Towards CRISPR/Cas crops – bringing together genomics and genome editing

Armin Scheben\textsuperscript{1\dagger}, Felix Wolter\textsuperscript{2\dagger}, Jacqueline Batley\textsuperscript{1}, Holger Puchta\textsuperscript{2}, David Edwards\textsuperscript{1*}

\textsuperscript{1}School of Plant Biology and Institute of Agriculture, University of Western Australia, Perth 6009, WA, Australia

\textsuperscript{2}Botanical Institute II, Karlsruhe Institute of Technology, Karlsruhe 76131, Germany

\dagger These authors contributed equally to the manuscript.

*Corresponding author: Tel +61 (0)8 6488 2415

Dave.Edwards@uwa.edu.au

Twitter name: @arminscheben

Running title: Towards CRISPR crops

Keywords: crop breeding; CRISPR/Cas; genome editing.

Total word count: 7,750

Section lengths: I (1,055), II (351), III (313), IV (592), V (1,666), VI (1,384), VII (2,160), VIII (201), Acknowledgements (28)

Number of figures: 6

Number of tables: 1

Supporting information: none
Summary

With the rapid increase in the global population and the impact of climate change on agriculture, there is a need for crops with higher yields and greater tolerance to abiotic stress. However, traditional crop improvement via genetic recombination or random mutagenesis is a laborious process and cannot keep pace with increasing crop demand. Genome editing technologies such as clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (CRISPR/Cas) allow targeted modification of almost any crop genome sequence to generate novel variation and accelerate breeding efforts. We expect a gradual shift in crop improvement away from traditional breeding towards cycles of targeted genome editing. Crop improvement using genome editing is not constrained by limited existing variation or the requirement to select alleles over multiple breeding generations. However, current applications of crop genome editing are limited by the lack of complete reference genomes, the sparse knowledge of potential modification targets, and the unclear legal status of edited crops. We argue that overcoming technical and social barriers to the application of genome editing will allow this technology to produce a new generation of high-yielding, climate ready crops.

Keywords: Breeding; Crops; CRISPR; Cas; Gene targeting; Genome editing; Targeted mutagenesis
## Box 1. Definitions used in this review

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genome editing</strong></td>
<td>Technique for generating site-specific insertions, deletions or substitutions in the genomes of living cellular organisms. Genome editing relies on programmable nucleases to cleave DNA while cellular DNA repair processes induce desired mutations. Depending on DNA repair pathway, mutations can be random or sequence-specific.</td>
</tr>
<tr>
<td><strong>Non-homologous end joining (NHEJ)</strong></td>
<td>Error-prone DNA repair pathway that mediates direct ligation of break ends without requiring a homologous template.</td>
</tr>
<tr>
<td><strong>Homology-directed repair (HDR)</strong></td>
<td>Highly accurate DNA repair pathway that mediates repair of break ends using a template homologous to the region of the break.</td>
</tr>
<tr>
<td><strong>Gene targeting</strong></td>
<td>Uses HDR for site-specific induction of specific insertions, deletions or substitutions in or around genes. Can be used in genome editing to delete or add genes and regulatory regions, and to introduce point mutations.</td>
</tr>
<tr>
<td><strong>CRISPR RNA (crRNA)</strong></td>
<td>Short RNA that determines the target specificity of CRISPR/Cas systems. It constitutes the variable part of the CRISPR array from which it is expressed as a pre-crRNA and processed. It consists itself of a constant and a variable part and associates with the DNA cleaving protein, depending on the system alone (Cpf1), in combination with tracrRNA (native type IIa system), or directly fused to tracrRNA.</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------</td>
</tr>
<tr>
<td>Trans-activating crRNA</td>
<td>Occurs only in type IIa CRISPR/Cas systems. Binds to pre-crRNA to</td>
</tr>
<tr>
<td>(tracrRNA)</td>
<td>form double stranded RNA which is cleaved by endogenous RNaseIII in</td>
</tr>
<tr>
<td></td>
<td>the presence of Cas9 to release functional crRNA-tracrRNA hybrids.</td>
</tr>
<tr>
<td>Single guide RNA (sgRNA)</td>
<td>Synthetic RNA consisting of a fusion between crRNA and tracrRNA to</td>
</tr>
<tr>
<td></td>
<td>reduce the number of RNAs that need to be expressed when using type</td>
</tr>
<tr>
<td></td>
<td>IIa CRISPR/Cas for genome editing purposes.</td>
</tr>
<tr>
<td>‘Dead’ Cas9 (dCas9)</td>
<td>Cas9 variant in which both nuclease domains (RuvC and HNH) are</td>
</tr>
<tr>
<td></td>
<td>deactivated by point mutations, yielding a programmable DNA binding</td>
</tr>
<tr>
<td></td>
<td>protein that can be fused to effector domains like transcriptional</td>
</tr>
<tr>
<td></td>
<td>regulators or fluorescent proteins.</td>
</tr>
<tr>
<td>Nickase</td>
<td>Enzyme generating a single-strand DNA break (nick). Cas9 variants</td>
</tr>
<tr>
<td></td>
<td>with a mutated nuclease domain are available that induce nicks rather</td>
</tr>
<tr>
<td></td>
<td>than double-strand DNA breaks.</td>
</tr>
<tr>
<td>Off-target effect</td>
<td>Mutations induced in unintended targets during genome editing.</td>
</tr>
</tbody>
</table>
I Introduction

1. Accelerating crop improvement to meet increased food demand under a changing climate

Rapid population growth and increased consumption of meat and dairy are putting pressure on agriculture to meet the rising demand for food, livestock feed and biofuels. By 2050, the global population will increase to more than 9 billion people and crop demand may increase by between 100–110% (Tilman et al., 2011). Yet at the current rate of improvement, yields of the staple crops maize, rice, wheat and soybean will increase by just ∼38-67% (Ray et al., 2013). At the same time, climate change may lead to crop yield loss due to increased frequencies of drought, flooding and pest incidence (IPCC, 2014). In recent history, increased crop production has been achieved through improved agronomic management such as the application of pesticides and fertiliser (Matson et al., 1997). However, these management practises rely on finite resources, occasionally have a negative impact on the environment and are unlikely to lead to further sizeable increases in yield in many regions (Edwards, 2016).

While yields of the major crops maize, rice, wheat and soybean continue to increase in many areas, yield is beginning to stagnate in parts of Europe, Asia and Africa (Ray et al., 2012). As the amount of arable land worldwide is limited and decreasing due to expanding urban areas and land degradation, improving crops to better utilize available resources and tolerate stress is an important approach complementing agronomic management. Improved pest-resistant and resource-efficient crops will also substantially reduce the environmental impact of agriculture by decreasing the amount of pesticide and fertiliser required for optimal yields.

While traditional breeding continues to deliver high-yielding crop varieties with enhanced traits, it relies on the crossing of germplasm or random mutagenesis which can take 7-12 years to produce an improved variety (Acquaah, 2012) and is unlikely to keep pace with the predicted demand for improved crops. While breeding efficiency can be improved using trait-linked genetic markers, the improvement of germplasm remains limited by the non-targeted
nature of recombination or random mutagenesis. Although transgenic genetically modified (GM) crops have delivered average yield benefits of up to 7% in industrial countries and up to 30% in developing countries (Carpenter, 2010), since 1995 there have been few traits broadly commercialised beside herbicide resistance and insect resistance, and strict regulation and public scepticism have slowed development of GM crops.

Over the last decade, targeted genome editing technologies have emerged, using hybrid enzymes or the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (Cas) system (Jinek et al., 2012). Genetic modification is performed by inducing DNA double-strand breaks (DSB) at specific genome locations and stimulating nonhomologous end joining (NHEJ) or homology-directed repair (HDR) to introduce specific DNA modifications into the genome. When applied to crop breeding, genome editing can rapidly generate transgene-free improved varieties. Approaches based on hybrid enzymes consisting of fused DNA-binding domains and nucleases such as zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) have been used to edit the genomes of many crops including soybean (Curtin et al., 2011) and wheat (Wang et al., 2014). However, a major drawback of hybrid enzymes is the high cost and complexity of the protein engineering required for their development.

The CRISPR/Cas system is superior to hybrid enzyme approaches because of its RNA based sequence specificity which provides high versatility and low cost, and its application can accelerate both basic research and crop improvement (Belhaj et al., 2015; Bortesi & Fischer, 2015; Puchta, 2017; Scheben & Edwards, 2017). CRISPR/Cas originates from the immune system of bacteria and archaea, carrying out RNA-guided cleavage of foreign DNA such as viruses or plasmids. Current classification differentiates two classes, six types and 19 subtypes of CRISPR (Shmakov et al., 2017), with the type II system of *Streptococcus pyogenes* currently most widely adopted for genome editing. The CRISPR/Cas system is
generally delivered into crops via DNA expression vectors, but can also be delivered as RNA or pre-formed ribonucleoprotein (RNP) complexes as was first demonstrated by Woo et al. (2015). The Cas9 nuclease, which is the enzyme commonly used, induces a DSB at a site specified by a target sequence of 19-22 nucleotides integrated into a single guide RNA (sgRNA) (Gasiunas et al., 2012; Jinek et al., 2012). The target site must be followed by a three nucleotide protospacer adjacent motif (PAM) which can be bound by Cas9. For the Streptococcus pyogenes system, the PAM site allowing highest binding is NGG, although NAG also binds at lower efficiency. In this review, we use the specific term CRISPR/Cas9 to refer to the editing system using the Cas9 nuclease and the general term CRISPR/Cas to refer to CRISPR editing systems using any Cas nuclease (see Box 1 for definitions used in this review).

CRISPR/Cas can directly introduce mutations into elite germplasm, and in the last three years, the CRISPR/Cas system has been shown to be effective in a wide range of major crop species, including maize (Svitashev et al., 2016), rice (Jiang et al., 2013; Wang et al., 2017) and wheat (Shan et al., 2013; Wang et al., 2014). Genomes of elite crop varieties can be edited in a targeted manner to produce new varieties with increased stress tolerance and nutrient-use efficiency, without being constrained by limited existing variation. In addition, the continuous improvement of elite varieties by genome editing does not introduce potentially deleterious alleles from crossing and recombination, or require time consuming repeat backcrossing to reconstruct the elite genetic background. In this review, we propose a shift in crop improvement towards cycles of targeted genome editing of elite varieties to produce improved varieties, with advanced traits or specific adaptation to local environments, replacing the cycles of crossing or untargeted mutagenesis and selection currently employed by breeders. This new crop improvement approach provides a powerful new breeding tool, but also brings substantial new requirements for genomic information and bioinformatics.
Genome editing based crop improvement schemes rely on functional genomic information to identify editing targets including genes and regulatory regions. Cycles of editing and selection of optimal variants during field trials allow further improvement of breeding outcomes. Here, we broadly outline the requirements for genome-wide data to enable the discovery of candidate genes and regulatory regions for genome editing. We also discuss applications of genome editing and methods to improve editing specificity, multiplexing and gene regulation using CRISPR/Cas.

II Genomic based crop improvement before CRISPR/Cas

Many current efforts to accelerate crop breeding rely on increasing selection efficiency using marker-assisted selection (MAS) (Collard & Mackill, 2008) and genomic selection (GS) (Desta & Ortiz, 2014). MAS uses linkage disequilibrium (LD) between genetic markers and quantitative trait loci (QTL) to select plants with traits of interest for breeding programs. Although the last two decades have seen substantial advances in marker technologies, polygenic quantitative traits are not readily amenable to MAS, and the complexity of genotype by environment interactions present a further challenge. In contrast to MAS, GS predicts breeding values for individual lines in a phenotyped and genotyped training population, allowing selection in breeding populations to be carried out using marker-based breeding values instead of phenotypes or gene associated markers (Desta & Ortiz, 2014). GS can accelerate the breeding cycle by reducing the need for phenotyping breeding populations and, unlike MAS, facilitates selection of complex polygenic traits. However, marker-based approaches such as MAS and GS are limited by their reliance on markers and the available genetic diversity in the population. As our understanding of crop genomes increases, marker-based breeding approaches may transition towards approaches based on functionally characterized candidate genes. Key functional genomics resources that provide insight into gene function using reverse genetics are mutant crop populations generated with approaches
such as targeting induced local lesions in genomes (TILLING) (McCallum et al., 2000). TILLING is used in breeding to generate and identify novel gene variants into crops by inducing random mutations in a population and subsequently detecting mutations in target genes with oligonucleotide probes or sequencing. Backcrossing to parental, unmutagenized lines then allows breeders to capture useful variants and introgress them into elite varieties. The power of such an approach in breeding is shown, for instance, by the pod shatter resistance trait introduced into commercial canola cultivars such as Bayer IH 51 RR via a single nucleotide genic mutation (Lambert et al., 2015). As breeding schemes increase their focus on candidate genes for traits, genome editing provides a more precise tool to produce novel allelic variants of these genes for agronomic assessment.

III Genome assembly as the starting point for genome editing

Annotated reference genome sequences underpin the targeted editing of crop genomes, and provide information on the base material for improvement. Reference genomes facilitate the discovery of functional regions implicated in agronomic traits such as genes, promoters and enhancers, and allow the design of sgRNA and alignment of sequencing reads used to validate edited positions of the genome. In 2015, there were more than 60 high-quality crop genomes available (Michael & VanBuren, 2015) and we can expect the decreasing cost of sequencing to soon facilitate the sequencing and assembly of all genomes of major and minor crops and their wild relatives. Transcriptomes may also play an important roles as starting points for genome editing because plant genome assemblies and their annotation often remain incomplete. Genome assemblies and their annotations are likely to improve in completeness and accuracy as third-generation sequencing technologies, which generate long reads, help address previous challenges of de novo genome assembly limited to using short reads (Yuan et al., 2017). The resequencing of crop genomes representing multiple diverse varieties is now also providing access to pangenomes, which represent the genomic diversity within a
species rather than the sequence of a representative individual (Golicz et al., 2016a). This is particularly useful for the development of improved varieties because within-species differences in the presence and absence of genes can be particularly high in plants, for instance affecting almost 20% of all genes in the crop Brassica oleracea (Golicz et al., 2016b). Pangene references make genome-wide variation in presence/absence and copy number accessible for editing. Knowledge of within-species diversity is also important for genome editing because it enables cultivar-specific design of sgRNAs, which is necessary when the 19-22 nucleotide target or PAM site differs between cultivars. The increasing number of crop pangene assemblies will be a valuable resource for efficient genome editing and breeding.

IV Identifying targets for genome editing using functional genomics

Since publication of the first Arabidopsis transcriptome (Weber et al., 2007), functional genomics has helped characterise gene function, expression and interaction in many plants. However, the functions, determinants of expression and interaction networks of most genes remain unclear and lack experimental characterisation (Radivojac et al., 2013). Moreover, functional sequences within the noncoding regions of the genome also remain poorly characterised, despite evidence indicating their important role in gene regulation (Haudry et al., 2013). Targeted genome editing relies on detailed knowledge of gene regulation to identify targets for editing, ensure genes are only expressed in relevant tissues and make manipulation of regulatory sequences feasible.

TILLING populations are an important resource for discovering and deploying functional gene variants, and have been developed for several crop plants, including rice (Till et al., 2007) and wheat (Uauy et al., 2009). These populations can help identify mutations and genes associated with agronomic traits. For example, TILLING was used to identify waxy gene variants controlling starch synthase in wheat (Slade et al., 2005) and mutations in
sorghum leading to economically valuable acyanogenic plants (Blomstedt et al., 2012). High-throughput mutation screens produced using lentiviral CRISPR-guide RNA libraries have also been used in human cells to interrogate gene function and identify enhancers (gene-distal regulatory DNA that is bound by transcription factors to enhance gene transcription) in the noncoding genome (Koike-Yusa et al., 2014; Sanjana et al., 2016), showing improvement over alternative approaches such as RNA interference by allowing complete target knockout and fewer off-target effects. A further line of evidence to characterise functional regions of the genome and their interactions is the increasing amount of data on the 3D structure of crop genomes and genome-wide chromatin accessibility, produced using methods such as Hi-C (Lieberman-Aiden et al., 2009), assay for transposase accessible chromatin with high-throughput sequencing (ATAC-seq) (Buenrostro et al., 2013), and chromatin immunoprecipitation sequencing (ChIP-seq) (Solomon et al., 1988). Employing these approaches for high-resolution, genome-wide mapping of epigenetic marks and binding sites for DNA-binding proteins such as transcription factors is crucial for understanding the processes underlying transcriptional regulation and harnessing them for crop improvement. For instance, Zhang et al. (2012) identified 97,975 deoxyribonuclease hypersensitive sites in rice seedlings, of which 42%-45% are potential regulatory elements located in intergenic regions.

Gene expression atlases with data from different tissues under different conditions are also an important resource for predicting the function of genes and the potential effects of editing. Dynamic gene expression during development and under different abiotic and biotic conditions provides insight into gene function and assists in modelling the repercussions of genome editing throughout the life of a crop. Expression atlases have been developed for many major crops such as maize (Sekhon et al., 2011), rice (Nobuta et al., 2007). Database projects integrating gene expression datasets for diverse species such as the PlexDB database
(Dash et al., 2012) are also providing platforms for better leveraging crop expression data. As protein-protein interaction data are poorly available for plants other than Arabidopsis, these expression atlases could help identification of functionally related genes by coexpression analysis (Ruprecht et al., 2017). Despite these advances, the lack of functional understanding of genes remains one of the biggest bottlenecks in crop improvement using genome editing. For instance, in maize, only ~1% of genes have functionally annotated using experimental data (Andorf et al., 2016). Comprehensive and integrated genomic databases including gene expression atlases together with more accurate in silico methods for gene function prediction based on homology and coexpression will be crucial for overcoming the gene function bottleneck.

V Using CRISPR/Cas to improve crops

1. Improving crops by disrupting genes with CRISPR/Cas

CRISPR/Cas dramatically increases the potential to improve traits in crops compared to conventional breeding approaches. Importantly for crop improvement, homozygous pyramiding of genes of interest into elite germplasm can be carried out within a single generation (Zhang et al., 2014). CRISPR/Cas is also a powerful tool for introducing heritable, trait-related mutations indistinguishable from natural allelic variants (Schaeffer & Nakata, 2015). Moreover, if the CRISPR/Cas system is transformed into the crop for delivery, as is commonly the case, plants are hemizygous for the transgenes while genome editing affects all alleles. This means that crossing or selfing these plants can generate transgene-free progeny, reducing potential risks and regulatory requirements that may apply to transgenic crops.

In its simplest form, genome editing can improve crops by knocking out genes conferring undesirable traits (Table 1). This approach has been used extensively to increase pathogen resistance in crops (Andolfo et al., 2016). For instance, CRISPR/Cas9 has been used to
enhance resistance against blast and bacterial blight in rice by disrupting the ERF transcription factor and the *SWEET* genes respectively (Jiang *et al.*, 2013; Wang *et al.*, 2016). The biotechnology industry is beginning to actively apply genome editing for crop improvement. In 2015, Cibus developed a transgene-free herbicide-tolerant canola using a proprietary genome editing system to carry out an amino acid exchange (Schinkel & Schillberg, 2016), with the cultivar now commercially available in the USA. Within five years, DuPont Pioneer aims to commercialise a high amylopectin corn with superior yield and food properties that was developed using CRISPR/Cas9 to disrupt the amylose biosynthesis gene *Wx1* (Waltz, 2016).

2. Precise gene modifications

Although disrupting genes can confer agronomically important traits, most traits can only be improved by precise gene modifications. These precise alterations can be achieved by gene targeting (GT), which refers to HDR mediated site-specific transgene integration or point mutations. GT can be performed by providing a template for DSB repair via HDR containing desired modifications at the same time as inducing the break. However, GT is still a challenge in plants, because NHEJ remains the dominant pathway of DSB repair in plants (Puchta, 2005), even when a repair template is available. A further challenge is that many crops still lack efficient transformation and regeneration procedures (Altpeter *et al.*, 2016). While isolated protoplasts can be transformed at high efficiency and allow GT (Townsend *et al.*, 2009), efficient regeneration of plants from protoplasts is difficult or not yet possible for many crops, particularly monocotyledons including wheat, rice, barley and sorghum (Eeckhaut *et al.*, 2013). Recently, a highly efficient and widely applicable transformation approach for monocot species was reported (Lowe *et al.*, 2016). The enhanced efficiency is based on overexpression of maize Baby boom (Bbm) and Wuschel2 (WS2) genes after *Agrobacterium*-mediated transformation of immature embryos, leading to a growth
stimulation of transformed tissue compared to non-transformed tissue. Furthermore, Lowe et al. successfully transformed alternative target tissues such as embryo slices from mature seeds with this approach. This is especially useful as supplying immature embryos for classical transformation is highly time and cost intensive and embryo slices from mature seeds can be supplied in much larger quantities due to automated preparation.

A GT method for crops that are recalcitrant to efficient transformation is in planta gene targeting (Fauser et al., 2012). Here, the CRISPR/Cas9 construct and the repair template are stably transformed into the plant genome and the nuclease not only cuts in the target, but also excises the repair template, activating it for HDR. Thus, GT events can happen during the life cycle of the plant, and upon entering the germ-line, can be harvested as seeds in the next generation, so in principle a single successful transformation event is sufficient. Using this method, Schiml et al. (2014) achieved site-specific insertion in Arabidopsis.

The limitation of reduced amounts of HDR template in transformed cells can be addressed by employing geminiviral replication mechanisms to increase repair template copy numbers (Baltes et al., 2014). Using this approach, it was possible to achieve efficient GT via Agrobacterium-mediated transformation of cotyledons or leaf explants and subsequent plant regeneration in tomato (Cermak et al., 2015). Biolistic transformation is another method that enables delivery of larger amounts of repair template compared to Agrobacterium-mediated transformation (Altpeter et al., 2016). Via biolistic transformation and regeneration of immature embryos, site-specific gene modifications and insertions were demonstrated in maize (Svitashev et al., 2016), soybean (Li et al., 2015) and rice (Sun et al., 2016). Gil-Humanes et al. (2017) combined viral replication with biolistic transformation of wheat cells and achieved multiplexed GT at all three homeoalleles, albeit without regeneration of edited plants.
Instead of HDR, NHEJ can also be harnessed for precise genome editing. For example, gene replacements and insertions were demonstrated in the rice 5-enolpyruvylshikimate-3-phosphate synthase gene (EPSPS) by targeting 2 adjacent introns and providing a repair template with point mutations in the intermediate exon via biolistic transformation (Li, J et al., 2016). Using the more dominant NHEJ repair pathway can help overcome the challenge of inefficient gene modification when relying on HDR. Finally, precise ‘base editing’ is now possible without cleaving double-stranded DNA, by using a cytidine deaminase enzyme (Komor et al., 2016). When fused to Cas9, it can target a specific site and effects the conversion of cytidine to uridine within a window dependent on the fusion protein linker, leading to the substitution of cytosine with thymine, or guanine with adenine. Using Cas9 nickase (an enzyme variant that induces single-stranded breaks or 'nicks') to cleave the non-edited strand strongly enhances efficiency compared to inactive Cas9 by promoting replacement of the non-edited base. Zong et al. (2017) successfully applied this technique in rice, wheat and maize with a high editing efficiency of up to 40% in transgenic plants. Using the XTEN linker to link Cas9 and cytidine deaminase, they observed an editing window covering 7nt of the protospacer.

3. Beyond Cas9: broadening the toolbox with Cpf1

Cpf1 (CRISPR from Prevotella and Francisella 1; recently renamed 'Cas12a' by Shmakov et al. (2017)), another DNA cleaving enzyme from a class II CRISPR system, is now also available for genome editing in plants. There are several key differences between Cas9 and Cpf1 (Fig. 1). Firstly, unlike Cas9, Cpf1 does not require trans-activating CRISPR RNA (tracrRNA) (Zetsche et al., 2015). Instead, it cleaves precursor crRNA (pre-CRISPR RNA) autonomously and associates with mature crRNA alone to cleave target DNA. Thus, only the short crRNA (c. 42nt) is required without the need for a long chimeric RNA as is the case for Cas9. This decreases the size of the construct that is delivered into plant cells, which can be
advantageous when multiplexing on a large scale. Furthermore, the pre-crRNA processing ability of Cpf1 can be harnessed to enhance multiplexing by constructing a polycistronic gene consisting of tandem repeats of a direct repeat alternating with target-specific spacers (Wang et al., 2017; Zetsche et al., 2017). Secondly, Cpf1 recognizes a T-rich PAM upstream instead of downstream of the target sequence (Zetsche et al., 2015). This provides a new range of target sites across plant genomes, particularly in promoter regions that are naturally AT-rich and are often difficult to target using the G-rich Cas9 PAM. Thirdly, Cpf1 cleaves each complementary strand of DNA at different sites, spaced five nucleotides apart, leading to sticky ends instead of the blunt ends produced by Cas9. This could prove useful because it might enhance the efficiency of genomic rearrangements such as precise chromosomal deletions or somatic crossovers. While blunt ends generated by two Cas9 cuts flanking the fragment targeted for deletion rarely spontaneously rejoin, compatible sticky ends generated by two Cpf1 cuts might have a higher likelihood of joining due to Watson-Crick base-pairing.

Three Cpf1 orthologues have been successfully tested in plants: Cpf1 from Francisella novicida (FnCpf1), Lachnospiraceae bacterium ND2006 (LbCpf1) and Acidaminococcus sp. BV3L6 (AsCpf1). Direct comparisons have shown that LbCpf1 has higher efficiency than AsCpf1 and FnCpf1 (Tang et al., 2017; Wang et al., 2017). Importantly, correct processing of the crRNA ends seems to be crucial for high Cpf1 activity in plants: Tang et al. (2017) dramatically improved Cpf1 efficiency by flanking the crRNA with ribozymes leading to a release of crRNAs without deleterious extensions such as the poly-U from polymerase III termination or the required ‘G’ for the initiation by the U6 promoter. They also demonstrated that Cpf1 has potential for transcriptional repression in plants. By fusing nuclease deficient Cpf1 to a repression domain, greater repression was achieved than previously with 'dead' Cas9 (dCas9) based fusions.

4. Chromosomal rearrangements and recombination control
CRISPR/Cas can enable large scale chromosomal rearrangements and provides opportunities for control of meiotic recombination. The ability to precisely control recombination would open a new dimension of possibilities to the plant breeder. While breaking genetic linkage between genes for beneficial and adverse traits is an obstacle for plant breeding, maintaining linkage between genes for beneficial traits is desirable. Both can be achieved by control of chromosomal translocations. By induction of a DSB on two different chromosomes, reciprocal exchanges of chromosome arms can be achieved (Pacher et al., 2007) (Fig. 2A). Similarly, linkage can be broken by inducing artificial crossovers (Sadhu et al., 2016). Induction of two DSBs on the same chromosome can lead to chromosomal deletions and inversions (Lee et al., 2012; Zhou et al., 2014; Ordon et al., 2016) (Fig. 2B). These chromosomal inversions can prevent meiotic recombination between homologues to stabilize linkage of positive traits (Puchta, 2016). It should be noted that Cpf1 may prove to be an efficient tool for these NHEJ based genomic rearrangements due to its ability to generate compatible overhangs. Alternatively, precise targeted recombination between parental genomes can be achieved by CRISPR/Cas induced manipulation of meiotic recombination. Using meiosis specific expression, DSBs can be induced during meiosis at specific sites, or effector proteins involved in initiation of meiotic recombination such as SPO11 can be targeted to specific sites to guide meiotic recombination, as has been shown in yeast (Peciña et al., 2002).

VI Improving the efficiency of the CRISPR/Cas system

1. CRISPR/Cas efficiency in plants

Initial demonstration of the functionality of the CRISPR/Cas9 system in plant cells was carried out in 2013 (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013) and the demonstration of stable inheritance of induced mutations followed one year later (Feng et al., 2014). The technology was rapidly adopted and targeted mutagenesis for disruption of gene
function is now a routine procedure. Numerous approaches to further enhance the efficiency of the system have been demonstrated, including the choice of the promoter to drive Cas9 (Wang et al., 2015; Yan et al., 2015; Eid et al., 2016) and improved design of the sgRNA scaffold (Dang et al., 2015). In the latter approach, the duplex part of the sgRNA was extended, mimicking the natural system, and the continuous stretch of T’s was mutated, as these function as a termination signal for RNA-Polymerase III, which commonly drives sgRNA expression. Concerning the choice of the target site, in contrast to the situation in animals, there are no robust tools for prediction of sgRNA efficiency in plants, and it is advisable to test sgRNA efficiencies in a transient protoplast system before engaging in the time and cost intensive procedure of transgenic plant production (Ishida et al., 2007).

2. Genome editing specificity

Early investigations of CRISPR/Cas9 genome editing specificity found a relatively high off-target potential in human cells (Fu et al., 2013) and extensive analyses of cleavage efficiency on mismatched targets with base specific resolution are available (Hsu et al., 2013). In plants, potential off-target sites identified by bioinformatics approaches have been sequenced but showed no off-target cleavage (Li et al., 2013; Nekrasov et al., 2013; Zhou et al., 2014; Peterson et al., 2016; Nekrasov et al., 2017; Ueta et al., 2017). Feng et al. (2014) demonstrated high specificity in plants as whole genome sequencing of CRISPR/Cas9 induced mutants could not detect any off-target effects. However, off-target cleavage has been reported in rice, soybean and maize, mainly occurring in gene paralogues with almost identical sequences to the targets (Shan et al., 2013; Jacobs et al., 2015; Svitashev et al., 2015; Li, M et al., 2016).

Editing specificity can be increased using computational sgRNA selection, protein engineering, RNA modifications and improved delivery systems (Tycko et al., 2016). The development and optimisation of bioinformatics tools for designing highly specific sgRNA
and detecting the entire repertoire of potential off-target editing sites allows substantial increases in target specificity (Rousseau et al., 2009; Heigwer et al., 2014; Montague et al., 2014). Prediction of highly specific genome-wide sgRNAs with minimized off-target effects in six crops and two further plant species indicated that 67.9%–96.0% of transcripts have at least 10 specific sgRNA designs, with the exception of maize where only 30% of transcripts allow 10 or more sgRNA designs (Xie et al., 2014). This exception is likely due to the large genome size and ancient polyploidy of maize, suggesting polyploid crops with large genomes such as wheat may also prove challenging to edit. Nevertheless, wheat homeoalleles have been successfully edited simultaneously (Wang et al., 2014), and in some cases the targeting of multiple related genes may be advantageous (Lawrenson et al., 2015).

Further solutions addressing the issue of off-target activity include the use of truncated sgRNAs (Fu et al., 2014) and paired nickases (Ran et al., 2013). The paired nickase approach makes use of the D10A mutant of Cas9. Here, one of the two nuclease domains, RuvC, is inactivated, converting the nuclease to a nickase that induces single strand breaks. Two nicks are then induced in close proximity, ultimately resulting in a mutagenic DSB. Nicks at potential off-target sites are repaired with high fidelity via base excision repair and are thus not mutagenic (Fauser et al., 2014). The system has already been used successfully in Arabidopsis (Schiml et al., 2014; Schiml et al., 2016), rice (Mikami et al., 2016), and maize (Wolter et al., 2017). The dosage of Cas9 and sgRNA delivered also affects off-target editing, with higher enzyme concentrations increasing off-target cleavage (Hsu et al., 2013; Pattanayak et al., 2013). For instance, Ranganathan et al. (2014) used the weaker H1 promoter to express sgRNAs, lowering off-target effects. Off-target cleavage is also affected by the delivery method. Recently, efficient genome editing was achieved in rice and wheat when Cas9 and gRNA were delivered as pre-assembled RNP complexes into immature embryos (Svitashev et al., 2016; Liang et al., 2017). Compared to conventional DNA
delivery, levels of on-target mutagenesis were comparable (up to 9% of regenerated plants), whereas off-target mutagenesis was drastically reduced and only detectable by amplicon deep sequencing. Finally, based on insights from the structure of Cas9 in complex with target DNA and sgRNA (Nishimasu et al., 2014), engineered modifications of Cas9 with dramatically enhanced specificity in human cells were developed (Kleinstiver et al., 2016; Slaymaker et al., 2016).

3. Multiplexing the CRISPR/Cas system

The nature of the CRISPR/Cas system renders it highly amenable for multiplexing approaches, targeting multiple sites in the genome simultaneously. While it is possible to use identical or multiple different promoters (Xing et al., 2014; Ma et al., 2015; Wang et al., 2015), this approach is only feasible for a small number of guides as efficiency drops with increasing construct size. Xie et al. (2015) used the endogenous tRNA-processing system to express multiple sgRNAs from a single synthetic gene consisting of multiple repeats of sgRNA and tRNA. Upon cleavage by the endogenous tRNA-processing RNases, the individual sgRNAs are released (Fig. 3). Using this approach, a maximum of eight genes have so far been edited simultaneously (Xie et al., 2015). Interestingly, Cas9 activity at the individual targets was only slightly reduced when eight genes were targeted compared to when only two or four genes were targeted, indicating low enzyme saturation. Concentration of free Cas9 was thus not a major limitation even when Cas9 was distributed among eight sgRNAs, which suggests that it is possible to target more sites simultaneously although this has not been demonstrated. The tRNA based approach has also been successfully applied in maize (Qi et al., 2016). Moreover, Tang et al. (2016) developed a multiplexing system where sgRNAs are expressed from the same PolIII promoter that drives expression of Cas9, along with a self-cleaving hammerhead ribozyme. After transcription, Cas9 and sgRNAs are separated by ribozyme cleavage sites, releasing functional Cas9 and sgRNAs. Recently, the
endogenous multiplexing capacity of Cpf1 was harnessed in plants, with a single polymerase III promoter used to generate a transcript consisting of several units of direct repeat of the crRNA and target sequence. Cpf1 recognizes the direct repeat sequence and cleaves upstream of the stem loop, releasing functional mature crRNAs (Wang et al., 2017).

4. Using CRISPR/Cas as a site-specific effector

Although the enzyme Cas9 has mainly been used for DNA cleavage, it is a versatile tool. Cas9 can be transformed into a DNA binding protein by mutating its two nuclease domains (Jinek et al., 2012). By fusing the enzyme to an effector domain, it can then be used to guide diverse enzymatic functions to any specific site in the genome (Gilbert et al., 2013) (Fig. 4). The effector domain can also be fused to an RNA binding protein that can interact with an aptamer sequence integrated into the sgRNA (Konermann et al., 2015). CRISPR/Cas9 can thus not only be employed to edit the genome but also to control expression of specific genes (Lowder et al., 2015) and induce sequence-specific epigenetic modifications (Hilton et al., 2015; Thakore et al., 2015). Interestingly, Cpf1 also appears to be a potent tool for effector fusions, as was shown for transcriptional repression in Arabidopsis (Tang et al., 2017).

The availability of various Cas9 orthologues offers the possibility to bring genome editing to a new level. Since different Cas9 orthologues only interact with their species-specific sgRNA (Steinert et al., 2015), diverse effects can be performed in the same cell simultaneously. For instance, one orthologue could guide transcriptional activators to one set of genes, while another orthologue guides transcriptional repressors to a different set of genes. In this way, the outcome of a specific DSB induced by a third orthologue could be influenced by manipulation of the protein machinery involved in the processing of the break, e.g. suppressing NHEJ while enhancing HDR for gene replacement (Puchta, 2016).

VII An emerging new breeding process
1. Breeding schemes using rapid cycles of genome editing

Since the first domestication of crops 10,000 years ago (Doebley et al., 2006), humans have relied on non-targeted approaches such as spontaneous mutation, random mutagenesis and genetic recombination to generate diversity for selecting improved crops. Selection from the genetic diversity generated in this way and subsequent fixation of desirable agronomic traits remain the cornerstone of crop breeding. While transgenic plants form a notable exception to the non-targeted production of diversity within the current breeding process, these crops suffer from a lack of public acceptance and are limited in the scope of improvement they can offer by incorporating novel genes that often have individual and specific effects such as herbicide resistance (Daniell, 2002). With rapid advances in crop genome sequencing, functional genomics and CRISPR/Cas efficiency, we expect a novel crop breeding process to play and increasingly important role (Fig. 5). As progress in sequencing technologies and genome assembly algorithms provide more high quality genome assemblies, and genome editing improves delivery and editing precision with fewer off-target effects, the limiting factor in large-scale crop improvement programs using genome editing will be deciding what to edit. When an integrated knowledge of functional genomics in plants can be incorporated into crop breeding schemes based on genome editing, we believe a fundamental shift in how crops are improved will follow.

Conventional breeding schemes generate genetic variants via recombination or untargeted mutagenesis, applying cycles of backcrossing, gene pyramiding, pedigree breeding or recurrent selection to introduce novel traits into elite varieties. As these approaches do not allow precise control of the genetic material introgressed, they can lead to unpredicted effects such as the introduction of deleterious material genetically linked to desirable traits via linkage drag. For this reason, current breeding and selection schemes are laborious and require multiple rounds of crossing and selection to generate improved varieties. In contrast,
genome editing provides an unparalleled level of control over the mutation process. For instance, Ito et al. (2015) point out that a deletion in the fruit-ripening regulator RIN used in conventional tomato breeding likely spans the regulatory region of a neighbouring gene, interfering with development of flowers and the pedicel abscission zone, which connects fruits to the main body of the plant. The authors disrupted the RIN gene using genome editing to delay fruit ripening without affecting neighbouring genes, illustrating the application of precise mutagenesis.

To generate diverse populations for selection, future breeders can apply genome editing of candidate genes and regulatory regions instead of relying on genetic recombination, random mutagenesis or somatic hybridization. By generating a comprehensive range of novel allelic variants aimed at modifying a specific trait, laboratory trials and subsequent multi-environment field trials will allow trait optimisation, rather than just trait improvement. To optimise traits, the multiplexing capability of the CRISPR/Cas system could be leveraged to harness epistasis (interaction between genes) and redundant pathways, which often present obstacles to conventional breeding. Under this new breeding process based on cycles of genome editing and selection, breeding outcomes are more predictable, and as there is no requirement to remove unfavourable alleles, the production of improved varieties is faster.

Genome editing represents a progression from genomics-assisted breeding approaches based on genetic markers such as MAS and genomic selection. By generating novel allelic variants that directly affect traits, genome editing overcomes the imprecision resulting from the use of markers linked to traits but not directly influencing them. Rather than genomic tools assisting in the breeding process, they will be able to play a more pivotal role as they become the means of identifying and generating variation. Furthermore, genome editing surpasses earlier GM approaches that introduced genes for herbicide resistance and pest resistance into staple crops, because it allows both the introduction of novel genes in precise genomic positions and
can modify the sequence and regulation of existing genes. Although we expect GM and marker-assisted approaches such as MAS and GS to offer further benefits to breeders in the coming years, these approaches will be superseded by genome editing as our functional knowledge of crop genomes grows.

2. Applications of genome editing beyond simple traits

During the history of crop breeding, many agronomic traits have been continuously selected in major crops and may now only allow incremental improvement using current breeding methods. In particular, simple monogenic traits that are easily targeted for breeding may no longer be efficient targets for further improvement as many have been optimised through traditional methods. Most recent applications of genome editing in crop improvement discussed in this review have focused on disrupting one or several genes linked to an agronomic trait, commonly pest resistance. This outcome may also be achieved using conventional breeding, albeit less efficiently. Genome editing, however, has the potential to target complex traits using combinations of gene editing, editing of regulatory elements and genomic rearrangement. As CRISPR/Cas multiplexing allows useful edits to be made simultaneously, multiple traits can be stacked in a new variety within a single generation. This provides a key advantage over conventional breeding methods, which rely on recombination over multiple generations or random simultaneous mutation events to stack traits. Genome editing will be most effectively used for crop improvement when it is used to generate traits that cannot be arrived at through non-specific mutations or the simple addition of a transgene.

Genome editing could be used to increase the efficiency of the key CO₂ fixing enzyme Rubisco and thus improve photosynthesis efficiency and yield (Sharwood, 2017). An ambitious goal is to increase photosynthesis efficiency in C₃ plants such as rice by making
their photosynthesis pathways more similar to those of C₄ plants such as maize. C₄ Rubisco has a faster carboxylation rate and thus higher efficiency than C₃ Rubisco, but is less specific and requires higher CO₂ concentrations, which C₄ plants provide via carbon concentration mechanisms. Important candidate genes for improving photosynthesis efficiency using genome editing are thus genes encoding for Rubisco and for carbon concentration mechanisms. Rubisco consists of large subunits and small subunits, encoded by the rbcL gene in the chloroplast genome and the RbcS multi-gene family in the nucleus. Catalytic switches responsible for faster carboxylation in C₄ plants are known, for instance the large subunit Met-309-Ile substitution (Whitney et al., 2011). A recent study found 88 candidate genes likely involved in the C₄ differentiation process including genes encoding carbon shuttle enzymes and key transporters such as PEPC, PPDK, NADP-ME and OMT and the less well known TPT and NHD-BASS2 (Huang et al., 2017). Although converting crops from C₃ to C₄ photosynthesis will be a challenge, genome editing could help overcome this challenge using multiplexed gene replacement to alter and optimise photosynthesis components. To circumvent the challenge of chloroplast transformation, genes could be encoded in the nuclear genome together with genes encoding transit peptides and membrane transporters for transfer into the chloroplast.

Another application of genome editing is biofortification of staple crops to improve human health. When pathways to accumulate micronutrients do not occur in a crop or do not allow micronutrient accumulation in edible parts of the crop, biofortification is generally not possible using conventional breeding because the changes required in the genome are too complex. Crop metabolic pathways generally involve suites of genes, for example, the pathway for production of the cancer-preventive glucoraphanin includes 13 genes (Mikkelsen et al., 2010). While inserting transgenes encoding metabolic pathways has been a key approach in biofortification, altering endogenous pathways with genome editing may provide
better micronutrient accumulation and control as endogenous genes and regulators are likely better optimised to function together. For instance, the beta-carotene (vitamin A precursor) pathway engineered into ‘Golden Rice’ using the two key transgenes psy and crtI (Ye et al., 2000) could potentially be achieved using editing of the endogenously occurring beta-carotene pathway.

Drought is an important driver in crop yield loss but enhancing drought tolerance is challenging because of the complex networks involved in regulating drought response (Hu & Xiong, 2014). An important target for increasing drought tolerance is the phytohormone ABA, which inhibits growth and helps regulate plant stress response. The discovery of 14 genes encoding receptors that bind to ABA and activate ABA signalling, PYR1 and 13 related PYR1-like genes (PYL) (Park et al., 2009), provides candidates for genome editing to adjust the ABA pathway to reduce yield loss under drought while maintaining plant growth during non-drought periods. Furthermore, a recent review listed 15 genes that are known to increase drought stress survival, biomass or yield in various crops during field trials (Mickelbart et al., 2015). Although there are likely to be trade-offs between drought tolerance genes, stacking combinations of edits of these genes may lead to overall improvements in drought tolerance.

3. Domesticating new crops using genome editing

The majority of calories consumed by humans are derived from only 20 crops, with rice, wheat and maize contributing the largest proportion to the global food supply (Massawe et al., 2016). Over the past decades, the improvement of these major crops has been a priority, and has been followed by an increasing homogenization of food production. However, improvement of major crops may not always be the most efficient approach for generating high-yielding, climate resilient crops. While monogenic traits have been targeted by breeders with considerable success, altering complex polygenic traits such as abiotic stress tolerance remains challenging. To overcome this difficulty, wild species or minor crops with a more
favourable genetic background for these traits could be rapidly domesticated using genome editing.

Recent publications have emphasised the potential for crop domestication using genome editing, listing candidate domestication genes involved in traits including seed shattering, growth architecture and flowering time (Osterberg et al., 2017; Zsögön et al., 2017). For example, in a pioneering study Soyk et al. (2017) characterised and edited the flowering repressor SP5G in tomato, reducing time to harvest by two weeks (Fig. 6). By fine-mapping a QTL region linked to photoperiod response, SP5G was found to cause delayed flowering during long days. Using genome editing to knockout combinations of SP5G and the dosage-dependent flowering activator SFT, the authors showed that SP5G is a likely repressor of SFT, making SP5G a candidate for helping to domesticate tomato relatives and other species by enabling them to grow in more northerly latitudes.

The ongoing domestication of numerous wild plants could be accelerated by targeting orthologs of domestication genes such as SP5G. For instance, in the distant Australian rice relative weeping grass (Microlaena stipoides), a potential crop that is resistant to a range of abiotic stresses, chemical mutagenesis of the rice orthologs qSH1 and sh4 decreased seed shattering (Shapter et al., 2013). Pennycress (Thlaspi arvense), a common weed in Eurasia and North-America, could be domesticated into a valuable, cold tolerant oilseed crop if traits such as oil quality and seed pod shatter could be improved (Sedbrook et al., 2014). Progenitors of staple crops such as the maize progenitor teosinte (Zea mays ssp. parviglumis), wild emmer wheat (Triticum dicoccoides) and common wild rice (Oryza rufipogon) could also be candidates for domestication as they may contain valuable adaptations lost during domestication bottlenecks. Genome editing was recently used in the potential rubber crop Taraxacum koksaghyz to disrupt an expected antagonist of rubber production (Iaffaldano et al., 2016). As more genomes of wild species and minor crops are sequenced, it will become
easier to identify orthologous domestication genes known from well-studied plants that can be targeted with genome editing. If candidate domestication genes are known and the challenges of epistasis and gene dosage can be overcome, genome editing in wild or minor crops will help substantially expand the crop germplasm pool.

4. Regulation of genome edited crops

Strict regulation and nationwide bans, particularly in Europe, have slowed the commercial development and widespread adoption of GM crops since they were first introduced in 1995. Research and development of non-transgenic genome edited crops may be similarly slowed if regulatory authorities treat them as GM crops. In contrast to transgenic GM crops, genome edited crops can be indistinguishable from crops produced by conventional breeding. To differentiate between conventional mutagenesis breeding and genome editing, legislation may therefore attempt to regulate the process of crop improvement rather than the result.

However, advances in the delivery of genome editing machinery such as the use of RNP complexes, which do not require even temporary introduction of foreign DNA into the genome, may rapidly undermine legislation regulating the crop improvement process (Wolter & Puchta, 2017). The European Court of Justice is due to decide in 2018 whether genome edited crops should be regulated as GM crops. The status of genome edited crops in many other countries remains unclear, as legislation lags behind scientific innovation. Regulation of genome edited crops must address rapid changes in the technology and distinguish between transgenic GM crops and non-transgenic genome edited crops. Canada provides an example of a pragmatic approach to biotechnology regulation by evaluating plants on a case-by-case basis with a focus on novel traits rather than the breeding process (Canadian Food Inspection Agency, 2016).

VIII Conclusion
Genome editing technology is likely to play an important role in addressing growing global crop demand in the face of population growth and predicted climate change. Genome editing can facilitate the production of crops with higher yield, improved nutrient content, resistance to pests and tolerance of abiotic stress by accelerating crop improvement schemes and increasing their effectiveness, and this can support food security in developed as well as developing countries. The main hurdles that remain to be overcome to establish a new genome editing process for crop improvement are the assembly of high-quality pangenome references, the systematic inference of candidate editing sites using functional genomics, the improved delivery of genome editing systems and the reduction of off-target editing. For this new crop improvement process to be successful, the accumulation and integration of knowledge in genomics, transcriptomics, phenomics and biotechnology will be essential. The regulation of genome edited crops must also be clarified to support the development of this technology and gain consumer acceptance. Despite the remaining technical and social challenges, just four years after the CRISPR/Cas9 system was first applied to edit a plant genome, large-scale crop improvement using genome editing is on the verge of becoming a reality.

Acknowledgements

Armin Scheben was supported by an IPRS awarded by the Australian government. This work was supported by the Australian Research Council (Projects LP160100030, LP140100537, LP130100061 and LP130100925).

References

**Acquaah G. 2012.** *Principles of Plant Genetics and Breeding.* Chichester: Wiley-Blackwell.


Zsögön A, Cermák T, Voytas D, Peres LEP. 2017. Genome editing as a tool to achieve the crop ideotype and *de novo* domestication of wild relatives: Case study in tomato. *Plant Science* 256: 120–130.
Figures

**Figure 1.** Key differences between the DNA cleaving nucleases clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 9 (Cas9) and CRISPR from *Prevotella* and *Francisella* 1 (Cpf1). Cpf1 recognizes a T-rich protospacer adjacent motif (PAM) (nucleotides highlighted in red) upstream of the target sequence and generates a staggered cut (indicated by red triangles) further from the PAM site than Cas9, leaving a sticky end. In addition, it associates with CRISPR-RNA (crRNA) alone, without trans-activating crRNA (tracrRNA) or a synthetic chimeric RNA. Lastly, the target sequence of Cpf1 is longer, usually 23-25bp as opposed to the 20bp required for Cas9.
Figure 2. Approaches to control recombination with chromosomal rearrangements using clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (CRISPR/Cas). Genetic linkage between traits A and B can be broken by (A) chromosome arm exchanges or (B) intrachromosomal inversions (vertical arrows indicating sequence orientation).
Figure 3. Multiplexing using the endogenous tRNA-processing system. Endogenous ribonucleases recognize and cleave the 5' and 3' ends of the tRNA sequence (shown in dark red) on the primary transcript, releasing functional single guide RNAs (sgRNAs) composed of a unique targeting sequence (shown in blue, purple and orange) and a scaffold sequence (shown in black).
Figure 4. Utilizing clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein (CRISPR/Cas) to regulate gene expression. Catalytically inactive ‘dead' Cas9 (dCas9) is fused to activator or repressor domains and guided to the promoter region of regulated genes. Using multiple Cas9 orthologues, both functions can be performed simultaneously.
Figure 5. Breeding scheme for crop improvement using genomic, transcriptomic and phenomic resources together with genome editing. By integrating various types of data on crop germplasm (wild species, landraces, and varieties), traits can be associated with genomic regions to help identify candidate genomic targets for editing. Genome editing is followed by crossing unwanted regions of the genome such as the CRISPR/Cas genes out of the first
generation transformant (T₀). Optimization of novel allelic variants is carried by selecting the best performing variants in multi-environment field trials. Cycles of editing and selection allow generation of novel elite varieties that can be added to crop germplasm resources.
Figure 6. Crop improvement strategy used by Soyk et al. (2017) to identify, characterise and edit the flowering repressor \textit{SP5G}, and introduce the quantitative trait earliness of yield into...
an elite cultivated tomato variety. Quantitative trait loci sequencing (QTL-seq) and fine mapping identified \( SP5G \), and genome editing and expression analysis characterised its function as a repressor of the florigen gene \( SFT \). Genome editing was then used to disrupt \( SP5G \) in a determinate background, stacking the determinate growth conferred by a mutation in the \( SP \) gene and the earliness of yield caused by disruption of \( SP5G \).
Table 1. Examples of genome editing studies with practical applications of genome editing for crop improvement.

<table>
<thead>
<tr>
<th>Crop</th>
<th>Editing system</th>
<th>Type of edit</th>
<th>Target gene(s)</th>
<th>Target traits</th>
<th>Results</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Brassica</em></td>
<td>CRISPR/Cas9</td>
<td>Gene disruption</td>
<td><em>Bolc.GA4.a</em></td>
<td>Dwarfing and fruit dehiscence</td>
<td>Significant dwarfing in all transformed lines compared to wild type and reduced fruit dehiscence</td>
<td>Lawrenson <em>et al.</em> (2015)</td>
</tr>
<tr>
<td><em>oleracea</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Citrus</em></td>
<td>CRISPR/Cas9</td>
<td>Promotor disruption</td>
<td><em>CsLOB1</em></td>
<td>Citrus canker resistance</td>
<td>Citrus canker disease severity was decreased by 83.2%–98.3% compared to the wild type</td>
<td>Peng <em>et al.</em> (2017)</td>
</tr>
<tr>
<td>(Orange)</td>
<td></td>
<td>(promoter)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Cucumber</em></td>
<td>CRISPR/Cas9</td>
<td>Gene disruption</td>
<td><em>eIF4E</em></td>
<td>Broad virus resistance</td>
<td>Immunity to cucumber vein yellowing virus</td>
<td>Chandrasekaran <em>et al.</em> (2016)</td>
</tr>
<tr>
<td>Plant</td>
<td>Genetic Modification</td>
<td>Treatment</td>
<td>Trait</td>
<td>Outcome</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>----------------------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>--------------------</td>
<td></td>
</tr>
<tr>
<td>Maize</td>
<td>CRISPR/Cas9 Promoter swap</td>
<td>ARGOS8</td>
<td>Drought tolerance</td>
<td>A total of 3.8% higher yield than the control under flowering drought stress</td>
<td>Shi et al. (2016)</td>
<td></td>
</tr>
<tr>
<td>Potato</td>
<td>TALEN</td>
<td>VInv</td>
<td>Low level of reducing sugars</td>
<td>Undetectable levels of reducing sugars that lead to low palatability and possible health risks</td>
<td>Clasen et al. (2016)</td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>CRISPR/Cas9 Promoter disruption</td>
<td>OsSWEET11, OsSWEET14</td>
<td>Bacterial blight resistance</td>
<td>The promoter of the blight susceptibility gene was disrupted</td>
<td>Jiang et al. (2013)</td>
<td></td>
</tr>
<tr>
<td>Crop</td>
<td>Technology</td>
<td>Disruption Method</td>
<td>Targets/Features</td>
<td>Outcome</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>CRISPR/Cas9</td>
<td>Gene disruption</td>
<td><em>Gn1a, GS3</em>,</td>
<td>More grains, larger grains, denser and erect panicles</td>
<td>Li, M et al. (2016)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>and <em>DEP1</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Grain number, grain size, panicle architecture</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soybean</td>
<td>TALEN</td>
<td>Gene disruption</td>
<td><em>FAD2-1A, FAD2-1B</em></td>
<td>The amount of the monounsaturated oleic acid increased from 20% to 80% and the polyunsaturated linoleic acid decreased from 50% to under 4%</td>
<td>Demorest et al. (2016)</td>
<td></td>
</tr>
<tr>
<td>Tobacco</td>
<td>CRISPR/Cas9</td>
<td>Viral gene disruption</td>
<td>43 regions in the viral genome</td>
<td>Resistance to the geminivirus beet severe curly top virus No severe leaf-curving symptoms were observed in systemic leaves, and virus accumulation in local leaves was reduced by</td>
<td>Ji et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>Plant</td>
<td>CRISPR/Cas9</td>
<td>Gene manipulation</td>
<td>Regions in the viral genome</td>
<td>Resistance to the geminivirus</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Tobacco</td>
<td>CRISPR/Cas9</td>
<td>Viral gene disruption</td>
<td>Six regions in the viral genome</td>
<td>Resistance to the bean yellow dwarf virus</td>
<td>Transgenic plants expressing CRISPR–Cas reagents and challenged with geminivirus had a ~5-87% reduced virus load and fewer symptoms</td>
<td>Baltes <em>et al.</em> (2015)</td>
</tr>
<tr>
<td>Tomato</td>
<td>CRISPR/Cas9</td>
<td>Viral gene disruption</td>
<td>Three regions in the viral genome</td>
<td>Resistance to the Tomato yellow leaf curl virus (TYLCV)</td>
<td>Accumulation of viral DNA was decreased, removing or significantly attenuating symptoms of infection</td>
<td>Ali <em>et al.</em> (2015)</td>
</tr>
<tr>
<td>Tomato</td>
<td>CRISPR/Cas9</td>
<td>Gene disruption/knockdown</td>
<td><em>SP5G</em></td>
<td>Time to harvest</td>
<td>Plants could be harvested 2 weeks earlier than controls,</td>
<td>Soyk <em>et al.</em> (2017)</td>
</tr>
<tr>
<td>Tomato</td>
<td>CRISPR/Cas9 Gene disruption</td>
<td>$RIN$</td>
<td>Fruit ripening (shelf life)</td>
<td>Delayed and permanently inhibited ripening, possibly without the side-effects of conventional $rin$ mutants</td>
<td>Ito et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------</td>
<td>-------</td>
<td>----------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>Tomato</td>
<td>CRISPR/Cas9 Gene disruption</td>
<td>$SlIAA9$</td>
<td>Parthenocarpy (leading to seedless fruit)</td>
<td>Fruit development was triggered before fertilization, generating seedless tomato fruits</td>
<td>Ueta et al. (2017)</td>
<td></td>
</tr>
<tr>
<td>Tomato</td>
<td>CRISPR/Cas9 Gene disruption</td>
<td>$SlMlo1$</td>
<td>Resistance to powdery mildew</td>
<td>Rapid generation of a transgene-free tomato fully resistant to powdery mildew</td>
<td>Nekrasov et al. (2017)</td>
<td></td>
</tr>
<tr>
<td>Wheat</td>
<td>TALEN and CRISPR/Cas9</td>
<td>Gene disruption</td>
<td>Resistance to powdery mildew</td>
<td>The number of mildew microcolonies formed on the leaves was significantly reduced against the control and no apparent fungal growth was observed on the leaves of edited plants</td>
<td>Wang et al. (2014)</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>TaMLO-A1, TaMLO-B1 and TaMLO-D1</td>
<td>Wheat TALEN and CRISPR/Cas9</td>
<td>Gene disruption</td>
<td>Resistance to powdery mildew</td>
<td>The number of mildew microcolonies formed on the leaves was significantly reduced against the control and no apparent fungal growth was observed on the leaves of edited plants</td>
<td>Wang et al. (2014)</td>
<td></td>
</tr>
</tbody>
</table>