basis, selection often leads to a differential loss of genetic diversity in targeted genomic regions, creating a molecular signature of selection (**Figure 1**). Thus, scanning for genomically localized genetic bottlenecks can provide clues as to the specific genes or mutations that underlie domestication-related traits, particularly when used in combination with other genetic analyses.

The term domestication syndrome is often used to describe the suite of traits arising during domestication that distinguish crops from their wild ancestors (52). In general, these phenotypes are those that render the crop more productive (e.g., more seed or bigger seed, these two often constituting alternatives), phenologically congruent with cropping practices (e.g., greater synchrony in flowering and/or fruiting), or easier to harvest (e.g., nonshattering, compactness) and consume (e.g., reduced toxicity). Some of the most common domestication traits include reduced seed dispersal (shattering), reduced seed dormancy, reduced branching with robust growth of the central stem (often leading to more erect plant architecture), determinate growth, uniform flowering and seed maturation (commonly associated with a loss of photoperiod sensitivity), increased resource allocation to the harvested plant part (e.g., fruits, seeds, roots, stems), and decreased chemical and morphological defenses (e.g., unpalatable secondary compounds, spines).

In considering the genetic basis of phenotypic changes during domestication, it is useful to draw a distinction between these *domestication* traits and other *improvement* traits that are subject to selection in crop varieties following initial domestication. Broadly speaking, domestication traits can be considered changes that occurred during the initial domestication process and are typically fixed within the crop

Wild ancestor

Domestication
bottleneck

Domesticated crop

Targeted gene
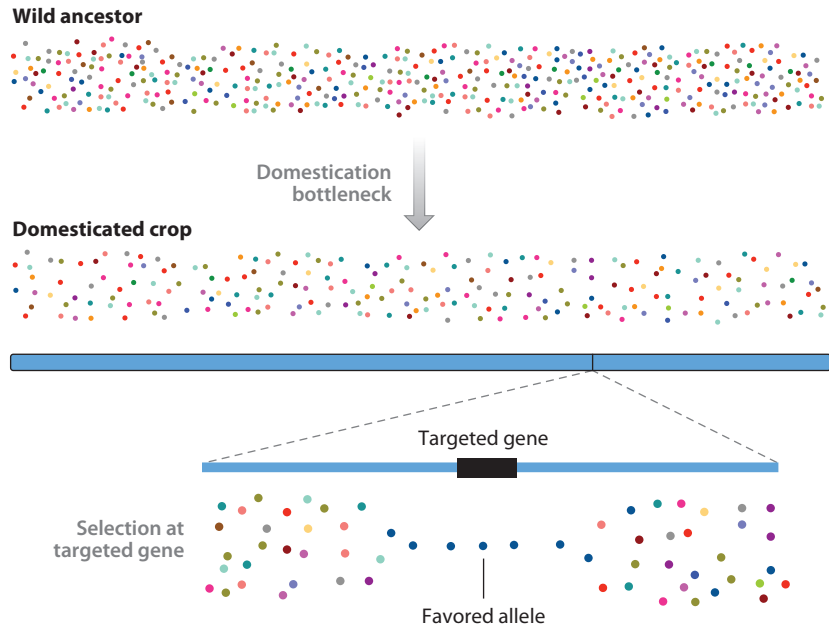
Selection at
targeted gene

Favored allele

Figure 1

The impact of domestication on genetic diversity. Colored dots represent neutral allelic diversity at genes across a chromosome (*blue bar*) in populations of a crop's wild ancestor (*top*) and in the crop itself (*middle*). Genetic drift acts strongly during the domestication bottleneck, when a subset of individuals in the wild species become the founders of the crop lineage; this is expected to result in a genome-wide reduction in genetic diversity. In contrast, selection is expected to differentially reduce diversity at the specific genes that control the traits subject to selection. As a favored allele is driven to high frequency, much of the standing genetic variation within and around the targeted gene (*black bar*) is removed from the population, creating a molecular signature of selection. The extent to which this deviation from neutrality is detectable depends on many factors, including the mating system, strength of selection, population structure, and recombination rates.

species. In contrast, crop improvement traits are typically variable among populations or cultivars of a crop; examples include characteristics such as adaptation to specific climates, seed starch composition, fruit pigmentation, and fruit morphology. These two sets of traits are highly variable among plants, reflecting both the multiplicity of phenotypes selected by humans and the range in domestication duration, dynamics, and population structures (among many variables) of different crop species.

In the following sections, we first review the methods currently employed in identifying the genes and mutations that have been targets of selection during crop domestication and improvement. We then discuss insights provided by recent studies into the genetic architecture

and molecular basis of phenotypic changes during domestication, including the roles of genomic structural variation and gene/genome duplications in this process. Our focus is on relatively recent studies (since approximately 2006); for earlier synopses, we refer readers to previous reviews (15, 32, 50, 86, 116).

ADVANCES IN METHODS FOR IDENTIFYING DOMESTICATION AND CROP IMPROVEMENT GENES

Biparental QTL Mapping

The traditional method for identifying genes underlying domestication-related phenotypes

is to perform quantitative trait locus (QTL) mapping in a population of recombinant inbred lines derived from a biparental cross. Domestication traits can be mapped using populations derived from wild-by-domesticate crosses, whereas improvement traits can be mapped using crosses of phenotypically distinct crop varieties. Identification of QTL regions is followed by fine mapping, cloning, and functional characterization of underlying genes, with the goal of identifying the specific mutation(s) selected during domestication. QTL mapping and fine mapping remain by far the most common approach for identifying domestication-related genes and causal mutations. Although applications of this approach have typically focused on obvious morphological or developmental traits, it is in principle broadly applicable, and hence may be used to examine subtler phenotypes such as domestication-related shifts in metabolic activity (e.g., photosynthetic rate, water use efficiency, fatty acid synthesis).

Genes underlying diverse domestication and improvement traits have been identified using biparental QTL mapping; **Table 1** highlights recent examples. These include traits involving (*a*) plant and inflorescence architecture [e.g., barley *Vrs1*, controlling two-rowed versus six-rowed inflorescence architecture (69); rice *SD1*, controlling rice culm length (a determinant of plant height) (3); rice *TAC1*, controlling the narrow tiller angle associated with erect plant growth in *japonica* varieties (144); and rice *PROG1*, controlling erect plant architecture (66, 115)], (*b*) yield [e.g., rice *OsSPL16* (*GW8*), controlling grain shape and size (130); rice *GW2*, controlling grain width and weight (107); rice *Ghd7*, controlling grain number, plant height, and heading date (141); rice *qSW5*, controlling grain width (101); tomato *fasciated*, controlling locule number (a determinant of fruit size) (24); and rice *GS5*, controlling rice grain size (74)], (*c*) pigmentation [e.g., rice *Bh4*, controlling rice hull color variation (150); sorghum *Tannin1*, controlling grain pigmentation (136); and rice *Phr1*, controlling differences between *indica* and *japonica*

varieties in grain discoloration (oxidation) during storage (146)], and (*d*) phenotypes targeted to enhance ease of planting or harvesting [e.g., sorghum *Sh1*, controlling loss of seed shattering (75); barley *Nud*, controlling free-threshing or "naked" (hulless) varieties (114); and rice *Sdr4*, controlling seed dormancy (110)].

This list of crop species and traits reveals some of the limitations of the traditional QTL mapping approach. Because the method requires populations derived from advanced generation crosses, it is limited largely to annual crops where one or more generations can be produced per year. Indeed, all but one of the examples cited above are in annual cereal crops. Equally important is that the phenotypes and genes surveyed are necessarily restricted to the two parental lines used in the cross. Thus, for traits controlled by multiple genes with small effects, this method is ineffective at capturing the range of genetic variation and gene-by-gene interactions that may contribute to a complex phenotype.

QTL Mapping Using Advanced Intercross Populations

One method for overcoming the genetically limited sampling of a biparental cross is to create a mapping population derived from multiple parental lines, potentially including both wild and domesticated plants. Because of the resulting increase in genetic heterogeneity in the mapping population, this approach requires dense marker coverage across the genome, which has become feasible only with the development of next-generation sequencing and single-nucleotide polymorphism (SNP) genotyping approaches (e.g., 39, 78). Advanced intercross populations may be derived using several different crossing designs (94). One of these designs, nested association mapping (NAM), has been applied in the study of domestication-related crop traits, specifically in a series of pioneering studies in maize (12, 13, 26, 61, 71, 87, 118).

In NAM, one or more reference parental lines are crossed with a panel of diverse strains

Table 1 Recent examples of crop domestication and improvement genes[a]

Crop	Gene	Gene category	Trait	Trait category	Causative change	Prevalence	Gene identification method	Selection evidence[b]	Reference(s)
Amaranths (three species)	GBSS1 (Waxy)	Enzyme (starch synthase)	Grain quality (glutinous phenotype)	Improvement	Premature stop codons	Subset of domesticates	Candidate gene	NT	82
Barley	Vrs1	Transcriptional regulator	Inflorescence architecture (two-versus six-rowed)	Domestication	Premature stop (indel or AA change)	Subset of domesticates	QTL mapping	Yes (independent allele evolution)	69
	Nud	Transcriptional regulator	Naked (free-threshing) grains	Domestication	Chromosomal deletion	Subset of domesticates	QTL mapping	NT	114
	INT-C (HvTB1)	Transcriptional regulator	Inflorescence architecture (two-versus six-rowed)	Domestication	Not definitively identified	Subset of domesticates	GWAS using population sample	NT	88
Broomcorn millet	GBSS1 (Waxy), two gene copies	Enzyme (starch synthase)	Grain quality (glutinous phenotype)	Improvement	Indels and AA change	Subset of domesticates	Candidate gene	NT	62
Grape	VvMybA gene family	Transcriptional regulator	Berry color	Improvement	TE insertion and AA change	Subset of domesticates	Candidate gene	Yes	47, 125
Maize	tb1 (teosinte branched1)	Transcriptional regulator	Plant architecture	Domestication	cis-Regulatory via TE insertion	All domesticates	QTL mapping	Yes	109, 129
	ra1 (ramosa1)	Transcriptional regulator	Inflorescence architecture	Domestication	Not definitively identified (likely cis-regulatory)	All domesticates	Candidate gene	Yes	102
	GBSS1 (Waxy)	Enzyme (starch synthase)	Grain quality (glutinous phenotype)	Improvement	Deletions	Subset of domesticates	Candidate gene	Yes	41, 42
	Tu (Tunicate)	Transcriptional regulator	Pod corn	Improvement	cis-Regulatory	Subset of domesticates	QTL mapping, candidate gene	NT	133
	ZmCCT	Transcriptional regulator	Flowering time	Improvement	Not definitively determined	Subset of domesticates	NAM, candidate gene	No	61

	Gene	Function	Trait	Domestication / Improvement	AA change (loss of function)	Domesticates	Method	Evidence	Ref.
Rice	PROG1	Transcriptional regulator	Plant architecture	Domestication	AA change (loss of function)	All domesticates	QTL mapping	Yes	66, 115
	SD1	Hormone synthesis	Culm length (plant height)	Domestication and improvement	AA changes	Subset of domesticates	QTL mapping	Yes	3
	GS5	Putative positive regulator of mitosis	Grain size	Improvement	cis-Regulatory	Subset of domesticates	QTL mapping	NT	74
	Sdr4	Transcriptional regulator	Seed dormancy	Domestication	AA changes	Subset of domesticates	QTL mapping	NT	110
	Bh4	AA transport	Hull color	Improvement	Deletions and premature stop codon	Subset of domesticates	QTL mapping	Marginal evidence	150
	GIF1	Cell wall invertase	Grain filling	Improvement	Probably cis-regulatory	All domesticates	Mutant screens, QTL mapping	Yes	126
	GS3	Putative negative regulator of ovule development	Grain size and length	Improvement	Premature stop	Subset of domesticates	QTL mapping	Yes	40, 113
	TAC1	Unknown	Tiller angle (erect growth)	Improvement	Intron splice site mutation	Subset of domesticates	QTL mapping	NT	144
	qSW5	Putative regulator of outer glume development	Grain width	Improvement	Deletion	Subset of domesticates	QTL mapping	Yes	101
	GW2	Ubiquitin ligase (putative repressor of cell division)	Grain width and weight	Improvement	Premature stop (deletion)	Subset of domesticates (survey incomplete)	QTL mapping	NT	107
	Ehd1	B-type response regulator of flowering	Photoperiod sensitivity (heading date)	Improvement	Premature stop (transposon insertion)	Subset of domesticates	QTL mapping	NT	35, 96
	BADH2	Enzyme (betaine aldehyde dehydrogenase)	Fragrance	Improvement	Premature stop (deletion or AA change)	Subset of domesticates	QTL mapping	Yes	11, 70
	Ghd7	Transcriptional regulator	Grain number, plant height, flowering time	Improvement	Premature stop, cis-regulatory	Subset of domesticates	QTL mapping	Inconclusive	76, 141
	Phr1	Enzyme (polyphenol oxidase)	Grain discoloration (oxidation)	Improvement	Premature stop (indel)	Subset of domesticates	QTL mapping	Yes	51, 146
	OsSPL16 (GW8)	Transcriptional regulator	Grain shape and size	Improvement	cis-Regulatory	Subset of domesticates	QTL mapping	NT	130

(Continued)

Table 1 *(Continued)*

Crop	Gene	Gene category	Trait	Trait category	Causative change	Prevalence	Gene identification method	Selection evidence[b]	Reference(s)
Sorghum	*Sh1 (Shattering1)*	Transcriptional regulator	Shattering	Domestication	*cis*-Regulatory and deletion	All domesticates	QTL mapping	NT	75
	Tannin1	WD40 protein (coordinates multiprotein complexes)	Grain pigmentation	Improvement	Frameshifts causing premature stop codons	Subset of domesticates	QTL mapping	Yes	136
Soybean	*Dt1 (GmTfl1)*	Transcriptional regulator	Determinate growth habit	Domestication	AA change	Subset of domesticates	Candidate gene	NT	119
Sunflower	*HaFT1*	Transcriptional regulator	Flowering time	Domestication	Frameshift (altered but functional protein)	All domesticates	Candidate gene, QTL mapping	Yes	10
Tomato	*fas (fasciated)*	Transcriptional regulator	Locule number (fruit size)	Domestication and improvement	*cis*-Regulatory	Subset of domesticates	QTL mapping	Yes	24
Wheat	*Q* and homeologs	Transcriptional regulator	Free threshing and other traits	Domestication	*cis*-Regulatory and AA change	Subset of domesticates (polyploids)	Mapping in deletion lines, candidate gene	NT	103, 148
	Vrn1	Transcriptional regulator	Vernalization requirement	Domestication and improvement	*cis*-Regulatory, including promoter duplication	Subset of domesticates	QTL mapping	NT	49, 143
	Vrn2 (ZCCT1 and *ZCCT1)*	Transcriptional regulator	Vernalization requirement	Domestication and improvement	AA change and gene deletions	Subset of domesticates	QTL mapping	NT	31, 142

[a]For the genes listed, the molecular genetic basis of the phenotypic variation has been well characterized, and the phenotype has likely been subjected to human selection either during initial domestication or through later breeding for crop improvement. Abbreviations: AA, amino acid; GWAS, genome-wide association study; indel, insertion or deletion; NAM, nested association mapping; QTL, quantitative trait locus; TE, transposable element.
[b]Evidence of positive selection based on patterns of genetic variation, population frequencies of genetic variants in domesticates and wild ancestors, and/or population genetic tests. NT indicates that the selection was not tested.

to generate the mapping population (77, 145); the maize NAM population used in studies to date consists of 5,000 recombinant inbred lines representing 25 families, with the genomic reference line B73 used as the common parent. Association analyses of this population have been used to identify candidate genes for a variety of domestication-related traits, including blight resistance (71, 87), kernel composition (26), leaf architecture (118), and flowering time (13, 61). For example, Hung and colleagues (61) identified 14 photoperiod-response QTLs and fine-mapped the QTL of largest effect to *ZmCCT*, a maize homolog of the rice photoperiod-response regulator *Ghd*7 (61). Interestingly, both maize *ZmCCT* and rice *Ghd*7 appear to have been targets of selection for reduced-function alleles as cultivation of these tropical grasses spread northward into temperate regions; these alleles confer photoperiod insensitivity, allowing for earlier flowering under long-day conditions and grain maturation in regions with shortened growing seasons (61, 76, 141). However, there are also cases of major-effect flowering-time genes that are not detected in the NAM analyses; this is the case for *dwarf8*, which in one study was found to account for 13–32% of the flowering-time variation in a sample of 92 inbred maize lines (117).

Association Mapping Using Unrelated Individuals

An alternative to mapping traits in pedigree populations is to perform linkage disequilibrium mapping using a population of unrelated individuals. First used in studies of human disease, this approach makes use of the natural history of recombination events in a population and tests for associations between phenotypes of interest and genetic markers. Although a specific genomic region of interest may be targeted for analysis, the method more commonly tests for associations with markers distributed throughout the genome [genome-wide association studies (GWAS)], in which case dense marker coverage is required to detect associations with statistical confidence. As with associ-

ation mapping in intercrossed populations, this approach has been greatly facilitated by recent next-generation advances in generating dense marker coverage.

To date, most GWAS research in crop population samples has been conducted in rice, where diverse QTLs and some candidate genes for flowering time and other agronomic traits have been identified (58, 59, 149). Domesticated rice is characterized by several genetically differentiated subgroups; by explicitly sampling across these subpopulations, recent studies have documented the importance of taking population structure into account in GWAS research to avoid spurious marker-trait associations caused by population structure (59, 149). In maize, a GWAS approach using unrelated population samples has recently been applied to identify QTLs and underlying candidate genes for leaf metabolite variation (92). This approach has also been applied in nonmodel crops, including sugar beet, where new QTLs for agronomic traits have been identified despite low marker coverage (137), and pearl millet (*Pennisetum glaucum*), where significant associations were detected between the *PHYC* photoreceptor gene and variation in flowering time and several morphological traits (95; see also 72). A GWAS analysis in barley has been used to identify *INTERMEDIUM-C*, an ortholog of the well-known maize domestication gene *tb1* (*teosinte branched 1*), as a modifier gene in the shift between two- and six-rowed barley cultivars (88).

Genome Resequencing and Screening for Selection Signatures

For crop species with well-characterized reference genomes, genome-wide screening for selection signatures offers a potentially powerful complementary approach to the genetic mapping strategies described above for identifying domestication-related genes. As alluded to above, selection typically winnows variation in genomic locations surrounding genes controlling targeted phenotypes, because only a portion of the standing variation in the

population will carry the alleles under positive selection, and hence only those alleles and alleles of genes in close linkage will be retained. Thus, neutral variation is reduced in these genomic locations relative to the remainder of the genome (**Figure 1**). Genome resequencing or SNP genotyping in a diverse population sample (including individuals of the crop lineage and its wild ancestor) can be used to identify specific genomic regions that bear signatures of domestication-related selection. Further study, using diverse approaches including comparative expression analysis and functional tests, can then support the roles of specific genes and mutations in domestication-related phenotypes.

A key advantage of genome-wide selection screening is that it requires no a priori assumptions about which traits or genes would have been targeted during domestication. Indeed, this approach makes it possible, in principle, to discover traits unknowingly targeted by selection without even having a phenotype in mind! Notwithstanding the proven utility of selection screening (see below), in many cases it will prove methodologically challenging in that putative selective sweeps may encompass dozens to hundreds of genes. Moreover, not all selective events during the domestication process necessarily create a signature of selection; this is particularly true in cases where selection acts on standing genetic variation rather than a newly arisen mutation, which can potentially generate an undetectable "soft sweep" (e.g., 90, 91), and in cases where issues of population structure or insufficient diversity lessen the power to detect sweeps (e.g., 46, 149).

Genome-wide selection screens in crop species to date have focused more on characterizing the genetic architecture of agronomic traits or genome-wide impacts of domestication than on definitively identifying the underlying targets of selection (e.g., 21, 65, 73, 122). Nonetheless, a number of domestication-related candidate genes and/or causative mutations have been identified by this approach. In an early study that relied on genome-wide genic simple sequence repeats, Chapman and colleagues (19) studied genetic

diversity in diverse sunflower accessions to detect candidate targets of selection during the initial domestication process as well as during later crop improvement. Genes involved in amino acid biosynthesis and protein catabolism were differentially identified, a pattern that, curiously, has also been observed in maize (135). In rice, genome sequencing of domesticated varieties and the wild progenitor has been used to identify candidate targets of selection (53, 140). By focusing on genomic regions that show signatures of selection in both domesticated rice subspecies (*indica* and *japonica*), He and colleagues (53) were able to identify 13 candidate domestication genes, including genes encoding regulatory and structural proteins as well as two putative retrotransposon proteins.

In an early genome resequencing study in maize, a comparison of the genome sequence of the Mexican popcorn landrace Palomero to the reference B73 sequence identified genes associated with heavy-metal tolerance as potential targets of selection in the popcorn landrace (124). A more recent genome screen of 75 accessions representing maize and its wild ancestor, teosinte, has revealed a number of important insights into maize domestication, including evidence of stronger selection during domestication than during subsequent improvement and of postdomestication adaptive introgression from teosinte into maize (60). Intriguingly, most of the candidate genes identified as potential targets of selection show much stronger selection signals than do the now-classical domestication genes in maize, such as *tb1* (34, 129). This finding suggests that selection during maize domestication has operated on many genes underlying many diverse biological functions that remain to be characterized.

Reverse Genetics and Candidate Gene Analysis

Insight into the genetic basis of domestication-related traits may in some cases be gleaned using translational approaches based on information from *Arabidopsis*, *Oryza*, or other model

systems. For example, glutinous (waxy) culti-vars have been selectively favored in a number of cereal and pseudocereal crop species, includ-ing rice, maize, barley, foxtail millet, and grain amaranths; in all cases, the trait has been shown to originate in loss-of-function mutations at the *Waxy* gene, which encodes a granule-bound starch synthase required for the production of amylose (see **Table 1** for recent examples). For more complex traits, candidate gene approaches can be particularly effective when combined with QTL mapping, screening of mutagenized lines, or selection screening, so that molecular analyses focus specifically on candidate genes that fall within a genomic region already impli-cated in domestication-related selection. The latter approach was recently used in sunflowers, for example, to identify several flowering-time genes that colocalize with flowering-time QTLs (9); several of these genes, all *Ara-bidopsis FT/TFL1* homologs, show molecular signatures consistent with selection during do-mestication or early improvement. Similarly, a combination of demographic modeling and diversity analysis was used, together with translational information derived largely from *Arabidopsis*, to demonstrate that flowering-time genes in pearl millet show evidence of selection during domestication or later improvement (23).

For phylogenetically closely related crop species such as cereal crops, the occurrence of a QTL across syntenic genomic regions can pro-vide additional support for the role of candi-date gene orthologs across multiple domesti-cation events. In the 1990s, comparative QTL mapping in maize, rice, and sorghum suggested that the reduced-seed-shattering domestication phenotype in these cereals might have occurred at least in part through parallel selection on an orthologous gene (85). The underlying candi-date gene, *Sh1*, has recently been cloned and functionally characterized in sorghum, where it was found to encode a *YABBY* transcription fac-tor. As predicted by the earlier QTL mapping, orthologs of this gene colocalize with shattering QTLs in rice and maize as well as foxtail millet (75). Moreover, the corresponding rice gene,

although not a major-effect shattering QTL, does emerge as a putative domestication gene in recent selection screens of the rice genome (see above) (53, 140).

ADVANCES IN UNDERSTANDING THE GENETIC BASIS OF DOMESTICATION-RELATED TRAITS

Genetic Architecture

One of the long-standing questions in stud-ies of crop domestication is whether the ge-netic architecture of traits evolving in response to artificial selection differs qualitatively from traits subject to natural selection in wild species. Whereas complex traits in wild species are typ-ically controlled by many genes with small effects (reviewed in 12), evidence until re-cently suggested that this was not the norm for domestication-related traits. In particular, QTL mapping studies using biparental crosses have generally pointed to a relatively few genes, each with large effects, as the determinants of many domestication-related traits (e.g., see 14). However, as the study of domestication phenotypes has moved toward more sophisti-cated association mapping and selection screens as well as quantitatively varying characters, a more complex picture has emerged, with many domestication-related traits being revealed as polygenic. It should be noted that because the identification of large and small effect sizes de-pends on the mapping population composition and statistical approaches employed, this shift is at least partly a direct reflection of changes in experimental design.

Evidence for many small-effect genes is provided by studies using the recently cre-ated maize NAM population. Dozens of small-effect QTLs have been detected for most maize traits examined to date, including flowering time (13), kernel composition (26), leaf mor-phology (118), and resistance to both northern leaf blight (87) and southern leaf blight (71). These small-effect genes are almost entirely ad-ditive in their effects, with little evidence of

epistasis (i.e., nonadditive gene-by-gene interactions). A somewhat different pattern is observed for maize inflorescence traits: Brown and colleagues (12) have observed larger-effect QTLs for inflorescence traits than for flowering time and leaf morphology, and larger-effect QTLs for female inflorescence (ear) traits than for male (tassel) traits. This pattern suggests a relationship between QTL effect size and lack of developmental stability associated with recently evolved traits; the maize ear evolved only recently, and so may be developmentally less canalized than traits such as flowering time (12).

In rice, recent GWAS research also suggests polygenic control of a number of selected agronomic traits, but with somewhat greater heterogeneity in genetic architecture among traits compared with maize. Polygenic rice traits include disease resistance and tolerance to drought and salt stress, whereas a smaller number of larger-effect QTLs are associated with flowering time (58, 59). There is also evidence for heterogeneity in the genetic architecture of traits among the different genetic subgroups within domesticated rice (149). Unlike maize, which is an outcrossing species, domesticated rice is largely self-fertilizing, with a genome characterized by low effective recombination and extensive linkage disequilibrium. This difference in population genomic structure, combined with the population substructure present in rice, would be expected to contribute to this greater observed heterogeneity in genetic architectures. Although we are clearly at the very early stages of understanding the genetic architecture of domestication-related traits, it seems likely that future studies in species beyond maize and rice will continue to reveal variability in genetic architectures among traits and among crop species.

Genomic Structural Variation

Broadly speaking, genomic structural variation describes segmental alterations of DNA that are greater than 1 kb (44); this includes gene copy number variation (CNV), presence/absence variation (PAV), and larger-scale chromosomal inversions, translocations, and segmental duplications. Structural variation is increasingly recognized as a common feature of organismal genomes, including in the human genome, where CNVs and associated gene dosage effects have been implicated in a number of diseases (25).

Among crop species, genomic studies in maize have led the way in generating insights into the role of genomic structural variation in shaping domestication-related phenotypes. Following the 2009 publication of the maize B73 reference genome sequence (99), a genome-wide comparison of two inbred lines was performed using microarray oligonucleotides for comparative genomic hybridization (CGH); hundreds to thousands of CNV/PAV features were detected, with estimates depending on the calculation method (108). Similarly, CGH among 14 inbred maize lines has revealed thousands of CNVs (6). An extended survey of 19 diverse inbred maize lines and 14 teosinte accessions indicated that most genic CNV and PAV features (~86%) are present in the wild ancestor; comparisons of these CNV/PAV frequencies in teosinte and maize suggested that selection on these ancestral CNVs has not played a major role in maize domestication or improvement (112). However, a recent genome-wide SNP screen of 103 diverse maize and teosinte lines (21) does suggest a correlation between genomic regions containing structural variation [detected as read-depth variants (RDVs) in genome resequencing] and QTLs for agronomic traits. Genomic regions containing QTLs for leaf architecture and resistance to northern and southern leaf blight are enriched for RDVs. This finding suggests a potential role for CNV/PAV in generating phenotypic variation for these agronomic traits.

CNV/PAV has been reported to be differentially represented among genes categorized as being involved in stress and stimulus response, perhaps in part because this category includes some large gene families (e.g., NBS-LRR genes). This pattern is detectable on a genome-wide scale in maize (21) and rice (140)

as well as in *Arabidopsis* (93). In addition, wheat CNVs at the photoperiod-response gene *Ppd-B1* and the vernalization-requirement gene *Vrn-A1* have been shown to underlie some of the variation among cultivars in flowering time (30). Similarly, CNVs at the barley *FR-2* locus are associated with varietal differences in freezing tolerance (67). The enrichment of maize RDVs at QTLs for northern and southern leaf blight resistance (21) is also consistent with this general pattern.

Transposable Elements

A major source of genomic structural variation comes from the massive proliferation of transposable elements (TEs) that characterizes the genomes of many plant species. A recent review (79) indicated that transposons constitute between 22% and 85% of the total genomic contents of 11 crop species examined. Gene insertions and other structural rearrangements caused by transposon activity can potentially provide a rich source of phenotypic diversity that can be selected on during domestication, either directly through mutagenesis or indirectly through their effects on gene expression (55). Perhaps most famous in this regard is maize, the organism in which TEs were discovered. Extraordinary diversity is known to occur among maize lines and alleles (128), largely reflecting the direct (insertion) or indirect (post-TE insertion recombination) effects of TE activity (147). Although maize has a particularly "active" genome, TE activity is implicated in generating diversity in many other crop plants. In grapes, for example, genome-wide surveys of class II transposons (8) and MITEs (miniature inverted-repeat transposable elements, a particular type of nonautonomous class II transposon) (7) indicate that these repetitive elements have been actively proliferating during vegetative propagation following domestication and subsequent breeding, contributing to the high genetic diversity in the domesticated grape genome.

Although most transposon-related repeat proliferation does not affect plant phenotypes,

there are a number of cases where domestication or improvement traits have arisen through insertions of TEs into genes or their *cis*-regulatory regions. In fact, a transposon-mediated insertion has been found to be the causal factor underlying perhaps the most renowned domestication gene, *tb1* in maize. The transition from the highly branched wild teosinte plant to the single-stemmed maize phenotype is controlled largely by increased expression of *tb1*, which encodes a transcriptional regulator that represses growth (33). By using combined inferences from fine mapping in maize-teosinte introgression lines and selective sweep mapping in maize landraces, Studer and colleagues (109) have recently shown that a *Hopscotch* retroelement insertion in the *tb1 cis*-regulatory region accounts for the increased *tb1* expression characterizing the maize plant phenotype. Remarkably, this insertion is located approximately 60 kb upstream of the *tb1* coding region, which draws added attention to the multiple possibilities for active TEs to cause dramatic morphological change. Molecular dating of this transposon insertion indicates that it predates the domestication of maize and therefore must have existed in the standing variation of the teosinte ancestor before being subject to selection.

A number of additional cases of TE involvement in domestication or improvement traits have also recently been documented. In maize, the tunicate (pod corn) phenotype arises through ectopic expression of *Tunicate* (*Tu*), leading to the development of papery glumes over individual kernels. This altered expression appears to arise in part from promoter sequence rearrangements facilitated by the insertion of a MuDR-like TE (133). In grapes, the white berry phenotype arose in part through insertion of a *Gret1* gypsy-type retrotransposon in the promoter of *VvMybA1*, a transcription factor controlling anthocyanin synthesis. Interestingly, this insertion was preceded by a single-nucleotide loss-of-function mutation in the adjacent *VvMybA2* gene, with the modern white grape berry arising through progressive selection events for the two nonfunctional

genes (47, 68, 125). A promoter transposon insertion is also the source of several loss-of-function mutations in wheat *Vrn1* homeologs; nonfunctional alleles of this vernalization gene are partly responsible for the phenotypic shift between winter and spring wheat varieties (49).

GENE AND GENOME DUPLICATIONS IN DOMESTICATION

A prominent feature of plant genomes is that they contain high levels of gene duplication. Nearly all genes exist as members of multigene families, with various copies (paralogs) related to one another through both ancient and more recent duplication events, often tracing back to before the origin of seed plants (64). These duplications have arisen through various mechanisms (45), the most prominent being whole-genome doubling, or polyploidy (36, 106, 132). It is now evident that the phylogenetic history of all plants includes multiple episodes of polyploidy (64), with each event doubling the degenerated duplicated genome surviving from the previous event, thereby creating a series of temporally nested gene duplications. In addition to polyploidy, which remains a prominent speciation process in many lineages today, more localized or tandem gene duplications are also characteristic features of plant genomes, contributing continually and contemporaneously to the duplicated genic content of plant genomes. These more localized duplications are generated by a diversity of mechanisms involving TEs or by small-scale genomic duplications arising from unequal crossing over and chromosomal anomalies (45).

Given that duplication is a prominent feature of the plant genomic architecture, there has long been an interest in understanding the forces and processes that dictate gene survivorship (e.g., 5, 84, 98) and adaptive genic diversification (36, 104, 105) following duplication. With respect to crop improvement, Paterson (83) and Udall & Wendel (120) reviewed several of the features of polyploid plants that are widely suspected to play a role in function and crop diversification, including allelic complementation, increased allelic diversity, environmental buffering, and possibilities for fine-tuning dosage leading to novel phenotypic variation (e.g., 81, 100). Also, evidence abounds of polyploidy playing an important role in the origin of certain crop traits, such as free threshing (38) and other traits (37) in hexaploid wheat. In most cases, however, the connections between traits and their molecular genetic determinants remain to be elucidated. This is an area in which rapid progress may be expected in the coming years, as empirical examples, at the molecular level, of a role for polyploidy or gene duplication in crop domestication and improvement begin to emerge.

One of the most remarkable examples of variation in a domestication trait relating to polyploidy is at the complex hardness locus (*Ha*) of wheat (18). This ∼60-kb compound locus contains several genes, present in the A, B, and D diploid wheat genomes, all leading to soft wheat grains. The hardness-locus genes were deleted from both wild and domesticated forms of tetraploid (AB genome) wheat, leading to the hard wheat grain trait, which is important in the pasta industry. Creation of hexaploid (ABD genome) wheat initially restored the missing genes owing to the addition of the D genome, which led to selection for the soft grains that are important for bread wheat. Subsequent mutations, deletions, and rearrangements of the hardness locus of the D genome of hexaploid wheat generated variation in hexaploid wheat seed quality, such as semihard wheats that were further subjected to human selection. These variations were shown to involve complex rearrangements and recombination between retroelements.

A second illustrative example from wheat concerns the *Q* locus, which confers free threshing in hexaploid wheat and is also pleiotropically implicated in a number of other important domestication and improvement traits, including plant height, inflorescence architecture, and flowering time. Zhang et al. (148) showed that the *Q* gene, which encodes a member of the AP2 family of transcription

factors, has experienced a fascinating and complex series of events during diploid and polyploid wheat divergence. This history includes duplication and subsequent loss of different paralogs at the diploid level, such that one gene was lost in A-genome diploids whereas the other gene was lost in the B- and D-genome diploids. Following polyploid formation, and the reunion of these now-diverged paralogs into a common nucleus, a single valine-to-isoleucine amino acid replacement in the A homeolog generated the *Q* phenotype, at least in part. The story is further complicated by two additional factors: the postpolyploidization pseudogenization (but continued transcription) of the surviving B-genome gene, which apparently contributes to homeolog expression regulation; and the D-genome homeolog, which contributes to the pleiotropic aspects of the free-threshing mutant phenotype. This example is remarkable in that it demonstrates a wholly unexpected avenue by which polyploidy may contribute to plant domestication, entailing a complex combination of ancient paralogy, subsequent gene loss, reunion of divergent paralogs, and continued interaction of subfunctionalized pseudohomeologs.

Recent studies have also revealed cases where selection during domestication has targeted paralogous genes derived from past gene duplication events, leading to sub- or neofunctionalization. Three examples illustrate the wide range of temporal scales for postduplication differentiation, ranging from recent to ancient. The first involves elongated-fruit varieties of tomato, in which a *Copia*-like long-terminal-repeat retrotransposon (*Rider*) was involved in a complex structural rearrangement that duplicated and retrotransposed a 24.7-kb genomic region from chromosome 7 to 10. This process generated a duplication of *Sun*, a major gene controlling elongated fruit shape, as well as increased *Sun* expression in varieties with elongated fruits, attributable in part to the duplicated gene copy having been placed in a location where it is upregulated through co-option of *cis*-regulatory factors that normally cause high expression of a different

gene (encoding a defensin protein) during fruit development (138).

A second example relates to paralogs of a flowering-time gene family in sunflower (9, 10). *FT* (*FLOWERING LOCUS T*) genes are photoperiod-induced positive regulators of reproductive meristem development. Blackman and colleagues were able to identify four paralogous copies in sunflower, three of which appear to be functional and colocalize with a major flowering-time QTL in mapping populations derived from wild-domesticate crosses. Additionally, these loci show molecular signatures consistent with selection during domestication. The three functional genes also show divergence in expression patterns between wild and domesticated sunflowers that are associated with the shift to earlier, long-day-responsive flowering in the domesticate. In addition, one of the paralogs, *HaFT1*, carries a protein-coding frameshift mutation in domesticates that alters developmental timing directly through interference with the expression of another paralog, *HaFT4*. The other functional paralog, *HaFT2*, shows altered expression arising through both *cis*- and *trans*-regulatory changes. This combination of features strongly suggests a key role for functional divergence in these *FT* paralogs in the sunflower domestication process, and further illustrates the multiplicity of molecular mechanisms by which altered phenotypes might arise.

A final example is from rice, involving two ancient paralogs (possibly dating to the origin of the grasses) that appear to have been independently targeted by selection during domestication. *OsCIN* is an ancient duplicate of *GIF1*, and both genes encode invertase enzymes that function during grain development, with tissue-specific differentiation in gene expression (127). Both genes also show molecular signatures of selection during domestication. For *GIF1*, selection for changes in the promoter region has led to decreased expression during grain development, which results in increased grain filling and yield (126). For *OsCIN1*, there appears to have been selection to fix a single protein-coding mutation during

domestication. Although the functional effects of this change in *OsCIN1*, if any, remain to be determined, the parallel selection signatures suggest that both paralogs have played a role in changes in rice grain development during rice domestication.

SYNTHESIS AND CONCLUSIONS

As the genetic bases of crop domestication and improvement traits have been resolved, it has become possible to assess whether generalizations can be drawn about the molecular underpinnings of phenotypic changes during domestication. Concomitantly, questions involving the relative roles of regulatory and structural genes in adaptive evolution have been a topic of considerable interest in the field of evolutionary biology in recent years (17, 54, 134). In their 2006 review on the molecular genetics of domestication, Doebley and colleagues (32) made several key observations and predictions regarding the genetic targets of selection during domestication and the molecular nature of the underlying mutations, including the following:

1. Changes in developmentally and morphologically complex traits, including many domestication traits (e.g., seed shattering and plant architecture), occur through selection on transcriptional regulators.
2. For simpler traits that involve specific metabolic pathways, such as carbohydrate or pigment biosynthesis, changes may occur through selection on structural genes within the pathway as well as selection on regulatory genes.
3. Changes that account for phenotypic differences among varieties of a crop commonly occur through loss-of-function mutations, although *cis*-regulatory changes can also play a role.

These patterns have generally been borne out by more recently published studies (**Table 1**). Among domestication-related genes characterized since 2007, those under-

lying complex phenotypes are primarily either transcriptional regulators or proteins known to regulate basic developmental processes such as hormone synthesis [e.g., rice *SD1*, controlling plant height (3)] and cell growth and division [e.g., rice *GS5*, *GIF1*, and *GW2*, controlling grain development (74, 107, 126)]. For less developmentally complex traits, both structural and regulatory genes have continued to be identified as the targets of selection, including enzymes that function within specific biosynthetic pathways [e.g., *Waxy*, controlling amylose synthesis (42, 62, 82)] and transcriptional regulators [e.g., the grape *VvMybA* gene family, controlling berry color (47, 125)]. In addition, among recently characterized genes that control varietal differences among crops, loss-of-function mutations have continued to account for much of the underlying genetic variation. Interestingly, a number of these mutations occur at genes controlling complex phenotypes, where pleiotropic effects might be expected to be strongly deleterious; examples include inflorescence architecture and grain development in barley [e.g., *Vrs1* and *Nud* (69, 114)] and grain development and flowering time in rice [e.g., *Ghd7* (76, 141)]. Still other cases represent examples implicating changes in microRNA genes or their expression levels in domestication traits, as suggested in rice (131) and maize (22).

Collectively, the examples introduced in **Table 1** and elsewhere in this review illustrate the idiosyncratic or perhaps opportunistic nature of human selection pressure with respect to the diversity of underlying causative mutations affecting phenotypes that are important in plant domestication and improvement. Not surprisingly, just as natural selection may variously entail fixation of amino acid–causing changes in protein-coding genes, as well as a multiplicity of genomic structural changes large and small that affect expression levels, this full spectrum of possibilities is being revealed in studies of crop domestication. Undoubtedly our understanding of the molecular underpinnings of domestication traits will continue to grow as the identification of domestication-related genes

begins to move toward phenotypically unbiased genomic selection screens (e.g., 60, 65, 122).

Moving forward, an equally important and complementary advance will be in the increasing application of genome-scale systems biology approaches to study domestication. Multiple "-omics" (e.g., genomics, transcriptomics, proteomics, metabolomics) and analyses of pathways and networks across various scales hold promise for revealing many of the intricacies of domestication and crop improvement and, by extension (echoing the words of Darwin invoked in the introduction to this review), the evolutionary process in general. Much of this review has focused on the mutations responsible for phenotypes found in crop plants, and to be sure, considerable progress has been made in this regard (**Figure 2**, **Table 1**). But there is a vast biology lying between genotype and phenotype, with the latter reflecting the end product of a complex transduction and propagation through the transcriptomic, proteomic, and metabolomic networks that lead to biosynthesis.

Recent forays into this arena have revealed astonishing complexity. In maize, for example, Hufford et al. (60) used a combination of genome resequencing and comparative expression profiling to reveal a surprisingly large number (1,179) of genomic regions that may have been subjected to selection during domestication (484) and crop improvement (695). Candidate domestication genes display greater gene expression change between maize and teosinte than do noncandidate genes, are on average expressed at higher levels, and have reduced expression variability; the latter is interpreted as potentially reflecting directional selection for reducing *cis*-regulatory variation. An extension of this work (111), involving comparative expression profiling of seedlings in 24 teosinte versus 38 maize accessions, led to the detection of more than 600 differentially expressed genes, many in genomic locations that were identified in population genomic diversity screens (60) to be targets of selection.

This evidence of large-scale rewiring of the transcriptome in response to domestication has
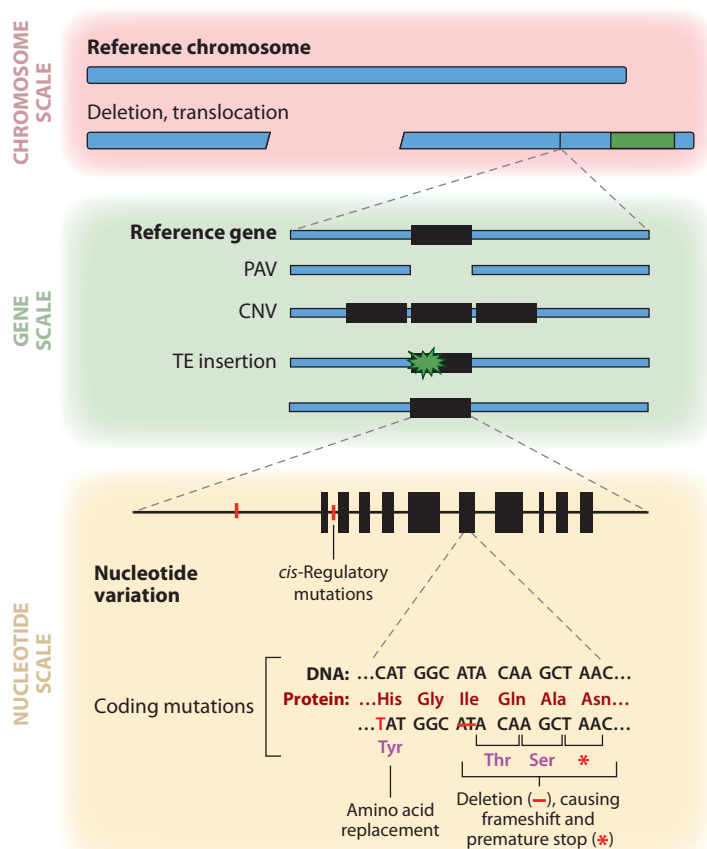
Figure 2

Mutations targeted by selection during domestication. Decreasing levels of genomic scale are indicated from top to bottom. Large-scale chromosomal rearrangements include large genomic deletions (indicated by a missing chromosomal segment) and transpositions from other chromosomes (indicated by the inserted green segment). Structural variation at the genic level includes gene presence/absence variation (PAV) and copy number variation (CNV) as well as insertions of transposable elements (TEs) into coding or *cis*-regulatory regions. At the nucleotide level, mutations may affect coding regions (e.g., substitutions or insertions/deletions that alter the protein sequence) or *cis*-regulatory sequences.

also been reported in cotton. Rapp et al. (89) studied the transcriptome of developing cotton (*Gossypium*) "fibers" (seed epidermal trichomes) in both wild and domesticated cotton during five stages representing primary and secondary wall synthesis, reporting significantly altered expression for 9,645 genes. And this is just for a single-celled structure! Other transcriptomic studies in the cotton model system

are revealing a comparably massive rewiring of the transcriptome accompanying domestication (20, 56, 57). As an example of the type of insight that can emerge from these studies, Bao et al. (4) used a combination of genomic and proteomic tools to investigate one of the protein families (profilin) implicated as highly upregulated during cotton domestication. Remarkably, it was not just a single gene that was upregulated; instead, all five profilin gene family members expressed in cotton fibers were simultaneously upregulated. This observation presumably reflects the downstream effects of upstream regulatory alterations of perhaps just a single mutation, which during cellular development propagates through the system to affect transcriptome and proteome levels for the whole profilin gene family. A final noteworthy and somewhat remarkable observation of this study is that the same phenomenon was observed in all three independently domesticated cotton species studied—two allopolyploids (*G. barbadense* and *G. hirsutum*) and one diploid (*G. herbaceum*). It will be of considerable interest to discover the specific genomic changes in each species that in parallel mediate these phenotypically homologous responses to domestication, and to assess their degree of genome-level similarities and differences.

These above examples foreshadow the types of insights into the origin of phenotypic innovation that will soon emerge from the application of an increasingly powerful suite of technologies to the study of crop plants and their living progenitors. This effort will yield novel perspectives on the myriad connections and networks that lie between the genomic landscape and the phenotypes that are targeted by human selection. In this sense, Darwin's promise that crop plants constitute our most gifted tutors for understanding the evolutionary process will continue to be fulfilled.

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

LITERATURE CITED

1. Akhunov ED, Akhunova AR, Anderson OD, Anderson J, Blake N, et al. 2010. Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genomics* 11:702
2. Arabidopsis Genome Init. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
3. Asano K, Yamasaki M, Takuno S. 2011. Artificial selection for a green revolution gene during *japonica* rice domestication. *Proc. Natl. Acad. Sci. USA* 108:11034–39
4. Bao Y, Hu G, Flagel LE, Salmon A, Bezanilla M, et al. 2011. Parallel up-regulation of the profilin gene family following independent domestication of diploid and allopolyploid cotton (*Gossypium*). *Proc. Natl. Acad. Sci. USA* 108:21152–57
5. Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, et al. 2008. Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. *Mol. Biol. Evol.* 25:2445–55
6. Beló A, Beatty MK, Hondred D, Fengler K, Li B, Rafalski A. 2010. Allelic genome structural variations in maize detected by array comparative genome hybridization. *Theor. Appl. Genet.* 120:355–67

7. Benjak A, Boué S, Forneck A, Casacuberta JM. 2009. Recent amplification and impact of MITEs on the genome of grapevine (*Vitis vinifera* L.). *Genome Biol. Evol.* 1:75–84

8. Benjak A, Forneck A, Casacuberta JM. 2008. Genome-wide analysis of the "cut-and-paste" transposons of grapevine. *PLoS ONE* 3:e3107

9. Blackman BK, Rasmussen D, Strasburg JL, Raduski AR, Burke JM, et al. 2011. Contributions of flowering time genes to sunflower domestication and improvement. *Genetics* 187:271–87

10. Blackman BK, Strasburg JL, Raduski AR, Michaels SD, Rieseberg LH. 2010. The role of recently derived *FT* paralogs in sunflower domestication. *Curr. Biol.* 20:629–35

11. Bradbury LMT, Fitzgerald TL, Henry RJ, Jin Q, Waters DLE. 2005. The gene for fragrance in rice. *Plant Biotechnol. J.* 3:363–70

12. Brown PJ, Upadyayula N, Mahone GS, Tian F, Bradbury PJ, et al. 2011. Distinct genetic architectures for male and female inflorescence traits of maize. *PLoS Genet.* 7:e1002383

13. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, et al. 2009. The genetic architecture of maize flowering time. *Science* 325:714–18

14. Burger JC, Chapman MA, Burke JM. 2008. Molecular insights into the evolution of crop plants. *Am. J. Bot.* 95:113–22

15. Burke JM, Burger JC, Chapman M. 2007. Crop evolution: from genetics to genomics. *Curr. Opin. Genet. Dev.* 17:525–32

16. Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, et al. 2007. Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genet.* 3:1745–56

17. Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25–36

18. Chantret N, Cenci A, Sabot F, Anderson O, Dubcovsky J. 2004. Sequencing of the *Triticum monococcum* hardness locus reveals good microcolinearity with rice. *Mol. Genet. Genomics* 271:377–86

19. Chapman MA, Pashley CH, Wenzler J, Hvala J, Tang S, et al. 2008. A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). *Plant Cell* 20:2931–45

20. Chaudhary B, Hovav R, Flagel L, Mittler R, Wendel JF. 2009. Parallel expression evolution of oxidative stress-related genes in fiber from wild and domesticated diploid and polyploid cotton (*Gossypium*). *BMC Genomics* 10:378

21. Chia J-M, Song C, Bradbury PJ, Costich D, de Leon N, et al. 2012. Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* 44:803–7

22. Chuck G, Cigan AM, Saeteurn K, Hake S. 2007. The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA. *Nat. Genet.* 39:544–49

23. Clotault J, Thuillet A-C, Buiron M, De Mita S, Couderc M, et al. 2012. Evolutionary history of pearl millet (*Pennisetum glaucum* [L.] R. Br.) and selection on flowering genes since its domestication. *Mol. Biol. Evol.* 29:1199–212

24. Cong B, Barrero LS, Tanksley SD. 2008. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nat. Genet.* 40:800–4

25. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, et al. 2010. Origins and functional impact of copy number variation in the human genome. *Nature* 464:704–12

26. Cook JP, McMullen MD, Holland JB, Tian F, Bradbury P, et al. 2012. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol.* 158:824–34

27. Cornille A, Gladieux P, Smulders MJM, Roldán-Ruiz I, Laurens F, et al. 2012. New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genet.* 8:e1002703

28. Darwin C. 1859. *On the Origin of Species*. London: Murray

29. D'Hont A, Denoeud F, Aury J-M, Baurens F-C, Carreel F, et al. 2012. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488:213–17

30. Díaz A, Zikhali M, Turner AS, Isaac P, Laurie DA. 2012. Copy number variation affecting the *Photoperiod-B1* and *Vernalization-A1* genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PLoS ONE* 7:e33234

31. Distelfeld A, Tranquilli G, Li C, Yan L, Dubcovsky J. 2009. Genetic and molecular characterization of the *VRN2* loci in tetraploid wheat. *Plant Physiol.* 149:245–57

32. Doebley JF, Gaut BS, Smith BD. 2006. The molecular genetics of crop domestication. *Cell* 127:1309–21

33. Doebley JF, Stec A, Hubbard L. 1997. The evolution of apical dominance in maize. *Nature* 386:485–88

34. Doebley JF, Stec A, Wendel J, Edwards M. 1990. Genetic and morphological analysis of a maize-teosinte F$_2$ population: implications for the origin of maize. *Proc. Natl. Acad. Sci. USA* 87:9888–92

35. Doi K, Izawa T, Fuse T, Yamanouchi U. 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.* 2:926–36

36. Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, et al. 2008. Evolutionary genetics of genome merger and doubling in plants. *Annu. Rev. Genet.* 42:443–61

37. Dubcovsky J, Dvorak J. 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316:1862–66

38. Dvorak J, Deal KR, Luo M-C, You FM, von Borstel K, Dehghani H. 2012. The origin of spelt and free-threshing hexaploid wheat. *J. Hered.* 103:426–41

39. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379

40. Fan C, Xing Y, Mao H, Lu T, Han B, et al. 2006. *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* 112:1164–71

41. Fan L, Bao J, Wang Y, Yao J, Gui Y, et al. 2009. Post-domestication selection in the maize starch pathway. *PLoS ONE* 4:e7612

42. Fan L, Quan L, Leng X, Guo X, Hu W, et al. 2008. Molecular evidence for post-domestication selection in the *Waxy* gene of Chinese waxy maize. *Mol. Breed.* 22:329–38

43. Feuillet C, Leach JE, Rogers J, Schnable PS, Eversole K. 2011. Crop genome sequencing: lessons and rationales. *Trends Plant Sci.* 16:77–88

44. Feuk L, Carson AR, Scherer SW. 2006. Structural variation in the human genome. *Nat. Rev. Genet.* 7:85–97

45. Flagel LE, Wendel JF. 2009. Gene duplication and evolutionary novelty in plants. *New Phytol.* 183:557–64

46. Flowers JM, Molina J, Rubinstein S, Huang P, Schaal BA, Purugganan MD. 2012. Natural selection in gene-dense regions shapes the genomic pattern of polymorphism in wild and domesticated rice. *Mol. Biol. Evol.* 29:675–87

47. Fournier-Level A, Lacombe T, Le Cunff L, Boursiquot J-M, This P. 2010. Evolution of the *VvMybA* gene family, the major determinant of berry colour in cultivated grapevine (*Vitis vinifera* L.). *Heredity* 104:351 –62

48. Garcia-Mas J, Benjak A, Sanseverino W, Bourgeois M, Mir G, et al. 2012. The genome of melon (*Cucumis melo* L.). *Proc. Natl. Acad. Sci. USA* 109:11872–77

49. Golovnina KA, Kondratenko EY, Blinov AG, Goncharov NP. 2010. Molecular characterization of vernalization loci *VRN1* in wild and cultivated wheats. *BMC Plant Biol.* 10:168

50. Gross BL, Olsen KM. 2010. Genetic perspectives on crop domestication. *Trends Plant Sci.* 15:529–37

51. Gross BL, Skare KJ, Olsen KM. 2009. Novel *Phr1* mutations and the evolution of phenol reaction variation in US weedy rice (*Oryza sativa*). *New Phytol.* 184:842–50

52. Hammer K. 1984. Das Domestikationssyndrom. *Kulturpflanze* 32:11–34

53. He Z, Zhai W, Wen H, Tang T, Wang Y, et al. 2011. Two evolutionary histories in the genome of rice: the roles of domestication genes. *PLoS Genet.* 7:e1002100

54. Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61:995–1016

55. Hollister JD, Smith LM, Guo Y-L, Ott F, Weigel D, Gaut BS. 2011. Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. *Proc. Natl. Acad. Sci. USA* 108:2322–27

56. Hovav R, Chaudhary B, Udall JA, Flagel L, Wendel JF. 2008. Parallel domestication, convergent evolution and duplicated gene recruitment in allopolyploid cotton. *Genetics* 179:1725–33

57. Hovav R, Udall JA, Chaudhary B, Hovav E, Flagel L, et al. 2008. The evolution of spinnable cotton fiber entailed prolonged development and a novel metabolism. *PLoS Genet.* 4:e25

58. Huang X, Wei X, Sang T, Zhao Q, Feng Q, et al. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42:961–67

59. Huang X, Zhao Y, Wei X, Li C, Wang A, et al. 2012. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44:32–39

60. Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia J-M, et al. 2012. Comparative population genomics of maize domestication and improvement. *Nat. Genet.* 44:808–11

61. Hung H-Y, 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. *Proc. Natl. Acad. Sci. USA* 109:E1913–21

62. Hunt HV, Denyer K, Packman LC, Jones MK, Howe CJ. 2010. Molecular basis of the waxy endosperm starch phenotype in broomcorn millet (*Panicum miliaceum* L.). *Mol. Biol. Evol.* 27:1478–94

63. Int. Rice Genome Seq. Proj. 2005. The map-based sequence of the rice genome. *Nature* 436:793–800

64. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, et al. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473:97–100

65. Jiao Y, Zhao H, Ren L, Song W, Zeng B, et al. 2012. Genome-wide genetic changes during modern breeding of maize. *Nat. Genet.* 44:812–15

66. Jin J, Huang W, Gao J-P, Yang J, Shi M, et al. 2008. Genetic control of rice plant architecture under domestication. *Nat. Genet.* 40:1365–69

67. Knox AK, Dhillon T, Cheng H, Tondelli A, Pecchioni N, Stockinger EJ. 2010. CBF gene copy number variation at *Frost Resistance-2* is associated with levels of freezing tolerance in temperate-climate cereals. *Theor. Appl. Genet.* 121:21–35

68. Kobayashi S, Goto-Yamamoto N, Hirochika H. 2005. Association of *VvMybA1* gene expression with anthocyanin production in grape (*Vitis vinifera*) skin-color mutants. *J. Jpn. Soc. Hortic. Sci.* 74:196–203

69. 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. *Proc. Natl. Acad. Sci. USA* 104:1424–29

70. Kovach MJ, Calingacion MN, Fitzgerald MA, McCouch SR. 2009. The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc. Natl. Acad. Sci. USA* 106:14444–49

71. Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, et al. 2011. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* 43:163–68

72. Lakis G, Navascués M, Rekima S, Simon M, Remigereau M-S, et al. 2012. Evolution of neutral and flowering genes along pearl millet (*Pennisetum glaucum*) domestication. *PLoS ONE* 7:e36642

73. Lam H-M, Xu X, Liu X, Chen W, Yang G, et al. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* 42:1053–59

74. Li Y, Fan C, Xing Y, Jiang Y, Luo L, et al. 2011. Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nat. Genet.* 43:1266–69

75. Lin Z, Li X, Shannon LM, Yeh C-T, Wang ML, et al. 2012. Parallel domestication of the *Shattering1* genes in cereals. *Nat. Genet.* 44:720–24

76. Lu L, Yan W, Xue W, Shao D, Xing Y. 2012. Evolution and association analysis of *Ghd7* in rice. *PLoS ONE* 7:e34021

77. 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

78. Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17:240–48

79. Morrell PL, Buckler ES, Ross-Ibarra J. 2011. Crop genomics: advances and applications. *Nat. Rev. Genet.* 13:85–96

80. Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, et al. 2011. Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci. USA* 108:3530–35

81. Ni Z, Kim E-D, Ha M, Lackey E, Liu J, et al. 2009. Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* 457:327–31

82. Park Y-J, Nemoto K, Nishikawa T, Matsushima K, Minami M, Kawase M. 2010. Waxy strains of three amaranth grains raised by different mutations in the coding region. *Mol. Breed.* 25:623–35

83. Paterson AH. 2005. Polyploidy, evolutionary opportunity, and crop adaptation. *Genetica* 123:191–96

84. Paterson AH, Chapman BA, Kissinger JC, Bowers JE, Feltus FA, Estill JC. 2006. Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis*, *Oryza*, *Saccharomyces* and *Tetraodon*. *Trends Genet.* 22:597–602

85. Paterson AH, Lin YR, Li Z, Schertz KF, Doebley JF, et al. 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269:1714–18

86. Pickersgill B. 2009. Domestication of plants revisited—Darwin to the present day. *Bot. J. Linn. Soc.* 161:203–12

87. Poland JA, Bradbury PJ, Buckler ES, Nelson RJ. 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc. Natl. Acad. Sci. USA* 108:6893–98

88. 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.* 43:169–72

89. Rapp RA, Haigler CH, Flagel L, Hovav RH, Udall JA, Wendel JF. 2010. Gene expression in developing fibres of upland cotton (*Gossypium hirsutum* L.) was massively altered by domestication. *BMC Biol.* 8:139

90. Raquin A-L, Brabant P, Rhoné B, Balfourier F, Leroy P, Goldringer I. 2008. Soft selective sweep near a gene that increases plant height in wheat. *Mol. Ecol.* 17:741–56

91. Remigereau M-S, Lakis G, Rekima S, Leveugle M, Fontaine MC, et al. 2011. Cereal domestication and evolution of branching: evidence for soft selection in the *Tb1* orthologue of pearl millet (*Pennisetum glaucum* [L.] R. Br.). *PLoS ONE* 6:e22404

92. Riedelsheimer C, Lisec J, Czedik-Eysenberg A, Sulpice R, Flis A, et al. 2012. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. USA* 109:8872–77

93. Rizzon C, Ponger L, Gaut BS. 2006. Striking similarities in the genomic distribution of tandemly arrayed genes in *Arabidopsis* and rice. *PLoS Comput. Biol.* 2:e115

94. Rockman MV, Kruglyak L. 2008. Breeding designs for recombinant inbred advanced intercross lines. *Genetics* 179:1069–78

95. Saïdou A-A, Mariac C, Luong V, Pham J-L, Bezançon G, Vigouroux Y. 2009. Association studies identify natural variation at *PHYC* linked to flowering time and morphological variation in pearl millet. *Genetics* 182:899–910

96. Saito H, Yuan Q, Okumoto Y, Doi K, Yoshimura A, et al. 2009. Multiple alleles at *Early flowering 1* locus making variation in the basic vegetative growth period in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 119:315–23

97. Sato S, Tabata S, Hirakawa H, Asamizu E, Shirasawa K, et al. 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485:635–41

98. Schnable JC, Wang X, Pires JC, Freeling M. 2012. Escape from preferential retention following repeated whole genome duplications in plants. *Front. Plant Sci.* 3:94

99. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, et al. 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112–15

100. Schranz ME, Osborn TC. 2004. De novo variation in life-history traits and responses to growth conditions of resynthesized polyploid *Brassica napus* (Brassicaceae). *Am. J. Bot.* 91:174–83

101. Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, et al. 2008. Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* 40:1023–28

102. Sigmon B, Vollbrecht E. 2010. Evidence of selection at the *ramosa1* locus during maize domestication. *Mol. Ecol.* 19:1296–311

103. Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai Y-S, et al. 2006. Molecular characterization of the major wheat domestication gene *Q*. *Genetics* 172:547–55

104. Soltis DE, Soltis PS, Tate JA. 2003. Advances in the study of polyploidy since *Plant Speciation*. *New Phytol.* 161:173–91

105. Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* 60:561–88

106. Soltis PS, Soltis DE. 2012. *Polyploidy and Genome Evolution*. New York: Springer

107. Song X-J, Huang W, Shi M, Zhu M-Z, Lin H-X. 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39:623–30

108. Springer NM, Ying K, Fu Y, Ji T, Yeh C-T, et al. 2009. Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS Genet.* 5:e1000734

109. Studer A, 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

110. 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. *Proc. Natl. Acad. Sci. USA* 107:5792–97

111. Swanson-Wagner RA, Briskine R, Schaefer R, Hufford MB, Ross-Ibarra J, et al. 2012. Reshaping of the maize transcriptome by domestication. *Proc. Natl. Acad. Sci. USA* 109:11878–83

112. Swanson-Wagner RA, Eichten SR, Kumari S, Tiffin P, Stein JC, et al. 2010. Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. *Genome Res.* 20:1689–99

113. Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, et al. 2009. Evolutionary history of *GS3*, a gene conferring grain length in rice. *Genetics* 182:1323–34

114. Taketa S, Amano S, Tsujino Y, Sato T, Saisho D, et al. 2008. Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. *Proc. Natl. Acad. Sci. USA* 105:4062–67

115. 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

116. Tang H, Sezen U, Paterson AH. 2010. Domestication and plant genomes. *Curr. Opin. Plant Biol.* 13:160–66

117. Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. *Nat. Genet.* 28:286–89

118. 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

119. Tian Z, Wang X, Lee R, Li Y, Specht JE, et al. 2010. Artificial selection for determinate growth habit in soybean. *Proc. Natl. Acad. Sci. USA* 107:8563–68

120. Udall JA, Wendel JF. 2006. Polyploidy and crop improvement. *Crop Sci.* 46(Suppl. 1):3–14

121. van Heerwaarden J, Doebley J, Briggs WH, Glaubitz JC, Goodman MM, et al. 2011. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc. Natl. Acad. Sci. USA* 108:1088–92

122. van Heerwaarden J, Hufford MB, Ross-Ibarra J. 2012. Historical genomics of North American maize. *Proc. Natl. Acad. Sci. USA* 109:12420–25

123. Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, et al. 2012. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* 30:83–89

124. Vielle-Calzada J-P, Martínez de la Vega O, Hernández-Guzmán G, Ibarra-Laclette E, Alvarez-Mejía C, et al. 2009. The Palomero genome suggests metal effects on domestication. *Science* 326:1078

125. Walker AR, Lee E, Bogs J, McDavid DAJ, Thomas MR, Robinson SP. 2007. White grapes arose through the mutation of two similar and adjacent regulatory genes. *Plant J.* 49:772–85

126. Wang E, Wang J, Zhu X, Hao W, Wang L, et al. 2008. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat. Genet.* 40:1370–74

127. Wang E, Xu X, Zhang L, Zhang H, Lin L, et al. 2010. Duplication and independent selection of cell-wall invertase genes *GIF1* and *OsCIN1* during rice evolution and domestication. *BMC Evol. Biol.* 10:108

128. Wang Q, Dooner HK. 2006. Remarkable variation in maize genome structure inferred from haplotype diversity at the *bz* locus. *Proc. Natl. Acad. Sci. USA* 103:17644–49

129. Wang RL, Stec A, Hey J, Lukens L, Doebley J. 1999. The limits of selection during maize domestication. *Nature* 398:236–69

130. Wang S, Wu K, Yuan Q, Liu X, Liu Z, et al. 2012. Control of grain size, shape and quality by *OsSPL16* in rice. *Nat. Genet.* 44:950–54

131. Wang Y, Shen D, Bo S, Chen H, Zheng J, et al. 2010. Sequence variation and selection of small RNAs in domesticated rice. *BMC Evol. Biol.* 10:119

132. Wendel JF. 2000. Genome evolution in polyploids. *Plant Mol. Biol.* 42:225–49

133. Wingen LU, Münster T, Faigl W, Deleu W, Sommer H, et al. 2012. Molecular genetic basis of pod corn (*Tunicate* maize). *Proc. Natl. Acad. Sci. USA* 109:7115–20

134. Wittkopp PJ, Kalay G. 2012. *cis*-Regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13:59–69

135. Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, et al. 2005. The effects of artificial selection on the maize genome. *Science* 308:1310–14

136. Wu Y, Li X, Xiang W, Zhu C, Lin Z, et al. 2012. Presence of tannins in sorghum grains is conditioned by different natural alleles of *Tannin1*. *Proc. Natl. Acad. Sci. USA* 109:10281–86

137. Würschum T, Maurer HP, Kraft T, Janssen G, Nilsson C, Reif JC. 2011. Genome-wide association mapping of agronomic traits in sugar beet. *Theor. Appl. Genet.* 123:1121–31

138. 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* 319:1527–30

139. Xu P, Wu X, Wang B, Luo J, Liu Y, et al. 2012. Genome wide linkage disequilibrium in Chinese asparagus bean (*Vigna. unguiculata* ssp. *sesquipedialis*) germplasm: implications for domestication history and genome wide association studies. *Heredity* 109:34–40

140. Xu X, Liu X, Ge S, Jensen JD, Hu F, et al. 2012. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30:105–11

141. 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

142. Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, et al. 2004. The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303:1640–44

143. Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J. 2003. Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. USA* 100:6263–68

144. Yu B, Lin Z, Li H, Li X, Li J, et al. 2007. *TAC1*, a major quantitative trait locus controlling tiller angle in rice. *Plant J.* 52:891–98

145. Yu J, Holland JB, McMullen MD, Buckler ES. 2008. Genetic design and statistical power of nested association mapping in maize. *Genetics* 178:539–51

146. Yu Y, Tang T, Qian Q, Wang Y, Yan M, et al. 2008. Independent losses of function in a polyphenol oxidase in rice: differentiation in grain discoloration between subspecies and the role of positive selection under domestication. *Plant Cell* 20:2946–59

147. Zerjal T, Rousselet A, Mhiri C, Combes V, Madur D, et al. 2012. Maize genetic diversity and association mapping using transposable element insertion polymorphisms. *Theor. Appl. Genet.* 124:1521–37

148. Zhang Z, Belcram H, Gornicki P, Charles M, Just J, et al. 2011. Duplication and partitioning in evolution and function of homoeologous Q loci governing domestication characters in polyploid wheat. *Proc. Natl. Acad. Sci. USA* 108:18737–42

149. Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, et al. 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2:467

150. Zhu B-F, Si L, Wang Z, Zhou Y, Zhu J, et al. 2011. Genetic control of a transition from black to straw-white seed hull in rice domestication. *Plant Physiol.* 155:1301–11

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Plant Biology* articles may be found at
http://www.annualreviews.org/errata/arplant