Trends in
Plant Science

CellPress
OPEN ACCESS

Review

Evolutionary dynamics and adaptive benefits of deleterious mutations in crop gene pools

Sangam L. Dwivedi,[1] Pat Heslop-Harrison,[2,3] Charles Spillane,[4] Peter C. McKeown,[4] David Edwards,[5] Irwin Goldman,[6] and Rodomiro Ortiz[7,*]

Mutations with deleterious consequences in nature may be conditionally deleterious in crop plants. That is, while some genetic variants may reduce fitness under wild conditions and be subject to purifying selection, they can be under positive selection in domesticates. Such deleterious alleles can be plant breeding targets, particularly for complex traits. The difficulty of distinguishing favorable from unfavorable variants reduces the power of selection, while favorable trait variation and heterosis may be attributable to deleterious alleles. Here, we review the roles of deleterious mutations in crop breeding and discuss how they can be used as a new avenue for crop improvement with emerging genomic tools, including HapMaps and pangenome analysis, aiding the identification, removal, or exploitation of deleterious mutations.

Highlights

Deleterious mutations are common in plant genomes and present both challenges and opportunities for domesticated crops.

While many deleterious mutations can negatively impact crop performance, some deleterious mutations are associated with breeding-relevant phenomena, such as heterosis, genetic variation, and key loss-of-function domestication traits that underpin crop production. Breeding may hasten the accumulation of deleterious variants.

Crop species vary in their pattern and distribution of deleterious mutations, and differences are noted among mating systems, ploidy levels, and propagation systems. Current efforts focus on predicting functional consequences of deleterious mutations on a genome-wide scale and comparing the effects across species.

Strategies for removal of deleterious variants with techniques such as gene editing may facilitate breeding and accelerate the redomestication of crops from wild species.

What are deleterious mutations?

Fitness is a relative term that describes differential reproductive success and survival. **Deleterious mutations** (see Glossary) cause loss or alteration of normal gene function, leading to reductions of organismal fitness. While fitness in nature and fitness under domestication often have overlapping criteria, the primary scope of this review concerns fitness under domestication. Deleterious mutations obviously have a role in both domesticated and natural populations, but the unique constraints of domesticated populations are our primary focus. Mutations with deleterious consequences in nature may be only conditionally deleterious in crop plants. Thus, while some genetic variants may reduce fitness under natural conditions, they may conversely remain under positive selection in domestication. In this way, conditionally deleterious alleles can be adaptive. Crop gene pools harbor a range of genetic variability with respect to reproductive traits, some of which may be deleterious to reproductive fitness. Paradoxically, some deleterious mutations may in fact be favorable in an agricultural context.

Deleterious mutations in coding sequences may induce frameshifts, premature stop codons, deletions and insertions, splice site disruptions, or nonsynonymous base changes (and suboptimal codon usage), and are commonly recessive [1]. Their phenotypes are exposed when homozygous because of inbreeding, as occurs in many crop species. Some examples of inbreeding are self-pollination, sib-mating, and mating among relatives in small populations. Assaf *et al.* [2] noted that recessive deleterious mutations can slow the process of adaptation in humans and *Drosophila*. While some loss-of-function mutations have been exposed through inbreeding and harnessed as domestication traits (see below), others remain cryptic (hidden).

Mutations in noncoding regions may also be deleterious by affecting promoters or other regulatory elements, such as transcription factor (TF) or miRNA-binding sites [3]. While some loss-of-function mutations have been exposed through inbreeding and harnessed as domestication traits (see below), others remain cryptic. Importantly, recent findings by Monroe *et al.* [4] demonstrated that

[1]Independent Researcher, Hyderabad, 500016, India
[2]Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, 510650, China
[3]Department of Genetics and Genome Biology, University of Leicester, Leicester, LE1 7RH, UK
[4]Agriculture and Bioeconomy Research Centre, Ryan Institute, University of Galway, University Road, Galway, H91 REW4, Ireland
[5]School of Biological Sciences and Institute of Agriculture, University of Western Australia, Perth, WA 6009, Australia

deleterious mutations occur less frequently in functionally constrained portions of the arabidopsis (*Arabidopsis thaliana)* genome, such as within genes. This phenomenon, which they called epigenome-associated mutation bias, suggests that deleterious mutations occur in greater frequency in noncoding regions of the plant genome.

Charlesworth [1] noted that, since deleterious mutations continuously arise and are subsequently eliminated by purifying selection if exposed by homozygosity, they can have important impacts on fitness, **inbreeding depression**, and genetic variability of populations. A loss-of-function mutation, in which the altered gene product lacks the function of the wild-type gene product, may be deleterious. Xu and Guo [5] commented that, while most loss-of-function mutations are deleterious or neutral, some are under positive selection and may contribute to adaptation. This paradoxical phenomenon arises because selection under domestication may favor the loss of a particular gene product due to human preferences. Little is known regarding the impact of deleterious mutations in most crops, but they probably influence both single-gene and polygenic characters under selection by breeders.

Deleterious mutations are more likely to be exposed in crops than in their wild relatives. Linkage disequilibrium leads to particular enrichment of deleterious alleles near loci subjected to intense positive selection during domestication or improvement [3] although selection for uniformity in elite cultivars may also purge deleterious alleles from the gene pool [6,7]. Here, we provide an overview of roles of deleterious mutations to mitigate their adverse impacts in crop breeding.

Complex traits harbor more deleterious mutations compared with other traits

Quantitative traits are genetically complex, commonly being multigenic and affected by copy number variation (CNV) or other structural effects [8]. Quantitative traits show continuous variation and are influenced by environmental factors and genotype-by-environment (G × E) interactions [9]. Deleterious (causal) variants can also impact quantitative traits. Fitness-related traits (i.e., biomass, plant height, specific leaf area, and total soluble carbohydrate in sorghum [10]; days to tassel, number of tassel branches, plant height, upper leaf angle, leaf length, and plant yield in maize [11]) harbor more deleterious mutations compared with traits that are not directly related to fitness [12].

Loss-of-function mutations are a particular class of deleterious mutations. In the context of plant breeding and crop improvement, many of the traits and alleles present in wild undomesticated progenitors or wild relatives can be considered as deleterious in the context of productivity [6,13–15], while the selection of some loss-of-function mutations may be regarded as beneficial in crops. Indeed, major loss-of-function mutations that could be expected to be deleterious in nature are often critical in plant breeding [16] and, therefore, are termed 'beneficial mutations'. For example, the six-rowed phenotype of the barley spikelet arises due to loss of function of *Vrs1*, which is normally expressed in lateral spikelet primordia of two-row barley, the loss of which converts rudimentary lateral spikelets into fertile spikelets of the six-row type [17]. Loss-of-function mutations have convergently occurred in other crops, such as *sh1* mutations in maize, rice, and sorghum, which eliminate rachis ('seed') shattering [18], and others, such as those underlying determinate flowering [19,20] or seed germination characters (e.g., quick imbibition and short, invariant dormancy [21,22]), as well as variants lacking light, smoke, or vernalization requirements [23,24]. Despite selection for these loss-of-function mutations, domestication often results in gene loss. Comestible traits may also be affected; for example, bitterness in cucumber is reduced through recessive loss-of-function mutations, such as *bi-1* [25], which likely block the formation of bitter terpenoids. Famously, loss of function of *GA20-OX* genes produces dwarf wheat and rice, the foundation of the **Green Revolution** cultivars [26]. Loss-of-function mutants in *GBSS1* lower amylose content. Starch is the major source of energy in the human diet, of which, resistant starch (RS) is recalcitrant to degradation, and is not absorbed by the small intestine. RS releases

[6]Department of Horticulture, College of Agricultural and Life Sciences, University of Wisconsin Madison, WI 53706, USA
[7]Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, SE 23053, Sweden

*Correspondence:
rodomiro.ortiz@slu.se (R. Ortiz).

glucose in small amounts and at slow rate to prevent sudden postprandial glucose responses. Hence, intake of RS prevents life style-related diseases and helps manage diabetes [27,28]. Some of the loss-of-function mutations [e.g., *tb1* and *tga1* in maize, *Rc* and *sh4* in rice, *Q* in wheat, *fw2.2* in tomato (*Solanum lycopersicum)*, and many others] may be of even greater significance in contemporary breeding efforts because the intense selection pressure associated with modern breeding tends to exploit large effect mutations [29].

Genome editing of deleterious genes and alleles

Deleterious alleles can arise in any genome due to mutations in protein-coding genes, noncoding (nc)RNA genes, or gene regulatory regions. In diploid plants, alleles that are deleterious during the gametophytic phase of the life cycle may be purged by purifying selection. However, recessive alleles that are deleterious during the sporophytic phase of the life cycle can be maintained in populations as heterozygotes, particularly in outcrossing crops. In inbreeding species (including model plants, such as arabidopsis), the mutation rate is lower in functionally constrained regions of the genome. Such mutation bias is associated with epigenomic and physical features of genetic regions under selection [4].

Genome editing (GE) technologies in plants have emerged as powerful tools for targeted mutation of genomes, including all those that can generate deleterious alleles. For instance, GE can generate base edits [30], indels, CNVs [31], and chromosomal rearrangements [32,33]. Where transformation systems are in place, GE technologies are applicable to many plant and crop species, allowing targeted generation or removal of deleterious DNA sequences within a single generation [34–37]. Additionally, multiplex targeting of multiple loci by GE augments the power of GE applications, such as CRISPR/Cas9 [38], along with strategies and protocols for polyploid genomes [39–41]. It has been demonstrated that strongly deleterious alleles are preferentially located on exposed chromatin in the wild Brassicaceae species *Capsella grandiflora*, indicating that nuclear organization and karyotype could have roles in how easily deleterious sequences can be targeted by CRISPR/Cas9 [42], and potentially requiring understanding of 3D-chromosomal conformations within polyploid nuclei to inform deletion or base-editing strategies [43].

The genomic era of crop improvement continues to generate an unprecedented wealth of whole-genome sequences for crops (primary gene pools) and wild relatives (secondary or tertiary gene pools) [44–46]. Such genomic data are combined with functional data on genes and genomic regions associated with traits arising from forward and reverse genetic approaches enabled by RNA sequencing and high-throughput phenotyping platforms [45,47,48].

In livestock breeding, simulations have been conducted to assess the feasibility of using GE to purge deleterious alleles from breeding pools [49]. For crop breeding, the prospects for using multiplex genome editing to purge deleterious alleles from breeding lines are contingent on developing bioinformatic approaches to identify and rank loci for purging through GE. Strategies used for prioritization of loci can differ depending on whether the alleles are codominant or recessive [49]. Indeed, identifying deleterious alleles genome wide in crops may be challenging, given their low frequency [11]. However, significant advances are being made in genome-wide identification of deleterious mutations in plant genomes [50], including in arabidopsis [14,51], rice [14], maize [11,52], barley [53], and sunflower (*Helianthus annuus*) [7].

Genome-wide scans in maize and sorghum have identified deleterious genes in both species that are in pericentromeric regions, enriched for non-syntenic genes and at low frequency [10,11] (although see the case of sorghum below). A substantial proportion of both trait variation and heterosis in maize is likely attributable to incomplete dominance of deleterious alleles [15]. In the

Glossary

Apomictic species: type of reproduction in which sexual organs of related structures take part but seeds are formed without union of gametes.

Coding gene: portion of the DNA or RNA of a gene that codes for protein.

Deleterious mutation: mutation in which the protein product of a gene is not produced, is produced but not functional, or is produced and interferes with normal function. Such mutations arise from single base changes or more extensive insertions, deletions, or frame shifts.

Domestication-cost hypothesis: increase in the number of deleterious genetic variants, fixed or segregating, in the genomes of domesticated species.

Dominance coefficient (*h*): expression of a trait or fitness in relation to an allele: that is, $h > 1$ when the fitness of both homozygotes is above that of the heterozygote (underdominance), $h < 0$ if the fitness of the heterozygote is superior to that of either homozygote (overdominance); and $0 < h < 1$ when the fitness of the heterozygote is only superior when the homozygote is deleterious (incomplete dominance).

Genomic prediction: prediction of the genetic value or phenotypic trait of an individual or population based on genome-wise scan of SNPs.

Green Revolution: agricultural development strategy based on the combined use of newly bred semidwarf cultivars of rice and wheat, fertilizers, irrigation water, and mechanization.

HapMaps: term coined by the 'International HapMap Project' to develop a haplotype map, which are variations in DNA sequences that are inherited together.

Heterosis: superior performance of a certain characteristic of a hybrid over the average of its parents (mid-parent heterosis), or its best parent (heterobeltiosis).

Inbreeding depression: reduced biological fitness of a given population or line because of selfing.

Incomplete dominance: phenomenon in which two true-breeding parents crossed to produce an intermediate offspring.

Linkage drag: reduced fitness of a cultivar due to deleterious genes introduced along with the beneficial genes during backcrossing.

Pangenome: entire set of genes within a species.

context of plant breeding and crop improvement, many of the traits and alleles present in wild undomesticated progenitors or wild relatives can be considered as deleterious in the context of crop production and agronomy.

In principle, where deleterious alleles are known, GE could be used to remove deleterious alleles of wild or landrace origin following or preceding introgression into elite germplasm, reducing the rounds of backcrossing needed, and allowing the elimination of deleterious alleles in tight linkage with beneficial ones. Emerging GE approaches for accelerated- and redomestication of crops from wild species target loci that are preferred for agriculture or consumption [54–57].

Identification and elimination of deleterious variants for crop improvement

Mutation load in crop genomes

Deleterious variants affect trait expression and organismal fitness. Variants at phylogenetically conserved sites are of interest, being more likely than those at phylogenetically variable sites to have deleterious effects on fitness and contribute to phenotypic variation [50]. Crops are considered to accumulate more deleterious mutations than their wild relatives (i.e., the cost of domestication) [58]. Understanding their pattern, distribution, and deleterious mutation load is critical to improve crop complex traits, such as yield and stress tolerance.

Recent advances in genome resequencing have led to increased interest in predicting the functional consequences of genetic variants genome wide. After the analysis of ~13 million variants from whole-genome sequence analysis of 499 sorghum lines and 25 million variants previously reported in 1218 maize lines, Lozano *et al.* [59] noted that such variants in both species were in **pericentromeric regions**, enriched in non-syntenic genes, and were recorded at low allele frequency. However, sorghum did not follow the **domestication-cost hypothesis** that was noted in maize following comparisons of deleterious variants. This hypothesis predicts a lower deleterious mutation burden in wild relatives than in the related crop. In a genome-wide variation map of 10.6 million SNPs and 1.4 million indels in 781 soybean (*Glycine max*) accessions involving domesticates, wild species, and natural hybrids, Kim *et al.* [6] found 183 domestication-selective sweeps and a set of 742 149 deleterious mutations in domesticated (landraces and improved lines) and wild soybean populations. A 7.1% decrease in overall deleterious alleles in landraces relative to wild soybean accessions and an additional 1.4% decrease was noted in improved lines. These detected domestication-selective sweeps also showed fewer deleterious alleles, likely due to artificial selection. The low frequency of deleterious alleles in sorghum and soybean relative to maize is due to mating system differences, with former two being inbreeders, while the latter outbreeds. However, a study involving cassava revealed that this vegetatively propagated crop has ~26% more predicted deleterious alleles compared with related wild *Manihot* species [60]. Its asexual propagation [61] could lead to increases in its mutation load in the cultigen pool. It has been also noted that keeping large recombination-suppressed introgressions from its wild relative *Manihot glaziovii* in the heterozygous state augmented deleterious mutations in cassava [62]. Hence, purging such a **linkage drag** as well as alleles contributing to genetic load may be interesting breeding targets, particularly when ‘cost of domestication’ restrains selection efficacy [63], thereby decreasing genetic gains. Hence, more research is necessary to document patterns, distribution, and purging of deleterious mutation load in crops differing in ploidy level, genome size, propagation, and breeding systems, and to assess the utility of such variants in crop improvement.

Genes favorable in the wild but deleterious in domesticated crops

Increased availability of plant genome sequences has improved our understanding of gene presence or absence within species. The study of **pangenomes** has shown that 40% of genes may be

Pericentromeric region: region around the centromere; contains repetitive sequences and transposable elements and adopts a chromatin state characterized by specific histone variants and post-translational modifications, leading to a transcriptionally repressive chromosomal environment.
Recombination: production of offspring with a combination of traits that differ from those of either parent.
Regulatory gene: gene involved in the production of a substance that controls or regulates the expression of one or more genes.
Selective sweeps: process by which a beneficial mutation eliminates or reduces variation in linked neutral sites as it increases in frequency in the population.
Structural variants (or copy number variants, CNVs): refers to genomic polymorphisms (DNA segments >1 kb in length), such as inversions, insertions, duplications, or deletions, resulting from gains or losses of DNA segments or chromosome rearrangements.
Superdomestication: process that leads to a domesticate with dramatically increased yield that could not be selected in natural environments from naturally occurring variation.
Telomere: compound structure of repetitive nucleotide sequences associated with specialized proteins at the end of a chromosome.

present in some individuals and not in others [64], with wild progenitors often more diverse than their domesticated relatives [35,44]. Reduction in gene content during domestication reflects both the selective loss of genes impacting adverse crop performance, and reduced diversity following domestication bottlenecks. In soybean, 1478 genes reduced in frequency following domestication, of which 98 were associated with defense responses [44]. By contrast, only 261 genes increased frequency (Table S1 in the supplemental information online). Requirements for diverse adaptation in wild species increase gene and allele diversity in populations, much of which is lost during selection for specific crop traits, including yield, resistance to pathogens, tolerance to abiotic stresses, and quality traits. Host plant resistance genes have undergone intense selection during breeding because they have either strong beneficial effects when providing resistance to relevant pathogens or potential negative impacts when associated with reduced crop performance [65,66], including taste and nutrition. There can also be trade-offs between these, including for selection against bitterness, which makes fruits more palatable to not only human consumers, but also other herbivores. The high degree of presence/absence variation for these genes also makes them more amenable to loss than many classes of gene [64]. Although many disease resistance genes have already been eliminated from cultivars, others are potential targets for removal [67]. In wild species, a wide flowering time window enables the response of the plant to prevailing conditions, whereas a cultivar requires uniform flowering, with associated selection for a limited repertoire of flowering time alleles. Similarly, selection for traits such as oil quality has led to the enrichment of the underlying genes and alleles in modern cultivars, with oil quality- and flowering time-associated gene frequency changes in soybean during breeding and adaptation [44,68].

Recombination rate impacts on number and distribution of deleterious mutations

Recombination generates novel combinations of genes and alleles that lead to phenotypes that may be under positive or negative selection. While selection can lead to a change in frequency of genes and alleles, or even the complete loss of unfavorable variants, the selection pressure of individual loci is balanced by selection of genetically linked loci [69]. Regions of the genome that experience high rates of recombination allow selection on much finer scales compared with regions of low recombination, with subsequent rapid and precise removal of deleterious variants. However, in regions of low recombination, the inability to separate unfavorable from favorable variants reduces the power of selection and removal of the unfavorable variant. On a practical basis, this makes it challenging to remove deleterious variants when they are in regions of low recombination, such as distant from **telomeres** or associated with major **structural variants**. Their removal may lead to fixation of large haplotype blocks, limiting the diversity available for further selection.

HapMaps for predicting loss-of-function mutations

The ability to investigate the genomic context of both favorable and deleterious genes and alleles supports their identification and selection, while potentially improving understanding of the mechanism of gene impact. The construction of **HapMaps** aids understanding of genome-wide variation, placing variants within the context of both the genomic and recombination landscapes [37,67,70,71]. By assessing the association of haplotypes presented in these HapMaps with phenotypic traits, one can identify haplotypes and possible causal variations associated with deleterious traits. Such fine-level haplotype analysis may also permit the dissection of genetically linked traits, either directly through study of haplotype variants and rare recombinants or through the identification of targets for GE-based validation [72,73]. Currently, few HapMaps are available for crops, but with the continued reduction in sequencing costs, the growth of pangenomes, and advances in bioinformatic analysis and visualization tools, HapMaps are rapidly growing in number and density, providing valuable tools for identifying and removing deleterious mutations.

Elimination of deleterious alleles to improve crops fitness and productivity

Weakly deleterious mutations may persist in populations [53] and contribute significantly to reductions in fitness [74]. Humans carry hundreds of loss-of-function variants and thousands of weakly deleterious variants in their genome [75], but similar estimates for plants have been limited. Three factors affecting the possible emergence and elimination of these variants are effective population size, the selective coefficient against homozygous individuals, and the **dominance coefficient**. On average, there are ~1000 deleterious variants per accession in a sample of barley genotypes and ~700 deleterious variants per accession in a sample of soybean genotypes [53]. These variants were identified using tools such as the Sorting Intolerant from Tolerant (SIFT) and Polymorphism Phenotyping (PolyPhen) programs. However, ~40% of the variants were associated with a single individual in each of the crop species, suggesting the potential for selection to reduce the number of deleterious mutations. It was also found that the proportion of deleterious variants was negatively correlated with recombination rate. Again, this demonstrates the role of recombination and selection in reducing the frequency of deleterious mutations in crop genomes.

If deleterious mutations are linked to alleles under selection during domestication, they may be present in cultivated genomes following **selective sweeps**. Breaking these linkages in breeding programs may be difficult. Outcrossing species, such as cassava, grape, maize, and sunflower have higher levels of deleterious mutations in their cultivated genomes, possibly due to selective sweeps under domestication [63], compared with their wild progenitors [60]. A study involving domestication induced selective sweeps in soybean, in contrast to outcrossing species, showed a reduction in deleterious mutations in cultivated genomes compared with wild-type, and further reductions in the number of deleterious mutations were found in improved accessions compared with landrace germplasm [6]. A decrease in mutational burden between landraces and modern lines was also observed in maize [15]. Artificial selection appears to have reduced deleterious mutations in the domesticated genome in this cross-pollinated species. Therefore, it is possible that efforts to reduce deleterious mutational load in breeding programs may focus on inbreeding methods or on those breeding methods that increase homozygosity. During the early 20th century, plant breeders began to isolate highly homozygous lines through inbreeding and sib-mating. These inbred lines were crossed in hybrid combinations to produce F_1 hybrids, which are now the predominant cultivar type for many crop species. This inbred–hybrid approach has dominated the past century of breeding for many crops and may represent a continued opportunity for purging deleterious mutations [76].

An even larger effect is noted for asexual or clonally propagated species. Fitness increases after a shift to clonal propagation because clonal propagation hides recessive deleterious mutations in a heterozygous state [77]. Thus, there is no additional genetic load and, as noted above, cassava has the highest number of deleterious alleles per genome of any crop examined so far.

Relationship of deleterious mutations to inbreeding depression and incomplete dominance in trait variation and hybrid vigor

Heterosis has been exploited by humans for maximizing crop productivity, but its genetic basis remains elusive. A likely explanation for at least some heterotic effects in crops is the dominance hypothesis, which posits that masking or complementation of deleterious alleles by dominant alleles contributes to superior performance of F_1 hybrids [78]. In such instances, heterosis, and its converse, inbreeding depression, are due to deleterious recessive alleles [79].

Strongly deleterious alleles were more likely to be recessive in a study measuring agronomic traits (such as grain yield) in both inbred parental genomes and hybrids of an elite maize partial diallel population [15]. Likewise, most deleterious alleles exhibited **incomplete dominance** for traits with high heterosis. It was further noted that **genomic prediction** of heterosis improved when incomplete dominance was taken into consideration in the model [15].

Modeling the impact of inbreeding depression caused by deleterious mutations in small populations reveals that inbreeding depression and deleterious mutations could be partly purged from the population through inbreeding, although mainly for large-effect mutations [80]. However, fitness decreased with inbreeding due to homozygosity of deleterious mutants. A study involving the effect of population size on inbreeding depression caused by weakly deleterious alleles and based on single locus models revealed that drift reduces the efficacy of selection in small populations, resulting in reduced inbreeding depression. By contrast, genetic load increases in small populations but decreases with more inbreeding [81]. Domestication has the effect of increasing genetic load by increasing mutations [82]. Taken together, the frequency and degree of dominance of deleterious mutations, along with population size, are important factors in determining how effective selection may be and how much heterosis may be expressed in populations.

Polyploidy, whole-genome duplications, chromosomal and genome rearrangements, and structural variation

Polyploidy (whole-genome duplications) along with segmental genome and gene duplications can generate the ‘redundancy space’ for recessive deleterious alleles to be maintained in populations. Given that many crops are polyploids, it can be expected that deleterious alleles will be less likely to be purged from such species, relative to diploid crops. In addition vegetatively propagated crops, such as potato [83], East African Highland banana (*Musa* spp.) [84], and yams (*Dioscorea spp.*) [85], can accumulate deleterious mutations over successive rounds of propagation. For instance, polyploidization of tetraploid potato is associated with the rapid accumulation of deleterious mutations [83]. There are efforts underway to characterize deleterious mutations in polyploid species, such as wheat [86,87]. While whole-genome duplications (polyploidy) can allow a greater proportion of recessive deleterious alleles to be maintained, segmental or gene duplications can also have a similar masking effect due to genetic redundancy [88]. While segmental or gene duplications can provide genetic robustness against deleterious mutations [89], they can also maintain deleterious alleles by masking them.

Large-effect mutations underlying adaptive traits and their vulnerability

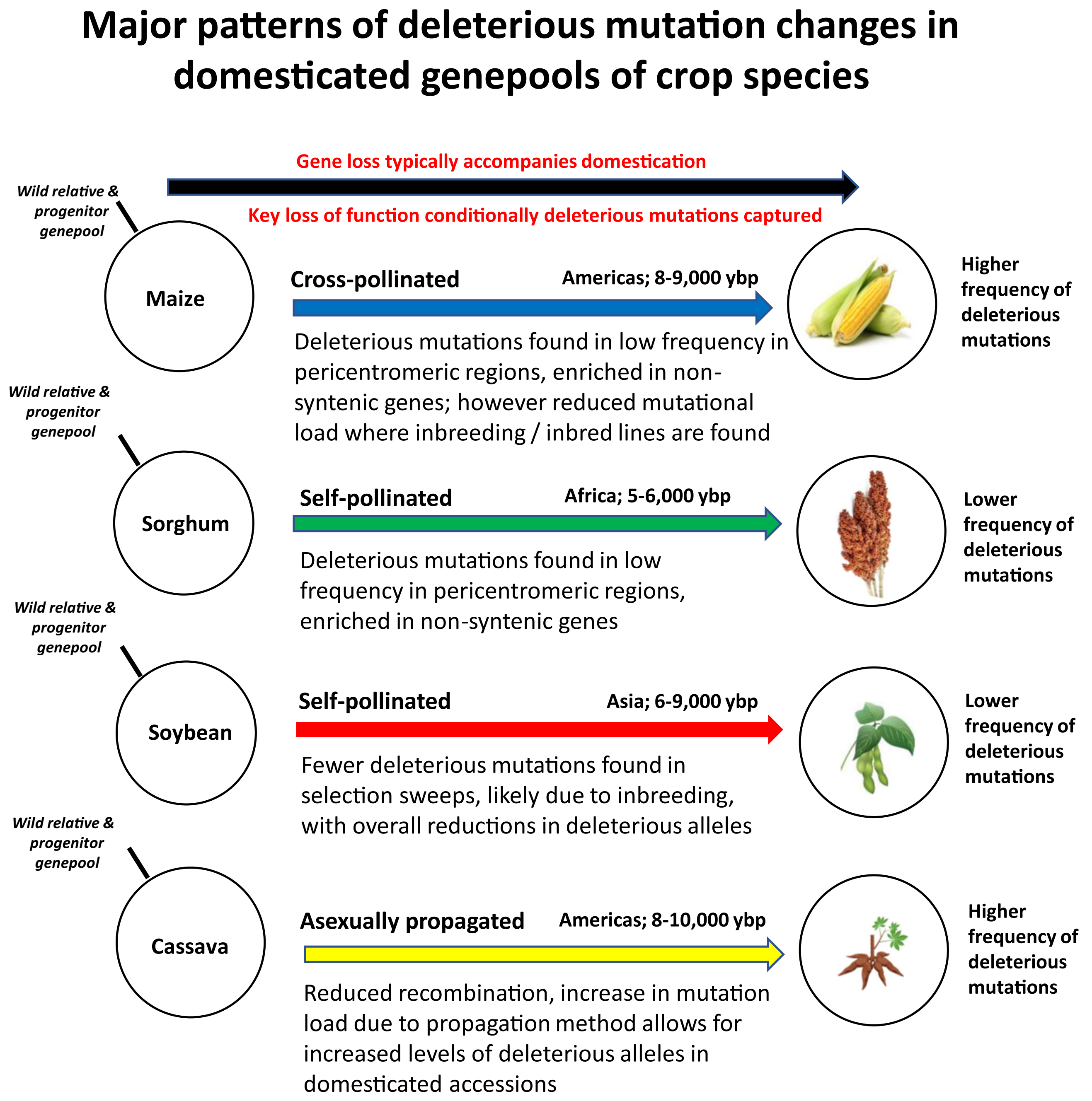
Deleterious alleles within germplasm are not only associated with simple traits, but also often contribute to the genetic architecture of complex traits. These associations may arise from genetic bottlenecks but may be more difficult to disentangle due to co-adaptations or epistasis between deleterious and beneficial alleles. Indeed, understanding effects of deleterious mutation in sorghum relies upon understanding the genetic architecture of complex traits, as well as varying between different lineages, possibly depending upon the history of the bottlenecks as they were experienced [10]. Regulatory interactions between genes are often complex (especially in the case of TFs or epigenetic regulators, which affect the expression of many genes), and deleterious effects of selected genes may lead to subsequent selection for compensatory mutations elsewhere, as shown for wheat semi-dwarfing genes [86]. Interactions can also occur between the encoded proteins. While editing genes to replace or substitute deleterious amino acid residues could allow redesigns to the function of multiprotein complexes, such editing would be reliant on the links between protein sequence and structure, function of multiprotein complexes, and resulting gene–trait mechanisms, being thoroughly understood. Despite such difficulties, from a horizon-scan viewpoint, breeding for crops to be more resilient to unprecedented climate stresses, which are likely to act in challenging combinations, may require such redesigns of crop genomes [90,91].

Comparison of deleterious variants across mating systems and propagation and end use of crop products

Most crops are propagated by seed, which may be from species that are inbred and self-pollinating (e.g., wheat and rice) or outcrossing. **Apomictic crop species**, propagated asexually by seed, include some important forage grasses [92]. Seed-propagated crops may also be based on F_1 seed [e.g., some maize, *Brassica* or tomato), where inbred parental lines must also be maintained. Other crops (including many root and tuber crops) and trees (fruits, nuts, or timber) are typically vegetatively propagated. The impact of deleterious mutations, present in populations, is different for each mating (or related propagation) system. For inbred crops, dominant deleterious mutations may be exposed in the homozygous state during selection, and breeders may purge them when plant performance is impacted. Outcrossing has been widely considered an adaptation to avoid inbreeding depression caused, at least partly, by deleterious alleles [93], although, even in outcrossing species the mutations are present at low allele frequencies and most loci are homozygous [94]. Given the need for vigorous, easily propagated plants as parents of F_1 hybrids, deleterious mutations may be selected against in parental breeding lines. In a recent direct comparison of sorghum (selfing) and maize (outcrossing), Lozano *et al.* [59] showed contrasts in the presence of deleterious alleles, with an excess in improved maize lines compared with wild relatives, while sorghum accumulated deleterious alleles in its wild relative [59]. In apomicts and vegetatively propagated plants, deleterious mutations are retained in the absence of recombination. Over time, the pressure of recurrent deleterious mutations and their accumulation in asexual populations may also reduce fitness, a phenomenon known as Muller's ratchet. Propagation methods (e.g., tissue culture) may also increase mutation pressure. In the vegetatively propagated crop cassava, it has been shown that recessive deleterious mutations have been kept in the heterozygous state, and purging these may be an important breeding goal [60].

De novo domestication approaches for conversion of wild relatives into domesticated prototypes and harnessing of adaptation genes from wild germplasm

Only a tiny proportion of the genetic variation present in the ~400 000 flowering plant species is used in crop production. The entire gene pool of angiosperms appears to be amenable for domestication, as suggested by both the common set of near-universal genes and the high frequency of transfer of nature and engineered genes between plant species [95]. Thus, research on the pangenome within species groups, and more widely, is identifying more nondeleterious mutations, which, within domesticated species, may have been lost in population or domestication bottlenecks. DNA sequencing and marker-based genetic approaches can identify desirable alleles, while artificial intelligence (AI) or machine learning-based classifiers are showing promise for the prediction of deleterious alleles [96]. Beyond sequence analysis, even in unimproved germplasm and crop wild relatives, genome-wide association studies (GWASs) can identify candidate genome regions for transfer into populations for breeding, allowing removal of deleterious alleles in **coding genes**, which may be **regulatory** or not. The race between plant hosts and their pathogens, as well as adapting to the changing climate, call for continuously introducing new variation, a task in which plant breeders have been succeeding for improving performance in almost all crops. **Superdomestication** relies on identification of traits that might be improved [97]. These traits may include deleterious alleles, which can be targets for removal. Nevertheless, there remains a question about the use of alleles that may be mildly deleterious in most conditions, but advantageous on occasion (i.e., tolerance to abiotic stresses, such as wind, drought or waterlogging, or biotic stress from a particular disease only sometimes prevalent): when does a deleterious mutation become favorable? An allele might also be mildly advantageous in most seasons, but sometimes catastrophically deleterious; for example, in ryegrass, growth at



Trends in Plant Science

Figure 1. Four major patterns of changes in deleterious mutations from domesticated gene pools to domesticated crops in different world regions. Each of these four crops was domesticated over millennia and represent three primary methods of reproduction: cross-pollination, self-pollination, and asexually reproduction. Cultivated maize generally displays increased levels of deleterious alleles in domesticated populations, with the exception of inbreeding programs, in which deleterious alleles may be purged. Maize is a cross-pollinated species, which may be partly responsible for the increase in deleterious alleles. The self-pollinated crops sorghum and soybean display lower levels of deleterious alleles in domesticated populations, likely due to inbreeding, which allows for purging of deleterious alleles. Cassava is typically propagated asexually and, therefore, may have accumulated a large mutational load due to this method of propagation. Gene loss typically accompanies crop domestication. Key loss-of-function mutations in traits that may have deleterious effects in nature, such as the shattering habit or pest/predator defense, have been captured and selected during domestication, and remain important features of modern crops.

cold temperatures can extend grazing seasons in cold climates in most years, but plants may die in a cold winter without snow cover [98].

Concluding remarks and future perspectives

Deleterious mutations reduce the fitness of cultivars to limit plant breeding efficacy. However, such mutations have also contributed to crop domestication traits. Transitioning from current breeding capabilities to an improved level requires adapting innovative approaches, including exploiting deleterious variants to foster food security and sustainable agriculture intensification under global warming [99,100]. Robust characterization and categorization of deleterious variants would enable improved understanding of their function in crop genomes and their potential, both positive and negative, in crop improvement (Figure 1).

Crop domesticates can contain more deleterious alleles compared with their wild relatives. Loss-of-function mutations are often deleterious (either conditionally or absolutely) or neutral, but some have contributed to adaptation and crop productivity in important ways. Many of these are responsible for important qualities in agriculture, such as seed traits [101] or non-shattering [18]. Complex traits have the potential to harbor more deleterious alleles compared with simpler traits and are more difficult to disentangle due to co-adaptation or epistasis between deleterious and beneficial alleles [10,11]. Apomicts and clonally propagated crops can accumulate more deleterious mutations in the absence of recombination, which reduce fitness [60,102].

Differences in mating systems, ploidy level, propagation systems, whole-genome duplication, chromosomal rearrangement, and recombination can all affect deleterious allele frequency. Deleterious recessive alleles are important in heterosis and inbreeding depression effects are of importance to crop breeding and production. For instance, masking or complementation of deleterious alleles by dominant alleles contributes to hybrid vigour in F_1.

Disentangling unfavorable variants from favorable variants is a challenge. Crossbreeding has been ineffective (or slow) in purging deleterious genetic variants, which are often recessive and can remain for generations in a heterozygous state. Advances in genome-wide sequencing and analysis, including analytical approaches to detect small-effect genes, allow prediction of the functional consequences of genetic variants.

Gene editing combined with bioinformatics can facilitate the generation, validation, and removal of deleterious alleles, while pangenome analyses can reveal the comparative loss of alleles in

Outstanding questions

How are deleterious mutations discovered and disentangled in complex and multigenic traits?

How deleterious are structural variants?

What role may transposons have in the generation of new deleterious 'mutations'?

How can we break linkage drags under linkage disequilibrium?

How can we assess deleterious mutations by integrating information from the pangenome?

How can we identify and remove deleterious mutations from vegetatively propagated crops?

How can we identify and remove deleterious mutations from polyploid crops?

What about future proofing of crops? Is it worth carrying deleterious mutations now that may be advantageous in the future?

Is there a global, optimum genotype, or should microbreeding of plants adapted to particularly small regions, be developed?

What quantitative tools are needed to identify deleterious mutations?

What are the impacts of deleterious mutations on crop sciences?

Box 1. Community resources for identifying putative deleterious genes and mutant alleles

Mutations in plant genomes continuously arise and accumulate due to population demographic history, domestication bottleneck, mating system, genetic drift, or linked selection due to genetic interactions. While most mutations are neutral, a small fraction of these genetic variants are deleterious due to the disruption of protein function or gene expression [10]. Discovering and removing deleterious mutations is a significant challenge in developmental plant biology. The substitution of amino acid effects on protein function was first predicted by the sorting intolerant from tolerant (SIFT) algorithm, with a SIFT score <0.05 defined as a putative deleterious mutation [106]. Genomic evolutionary rate profiling (GERP), obtained from a multispecies whole-genome alignment, >2 is another approach to identifying a high confidence set of deleterious mutations [107]. Combining both approaches (GERP >2 and SIFT <0.05) provides a more conservative estimate of deleterious mutations. Both these approaches were used for identifying putative deleterious variants in important crops, such as maize and sorghum [10,11], while Random Forest Classifiers, a machine learning approach based on a data set of deleterious and neutral mutations in arabidopsis, efficiently detected deleterious variants in pea (*Pisum sativum*) and rice, with an accuracy of 93% and 87%, respectively, which are higher than those obtained by PolyPhen-2 [96].

A database platform of characterized large-effect beneficial mutations (e.g., Gephebase [108]; www.gephebase.org) would be an important genotype–phenotype information resource, but does not currently exist for characterizing deleterious genetic variants in eukaryotes. Hence, there is a need to also develop such a database for deleterious variants.

domesticates and their wild relatives. In principle, simultaneous editing of multiple variants could be used to enhance both the fitness and productivity of crops [49,103].

Pericentromeric regions accumulate more deleterious mutation due to less selection efficiency. Enhancing the recombination rate by deploying anti-crossover genes (e.g., *FANCM*, *RECQ4*, or *FIGL1*) could greatly elevate the number of crossovers in crop genome to facilitate purging of purge deleterious variants [104,105]. In the face of the climate and sustainability challenges facing crop science and agriculture, documenting and purging deleterious alleles in crops differing in ploidy level, genome size, propagation systems, and breeding systems should be given priority. As the genomes of more crop species are characterized, a platform that houses, characterizes, and catalogs deleterious alleles in crop plants should be pursued (Box 1; see also Outstanding questions).

Acknowledgments

P.H-H. thanks the Overseas Distinguished Scholar Project of South China Botanical Garden, Chinese Academy of Sciences (Y861041001) and United Kingdom-India Pulses and Oilseeds Research Initiative, Official Development Assistance Newton Fund awarded by the Biotechnology and Biological Sciences Research Council (BB/R019819/1) for funding. R.O. acknowledges grant funding from both Formas and Vetenskapsråde for projects 'Genomisk Prediktion i Kombination med Högkapacitetsfenotypning för att Öka Potatisens Knölskörd i ett Föränderligt Klimat' (FR-2019/0007) and 'Framtagning av värmetoleranta alleler för att öka inkomsterna från jordbruket längs Senegalfloden' (VR 2021–04956), respectively.

Declaration of interests

No interests are declared.

Supplemental information

Supplemental information to this article can be found online at https://doi.org/10.1016/j.tplants.2023.01.006.

References

1. Charlesworth, B. (2012) The effects of deleterious mutations on evolution at linked sites. *Genetics* 190, 5–22
2. Assaf, Z.J. *et al.* (2015) Obstruction of adaptation in diploids by recessive, strongly deleterious alleles. *Proc. Natl. Acad. Sci. U. S. A.* 112, E2658–E2666
3. Zhao, J. *et al.* (2017) Quantifying the impact of non-coding variants on transcription factor-DNA binding. *Res. Comput. Mol. Biol.* 10229, 336–352
4. Monroe, J.G. *et al.* (2022) Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature* 602, 101–105
5. Xu, Y.-C. and Guo, Y.-L. (2020) Less is more, natural loss-of-function mutation is a strategy for adaptation. *Plant Commun.* 1, 100103
6. Kim, M.-S. *et al.* (2021) The patterns of deleterious mutations during the domestication of soybean. *Nat. Commun.* 12, 97
7. Renaut, S. and Rieseberg, L.H. (2015) The accumulation of deleterious mutations as a consequence of domestication and improvement in sunflowers and other Compositae crops. *Mol. Biol. Evol.* 32, 2273–2283
8. Żmieńko, A. *et al.* (2014) Copy number polymorphism in plant genomes. *Theor. Appl. Genet.* 127, 1–18
9. Mackay, T.F.C. (2009) Q&A: genetic analysis of quantitative traits. *J. Biol.* 8, 23
10. Valluru, R. *et al.* (2019) Deleterious mutation burden and its association with complex traits in sorghum (*Sorghum bicolor*). *Genetics* 211, 1075–1087
11. Mezmouk, S. and Ross-Ibarra, J. (2014) The pattern and distribution of deleterious mutations in maize. *G3* 4, 163–171
12. Park, J.H. *et al.* (2011) Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. *Proc. Natl. Acad. Sci. U. S. A.* 108, 18026–18031
13. Makino, T. *et al.* (2018) Elevated proportions of deleterious genetic variation in domestic animals and plants. *Genome Biol. Evol.* 10, 276–290
14. Günther, T. and Schmid, K.J. (2010) Deleterious amino acid polymorphisms in *Arabidopsis thaliana* and rice. *Theor. Appl. Genet.* 121, 157–168
15. Yang, J. *et al.* (2017) Incomplete dominance of deleterious alleles contributes substantially to trait variation and heterosis in maize. *PLoS Genet.* 13, e1007019
16. Monroe, J.G. *et al.* (2020) The lowest hanging fruit: beneficial gene knockouts in past, present, and future crop evolution. *Curr. Plant Biol.* 24, 100185
17. Komatsuda, T. *et al.* (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1424–1429
18. Lin, Z. *et al.* (2012) Parallel domestication of the *Shattering1* genes in cereals. *Nat. Genet.* 44, 720–724
19. Foucher, F. *et al.* (2003) *DETERMINATE* and *LATE FLOWERING* are two *TERMINAL FLOWER1/CENTRORADIALIS* homologs that control two distinct phases of flowering initiation and development in pea. *Plant Cell* 15, 2742–2754
20. Krylova, E.A. *et al.* (2020) Determinate growth habit of grain legumes: role in domestication and selection, genetic control. *Ekol. Genet.* 18, 43–58
21. Abe, F. *et al.* (2019) Genome-edited triple-recessive mutation alters seed dormancy in wheat. *Cell Rep.* 28, 1362–1369
22. Chai, M. *et al.* (2016) A class *II KNOX* gene, *KNOX4*, controls seed physical dormancy. *Proc. Natl. Acad. Sci. U. S. A.* 113, 6997–7002
23. Michaels, S.D. and Amasino, R.M. (2001) Loss of *FLOWERING LOCUS C* activity eliminates the late-flowering phenotype of *FRIGIDA* and autonomous pathway mutations but not responsiveness to vernalization. *Plant Cell* 13, 935–941
24. Mizuno, N. *et al.* (2016) Loss-of-function mutations in three homoeologous *PHYTOCLOCK 1* genes in common wheat are associated with the extra-early flowering phenotype. *PLoS One* 11, e0165618

25. Andeweg, J.M. and De Bruyn, J.W. (1959) Breeding of non-bitter cucumbers. *Euphytica* 8, 13–20
26. Barboza, L. *et al.* (2013) *Arabidopsis* semidwarfs evolved from independent mutations in *GA20ox1*, ortholog to green revolution dwarf alleles in rice and barley. *Proc. Natl. Acad. Sci. U. S. A.* 110, 15818–15823
27. Birt, D.F. *et al.* (2013) Resistant starch: promise for improving human health. *Adv. Nutr.* 4, 587–601
28. Miura, S. *et al.* (2022) Improving agricultural traits while maintaining high resistant starch content in rice. *Rice* 15, 28
29. Doebley, J.F. *et al.* (2006) The molecular genetics of crop domestication. *Cell* 127, 1309–1321
30. Molla, K.A. *et al.* (2021) Precise plant genome editing using base editors and prime editors. *Nat. Plants* 7, 1166–1187
31. Lopez, F.B. *et al.* (2021) Gene dosage compensation of rRNA transcript levels in *Arabidopsis thaliana* lines with reduced ribosomal gene copy number. *Plant Cell* 33, 1135–1150
32. Rönspies, M. *et al.* (2021) CRISPR-Cas-mediated chromosome engineering for crop improvement and synthetic biology. *Nat. Plants* 7, 566–573
33. Gehrke, F. *et al.* (2022) Nonhomologous end joining as key to CRISPR/Cas-mediated plant chromosome engineering. *Plant Physiol.* 188, 1769–1779
34. Chen, K. *et al.* (2019) CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant Biol.* 70, 667–697
35. Khan, A.W. *et al.* (2020) Super-pangenome by integrating the wild side of a species for accelerated crop improvement. *Trends Plant Sci.* 25, 148–158
36. Deb, S. *et al.* (2022) Applications of CRISPR/Cas9 technology for modification of the plant genome. *Genetica* 150, 1–12
37. Chia, J.M. *et al.* (2012) Maize HapMap2 identifies extant variation from a genome in flux. *Nat. Genet.* 44, 803–807
38. Abdelrahman, M. *et al.* (2021) Multiplex genome-editing technologies for revolutionizing plant biology and crop improvement. *Front. Plant Sci.* 12, 721203
39. Li, Y. *et al.* (2018) DeepSimulator: a deep simulator for nanopore sequencing. *Bioinformatics* 34, 2899–2908
40. Ryder, P. *et al.* (2017) Generation of stable nulliplex autopolyploid lines of *Arabidopsis thaliana* using CRISPR/Cas9 genome editing. *Plant Cell Rep.* 36, 1005–1008
41. Mousavi-Derazmahalleh, M. *et al.* (2018) Exploring the genetic and adaptive diversity of a pan-Mediterranean crop wild relative: narrow-leafed lupin. *Theor. Appl. Genet.* 131, 887–901
42. Horvath, R. *et al.* (2021) Selection on accessible chromatin regions in *Capsella grandiflora*. *Mol. Biol. Evol.* 38, 5563–5575
43. Wang, M. *et al.* (2018) Evolutionary dynamics of 3D genome architecture following polyploidization in cotton. *Nat. Plants* 4, 90–97
44. Bayer, P.E. *et al.* (2021) Sequencing the USDA core soybean collection reveals gene loss during domestication and breeding. *Plant Genome*, e20109
45. Zenda, T. *et al.* (2021) Omics-facilitated crop improvement for climate resilience and superior nutritive value. *Front. Plant Sci.* 12, 774994
46. Pramanik, D. *et al.* (2021) CRISPR-mediated engineering across the central dogma in plant biology for basic research and crop improvement. *Mol. Plant* 14, 127–150
47. Van Tassel, D.L. *et al.* (2022) Re-imagining crop domestication in the era of high throughput phenomics. *Curr. Opin. Plant Biol.* 65, 102150
48. Jiang, Y. and Li, C. (2020) Convolutional neural networks for image-based high-throughput plant phenotyping: a review. *Plant Phenomics* 2020, 4152816
49. Johnsson, M. *et al.* (2019) Removal of alleles by genome editing (RAGE) against deleterious load. *Genet. Sel. Evol.* 51, 14
50. Kono, T.J.Y. *et al.* (2018) Comparative genomics approaches accurately predict deleterious variants in plants. *G3* 8, 3321–3329
51. Cao, J. *et al.* (2011) Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nat. Genet.* 43, 956–963
52. Rodgers-Melnick, E. *et al.* (2015) Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3823–3828
53. Kono, T.J.Y. *et al.* (2016) The role of deleterious substitutions in crop genomes. *Mol. Biol. Evol.* 33, 2307–2317
54. Zsögön, A. *et al.* (2018) *De novo* domestication of wild tomato using genome editing. *Nat. Biotechnol.* 36, 1211–1216
55. Zsögön, A. *et al.* (2017) Genome editing as a tool to achieve the crop ideotype and de novo domestication of wild relatives: case study in tomato. *Plant Sci.* 256, 120–130
56. Rehman, F. *et al.* (2022) CRISPR gene editing of major domestication traits accelerating breeding for Solanaceae crops improvement. *Plant Mol. Biol.* 108, 157–173
57. Razzaq, A. *et al.* (2021) Next-generation breeding strategies for climate-ready crops. *Front. Plant Sci.* 12, 620420
58. Bitocchi, E. *et al.* (2016) High level of nonsynonymous changes in common bean suggests that selection under domestication increased functional diversity at target traits. *Front. Plant Sci.* 7, 2005
59. Lozano, R. *et al.* (2021) Comparative evolutionary genetics of deleterious load in sorghum and maize. *Nat. Plants* 7, 17–24
60. Ramu, P. *et al.* (2017) Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. *Nat. Genet.* 49, 959–963
61. Hojsgaard, D. and Hörandl, E. (2015) A little bit of sex matters for genome evolution in asexual plants. *Front. Plant Sci.* 6, 82
62. Wolfe, M.D. *et al.* (2019) Historical introgressions from a wild relative of modern cassava improved important traits and may be under balancing selection. *Genetics* 213, 1237–1253
63. Moyers, B.T. *et al.* (2017) Genetic costs of domestication and improvement. *J. Hered.* 109, 103–116
64. Bayer, P.E. *et al.* (2021) Modelling of gene loss propensity in the pangenomes of three *Brassica* species suggests different mechanisms between polyploids and diploids. *Plant Biotechnol. J.* 19, 2488–2500
65. Bayer, P.E. *et al.* (2019) Variation in abundance of predicted resistance genes in the *Brassica oleracea* pangenome. *Plant Biotechnol. J.* 17, 789–800
66. Dolatabadian, A. *et al.* (2020) Characterization of disease resistance genes in the *Brassica napus* pangenome reveals significant structural variation. *Plant Biotechnol. J.* 18, 969–982
67. Hulse-Kemp, A.M. *et al.* (2016) A HapMap leads to a *Capsicum annuum* SNP infinium array: a new tool for pepper breeding. *Hortic. Res.* 3, 16036
68. Bayer, P.E. *et al.* (2021) Yield is negatively correlated with nucleotide-binding leucine-rich repeat gene content in soybean. *bioRxiv* Published online December 13, 2021. https://doi.org/10.1101/2021.12.12.472330
69. Birky, C.W. and Walsh, J.B. (1988) Effects of linkage on rates of molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* 85, 6414–6418
70. Kumar, V. *et al.* (2016) First-generation HapMap in *Cajanus* spp. reveals untapped variations in parental lines of mapping populations. *Plant Biotechnol. J.* 14, 1673–1681
71. Kang, Y. *et al.* (2015) Genome-wide association of drought-related and biomass traits with HapMap SNPs in *Medicago truncatula*. *Plant Cell Environ.* 38, 1997–2011
72. Scheben, A. and Edwards, D. (2017) Genome editors take on crops. *Science* 355, 1122–1123
73. Scheben, A. *et al.* (2017) Towards CRISPR/CAS crops – bringing together genomics and genome editing. *New Phytol.* 216, 682–698
74. Eyre-Walker, A. *et al.* (2006) The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics* 173, 891–900
75. MacArthur, D.G. and Tyler-Smith, C. (2010) Loss-of-function variants in the genomes of healthy humans. *Hum. Mol. Genet.* 19, R125–R130
76. Labroo, M.R. *et al.* (2021) Heterosis and hybrid crop breeding: a multidisciplinary review. *Front. Genet.* 12, 643761
77. Gaut, B.S. *et al.* (2018) Demography and its effects on genomic variation in crop domestication. *Nat. Plants* 4, 512–520
78. Crow, J.F. (1998) 90 years ago: the beginning of hybrid maize. *Genetics* 148, 923–928
79. Charlesworth, D. and Willis, J.H. (2009) The genetics of inbreeding depression. *Nat. Rev. Genet.* 10, 783–796
80. Wang, J. *et al.* (1999) Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. *Genet. Res.* 74, 165–178
81. Bataillon, T. and Kirkpatrick, M. (2000) Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. *Genet. Res.* 75, 75–81

82. Bosse, M. *et al.* (2019) Deleterious alleles in the context of domestication, inbreeding, and selection. *Evol. Appl.* 12, 6–17
83. Lian, Q. *et al.* (2019) Acquisition of deleterious mutations during potato polyploidization. *J. Integr. Plant Biol.* 61, 7–11
84. Kitavi, M. *et al.* (2016) The triploid East African Highland Banana (EAHB) gene pool is genetically uniform arising from a single ancestral clone that underwent population expansion by vegetative propagation. *Theor. Appl. Genet.* 129, 547–561
85. Girma, G. *et al.* (2014) Next-generation sequencing based genotyping, cytometry and phenotyping for understanding diversity and evolution of Guinea yams. *Theor. Appl. Genet.* 127, 1783–1794
86. Raherison, E. *et al.* (2020) Evidence for the accumulation of nonsynonymous mutations and favorable pleiotropic alleles during wheat breeding. *G3* 10, 4001–4011
87. Krasileva, K.V. *et al.* (2017) Uncovering hidden variation in polyploid wheat. *Proc. Natl. Acad. Sci. U. S. A.* 114, E913–E921
88. Otto, S.P. and Yong, P. (2002) The evolution of gene duplicates. *Adv. Genet.* 46, 451–483
89. Tanaka, K.M. *et al.* (2009) Enhanced fixation and preservation of a newly arisen duplicate gene by masking deleterious loss-of-function mutations. *Genet. Res.* 91, 267–280
90. Pourkheirandish, M. *et al.* (2020) Global role of crop genomics in the face of climate change. *Front. Plant Sci.* 11, 922
91. Massel, K. *et al.* (2021) Hotter, drier, CRISPR: the latest edit on climate change. *Theor. Appl. Genet.* 134, 1691–1709
92. Spillane, C. *et al.* (2004) Apomixis technology development-virgin births in farmers' fields? *Nat. Biotechnol.* 22, 687–691
93. Charlesworth, B. and Charlesworth, D. (1998) Some evolutionary consequences of deleterious mutations. *Genetica* 102–103, 3–19
94. Jaramillo-Correa, J.P. *et al.* (2010) The contribution of recombination to heterozygosity differs among plant evolutionary lineages and life-forms. *BMC Evol. Biol.* 10, 22
95. Wickell, D.A. and Li, F.-W. (2020) On the evolutionary significance of horizontal gene transfers in plants. *New Phytol.* 225, 113–117
96. Kovalev, M.S. *et al.* (2018) A pipeline for classifying deleterious coding mutations in agricultural plants. *Front. Plant Sci.* 9, 1734
97. Vaughan, D.A. *et al.* (2007) From crop domestication to super-domestication. *Ann. Bot.* 100, 893–901
98. Weckwerth, W. *et al.* (2020) PANOMICS meets germplasm. *Plant Biotechnol. J.* 18, 1507–1525
99. Varshney, R.K. *et al.* (2021) Rapid delivery systems for future food security. *Nat. Biotechnol.* 39, 1179–1181
100. Wallace, J.G. *et al.* (2018) On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. *Annu. Rev. Genet.* 52, 421–444
101. Lora, J. *et al.* (2011) Seedless fruits and the disruption of a conserved genetic pathway in angiosperm ovule development. *Proc. Natl. Acad. Sci. U. S. A.* 108, 5461–5465
102. Lovell, J.T. *et al.* (2017) Mutation accumulation in an asexual relative of *Arabidopsis*. *PLoS Genet.* 13, e1006550
103. Armario Najera, V. *et al.* (2019) Applications of multiplex genome editing in higher plants. *Curr. Opin. Biotechnol.* 59, 93–102
104. de Maagd, R.A. *et al.* (2020) CRISPR/Cas inactivation of RECQ4 increases homeologous crossovers in an interspecific tomato hybrid. *Plant Biotechnol. J.* 18, 805–813
105. Mieulet, D. *et al.* (2018) Unleashing meiotic crossovers in crops. *Nat. Plants* 4, 1010–1016
106. Vaser, R. *et al.* (2016) SIFT missense predictions for genomes. *Nat. Protoc.* 11, 1–9
107. Davydov, E.V. *et al.* (2010) Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput. Biol.* 6, e1001025
108. Courtier-Orgogozo, V. *et al.* (2020) Gephebase, a database of genotype-phenotype relationships for natural and domesticated variation in eukaryotes. *Nucleic Acids Res.* 48, D696–D703