SPECIAL ISSUE ON GENOME EDITING





Generating novel plant genetic variation via genome editing to escape the breeding lottery

Nathaniel Schleif¹ · Shawn M. Kaeppler^{1,2} · Heidi F. Kaeppler^{1,2}

Received: 22 June 2021 / Accepted: 24 June 2021 / Published online: 8 September 2021 / Editor: David Songstad © The Author(s) 2021

Abstract

Plant breeding relies on the presence of genetic variation, which is generated by a random process of mutagenesis that acts on existing gene pools. This variation is then recombined into new forms at frequencies impacted by the local euchromatin and heterochromatin environment. The result is a genetic lottery where plant breeders face increasingly low odds of generating a "winning" plant genotype. Genome editing tools enable targeted manipulation of the genome, providing a means to increase genetic variation and enhancing the chances for plant breeding success. Editing can be applied in a targeted way, where known genetic variation that improves performance can be directly brought into lines of interest through either deletion or insertion. This empowers approaches that are traditionally difficult such as novel domestication and introgression of wild accessions into a germplasm pool. Furthermore, broader editing-mediated approaches such as recombination enhancement and targeted random mutagenesis bring novel ways of variation creation to the plant breeding toolbox. Continued development and application of plant genome editing tools will be needed to aid in meeting critical global crop improvement needs.

Keywords Genome editing · Plant breeding · Genetic variation · Crop improvement

Overview

Plant breeding has traditionally been an iterative process of combining favorable alleles from two or more parents into individual progeny, improving their phenotypic characteristics. It is a biological lottery with poor odds, dependent primarily on segregation and recombination of alleles. For example, the probability of recovering a homozygous inbred individual combining favorable alleles at 10 unlinked loci is one in one million. The probability decreases to one in 10 billion or worse if the loci are physically located near each other on one or more chromosomes. Given that many species have 25,000 to 30,000 genes, with multiple alleles present for each, the challenge of relying on probability clearly produces incremental progress at best.

In addition to the probabilistic challenge of optimal individuals being present in a population, superior cultivars need to be identified based on phenotypic observation. If the lottery winner happens to be grown in poor field plots, it will be measured as underperforming relative to others and never be identified. Heritabilities for many relevant traits range from 30 to 80% on average, meaning that for traits such as grain yield, more than one-half of the observed variation may be nongenetic, which masks the true value of individuals. The statistical improbability of generating genetically optimal individuals along with challenges in measuring performance has driven the search for technologies to improve cultivar development. It should be noted that, despite the aforementioned difficulties, crop domestication and formalized breeding efforts have produced the species and varieties that support civilization today through a large-scale composite effort across thousands of years. Time is no longer on our side, however, and breeding progress must be increased to address current and future agricultural production needs in the face of population increase and changing climate conditions.

Genetic technologies play an increasingly important role in the breeding process: Genomic sequencing and genetic marker analysis have been harnessed to identify genetically superior individuals based on their DNA composition and



Heidi F. Kaeppler hfkaeppl@wisc.edu

Department of Agronomy, University of Wisconsin-Madison, 1575 Linden Drive, Madison, WI, USA

Wisconsin Crop Innovation Center, University of Wisconsin-Madison, Middleton, WI, USA

minimizing reliance on variable field evaluations. In addition to improving the probability of identifying genetically superior individuals, generation time can be accelerated, and cost reduced, by selecting on genetic marker profiles rather than measured performance in production environments. Genomic and marker-assisted selection are now common in many breeding programs for many species. Despite the utility of these approaches, the process still fundamentally relies on the genetic segregation and recombination lottery previously described while improving the odds of identifying superior individuals, if they exist. The process also is limited to a large extent by the allelic variation present within the scope of accessible sexual crosses. Precise, targeted modification of DNA sequences to accelerate accumulation of favorable alleles, and to create variation that is not available through sexual crosses, has the potential to dramatically change the process, pace, and efficiency of cultivar development (Arora and Narula 2017; Scheben et al. 2017; Wang et al. 2019a, b). New breeding technologies, including transgenics, cisgenics, and geneediting, provide tools to raise the process of crop improvement to a new level.

Specific Challenges and Opportunities in Crop Breeding

Positive allelic variation is necessary for breeding progress; major effect alleles are preferred The process of breeding relies on having sequence variation available to select improved individuals. Sequence variation exists on a continuum of complexity, ranging from single nucleotide changes to translocations thousands of base-pairs in length. The simplest form of sequence variation is single nucleotide polymorphisms (SNPs), which is the change of one base to another within a sequence. More complex types of sequence variation include insertions and deletions, commonly called InDels, which are typically defined as ranging from one base pair to 50 base pairs in length. Also in this category are copy number variants (CNVs) and presence/absence variants (PAVs). CNVs are variations in the number of copies of a sequence in a genome while PAVs are differences in the presence or absence of the sequence entirely. Complementing these are larger structural variants: inversions and translocations, which result in large-scale rearrangement of portions of chromosomes.

Many genetic variants are "neutral changes" in sequence that do not result in observable phenotypic variance (Nei *et al.* 2010); however, the genetic changes most interesting and valuable to plant breeders are those that alter phenotypes for traits of interest. Consequently, understanding how sequence variation manifests into phenotypic variation is vital to being able to manipulate variation intelligently. SNPs and InDels can induce a variety of changes. The coding sequence of a

protein can be changed, resulting in different structures and function (Wu *et al.* 2017). Alternatively, translation start sites, premature stop codons, RNA stability, and even translation efficiency can all be impacted by SNPs and InDels. Changes can also be made outside of coding regions such as at cisregulatory elements (Wang *et al.* 2019a, b), changing the expression pattern of an affected gene.

Larger modifications can induce all the above changes as well as novel outcomes: CNVs over evolutionary time can lead to neo-functionalization and novel adaptations. For example, the rice gene, TB2, is a duplicate gene of TB1 and has been found to have a novel role in positive regulation of tillering, opposite to the negative regulation by TB1. Consequently, variants of TB2 which cause heightened expression have been selected for in upland rice for enhanced tiller number (Lyu et al. 2020). PAVs can modify whole networks of genes through the introduction of novel genes or modification of existing ones. This is exemplified in the case of the formation of the VGT1 locus, which is caused by a miniature inverted-repeat transposable element (MITE) insertion into an upstream regulatory element, resulting in differences in flowering time in maize (Castelletti et al. 2014). In this case, the MITE insertion resulted in a change in the epigenomic landscape, which is another way for variations to manifest into phenotypic outcomes. Ultimately, there are many different outcomes that sequence variation can create, each of which is potentially useful for generating the next productive crop variety.

Equally important to understanding the kinds of variation is understanding how variation has been generated historically. Much of the crop genetic variation accessed today is historical in origin and was accumulated over thousands to millions of years of evolution. For many species, trillions or more individuals grow each year, and mutations occur at some frequency in all of them. In Arabidopsis thaliana, the rate of single site mutation has been found to be only 7×10^{-9} base substitutions per site per generation, which, given a genome size of 120 million base pairs, means only a single base change per generation (Ossowski et al. 2010). In addition to mutagenic forces, sequence variants arise from recombination errors, DNA repair errors, and replication errors (Gabur et al. 2019). Most sequence variations are neutral or negative in effect, with the negative variations being rapidly purged from populations due to negative selection. Some mutations are positive, and a small fraction of those positive mutations have large effects on plant phenotype. These large effect positive alleles are usually quickly fixed in populations through selection (Stetter et al. 2018), leaving breeders to sort through many small effect alleles across many loci, a relatively inefficient substrate to make breeding progress.

Unfortunately, the processes of novel variation generation in plant breeding have become substantially limited. For each crop that was domesticated in human history, billions to



trillions of plants were grown by countless small-scale farmers, and each genetically unique plant is manually observed by a human (Meyer and Purugganan 2013). Currently, in the USA alone, trillions of crop plants are grown each year, but the vast majority are bulk harvested for consumption, not observed, and not available to transmit new variation to the next generation. As a result, in current farming systems, the genetic variation-generating "machine" has largely been taken out of operation. Approaches to increase or produce positive allelic variation, and ideally large effect positive alleles, are critically needed to replace that machinery.

Targeted alteration of recombination frequencies and patterns could unlock cryptic variation It is not sufficient to merely possess favorable alleles in a breeding population; they need be combined into a single individual, which is complicated by the nature of genetic linkage. The proximity of genes on the same chromosome that are contributing to phenotypic variation can be either positive or negative. If favorable alleles are genetically linked, they will disproportionately be transmitted together through sexual generations increasing the chance that offspring will inherit the favorable combination. In many cases, favorable and unfavorable alleles are genetically linked, locking up variation that can only be expressed and resolved following recombination, which is a challenging process to rely on. Recombination occurs infrequently, with rarely more than three cross-over (CO) events occurring per pair of chromosomes per generation (Christine et al. 2015). Cross-over number increases with size of the chromosome within a species, but that does not apply across all species. Furthermore, recombination does not take place uniformly across the genome with 80% of COs concentrated into only 25% of the genome in most plant species (Blary and Jenczewski 2019).

One repercussion of this low frequency of recombination is that undesirable alleles accrue in those regions because of the selection of linked alleles (Cutter and Payseur 2013). Genetic drift is also more likely to cause fixation in these regions, even in the absence of selection. The specificity of where crossover events occur is a result of differentially accessible regions of the genome, euchromatin and heterochromatin, for those that are more or less accessible, respectively. Creation of these regions has been found to be influenced by methylation and chromatin accessibility, though this is still an open area of investigation (Blary and Jenczewski 2019). The outcome of this phenomenon is that the exploitable sequence variation space is severely limited. For example, in maize, up to 20% of genes are located in regions largely untouched by recombination (Bauer et al. 2013). This impacts the ability to make genetic gain within crops as well as the ability to deploy approaches such as QTL mapping and genomic selection models, both of

which rely on recombination. Overcoming this limitation represents a large potential gain for breeding progress.

While major effect alleles have been impactful on crop development, much of the phenotypic variation resides outside of detectable QTL regions, indicating that many small effect alleles working in concert underlie many traits (Huang et al. 2012; Boyle et al. 2017; Miao et al. 2019; Chen et al. 2020). Organizing these small-effect alleles into a maximized configuration is the modern plant breeding struggle, which is currently limited by the number of COs per meiotic event. Even in ideal conditions where breeders have perfect information on the best alleles, it would take dozens of generations to arrive at an ideal combination. Breeding is ultimately a numbers game and recombination is a multiplier on the efficiency of a program. Increasing this multiplier can now be achieved through application of editing technologies.

Combining favorable genes and alleles from wild progenitors and across species boundaries Integration of novel or exotic breeding material into the germplasm of a breeding program is generally limited by both biological boundaries which prevent effective sexual crosses as well as the barrier of geneticmaladaption of exotic germplasm. In the first case, genomic information coupled with knowledge of local adaptation, evolutionary history, and direct performance measurements now supports identification of favorable alleles within a diverse plant species. This information is commonly found to be relevant to other species, even those across sexual boundaries. The greatest example of this is the past and present use of model species to learn more about plants broadly. For instance, much of the understanding of flowering time in maize and rice originated in Arabidopsis and has been transferred through laborious experimentation (Hill and Li 2016). In the second case, exotic lines and wild progenitors carry negative traits along with potentially useful sequence variation. Because variation tends to be reduced in modern crop breeding pools compared to exotic germplasm and wild progenitors, breeders may be missing out on potentially useful variation. For example, the genetic regions that most strongly differentiate maize from its wild progenitor, teosinte, explain the least quantitative variation, demonstrating fixation of alleles through domestication. Overall, the number of detectable quantitative trait loci (QTLs) in maize compared to teosinte are much smaller, indicating additional evidence of fixation (Chen et al. 2020). Nucleotide diversity in maize has been reduced by 30% overall, which is even greater in regions under selection (Wright et al. 2005; Hufford et al. 2012). Similar patterns of variation loss have been noted in other crops such as wheat and soybean (Sedivy et al. 2017; Maccaferri et al. 2019). Without new sources of variation, increased selection pressure will only further deplete existing variation, causing breeding gains to plateau over time (Jannink 2010).



Traditional breeding is able overcome both barriers in some capacity: Cross-species information can be applied by looking for existing variation within the shared genetic regions, though there is no guarantee that the necessary materials will be available. Prebreeding techniques can be used to make exotic and wild germplasm more amendable for use; however, this is time and resource intensive. Editing technologies have the potential to lower the barrier for bringing to bear the encyclopedia of evolution across all species by enabling direct modification of existing sequences. These modifications can duplicate findings in other species and distant relatives, enabling faster cycles of development at reduced cost. Furthermore, editing technologies may allow creation of wholly designed sequences, potentially better than any that could be found within any existing individual.

The Current Tools of Gene Editing

First-generation editing tools The ability to selectively identify and manipulate DNA sequences can significantly impact advancements made in the areas of crop functional genomics research and crop improvement. Gene editing systems now provide the molecular tools to change DNA sequences in specific and controlled ways. The first gene editing systems developed were based on meganucleases as well as use of zincfinger nucleases (ZFNs) to locate, cut, and create doublestranded breaks (DSB) in targeted DNA sequences (Carroll 2008). ZFNs systems use the FokI nuclease cleavage domain paired with zinc finger DNA-binding domain protein motifs (ZFP) which recognize 3 base-pair DNA sequences, with every possible 3 base-pair combination now available for targeting of typically 20-30bp DNA sequences through linking of 6-10 ZFN units. Another system for genome editing was subsequently developed based on utilization of transcription activator-like (TAL) effector proteins to target nuclease activity, called transcription activator-like effector nucleases (TALENs), to specific DNA sequences (Bedell et al. 2012). Each TAL effector targets a single base pair and effectors can be connected to generate targeted editing of specific DNA sequences. Similar to ZFNs, the TAL effectors are paired with FokI nuclease to provide the sequence recognition and DNA cleavage functions, respectively. In both systems, to target a different DNA sequence for cleavage, new vectors encoding ZFN and TALEN arrays able to recognize the new target sequence must be built. Additionally, since the FokI nuclease is used for DNA cleavage, vectors encoding the arrays must be designed as dimers targeting both the forward and reverse strands. The complexities of array and vector design in the ZFN and TALEN systems, coupled with the laborious effort required to redesign vectors for each new DNA sequence target, have impeded the widespread use of those systems for plant genome editing.

CRISPR/Cas9-based editing systems CRISPR/Cas9-based genome editing systems (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated protein 9) have been utilized for 8 yr in plants (Li et al. 2013; Nekrasov et al. 2013; Shan et al. 2013) and has continued to expand in applications and approaches. CRISPR/Cas9-based gene editing systems present a significant advancement over previous systems due to simplicity and versatility of vector design and construction, which has resulted in a wider array of editing tools being developed utilizing this system (Chen et al. 2019; Aamir et al. 2020; Zhu et al. 2020; Nadakuduti and Enciso-rodríguez 2021). In the most common implementation of the system, the Cas9 nuclease enzyme targets a specific DNA sequence for cleavage and creation of double-strand breaks (DSBs) as specified by the single guide RNA sequence (sgRNA), which is made up of a variable sequence targeting region and a fixed Cas9-associated region. This sgRNA contains a protospacer-adjacent motif (PAM) that is part of the fixed region and specifies a 5'-NGG-3' sequence along with the variable targeting region (Jinek et al. 2012; Cong et al. 2013; Mali et al. 2013a, b). Modifications of the original system have overcome this limitation however, with the part of the Cas protein that confers PAM specificity being changed to provide a range of possible PAM sites (Kleinstiver et al. 2015; Hu et al. 2018; Nishimasu et al. 2018). A recent development in this vein has been the "SpRY" engineered SpCas9 which has been shown to have a highly flexible PAM recognition in rice (Xu et al. 2021).

Changing the DNA sequence to be targeted for editing requires simply changing the guide sequence of the sgRNA instead of engineering a new protein, as in previous editing systems. Additionally, since the platform is not reliant on expression of sequence-recognition proteins such as those in ZFN and TALENS systems, it makes deployment against many targets at once (multiplex editing) much easier. One drawback of the system is that off-target edits have been observed, though this effect seems to be smaller in plants compared to other species (Fu et al. 2013; Zhang et al. 2018a, b). Cas9 is not the only protein utilized for editing as it is a member of a broader family of related enzymes that have been derived from other CRISPR systems, including Cas12a, $Cas\Phi$, and Cms1, each with slightly different capabilities (Zetsche et al. 2015; Begemann et al. 2017; Pausch et al. 2020). Cas12a proteins have also been modified to have altered PAM recognition sites (Gao et al. 2017; Li et al. 2018; Zhong *et al.* 2018).

The double-stranded cut created by the Cas9 nuclease creates genetic modifications by engaging the natural cellular DNA repair pathways of nonhomologous end joining (NHEJ) and homology-directed repair (HDR) (Symington and Gautier 2011; Ceccaldi *et al.* 2016; Kosicki *et al.* 2018). NHEJ functions in plants through a classical Ku-dependent manner as well as many backup systems, which results in



small insertions and deletions (indels) at the site of repair, though weighted towards deletions over insertions (Shen *et al.* 2017). HDR employs a template sequence which it uses to repair break sites, repairing the damaged site as a copy of the template. This template sequence can come from a homologous chromatid or exogenous sources delivered to the cell and can result in sequence replacements up to several hundred base pairs in length (Shi *et al.* 2016). HDR is limited, however, to the S and G2 phases of the cell cycle which makes it exceedingly less common than NHEJ, with some estimates of the natural rate of HDR as low as 1 in 10,000 (Puchta 2005; Salsman and Dellaire 2017).

Expanded applications of Cas9 for editing beyond doublestrand breaks The establishment of CRISPR/Cas9-based editing systems and the associated advantages of those systems has led to rapid, widespread adoption for editing applications in a broad array of species, including plants. This has resulted in steady development of modified and improved forms of Cas9 and Cas9-like nucleases and associated RNA guides leading to novel editing functions. These novel functions and applications are summarized in Table 1, both as applied in plants as well as their non-plant origins. Deactivating one of the two nuclease domains of Cas9 results in the formation of a nickase (nCas9) which cuts only one strand of the double-stranded DNA target sequence (Jiang and Doudna 2017). This can then be used in a paired fashion to create DSBs that have enhanced specificity compared to standard Cas9 (Ran et al. 2013). If both nucleases are deactivated (dCas9), the enzyme will still retain its ability to associate with guide RNAs and can thus be directed to specific DNA sequence locations. The dCas9 can, for example, be targeted to cis-regulatory elements (CREs) and RNA polymerase binding sites to disrupt gene expression, referred to as CRISPR interference (Oi et al. 2013). dCas9 can also be paired with other proteins to target them to specific regions of the genome. Transcription factors can be attached to the Cas9 protein and brought to sites of interest, such as promoter regions, enhancing or repressing expression. For example, a fusion of dCas9 with the transcription factors EDLL and MS:VPR was found to be effective for targeted strong activation of low-expressed genes in tomato (Vazquez-vilar et al. 2019). A final application of dCas9 fusions has been with methylation and demethylation-related proteins and has been used to epigenetically modulate gene expression (Vojta et al. 2016; Xu et al. 2016).

One suite of tools enabled by linking other proteins to nCas9 proteins is base editing, in which one base can be specifically exchanged for another. C:G to T:A base transitions, known as cytosine base editing (CBE), can be accomplished by pairing nCas9 with a cytidine deaminase and a uracil DNA glycosylase (Komor *et al.* 2016). While originally developed in mammalian models, it has been optimized and tested in

crops (Zong et al. 2017, 2018). CBE is especially useful because it can be used to create precise DNA deletion sites through the activity of uracil DNA glycosylase and a coexpressed AP lyase, which cuts at a specific location in the protospacer (Wang et al. 2020). A:T to G:C transitions are also possible, termed adenine base editing (ABE), and are enabled by nCas9 fusions with adenosine deaminase (Gaudelli et al. 2017).

Non-specific changes are another vital tool enabled by Cas9-mediated sequence targeting. One such system is "saturated targeted endogenous mutagenesis editor" (STEME) which consists of a fusion of adenosine deaminase, nCas9, and a uracil DNA glycosylase inhibitor. As opposed to ABEs, STEMEs are capable of generating dual transitions of C:G to T:A and A:T to G:C, creating diverse changes at a given site (Li *et al.* 2020a, b). Another system, termed Evolvr, is based on nCas9 paired with an error-prone DNA polymerase providing the capability of creating all 12 substitutions (Halperin *et al.* 2018). These methods are exciting as they enable targeted random mutagenesis, though they have yet to be applied in plants. Notably, though, the editing systems described above do not insert desired sequences, but merely change what is already present.

Controlled insertion of DNA at specified locations through editing One of the most exciting applications of editing technologies is to precisely insert/replace specified sequences into genomes. This has been enabled by leveraging the HDR pathway, despite its rare occurrence. In this system, the standard Cas9 enzyme with a sgRNA is paired with a homology repair DNA template. The Cas9 cleaves the DNA at a specific site and then the provided repair template is ideally used to make the repair (Huang and Puchta 2019). Because the process is rare, many techniques have been developed to try to enhance occurrence of sequence insertion. Modulating the expression of genes involved in the NHEJ and HDR processes has been fruitful. In rice, removal of the ligase 4 gene involved in the NHEJ mechanism led to enhancement of HDR-directed editing (Endo et al. 2016). Additionally, in Arabidopsis, expression of yeast-derived RAD54 has been found to enhance HDR as well as successful insertion events (Even-Faitelson et al. 2011). Another successful approach has been to enhance expression of the template sequence through use of geminivirus replicons, which has been successfully deployed in both monocot and dicot targets (Baltes et al. 2014; Wang et al. 2017b; Hahn et al. 2018). Given the importance of controlled insertions as a goal, new systems are continually being developed.

Potentially more useful than HDR-mediated approaches is the recently developed "prime editing" system which utilizes nCas9 fused with reverse transcriptase and prime editing guide RNAs (pegRNA). The nickase creates a site of singlestranded DNA which matches the 3' end of the pegRNA,



 Table 1. Fundamental gene editing studies in plant and non-plant organisms

Species	Technology	Application	Description	Reference
Plant applications of editing				
Solanum lycopersicum	Prime editing	Sequence replacement	Demonstrated efficacy of prime editing in tomato through restoring luciferase activity	Lu et al. (2021)
Arabidopsis / Brassica napus	ABE	Base change	Demonstrated efficacy of ABE editing in generating germline edits in <i>Arabidopsis</i> and rapeseed	Kang <i>et al.</i> (2018)
Zea mays	SpCas9 HDR	Base change	Application of Cas9-mediated HDR to characterize gene function in maize	Jaqueth <i>et al</i> . (2020)
Zea mays	SpCas9	Expression disruption at scale	Large-scale screen of potential agriculturally relevant phenotype causing genes in maize	Liu et al. (2020)
Solanum lycopersicum	SpCas9	Mutagenesis	Generating novel variation in a promoter sequence through non-specific indel creation in tomato.	Rodríguez-Leal et al. (2017)
Oryza sativa	SpCas9	Mutagenesis	Variation creation within the coding sequence of a spliceosome component of rice	Butt et al. (2019)
Oryza sativa	SpCas9	Expression disruption at scale	Large-scale screen of genes that underlie agricultural phenotypes through indel creation	Meng <i>et al.</i> (2017)
Arabidopsis thaliana / Lactuca sativa	SpCas9	Mutagenesis	Variation creation within the uORF of genes important to the brassinosteroid and ascorbic acid pathways through indel creation	Zhang <i>et al.</i> (2018a, b)
Nicotiana tabacum	SpCas9	Expression disruption	Demonstrated heritable multiplex editing using an RNA virus and mobile sgRNA in tobacco.	Ellison <i>et al</i> . (2020)
Oryza sativa	Varied	Varied	Demonstrated simultaneous indel and base editing in a multiplexed fashion in rice	Li <i>et al.</i> (2020a, b)
Zea mays	SpCas9	Editing elite lines	Demonstrates the efficacy of applying editing to enhance performance in elite lines of maize.	Gao et al. (2020)
Solanum lycopersicum	SpCas9	Mutagenesis	Domestication of wild tomato species through modifying expression patterns	Li et al. (2018)
Oryza sativa	Varied	Base change	Developed novel base editing options for both ABE and CBE	Hua et al. (2019)
Oryza sativa	SpCas9 HDR	Sequence replacement	Demonstrated the ability to replace gene sequences in rice	Li et al. (2016)
Oryza sativa /Solanum lycopersicum	CBE	Base change	Demonstrated efficacy of using activation-induced cytidine deaminase with Cas9 to induce point mutations in rice and tomato	Shimatani <i>et al.</i> (2017a, b)
Oryza sativa/Pisum sativum/Solanum lycopersicum	SpCas9	Recombination enhancement	Demonstrated efficacy of disrupting recombination genes to enhance recombination in pea, rice, and tomato	Mieulet <i>et al.</i> (2018)
Zea mays/Triticum aestivum	SpCas9		Demonstrated efficacy of editing with a haploid-inducing line in maize and wheat	Kelliher <i>et al.</i> (2019)
Triticum aestivum	TALEN/Cas9	Expression Disruption	Demonstrated editing homoeoalleles in polyploid wheat	Wang <i>et al</i> . (2014)
Zea mays	Prime editing	Sequence Replacement	Utilized prime editing to convert two maize ALS genes to a resistant form, conferring herbicide resistance	Jiang <i>et al.</i> (2020)
Oryza sativa/Triticum aestivum	Prime editing	Sequence Replacement	First study demonstrating prime editing in crops, specifically in rice and wheat	Lin et al. (2020)
Non-plant applications of edition	ng	•	1	
Homo sapiens	Prime editing	Sequence replacement	Demonstrated efficacy of prime editing in human cells.	Anzalone <i>et al.</i> (2019)
Homo sapiens	CBE	Base change	Demonstrated efficacy of targeted C to U conversion in human cells.	Komor <i>et al.</i> (2016)
Homo sapiens	ABE	Base change	Demonstrated efficacy of targeted A to G conversion in human cells	Gaudelli <i>et al.</i> (2017)
Escherichia coli	Evolvr	Mutagenesis	Demonstrated efficacy of Evolvr for inducing non-specific mutations in <i>E. coli</i>	Halperin <i>et al.</i> (2018)
Homo sapiens/Mus musculus	nCas9	Expression disruption	Demonstrated functionality of the nickase form of SpCas9 in mouse and human cells	Cong <i>et al</i> . (2013)



Table 1. (continued)

Species	Technology	Application	Description	Reference
Homo sapiens	Cas9 fusion	Transcription enhancement	Utilized a deactivated Cas9 for inducing targeted expression and showed the higher specificity of paired nickases in human cells	Mali <i>et al.</i> (2013a, b)

which primes the reaction of reverse transcriptase. The reverse transcriptase then synthesizes the DNA sequence specified by the pegRNA, and the DNA sequence is then integrated into the genome through ligation and repair (Anzalone *et al.* 2019). This system enables any base substitution to be performed as well as insertions currently up to 44 bp and deletions of 80 bp. While having been tested in rice, maize, wheat, and tomato, more work still needs to be conducted in testing, optimizing, and applying the technology to diverse crops (Jiang *et al.* 2020; Lin *et al.* 2020; Lu *et al.* 2020). As a new technology, prime editing likely holds great unexplored potential.

Applying Genome Editing to the Creation of Genetic Variation

There are two interconnected approaches to create genetic variation and diversity with gene editing tools: targeted and broad spectrum. In the first case, genes or genome regions of interest that have been identified can be targeted for specific changes, for example, the deletion of a gene or sequence of interest, or the recombination of a known sequence into backgrounds not previously explored. In the second approach, broad areas of the genome are targeted for nonspecific genetic variation creation, for example, enhancement of recombination location/frequency or mutation of a targeted region. While targeted approaches have the drawback of relying on detailed pre-existing sequence knowledge, they would generate predictable outcomes. Broad spectrum approaches can be more generally applied, but, likewise, can be strengthened by application of specific information. These approaches exist on a continuum and some approaches will fall into both categories.

Worth addressing is the relative value and power of the different forms of genetic variation that can be created via editing. The preponderance of studies, thus far, have used deletion to disrupt gene function. This has proven very powerful for plant functional genomics research and breeding applications; however, fundamentally subtractive processes will only go so far in making breeding advancements. Much more challenging is creating novel additive variation because it either (A) requires foreknowledge about what variation is beneficial, as in the case of knock-in experiments, or (B) requires

screening for rare instances of novel creation of variation when using random mutators like Evolvr.

Targeted insertions and deletions for variation expansion The ability to identify the genetic mechanisms underlying agriculturally relevant traits has greatly increased with the advent of low-cost genome sequencing technologies. QTL and genome-wide association studies (GWAS) analysis have proven very effective, for example, in finding causative alleles in flowering time and stalk biomass in maize (Buckler et al. 2009; Mazaheri et al. 2019), yield in rice (Li et al. 2018a), salt tolerance in soybean (Zeng et al. 2017), and many other traits worth interrogating. Added to this is the diverse body of mutation screens, expression assays, and many other approaches that enable connection of specific traits with causative genes. Editing technologies enable direct targeting of these identified genes, either replicating known useful alleles or trying to determine what the optimal allele would be.

The most apparent application of editing in plant breeding applications is to knock-out genes that are traitrelevant to study gene function and effects on phenotype. This approach has been applied in species ranging from apples to wheat for traits ranging from disease resistance to yield enhancement (Shelake et al. 2019). Advances in multiplexing technology has enabled many genes to be targeted at once such as multiplex targeting of multiple grain-related genes in rice (Zhou et al. 2019). This concept has been advanced to a larger scale in maize where 743 genes linked with agronomically relevant traits were targeted for editing via pooled and multiplexed Cas9 constructs, resulting in 118 genes being identified with phenotypic outcomes (Liu et al. 2020). This subtractive process can also be used to manifest an enhancement of expression.

Upstream open reading frames (uORFs) are alternative reading frames that are highly abundant in the genomes of plants, with 30 to 40% of genes having them (Von Arnim et al. 2014). The outcomes of expression of uORFs are a general attenuation of expression at the translation step as the uORFs "take up" the time of the translational machinery (Kozak 2001, 2002). It has been demonstrated that editing these sites can result in an enhancement and modulation of the expression level of mRNA in Arabidopsis and lettuce (Zhang et al. 2018a). Specifically, the disruption of the



uORF in LsGGP2 resulted in enhancement of both ascorbate content and oxidation stress tolerance. This technique provides another form of variation creation, one focused on alteration of expression levels.

Due to the challenging nature of generating sequence knock-ins, it has been used much less frequently to introduce genetic variation compared to knockouts. One application has been in maize where the promoter of the maize gene ARGOS8 was replaced with the promoter of GOS2, a different though related endogenous gene (Shi et al. 2017). ARGOS8 regulates ethylene response negatively and generally is expressed at a low level. By exchanging the promoter and raising expression of the gene, the researchers were able to suppress the ethylene response. The outcome of this modification was a 5 bushel per acre increase under drought stress conditions. Another instance has been in tomato where a single nucleotide change mediated through HDR was made to the ALC gene (Yu et al. 2017). The resulting mutant had excellent storage performance and the edited sequence was inherited in future generations. Another additive-variation tool, base editing, has also been used in a limited capacity for crop improvement. The majority of efforts have centered on modifying genes to confer herbicide resistance where, in many cases, a single point mutation is sufficient to confer resistance (Shimatani et al. 2017a, b; Kuang et al. 2020).

While very new, prime editing has already been used in many crop species to introduce specific variation. Among crops, the first application was in rice and wheat, where insertions up to 15bp were demonstrated (Lin et al. 2020). Subsequent studies have shown even larger insertions, with sequences as long as 2,049bp being successfully inserted in rice (Lu et al. 2020). Prime editing has also been demonstrated in maize, where specific point mutations were made in the ALS genes, with efficiencies of 53.2% and 6.3% for their two targets (Jiang et al. 2020). Dicots have proven more challenging to edit in this manner, though success in tomato has been shown through codon and promoter optimization, with editing frequencies reported between 0.025 and 1.66% (Lu et al. 2020). Increasing the efficiency of prime editing is a primary goal, and strategies such as deploying paired and optimized pegRNAs to enhance prime editing efficiency have been demonstrated in rice protoplasts (Lin et al. 2021); further enhancements will continue to be made as the technology is explored. Prime editing has the potential to open the door for precise but wide-scale engineering of plants as the technology advances.

Targeted but non-specific variation creation The above applications are all dependent on the quality of target sequence information available and knowing what genetic variation state is best. It is common to simply know that a gene or region is impactful to a trait but not what configuration would be "optimal." Targeting non-specific editing to these sites to generate novel variation could potentially generate new useful

alleles. The most straight-forward way to do this is by targeting promoter regions specifically, since changing gene expression patterns is a powerful way to modulate plant traits. An example of successful application of this approach using traditional transformation technologies was with modulation of expression of the plant development gene ZMM28 in maize. The gene is naturally expressed at a low level during the stages of development past V6, with expression being undetectable at early phases. Moderate level constitutive expression of the gene across all developmental stages using the Gos2 promoter resulted in a significant gain in yield through enhancing nitrogen utilization, carbon assimilation, and plant growth overall (Wu *et al.* 2019). Editing tools can still be used in this case to bring in novel variation through targeted, though nonspecific, variation creation.

One way that this can be achieved is by using Cas9 to target the promoter region of a gene of interest. This method has been applied in tomato to disrupt expression of the SICLV3 gene for fruit locule number as well as the S and SP genes for plant architecture (Rodríguez-Leal *et al.* 2017). A multiplexed Cas9 construct was used to target the 2kb upstream region without concern for any predicted CREs. The outcome was a range of phenotypic traits created for each gene, which were able to be fixed in the F2 generation. Significantly, they were not able to correlate phenotypic impact with expression level, revealing the complicated, non-linear nature of dose-sensitive genes in plants (Birchler *et al.* 2016). Furthermore, they discovered that expression levels could not be explained by the simple action of predicted CREs, revealing the still-present gaps in our understanding of promoter structure, function, and regulation.

As with the other approaches, this process is subtractive in nature. The changes created by the edits were predominantly deletions of various regions, but also some inversions and small indels within the promoter. However, expression of genes is controlled by both positive and negative regulators. One such negative regulator is the lc QTL in the WUS promoter of tomato, where targeting it for mutation via editing was shown to enhance expression of the gene (Rodríguez-Leal *et al.* 2017). Through this mechanism, expression can be changed in both directions. However, it is still dependent on sequences already present in the genome, instead of what could potentially be there. As of yet, no study has looked at creating additive variations within the promoter region.

Less commonly targeted for additive variation creation are proteins themselves. This is due in part to the difficulty in understanding the structural features that give proteins their function. Ever declining sequencing costs have made manipulating expression levels by changing regulatory sequences much easier in comparison. However, there are instances of agronomically important alleles such as sd1, a GA-related gene related to a dwarf phenotype in rice, being determined to be a result of protein structural changes (Hedden 2003). Many of the herbicide-resistant plants that have been



developed fall into this category. Targeting variation to protein sequences is a routine process in bacterial research, but that is typically in the context of directed evolution, which is challenging to do in crops given their long cycle time. Greater advancement in this area remains to be made.

Novel domestication to enhance variation Targeted approaches can also be leveraged to rapidly domesticate new species. The domestication process has led to utilization of a relatively small group of plant species for agronomically and economically viable food, feed, fiber, and fuel options compared to the vast number of undomesticated species. Many of these undomesticated species carry readily apparent useful traits such as resistance to disease and harsh conditions (Fernie and Yan 2019). Additionally, there is no reason to assume that the crops that have thus far been domesticated represent ideal performance for various traits across environments. There could be many plant species that have a higher upward limit on productivity that have not yet been tested. Domestication genes tend to be shared across species, meaning that knowledge gained from domesticated crops can be applied to wild species through application of editing technologies.

The main efforts, to date, in directed crop domestication have been largely in Solanum species. Three recent studies focused on wild species related to tomatoes for targeting via editing. Two of the studies focused on Solanum pimpinellifolium, known for its stress resistance against both disease and environmental conditions. One research group focused on a single accession and targeted 6 sites for editing (Zsögön et al. 2018), while in the second study, another group focused on four accessions but only four target sites (Li et al. 2018b). Interestingly, there was only an overlap of a single gene targeted between the two studies: selfpruning, which modifies plant architecture. The unshared genes were known to affect yield, nutritional quality, and flowering time. The most impressive editing-based domestication research to date has been in ground cherry (Physalis pruinosa), which included the development of a transformation protocol and whole-genome sequencing, as well as the demonstration of editing of domestication genes (Lemmon et al. 2018). Ground cherry suffers from negative traits similar to S. pimppinellifolium: bushy growth, undesirable flowering time, and small/few fruits. The genes chosen for editing overlapped with the previous studies, targeting the SP, SP5G, and SICLV1 genes. Interestingly, the null alleles were found to be too strong and did not achieve the stated goals. Taken together, these three studies provide a roadmap for other plants to be domesticated using editingbased approaches.

Integration of exotic germplasm into existing breeding populations Related to the issues of rapid domestication are the challenges of integrating exotic germplasm into plant breeding programs. These exotic materials are useful because they can bring in novel genetic variation that can enhance valuable

traits; however, they also tend to bring many undesirable alleles through linkage (Wang *et al.* 2017). Enhancing recombination can assist with this; however, more direct approaches can be taken. If the specific genetic modification is identified, it can be immediately incorporated into more amenable germplasm or, in some cases, directly into elite lines through editing.

Targeted introduction of known genetic variation helps to address issues of scale in plant breeding. It might be reasonable to conduct physical crosses if one was working with a single pair of lines; however, the amount of work quickly escalates when trying to evaluate an allele within many elite genetic backgrounds. Furthermore, simultaneous introductions of alleles from many parents inevitably takes several generations and many plants to evaluate. In these cases, editing can be a more economical choice even when dealing with time intensive transformation protocols. Speed gains can also come from being able to instantly test and apply information from related species, broadening the base of available information for utilization in plant breeding applications.

Genome-wide recombination enhancement to increase variation Among broad spectrum approaches to increasing genetic variation, enhancing recombination is alluring because it ties in naturally with existing breeding programs. Native recombination is not uniform and does not occur frequently, which results in underutilization of the variation that breeders have available to them. This limits the ability to recombine material and in turn limits integration of exotic germplasm, prebreeding efforts, and arrival at ideal genotypes, each of which would be enhanced by increasing recombination rates (Taagen *et al.* 2020). Editing tools also enable the possibility of achieving targeted recombination; estimates of the potential gains from targeted recombination have been shown to range from 60 to 400% over nontargeted recombination, depending on the trait (Bernardo 2017; Brandariz and Bernardo 2019).

Many factors contribute to where and how frequently recombination occurs and manipulating these can enhance overall recombination rate in plants (Choi 2017; Fayos et al. 2019; Taagen et al. 2020). The most straightforward approach to enhancing recombination rates in plants is to modify genes involved in the recombination process, eliminating those that suppress the CO process and enhance expression of those that stimulate it (Taagen et al. 2020). The first approach was tested in rice, pea, and tomato through targeting orthologues of the FANCM, RECQ4, and FIGL1 genes originally identified in Arabidopsis, each of which are involved in the recombination process. Among those gene targets, RECQ4 proved to be the most impactful, with editing-mediated deletions of the gene resulting in a three-fold increase in number of COs (Mieulet et al. 2018). Editing-based alteration of these types of genes can also have far reaching impacts, however, such as the case where knockouts of FIGL1 resulted in sterility in tomato and



pea. Because COs are an important element of the meiotic process, interfering with the CO mechanism can result in sterility. This phenomenon was found to be true for another antagonist of COs, MEICA1 in rice. Knocking out the gene increased CO frequency but also resulted in high sterility rates (Hu *et al.* 2017).

Another route to broad recombination enhancement is through modification of the epigenetic landscape of a genome. In Arabidopsis, plants with mutations in the decreased DNA methylation 1 (ddm1) gene were found to have enhanced recombination in euchromatic regions (Melamed-Bessudo and Levy 2012; Yelina et al. 2012). Unfortunately, this did not impact heterochromatic pericentric regions, revealing additional recombination control mechanisms at work. Impacting these heterochromatic regions necessitated the dual targeting of both H3K9 methyltransferase genes and the non-GC methylation genes, which, in Arabidopsis, was sufficient to induce an increase in recombination in heterochromatic pericentric regions (Underwood et al. 2017). As with targeting of CO regulators, targeting epigenetic changes can have complicated whole genome impacts that impede their application to crops. In tomato, orthologs of the cytosine methylation gene, ddm1, were targeted for editing, which resulted in expected changed methylation patterns but also negative impacts on growth and development of the resulting plants (Corem et al. 2018). Further research will be necessary to make this a more viable approach for enhancement of recombination in crops.

Pooled guide approaches toward creation of genetic variation Some approaches toward genetic variation creation bridge the gap between targeted and broad approaches, applying specific knowledge at a broad scale. A recent editing-mediated technique utilized in generation of genetic variation has been to use pools of guide RNAs that are targeted to a wide number of genome sites. Successful guides can be identified through sequencing the resultant mutants, enabling sorting of results despite the pool being mixed. In tomato, a total of 72 genes were targeted via editing, resulting in 30 recovered mutants displaying aberrant phenotypes ranging from boron deficiency to modified leaves (Jacobs et al. 2017). The largest scale experiment, to date, has been in rice where all 39,045 rice genes were targeted for mutagenesis, resulting in generation of 91,004 mutants valuable for functional genomics research and breeding purposes (Lu et al. 2017; Meng et al. 2017). This pooled approach has also been successfully deployed in maize and soybean, each targeting novel QTLs for discovery purposes (Bai et al. 2020; Liu et al. 2020).

The strength of this combined approach is in being able to target more sites at once by eliminating the need for conducting many individual experiments. In a transformation-based editing pipeline, each distinct experiment needs to be conducted, tracked, and monitored separately. This costs, time, and resources, both through the physical

effort involved as well as the cognitive work of managing each experiment. These time and resource costs can be minimized in a pooled approach because only the events that resolve into a useful phenotype are traced back via sequencing to the genome sequence change that caused it. Even the labor of tracing could be eliminated in situations where the only goal is to affect the end product of performance; no sequencing needs to be done if it is not a goal to figure out the underlying genetic cause of the change in phenotype. This approach will be most powerful in situations where the expected percentages of "positives" is low since there will be the greatest time savings in avoiding negatives. If only 10% of experiments would result in a phenotypic positive, then 90% of overhead costs can be avoided through using a pooled approach.

Controlled nonspecific recombination enhancement as a means to increase variation Another strategy that falls into a middle path between targeted and broad approaches is targeting of specific genome regions for recombination enhancement. This has been demonstrated in plants through the use of somatic homologous recombination (HR), spo11 fusions, and epigenetic modification (Hayut et al. 2017; Sarno et al. 2017; Taagen et al. 2020). Somatic HR pathways take advantage of the same DSB repair pathway used by HDRmediated insertions, except to cause repair by the other chromatid. A DSB is induced by Cas9 which is then repaired by the sister chromatid. This approach has been demonstrated in tomato through targeting the phytoene synthase gene to cause reversion to wildtype red sectors in fruit (Hayut et al. 2017). To date, this approach has not yet been explored widely and further research is needed to define its value as a technique to enhance genetic variation.

A potentially exciting approach may be to target the initiators of the CO process directly to sites of interest. This has been tested in yeast where spo11, a conserved gene which generates the DSB at the beginning of CO, was targeted to different genetic regions by pairing it with dCas9 and found to enhance COs in otherwise "cold" regions of the genome (Sarno *et al.* 2017). This approach has not yet been reported in crops but could be useful to consider. Finally, epigenomic modification may be a future route toward controlled COs, given that methylation and histone modification impacts CO occurrence. This approach has yet to be combined with the current epigenetic editing tools that have been developed (Taagen *et al.* 2020).

Expanding the Crop Genetic Terrain into New Areas

Each of the previous examples of genetic variation creation work within the context of the existing breeding landscape



and pangenome. One informative way to think about the pangenome is as a multidimensional space. The various axes are different regions of the genome and moving along these axes equates to exploring the genetic variations within these regions. Setting these axes to a certain fixed state, for example, in a given crop variety, results in informative outputs: yield, pest resistance, drought hardiness, nutritional value, etc. This is a useful metaphor because it reframes the task of breeding to one of exploration of the variation-space that a crop contains. Areas of depleted variation become unexplored genetic terrain. Editing technologies will allow both understanding and rapid shifting of this landscape, resulting in greater variation and diversity.

One impact of this landscape is that allelic effects depend on the background they are in, both due to epistatic interactions and the combination of many small-effect alleles. Different genetic backgrounds occupy different regions of genetic space and few changes to these spaces will have uniform outcomes. This is especially true of resistance alleles. For example, an analysis of 19 crop QTL studies on resistance alleles found that a majority of them displayed epistatic interactions with other alleles, resulting in QTLs performing differently across different backgrounds (Gallois et al. 2018). In maize, it was observed that flowering time alleles can vary in their impact depending on the genetic background. Specifically, the MADS69 QTL was shown to have a large impact in dent backgrounds but small effect in flint backgrounds (Rio et al. 2020). Additionally in maize, the impact of genes involved in the hypersensitive response has been shown to be dependent on the genetic background of the plant (Chaikam et al. 2011). Manipulating and understanding these background effects are empowered by editing techniques which enable alleles to be tested in more diverse backgrounds than would be otherwise feasible (Gao et al. 2020).

Traditional breeding approaches toward understanding and leveraging QTLs are further hampered by linkage drag and the unfavorable statistics of multiple segregating alleles. Linkage drag causes unwanted alleles to accompany the QTL of interest, obscuring its actual impact (Zhou et al. 2019). Bringing together multiple favorable QTLs into a single line is labor intensive: combining a favorable allele from 6 parents would take at least 3 generations in a diploid crop, ignoring potential time it would take to segregate away unfavorable alleles. Editing bypasses each of these problems as it is precise, can be completed in fewer generations, and can be deployed in many backgrounds at once. This has been illustrated in rice where three yield-related QTLs were targeted in elite rice varieties (Zhou et al. 2019). By the T₂ generation, a wide variety of single to triple mutants had been generated and confirmed to be transgene free. Gains of 68% and 30% were made in yield per panicle in the two lines studied most thoroughly, J809 and L237, respectively. This difference in impact was also found in the single mutants, with GS3 mutants having a greater effect on grain length in J809 compared to L237. An equivalent study using non-editing approaches would have taken many more resources and years to conduct.

Establishment and manipulation of large effect genes Within variation-space, axes are not uniformly impactful. Those genes with a disproportionate effect on trait performance, called large effect genes (LEGs), can have a very strong impact on the traits and performance of a crop. Domestication genes are generally of this type, as has been found in many crops (Meyer and Purugganan 2013). These genes radically change a species to make it amenable to human cultivation, and commonly are shared across many crop species such as the Shattering1 (Sh1) gene, which is shared as a domestication target across rice, maize, and sorghum (Lin et al. 2012). Modern breeding efforts have also generated LEGs. For example, rht in wheat and sd1 in rice cause semi-dwarf phenotypes and were one of the innovations of the Green Revolution (Hedden 2003). Another example are the brown midrib genes in grasses, which cause an alteration in lignin content, making them much more digestible by ruminants (Barrière and Argillier 1993).

Editing can be used both to modify existing LEGs as well as create new ones. Many domestication genes and LEGs have been subject to fixation due to selection, resulting in a depletion of variation (Meyer and Purugganan 2013). It would be naive to assume that most of these genes exist in their ideal configuration given that they were selected for during a period of agriculture that looked vastly different than today. Even for non-fixed genes, it has been demonstrated that alleles superior to naturally occurring ones can be manufactured. Zmm28 is an AP1-FUL MADS-box gene in maize which, when expressed under the moderately constitutive gos2 promoter, had higher grain yield compared to wild type (Wu et al. 2019). Modification of these selected regions could be done by first looking for existing variation found in wild relatives which have not undergone selection and re-introduction of that variation through targeted editing or recombination-enhanced breeding. Alternatively, novel variation could be generated through using targeted mutagenesis tools.

Transfer of genetic enhancement across species borders In addition to LEGs, evolution has led to the development of many processes of potential value, but that are isolated within selected species. Nitrogen fixation and photosynthesis optimization in particular stand out as targets to make large-scale gains in agriculture. Theoretically these traits could be selected for using traditional breeding approaches, but significant challenges mean that achieving the desired breeding outcome would be an extremely lengthy process. Specifically, the number of genetic changes that would be necessary, paired with the lack of needed sequence variation to make those changes, makes those traits very difficult to address through traditional



breeding approaches. For instance, photosynthetic conversion efficiency has improved in soybean over the last 80 yr, but that is only under well irrigated conditions and has not come with improvements in other aspects of photosynthetic efficiency (Koester *et al.* 2016). Editing would enable the direct change of genetic factors underlying these processes in less time and more broadly, as findings in one species could be applied to another.

Introduction of nitrogen fixation capabilities into nonnitrogen fixing crops would be a significant achievement for modern agriculture that is currently dependent on added nitrogen, largely derived from the energy-intensive Haber-Bosch process. This could theoretically be achieved through either engineering the legume symbiosis pathway or the nitrogenase-related genes directly into plants (Oldroyd and Dixon 2014). Equally significant, would be improvements in the photosynthesis machinery as it would mean more available energy for yield. The available routes that could be taken are very diverse, but one of the main approaches investigated is the introduction C4 photosynthesis into C3 crops such as wheat (Kubis and Bar-Even 2019). Both nitrogen-fixation and photosynthetic enhancements are active areas of research, though not yet at the implementation phase. Either of these developments will only be made possible through modern editing technologies as existing pathways will need to be eliminated and modified to accommodate the new ones. Altering metabolite flux will require fine tuning through novel variation creation in contributing genes and their promoters. The process will not be as simple as transforming in new genes because these processes interact with existing regulatory pathways and product stoichiometry can be critical.

Evolutionary history can give us more insights into potentially worthwhile breeding goals toward which to apply editing-based technologies. Important factors in plant evolution have been the twin forces of whole genome duplication (WGD) and tandem duplication (TD). Both introduce new genes into the genome which are subsequently selected for deletion, subfunctionalization, and, most relevantly, neofunctionalization. This is the process by which genes take on new functions, which has been linked to plant adaptations to new environments as well as many agronomic traits (Lyu et al. 2020). Replicating these processes could be an avenue of future advancement in crops, either by creating TDs through insertion or through synthetic chromosome engineering. Continued development in the power of editing technologies will be needed to make this a reality, however.

Variation creation via exploration of insignificant data A currently underexplored application of editing for variation creation is in leveraging "insignificant" data sources. QTL analysis, GWAS, gene regulatory network (GRN) inference, and other bioinformatic analysis techniques all rely on defining

cutoffs of statistical "significance" based on acceptable values of error. QTL and GWAS studies generally begin with a standard p-value of 0.05 which is then further modified to account for multiple comparison, population structure, and recombination rates (Liu and Yan 2019). Given the large effort involved in proving a single QTL as causative of variation in a trait, it makes sense to have high standards in what is accepted as significant. However, this approach removes from consideration many potential QTLs that may have been important but did not meet the necessary level of significance. GRN inference suffers from similar issues in trying to balance precision and accuracy in reporting of interactions. In maize, the GENIE3 (GEne Network INference with Ensemble of trees) genetic network mapping tool was used to create a gene network from protein and mRNA expression data, resulting in an area under the receiver operating characteristic (ROC) curve of 0.717 (Walley et al. 2016). Said another way, the network would predict 50% of true-positives with a false-positive rate of 20%. Even single cell expression datasets have challenges in the inference task despite the high density of their data (Chen and Mar 2018). As of yet, GRNs broadly have not been widely used for improvement of crops due to these challenges.

The result of this cut-off approach is that the most impactful genes with sufficient variation are the ones that receive investigation. There is still useful information that is of "insignificant" quality within these data that could be useful in approaches similar to Liu *et al.* which utilized pooled guides to target vast numbers of genomic sites for editing. In QTL and GWAS type data, this would be to directly generate variation within low-significance targets. In GRNs, this would be used in targeting the "spokes" of known important hub genes. This has not yet been reported in crops, but as the technological ability to perform editing at large scales expands, so too will be the acceptable range of error.

The Future of Editing Technologies

While gene editing technologies have advanced rapidly, there is still room for improvement. Principally, while there are many tools for making targeted genetic modifications, it is still difficult to make modifications at a large scale and target many sequences at once. One component of this bottleneck is that, currently, genome editing in crops relies on crop transformation systems. Many crops are still challenging to transform, and current transformation protocols tend to be genotype-dependent, inefficient, expensive, time-consuming, and/or subject to regulatory or intellectual property issues (Altpeter *et al.* 2016). Development of improved transformation systems that overcome current limitations, such as morphogenic gene-based systems, meristem-based systems, and nanoparticle-mediated protocols, will help in expanding the use, and reducing the time and cost, of editing-based genetic



study and improvement of crops (Altpeter *et al.* 2016; Kausch *et al.* 2019; Nadakuduti and Enciso-rodríguez 2021). This should enable application of gene editing technologies at a much larger scale to more species and genotypes.

The other side of this bottleneck is that the solutions for deploying large numbers of guides are still under-explored in plants. Pooled-guide approaches have been successful but have featured a low number of guides per construct. This fits the objective of the studies in identifying causative QTLs. However, in cases where phenotypic gains are more important than basic genetic research, it would be advantageous to launch multiplex guide cassettes to multiply the number of variation sites. A promising lead in this area is work with Cas12 paired with utilization of a pol II promoter to launch a large number of guides, though the approach has not yet been deployed to specifically generate novel variation (Campa *et al.* 2019). Combined with different molecular machinery that can create diverse forms of variation, this could be a very powerful approach in the future.

Even in the era of big data, information gaps abound

Advancements in molecular tools need to be paired with greater information to be useful. Although this is considered the era of "big data," there are still challenges with identifying actionable biological information from that data, beyond simply large effects. This is in part due to the magnitude of questions asked of the data. The genomes of plant species each contain tens of thousands of genes, which in turn can have hundreds of variations across the pangenome, each of which modifying developmental pathways by incremental amounts. There is a near infinite developmental and genetic space that scientists are trying to describe with big data approaches, which results in only the most impactful rising to the top. Another complicating issue is the current difficulties in integrating diverse datasets. There are many ways that one can, for instance, attempt to identify the location and size of the promoter sequence of a gene. This is relevant to know to be able to deploy editing components to the correct location; however, deciding whether to use ATAC-seq or DNase-seq or even how to integrate both is not presently clear. Supporting these objectives will require the continued advancement of sequencing methods to further reduce costs. It will also require the development and maintenance of standard datasets against which aggregate models can be evaluated.

In addition to the considerations raised above is the necessary sustained development of trust between the creators of molecular technologies and those people affected by them. While some editing technologies have been deemed unregulated in the USA as well as many South American countries, they have been mostly prohibited in the European Union and beyond. If these technologies are to be used to address current and growing global agronomic needs, there continues to be the need to both develop safe technology as well as

communicate the benefits of this technology. This cannot be couched solely in discussion of facts, as is the tendency, as this has been proven ineffective (Kahan *et al.* 2012). Instead, it needs to come from a place of shared values and goals and knowledge of how editing and other molecular techniques can meet crop improvement needs and benefit society.

Genome editing is a very powerful technique that can be used to bring much needed genetic variation into plant breeding programs and has already delivered many gains. Sustained development in editing technologies should allow advancements beyond mainly subtractive processes into *de novo* creation of novel diversity that can be leveraged to improve crop yield, quality, and abiotic and biotic stress tolerances, and reduce the environmental footprint of crop production.

Acknowledgements This work was partially supported by grants from the National Science Foundation (IOS-1917138 and IOS-1844701).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aamir M, Ahmad M, Kumar V, Meng X, Yu H, Zhang Y, Zhuang F, Song X, Gao SG, Caixa G, Li J (2020) The era of editing plant genomes using CRISPR/Cas: a critical appraisal. J Biotechnol 324:34–60. https://doi.org/10.1016/j.jbiotec.2020.09.013

Altpeter F, Springer NM, Bartley LE, Blechl AE, Brutnell TP, Citovsky V, Conrad LJ, Gelvin SB, Jackson DP, Kausch AP, Lemaux PG, Medford JI, Orozco-Cárdenas ML, Tricoli DM, Van Eck J, Voytas DF, Walbot V, Wang K, Zhang ZJ, Neal Stewart C (2016) Advancing crop transformation in the era of genome editing. Plant Cell 28:1510–1520. https://doi.org/10.1105/tpc.16.00196

Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 576:149–157. https://doi.org/10.1038/s41586-019-1711-4

Arora L, Narula A (2017) Gene editing and crop improvement using CRISPR-cas9 system. Front Plant Sci 8:1932. https://doi.org/10.3389/fpls.2017.01932

Bai M, Yuan J, Kuang H, Gong P, Li S, Zhang Z, Liu B, Sun J, Yang M, Yang L, Wang D, Song S, Guan Y (2020) Generation of a multiplex mutagenesis population via pooled CRISPR-Cas9 in soya bean. Plant Biotechnol J 18:721–731. https://doi.org/10.1111/pbi.13239

Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF (2014) DNA replicons for plant genome engineering. Plant Cell 26:151–163. https://doi.org/10.1105/tpc.113.119792



Barrière Y, Argillier O (1993) Brown-midrib genes of maize. EDP Sci 13: 865–876

- Bauer E, Falque M, Walter H, Bauland C, Camisan C, Campo L, Meyer N, Ranc N, Rincent R, Schipprack W, Altmann T, Flament P, Melchinger AE, Menz M, Moreno-gonzález J, Ouzunova M, Revilla P, Charcosset A, Martin OC, Schön C (2013) Intraspecific variation of recombination rate in maize. Genome Biol 14(9):1–17. https://doi.org/10.1186/gb-2013-14-9-r103
- Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG, Tan W, Penheiter SG, Ma AC, Leung AYH, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC (2012) In vivo genome editing using a high-efficiency TALEN system. Nature 491:114–118. https://doi.org/10.1038/nature11537
- Begemann M, Gray B, January E, Singer A, Kesler D, He Y, Liu H, Guo H, Jordan A, Brutnell T, Mockler T, Oufattole M (2017) Characterization and validation of a novel group of type V, class 2 nucleases for in vivo genome editing. bioRxiv 192799. https://doi.org/10.1101/192799
- Bernardo R (2017) Prospective targeted recombination and genetic gains for quantitative traits in maize. Plant Genome 10:1–9. https://doi.org/10.3835/plantgenome2016.11.0118
- Blary A, Jenczewski E (2019) Manipulation of crossover frequency and distribution for plant breeding. Theor Appl Genet 132:575–592. https://doi.org/10.1007/s00122-018-3240-1
- Boyle EA, Li YI, Pritchard JK (2017) An expanded view of complex traits: from polygenic to omnigenic. Cell. 169:1177–1186. https:// doi.org/10.1016/j.cell.2017.05.038
- Brandariz SP, Bernardo R (2019) Predicted genetic gains from targeted recombination in elite biparental maize populations. Plant Genome 12:180062. https://doi.org/10.3835/plantgenome2018.08.0062
- Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harje C, McMullen MD (2009) The genetic architecture of maize flowering time. Science. 325:714–718. https://doi.org/10.1534/genetics.113.159152
- Butt H, Eid A, Momin AA, Bazin J, Crespi M, Arold ST, Mahfouz MM (2019) CRISPR directed evolution of the spliceosome for resistance to splicing inhibitors. Genome Biol 20:1–9. https://doi.org/10.1186/ s13059-019-1680-9
- Campa CC, Weisbach NR, Santinha AJ, Incarnato D, Platt RJ (2019) Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. Nat Methods 16:887–893. https:// doi.org/10.1038/s41592-019-0508-6
- Carroll D (2008) Progress and prospects: zinc-finger nucleases as gene therapy agents. Gene Ther 15:1463–1468. https://doi.org/10.1038/ gt.2008.145
- Castelletti S, Tuberosa R, Pindo M, Salvi S (2014) A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL Vgt. Genes Genomes Genet 4:805–812. https://doi.org/10.1534/g3.114.010686
- Ceccaldi R, Rondinelli B, D'andrea AD (2016) Repair pathway choices and consequences at the double-strand break mechanisms of DNA DSB repair. Trends Biochem. Sci. 26:52–64. https://doi.org/10.1016/j.tcb.2015.07.009.Repair
- Chaikam V, Negeri A, Dhawan R, Puchaka B, Ji J, Chintamanani S, Gachomo EW, Zillmer A, Doran T, Weil C, Balint-Kurti P, Johal G (2011) Use of mutant-assisted gene identification and characterization (MAGIC) to identify novel genetic loci that modify the maize hypersensitive response. Theor Appl Genet 123:985–997. https://doi.org/10.1007/s00122-011-1641-5
- Chen K, Wang Y, Zhang R, Zhang H, Gao C (2019) CRISPR/Cas genome editing and precision plant breeding in agriculture. Annu Rev Plant Biol 70:667–700
- Chen Q, Samayoa LF, Yang CJ, Bradbury PJ, Olukolu BA, Neumeyer MA, Romay MC, Sun Q, Lorant A, Buckler ES, Ross-Ibarra J, Holland JB, Doebley JF (2020) The genetic architecture of the maize

- progenitor, teosinte, and how it was altered during maize domestication. PLoS Genet 16:1–21. https://doi.org/10.1371/journal.pgen.
- Chen S, Mar JC (2018) Evaluating methods of inferring gene regulatory networks highlights their lack of performance for single cell gene expression data. BMC Bioinform 19:1–21. https://doi.org/10.1186/s12859-018-2217-z
- Choi K (2017) Advances towards controlling meiotic recombination for plant breeding. Mol Cell 40:814–822. https://doi.org/10.14348/ molcells.2017.0171
- Christine M, Jenczewski E, Macaisne N, Grelon M (2015) The molecular biology of meiosis in plants. Annu Rev Plant Biol 66:297–327. https://doi.org/10.1146/annurev-arplant-050213-035923
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. Science. 80:819–824. https:// doi.org/10.1126/science.1231143
- Corem S, Doron-Faigenboim A, Jouffroy O, Maumus F, Arazi T, Bouché N (2018) Redistribution of CHH methylation and small interfering RNAs across the genome of tomato ddm1 mutants. Plant Cell 30: 1628–1644. https://doi.org/10.1105/tpc.18.00167
- Cutter AD, Payseur BA (2013) Genomic signatures of selection at linked sites: unifying the disparity among species. Nat Rev Genet 14:262– 274. https://doi.org/10.1038/nrg3425
- Ellison EE, Nagalakshmi U, Gamo ME, Huang PJ, Dinesh-Kumar S, Voytas DF (2020) Multiplexed heritable gene editing using RNA viruses and mobile single guide RNAs. Nature Plants 6:620–624. https://doi.org/10.1038/s41477-020-0670-y
- Endo M, Mikami M, Toki S (2016) Biallelic gene targeting in rice. Plant Physiol 170:667–677. https://doi.org/10.1104/pp.15.01663
- Even-Faitelson L, Samach A, Melamed-Bessudo C, Avivi-Ragolsky N, Levy AA (2011) Localized egg-cell expression of effector proteins for targeted modification of the Arabidopsis genome. Plant J 68: 929–937. https://doi.org/10.1111/j.1365-313X.2011.04741.x
- Fayos I, Mieulet D, Petit J, Meunier AC, Périn C, Nicolas A, Guiderdoni E (2019) Engineering meiotic recombination pathways in rice. Plant Biotechnol J 17:2062–2077. https://doi.org/10.1111/pbi.13189
- Fernie AR, Yan J (2019) De novo domestication: an alternative route toward new crops for the future. Mol Plant 12:615–631. https://doi.org/10.1016/j.molp.2019.03.016
- Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31(9):822–826. https://doi.org/10.1038/nbt.262
- Gabur I, Chawla HS, Snowdon RJ, Parkin IAP (2019) Connecting genome structural variation with complex traits in crop plants. Theor Appl Genet 132:733–750. https://doi.org/10.1007/s00122-018-3233-0
- Gallois J, Moury B, German-retana S (2018) Role of the genetic background in resistance to plant viruses. Int J Mol Sci 19(10):2856. https://doi.org/10.3390/ijms19102856
- Gao H, Gadlage MJ, Lafitte HR, Lenderts B, Yang M, Schroder M, Farrell J, Snopek K, Peterson D, Feigenbutz L, Jones S, St Clair G, Rahe M, Sanyour-Doyel N, Peng C, Wang L, Young JK, Beatty M, Dahlke B, Hazebroek J, Greene TW, Cigan AM, Chilcoat ND, Meeley RB (2020) Superior field performance of waxy corn engineered using CRISPR-Cas9. Nat Biotechnol 38: 579-581. https://doi.org/10.1038/s41587-020-0444-0
- Gao L, Cox DBT, Yan WX, Manteiga JC, Schneider MW, Yamano T, Nishimasu H, Nureki O, Crosetto N, Zhang F (2017) Engineered Cpf1 variants with altered PAM specificities. Nat Biotechnol 35: 789–792. https://doi.org/10.1038/nbt.3900
- Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR (2017) Programmable base editing of A T to G C in genomic DNA without DNA cleavage. Nature. 551:464–471. https://doi.org/ 10.1038/nature24644



- Hahn F, Eisenhut M, Mantegazza O, Weber APM (2018) Homologydirected repair of a defective glabrous gene in Arabidopsis with cas9-based gene targeting. Front Plant Sci 9:1-13. https://doi.org/ 10.3389/fpls.2018.00424
- Halperin SO, Tou CJ, Wong EB, Modavi C, Schaffer DV, Dueber JE (2018) CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. Nature. 560:248-252. https:// doi.org/10.1038/s41586-018-0384-8
- Hayut SF, Bessudo CM, Levy AA (2017) Targeted recombination between homologous chromosomes for precise breeding in tomato. Nat Commun 8:1-9, https://doi.org/10.1038/ncomms15605
- Hedden P (2003) The genes of the Green Revolution. Trends Genet 19:5-9. https://doi.org/10.1016/S0168-9525(02)00009-4
- Hill CB, Li C (2016) Genetic architecture of flowering phenology in cereals and opportunities for crop improvement. Front Plant Sci 7: 1-23. https://doi.org/10.3389/fpls.2016.01906
- Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X, Rees HA, Lin Z, Liu DR (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature. 556: 57-63. https://doi.org/10.1038/nature26155
- Hu Q, Li Y, Wang H, Shen Y, Zhang C, Du G, Tang D, Cheng Z (2017) Meiotic chromosome association 1 interacts with TOP3α and regulates meiotic recombination in rice. Plant Cell 29:1697-1708. https://doi.org/10.1105/tpc.17.00241
- Hua K, Tao X, Zhu JK (2019) Expanding the base editing scope in rice by using Cas9 variants. Plant Biotechnol J 17:499-504. https://doi.org/ 10.1111/pbi.12993
- Huang TK, Puchta H (2019) CRISPR/Cas-mediated gene targeting in plants: finally a turn for the better for homologous recombination. Plant Cell Rep 38:443-453. https://doi.org/10.1007/s00299-019-02379-0
- Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C, Fan D, Lu Y, Weng Q, Liu K, Zhou T, Jing Y, Si L, Dong G, Huang T, Lu T, Feng Q, Qian Q, Li J, Han B (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet 44:32–39. https://doi. org/10.1038/ng.1018
- Hufford MB, Xu X, Van Heerwaarden J, Pyhäjärvi T, Chia JM, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaeppler SM, Lai J, Morrell PL, Shannon LM, Song C, Springer NM, Swanson-Wagner RA, Tiffin P, Wang J, Zhang G, Doebley J, McMullen MD, Ware D, Buckler ES, Yang S, Ross-Ibarra J (2012) Comparative population genomics of maize domestication and improvement. Nat Genet 44:808-811. https://doi.org/10.1038/ng.2309
- Jacobs TB, Zhang N, Patel D, Martin GB (2017) Generation of a collection of mutant tomato lines using pooled CRISPR libraries. Plant Physiol 174:2023–2037. https://doi.org/10.1104/pp.17.00489
- Jannink J (2010) Dynamics of long-term genomic selection. Genetics Select. Evol. 42(1):1-11. https://doi.org/10.1186/1297-9686-42-35
- Jaqueth JS, Hou Z, Zheng P, Ren R, Nagel BA, Cutter G, Niu X, Vollbrecht E, Greene TW, Kumpatla SP (2020) Fertility restoration of maize CMS-C altered by a single amino acid substitution within the Rf4 bHLH transcription factor. Plant J 101:101–111. https://doi. org/10.1111/tpj.14521
- Jiang F, Doudna JA (2017) CRISPR Cas9 structures and mechanisms. Annu Rev Biophys 46:505-529. https://doi.org/10.1146/annurevbiophys-062215-010822
- Jiang Y-Y, Chai Y-P, Lu M-H, Han X-L, Lin Q, Zhang Y, Zhang Q, Zhou Y, Wang X-C, Gao C, Chen Q-J (2020) Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. Genome Biol 212(1):1–10. https://doi.org/10.1101/ 2020.07.06.188896
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 337:816-822. https://doi.org/ 10.1126/science.1225829

- Kahan DM, Peters E, Wittlin M, Slovic P, Ouellette LL, Braman D, Mandel G (2012) The polarizing impact of science literacy and numeracy on perceived climate change risks. Nat Clim Chang 2: 732–735. https://doi.org/10.1038/nclimate1547
- Kang BC, Yun JY, Kim ST, Shin YJ, Ryu J, Choi M, Woo JW, Kim JS (2018) Precision genome engineering through adenine base editing in plants. Nat Plants 4:427-431. https://doi.org/10.1038/s41477-018-0178-x
- Kausch AP, Nelson-vasilchik K, Hague J, Mookkan M, Quemada H, Dellaporta S, Fragoso C, Zhang ZJ (2019) Edit at will: genotype independent plant transformation in the era of advanced genomics and genome editing. Plant Sci 281:186-205. https://doi.org/10. 1016/j.plantsci.2019.01.006
- Kelliher T, Starr D, Su X, Tang G, Chen Z, Carter J, Wittich PE, Dong S, Green J, Burch E, McCuiston J, Gu W, Sun Y, Strebe T, Roberts J, Bate NJ, Que Q (2019) One-step genome editing of elite crop germplasm during haploid induction. Nat Biotechnol 37:287-292. https://doi.org/10.1038/s41587-019-0038-x
- Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales APW, Li Z, Peterson RT, Yeh JRJ, Aryee MJ, Joung JK (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 523:481-485. https://doi.org/10.1038/nature14592
- Koester R, Nohl B, Diers B, Ainsworth E (2016) Has photosynthetic capacity increased with 80 years of soybean breeding? An examination of historical soybean cultivars. Plant Cell Environ 39:1058-1067. https://doi.org/10.1111/pce.12675
- Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 533:420-424. https://doi. org/10.1038/nature17946
- Kosicki M, Tomberg K, Bradley A (2018) Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. Nat Biotechnol 36(8):765-771. https://doi.org/10. 1038/nbt.4192
- Kozak M (2001) Constraints on reinitiation of translation in mammals. Nucleic Acids Res 29:5226-5232. https://doi.org/10.1093/nar/29. 24.5226
- Kozak M (2002) Pushing the limits of the scanning mechanism for initiation of translation. Gene. 299:1-34. https://doi.org/10.1016/S0378-1119(02)01056-9
- Kuang Y, Li S, Ren B, Yan F, Spetz C, Li X, Zhou X, Zhou H (2020) Base-editing-mediated artificial evolution of OsALS1 in planta to develop novel herbicide-tolerant rice germplasms. Mol Plant 13: 565-572. https://doi.org/10.1016/j.molp.2020.01.010
- Kubis A, Bar-Even A (2019) Synthetic biology approaches for improving photosynthesis. J Exp Bot 70:1425-1433. https://doi.org/10.1093/ jxb/erz029
- Lemmon ZH, Reem NT, Dalrymple J, Soyk S, Swartwood KE, Rodriguez-Leal D, Van Eck J, Lippman ZB (2018) Rapid improvement of domestication traits in an orphan crop by genome editing. Nat Plants 4:766–770. https://doi.org/10.1038/s41477-018-0259-x
- $Li\ C,\ Zhang\ R,\ Meng\ X,\ Chen\ S,\ Zong\ Y,\ Lu\ C,\ Qiu\ JL,\ Chen\ YH,\ Li\ J,$ Gao C (2020a) Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. Nat Biotechnol 38:875-882. https://doi.org/10.1038/s41587-019-0393-7
- Li C, Zong Y, Jin S, Zhu H, Lin D, Li S, Qiu JL, Wang Y, Gao C (2020b) SWISS: multiplexed orthogonal genome editing in plants with a Cas9 nickase and engineered CRISPR RNA scaffolds. Genome Biol 21:1-15. https://doi.org/10.1186/s13059-020-02051-x
- Li F, Xie J, Zhu X, Wang X, Zhao Y, Ma X, Zhang Z, Rashid MAR, Zhang Z, Zhi L, Zhang S, Li J, Li Z, Zhang H (2018a) Genetic basis underlying correlations among growth duration and yield traits revealed by GWAS in rice (Oryza sativa L.). Front Plant Sci 9:1-13. https://doi.org/10.3389/fpls.2018.00650
- Li J, Meng X, Zong Y, Chen K, Zhang H, Liu J, Li J, Gao C (2016) Gene replacements and insertions in rice by intron targeting using





CRISPR-Cas9. Nat Plants 2:1–6. https://doi.org/10.1038/nplants. 2016.139

- Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013) Multiplex and homologous recombination mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat Biotechnol 31:8–11. https://doi.org/10.1038/nbt.2650
- Li S, Zhang X, Wang W, Guo X, Wu Z, Du W, Zhao Y, Xia L (2018) Expanding the scope of CRISPR/Cpf1-mediated genome editing in rice. Mol Plant 11:995–998. https://doi.org/10.1016/j.molp.2018. 03.009
- Li T, Yang X, Yu Y, Si X, Zhai X, Zhang H, Dong W, Gao C, Xu C (2018b) Domestication of wild tomato is accelerated by genome editing. Nat Biotechnol 36:1160–1163. https://doi.org/10.1038/nbt. 4273
- Lin Q, Zong Y, Xue C, Wang S, Jin S, Zhu Z, Wang Y, Anzalone AV, Raguram A, Doman JL, Liu DR, Gao C (2020) Prime genome editing in rice and wheat. Nat Biotechnol 38:582–585. https://doi. org/10.1038/s41587-020-0455-x
- Lin Q, Jin S, Zong Y, Yu H, Zhu Z, Liu G, Kou L, Wang Y, Qiu JL, Li J, Gao C (2021) High-efficiency prime editing with optimized, paired pegRNAs in plants. Nat Biotechnol 25:1–5. https://doi.org/10.1038/s41587-021-00868-w
- Lin Z, Li X, Shannon LM, Yeh CT, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, Doebley J, Schnable PS, Tuinstra MR, Tesso TT, White F, Yu J (2012) Parallel domestication of the Shattering1 genes in cereals. Nat Genet 44:720–724. https://doi.org/10.1038/ng.2281
- Liu HJ, Jian L, Xu J, Zhang Q, Zhang M, Jin M, Peng Y, Yan J, Han B, Liu J, Gao F (2020) High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize. Plant Cell 32(5):1397– 1413. https://doi.org/10.1105/tpc.19.00934
- Liu HJ, Yan J (2019) Crop genome-wide association study: a harvest of biological relevance. The Plant J 97(1):8–18. https://doi.org/10.1111/tpj.14139
- Lu Y, Tian Y, Shen R, Yao Q, Wang M, Chen M, Dong J, Zhang T, Li F, Lei M, Zhu JK (2020) Targeted, efficient sequence insertion and replacement in rice. Nat Biotechnol 38:1402–1407. https://doi.org/ 10.1038/s41587-020-0581-5
- Lu Y, Tian Y, Shen R, Yao Q, Zhong D, Zhang X, Zhu JK (2021) Precise genome modification in tomato using an improved prime editing system. Plant Biotechnol J 19(3):415–417. https://doi.org/10.1111/ pbi.13497
- Lu Y, Ye X, Guo R, Huang J, Wang W, Tang J, Tan L, Zhu JK, Chu C, Qian Y (2017) Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system. Mol Plant 10:1242–1245. https://doi.org/10. 1016/j.molp.2017.06.007
- Lyu J, Huang L, Zhang S, Zhang Y, He W, Zeng P, Zeng Y, Huang G, Zhang J, Ning M, Bao Y, Zhao S, Fu Q, Wade LJ, Chen H, Wang W, Hu F (2020) Neo-functionalization of a Teosinte branched 1 homologue mediates adaptations of upland rice. Nat Commun 11: 1–13. https://doi.org/10.1038/s41467-019-14264-1
- Maccaferri M, Harris NS, Twardziok SO, Pasam RK, Gundlach H, Spannagl M, Ormanbekova D, Lux T, Prade VM, Milner SG, Himmelbach A, Mascher M, Bagnaresi P, Faccioli P, Cozzi P, Lauria M, Lazzari B, Stella A, Manconi A, Gnocchi M, Moscatelli M, Avni R, Deek J, Biyiklioglu S, Frascaroli E, Corneti S, Salvi S, Sonnante G, Desiderio F, Marè C, Crosatti C, Mica E, Özkan H, Kilian B, De Vita P, Marone D, Joukhadar R, Mazzucotelli E, Nigro D, Gadaleta A, Chao S, Faris JD, Melo ATO, Pumphrey M, Pecchioni N, Milanesi L, Wiebe K, Ens J, MacLachlan RP, Clarke JM, Sharpe AG, Koh CS, Liang KYH, Taylor GJ, Knox R, Budak H, Mastrangelo AM, Xu SS, Stein N, Hale I, Distelfeld A, Hayden MJ, Tuberosa R, Walkowiak S, Mayer KFX, Ceriotti A, Pozniak CJ, Cattivelli L (2019) Durum wheat genome highlights past domestication signatures and future improvement targets. Nat Genet 51: 885–895. https://doi.org/10.1038/s41588-019-0381-3

- Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM (2013a) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31:833–838. https://doi.org/10.1038/pbt.2675
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013b) RNA-guided human genome engineering via Cas9. Science. 339:823–827. https://doi.org/10.1126/science. 1232033
- Mazaheri M, Heckwolf M, Vaillancourt B, Gage JL, Burdo B, Heckwolf S, Barry K, Lipzen A, Ribeiro CB, Kono TJY, Kaeppler HF, Spalding EP, Hirsch CN, Robin Buell C, de Leon N, Kaeppler SM (2019) Genome-wide association analysis of stalk biomass and anatomical traits in maize. BMC Plant Biol 19:1–17. https://doi.org/10.1186/s12870-019-1653-x
- Melamed-Bessudo C, Levy AA (2012) Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in Arabidopsis. Proc Natl Acad Sci U S A 109(16):E981–E988. https://doi.org/10.1073/pnas.1120742109
- Meng X, Yu H, Zhang Y, Zhuang F, Song X, Gao S, Gao C, Li J (2017) Construction of a genome-wide mutant library in rice using CRISPR/Cas9. Mol Plant 10:1238–1241. https://doi.org/10.1016/j. molp.2017.06.006
- Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. Nat Rev Genet 14:840–852. https://doi.org/10.1038/nrg3605
- Miao C, Yang J, Schnable JC (2019) Optimising the identification of causal variants across varying genetic architectures in crops. Plant Biotechnol J 17:893–905. https://doi.org/10.1111/pbi.13023
- Mieulet D, Aubert G, Bres C, Klein A, Droc G, Vieille E, Rond-Coissieux C, Sanchez M, Dalmais M, Mauxion JP, Rothan C, Guiderdoni E, Mercier R (2018) Unleashing meiotic crossovers in crops. Nat Plants 4:1010–1016. https://doi.org/10.1073/pnas.1713078114
- Nadakuduti SS, Enciso-rodríguez F (2021) Advances in genome editing with CRISPR systems and transformation technologies for plant DNA manipulation. Front. Plant Sci 11:1–9. https://doi.org/10.3389/fpls.2020.637159
- Nei M, Suzuki Y, Nozawa M (2010) The neutral theory of molecular evolution in the genomic era. Annu Rev Genomics Hum Genet 11: 265–289. https://doi.org/10.1146/annurev-genom-082908-150129
- Nekrasov V, Staskawicz B, Weigel D, Jones JDG, Kamoun S (2013) Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat Biotechnol. 31:691– 693. https://doi.org/10.1038/nbt.2654
- Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, Noda T, Abudayyeh OO, Gootenberg JS, Mori H, Oura S, Holmes B, Tanaka M, Seki M, Hirano H, Aburatani H, Ishitani R, Ikawa M, Yachie N, Zhang F, Nureki O (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science. 361:1259–1262. https://doi.org/10.1126/science.aas9129(2018)
- Oldroyd GED, Dixon R (2014) Biotechnological solutions to the nitrogen problem. Curr Opin Biotechnol 26:19–24. https://doi.org/10.1016/j.copbio.2013.08.006
- Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M (2010) The rate and molecular spectrum of spontaneous mutations in Arabidopsis thaliana. Science. 327:92–94. https://doi.org/10.1126/science.1180677
- Pausch P, Al-Shayeb B, Bisom-Rapp E, Tsuchida C, Li Z, Cress B, Knott G, Jacobsen S, Banfield J, Doudna J (2020) Crispr-casΦ from huge phages is a hypercompact genome editor. Science. 369:333–337. https://doi.org/10.1126/science.abb1400
- Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J Exp Bot 56:1–14. https://doi.org/10.1093/jxb/eri025



- Oi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell. 152:1173-1183. https://doi.org/10.1016/j.cell.2013.02.022
- Ran FA, Hsu PD, Lin C, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 155:479-480. https://doi.org/10.1016/j. cell.2013.09.040
- Rio S, Mary-Huard T, Moreau L, Bauland C, Palaffre C, Madur D, Combes V, Charcosset A (2020) Disentangling group specific QTL allele effects from genetic background epistasis using admixed individuals in GWAS: an application to maize flowering. PLoS Genet 16(3):e1008241. https://doi.org/10.1371/journal.pgen. 1008241
- Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett M, Lippman Z (2017) Engineering quantitative trait variation for crop improvement by genome editing. Cell 171:470–480. https://doi.org/10.1016/j.cell. 2017.08.030
- Salsman J, Dellaire G (2017) Precision genome editing in the CRISPR era. Biochem Cell Biol 95(2):187-201. https://doi.org/10.1139/bcb-
- Sarno R, Vicq Y, Uematsu N, Luka M, Lapierre C, Carroll D, Bastianelli G, Serero A, Nicolas A (2017) Programming sites of meiotic crossovers using Spo11 fusion proteins. Nucleic Acids Res 45:1-2. https://doi.org/10.1093/nar/gkx739
- Scheben A, Wolter F, Batley J, Puchta H, Edwards D (2017) Towards CRISPR/CAS crops - Bringing together genomics and genome editing. New Phytol 216:682-698. https://doi.org/10.1111/nph.
- Sedivy EJ, Wu F, Hanzawa Y (2017) Soybean domestication: the origin, genetic architecture and molecular bases. New Phytol 214:539-553. https://doi.org/10.1111/nph.14418
- Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu J-L, Gao C (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31:8-10. https://doi. org/10.1038/nbt.2652
- Shelake RM, Pramanik D, Kim JY (2019) Evolution of plant mutagenesis tools: a shifting paradigm from random to targeted genome editing. Plant Biotechnol Rep 13:423-445. https://doi.org/10.1007/s11816-019 - 00562 - 7
- Shen H, Strunks GD, Klemann BJPM, Hooykaas PJJ, De Pater S (2017) CRISPR / Cas9-induced double-strand break repair in Arabidopsis nonhomologous end-joining mutants. Genes Genomes Genet 7: 193-202. https://doi.org/10.1534/g3.116.035204
- Shi J, Gao H, Wang H, Lafitte HR, Archibald R, Yang M, Hakimi S, Mo H, Habbe J (2016) ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. Plant Biotechnol J. 15:207–216. https://doi.org/10.1111/pbi.12603
- Shimatani Z, Kashojiya S, Takayama M, Terada R, Arazoe T, Ishii H, Teramura H, Yamamoto T, Komatsu H, Miura K, Ezura H, Nishida K, Ariizumi T, Kondo A (2017a) ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. Plant Biotechnol J 15:207–216. https://doi.org/10.1111/ pbi.12603
- Shimatani Z, Kashojiya S, Takayama M, Terada R, Arazoe T, Ishii H, Teramura H, Yamamoto T, Komatsu H, Miura K, Ezura H, Nishida K, Ariizumi T, Kondo A (2017b) Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. Nat Biotechnol 35:441-443. https://doi.org/10.1038/nbt.3833
- Stetter MG, Thornton K, Ross-Ibarra J (2018) Genetic architecture and selective sweeps after polygenic adaptation to distant trait optima. PLoS Genet 14:1-24. https://doi.org/10.1371/journal.pgen.1007794
- Symington LS, Gautier J (2011) Double-strand break end resection and repair pathway choice. Annu Rev Genet 45:247-271. https://doi. org/10.1146/annurev-genet-110410-132435

- Taagen E, Bogdanove AJ, Sorrells ME (2020) Counting on crossovers: controlled recombination for plant breeding. Trends Plant Sci 25: 455-465. https://doi.org/10.1016/j.tplants.2019.12.017
- Underwood CJ, Choi K, Lambing C, Zhao X, Serra H, Borges F, Simorowski J, Ernst E, Jacob Y, Henderson IR, Martienssen RA (2017) Epigenetic activation of meiotic recombination in Arabidopsis centromeres via loss of H3K9me2 and non-CG DNA methylation. bioRxiv 519-531. https://doi.org/10.1101/160929
- Vazquez-vilar M, Diego-martin B, Ajenjo M, Selma S, Bernab JM, Garcia-carpintero V, Granell A, Orzaez D (2019) Strong gene activation in plants with genome-wide specificity using a new orthogonal CRISPR / Cas9-based programmable transcriptional activator. Plant Biotechnol J 17(9):1703–1705. https://doi.org/10.1111/pbi.
- Vojta A, Dobrinic P, Tadic V, Bockor L, Korac P, Julg B, Klasic M, Zoldos V (2016) Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Res 44:5615-5628. https://doi. org/10.1093/nar/gkw159
- Von Arnim AG, Jia Q, Vaughn JN (2014) Regulation of plant translation by upstream open reading frames. Plant Sci 214:1-12. https://doi. org/10.1016/j.plantsci.2013.09.006
- Walley J, Sartor R, Shen Z, Schmitz R, Wu K, Urich M, Nery J, Smith L, Schnable J, Echer J, Briggs S (2016) Integration of omic networks in a developmental atlas of maize. Science 353:1-5. https://doi.org/10. 5061/dryad.v8969
- Wang C, Hu S, Gardner C, Lübberstedt T (2017) Emerging avenues for utilization of exotic germplasm. Trends Plant Sci 22:624-637. https://doi.org/10.1016/j.tplants.2017.04.002
- Wang M, Lu Y, Botella JR, Mao Y, Hua K, Zhu JK (2017b) Gene targeting by homology-directed repair in rice using a geminivirusbased CRISPR/Cas9 system. Mol Plant 10:1007-1010. https://doi. org/10.1016/j.molp.2017.03.002
- Wang P, Xiong Y, Gong R, Yang Y, Fan K, Yu S (2019a) A key variant in the cis-regulatory element of flowering gene Ghd8 associated with cold tolerance in rice. Sci Rep 9:1-14. https://doi.org/10. 1038/s41598-019-45794-9
- Wang S, Zong Y, Lin O, Zhang H, Chai Z, Zhang D, Chen K, Qiu JL, Gao C (2020) Precise, predictable multi-nucleotide deletions in rice and wheat using APOBEC-Cas9. Nat Biotechnol 38:1460-1465. https://doi.org/10.1038/s41587-020-0566-4
- Wang T, Zhang H, Zhu H (2019b) CRISPR technology is revolutionizing the improvement of tomato and other fruit crops. Hortic Res 6(1):1-3. https://doi.org/10.1038/s41438-019-0159-x
- Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, Qiu JL (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotechnol 32:947-951. https://doi.org/10.1038/nbt.2969
- Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley J, Mcmullen M, Gaut B (2005) The effects of artificial selection on the maize genome. Science. 308:1310-1314. https://doi.org/10.1126/science. 1107891
- Wu J, Lawit SJ, Weers B, Sun J, Mongar N, van Hemert J, Melo R, Meng X, Rupe M, Clapp J, Collet KH, Trecker L, Roesler K, Peddicord L, Thomas J, Hunt J, Zhou W, Hou Z, Wimmer M, Jantes J, Mo H, Liu L, Wang Y, Walker C, Danilevskaya O, Lafitte RH, Schussler JR, Shen B, Habben JE (2019) Overexpression of zmm28 increases maize grain yield in the field. Proc Natl Acad Sci U S A 116: 23850–23858. https://doi.org/10.1073/pnas.1902593116
- Wu W, Liu X, Wang M, Meyer RS, Luo X, Ndjiondjop MN, Tan L, Zhang J, Wu J, Cai H, Sun C, Wang X, Wing RA, Zhu Z (2017) A single-nucleotide polymorphism causes smaller grain size and loss of seed shattering during African rice domestication. Nat Plants 3:1-7. https://doi.org/10.1038/nplants.2017.64
- Xu X, Tao Y, Gao X, Zhang L, Li X, Zou W, Ruan K, Wang F, Xu GL, Hu R (2016) A CRISPR-based approach for targeted DNA





demethylation. Cell Discov 2(1):1–12. https://doi.org/10.1038/celldisc.2016.9

- Xu Z, Kuang Y, Ren B, Yan D, Yan F, Spetz C, Sun W, Wang G, Zhou X, Zhou H (2021) SpRY greatly expands the genome editing scope in rice with highly flexible PAM recognition. Genome Biol 22:1–15. https://doi.org/10.1186/s13059-020-02231-9
- Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, Miller N, Drouaud J, Grelon M, Copenhaver GP, Mezard C, Kelly KA, Henderson IR (2012) Epigenetic remodeling of meiotic crossover frequency in Arabidopsis thaliana DNA methyltransferase mutants. PLoS Genet 8(8):e1002844. https://doi.org/10.1371/journal. pgen.1002844
- Yu QH, Wang B, Li N, Tang Y, Yang S, Yang T, Xu J, Guo C, Yan P, Wang Q, Asmutola P (2017) CRISPR/Cas9-induced targeted mutagenesis and gene replacement to generate long-shelf life tomato lines. Sci Rep 7:1–9. https://doi.org/10.1038/s41598-017-12262-1
- Zeng A, Chen P, Korth K, Hancock F, Pereira A, Brye K, Wu C, Shi A (2017) Genome-wide association study (GWAS) of salt tolerance in worldwide soybean germplasm lines. Mol Breed 37(3):1–14. https:// doi.org/10.1007/s11032-017-0634-8
- Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, Van Der Oost J, Regev A, Koonin EV, Zhang F (2015) Cpf1 Is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. Cell. 163:759–771. https://doi.org/10.1016/j.cell.2015.09.038
- Zhang H, Si X, Ji X, Fan R, Liu J, Chen K, Wang D, Gao C (2018a) Genome editing of upstream open reading frames enables translational control in plants. Nat Biotechnol 36:894–900. https://doi.org/10.1038/nbt.4202

- Zhang Q, Xing HL, Wang ZP, Zhang HY, Yang F, Wang XC, Chen QJ (2018b) Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. Plant Mol Biol 96: 445–456. https://doi.org/10.1007/s11103-018-0709-x
- Zhong Z, Zhang Y, You Q, Tang X, Ren Q, Liu S, Yang L, Wang Y, Liu X, Liu B, Zhang T, Zheng X, Le Y, Zhang Y, Qi Y (2018) Plant genome editing using FnCpfl and LbCpfl nucleases at redefined and altered PAM sites. Mol Plant 11:999–1002. https://doi.org/10.1016/j.molp.2018.03.008
- Zhou J, Xin X, He Y, Chen H, Li Q, Tang X, Zhong Z, Deng K, Zheng X, Akher SA, Cai G, Qi Y, Zhang Y (2019) Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. Plant Cell Rep 38:475–485. https://doi.org/10.1007/s00299-018-2340-3
- Zhu H, Li C, Gao C (2020) Applications of CRISPR—Cas in agriculture and plant biotechnology. Nat Rev Mol Cell Biol 21:661–677. https://doi.org/10.1038/s41580-020-00288-9
- Zong Y, Song Q, Li C, Jin S, Zhang D, Wang Y, Qiu JL, Gao C (2018) Efficient c-to-t base editing in plants using a fusion of ncas9 and human apobec3a. Nat Biotechnol 36(10):950–953. https://doi.org/ 10.1038/nbt.4261
- Zong Y, Wang Y, Li C, Zhang R, Chen K, Ran Y, Qiu JL, Wang D, Gao C (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nat Biotechnol 35:438–440. https://doi.org/10.1038/nbt.3811
- Zsögön A, Čermák T, Naves ER, Notini MM, Edel KH, Weinl S, Freschi L, Voytas DF, Kudla J, Peres LEP (2018) De novo domestication of wild tomato using genome editing. Nat Biotechnol 36:1211–1216. https://doi.org/10.1038/nbt.4272