Planta (2014) 239:921–939
DOI 10.1007/s00425-014-2029-y

EMERGING TECHNOLOGIES

Precision genetic modifications: a new era in molecular biology and crop improvement

Franziska Fichtner · Reynel Urrea Castellanos · Bekir Ülker

Received: 29 November 2013 / Accepted: 6 January 2014 / Published online: 8 February 2014

Abstract Recently, the use of programmable DNA-binding proteins such as ZFP/ZFNs, TALE/TALENs and CRISPR/Cas has produced unprecedented advances in gene targeting and genome editing in prokaryotes and eukaryotes. These advances allow researchers to specifically alter genes, reprogram epigenetic marks, generate site-specific deletions and potentially cure diseases. Unlike previous methods, these precision genetic modification techniques (PGMs) are specific, efficient, easy to use and economical. Here we discuss the capabilities and pitfalls of PGMs and highlight the recent, exciting applications of PGMs in molecular biology and crop genetic engineering. Further improvement of the efficiency and precision of PGM techniques will enable researchers to precisely alter gene expression and biological/chemical pathways, probe gene function, modify epigenetic marks and improve crops by increasing yield, quality and tolerance to limiting biotic and abiotic stress conditions.

Keywords TALE · TALEN · ZFN · CRISPR · Cas9 · Gene targeting

Electronic supplementary material The online version of this article (doi:10.1007/s00425-014-2029-y) contains supplementary material, which is available to authorized users.

F. Fichtner · R. Urrea Castellanos · B. Ülker (✉)
Plant Molecular Engineering Group, Institute of Cellular and Molecular Botany, University of Bonn, Kirschallee 1, 53115 Bonn, Germany
e-mail: ulker@uni-bonn.de

Introduction

Steady growth in global population and food consumption place unprecedented demands on agriculture and natural resources (Griggs et al. 2013). Besides their vital uses as food and feed, plants have become increasingly important as renewable materials and biofuel feedstocks. Plants also have enormous potential as bioreactors for pharma and nutraceuticals. Responding to these current and future demands and coping with changing climate require effective, versatile, safe, rapid and environmentally friendly technologies to improve yield, stress tolerance, disease resistance and biofuel and biomaterial properties of crops.

Classical plant breeding and transgenic geneticmodification (GM) techniques have driven increasing crop yield; however, these techniques lack precision and require substantial time for development of crop varieties. Finally, concerns on gene flow, ecological consequences, toxicity and allergenicity of GM crops have fuelled GM critics. Recent advances in gene targeting and genome editing provide welcome developments in precision crop genome engineering to address shortcomings in crop improvement.

Precision genetic modification techniques

Using precision genetic modification (PGM) techniques, researchers can target specific sequences in complex genomes and precisely introduce modifications including single nucleotide changes. Implementation of these techniques required several breakthroughs, including the ability to engineer recognition of a specific DNA sequence and generate a specific double-stranded DNA break (DSB). Homology-directed repair (HDR) of this site-specific DSB can then produce various alterations in the sequence

for genome editing purposes. The ability to target defined sequences resulted from improved understanding of DNA-binding proteins such as zinc finger type transcription factors (ZFPs) and bacterial transcription activator-like effectors (TALEs). This was followed by the recent discovery of Cas9, a bacterial small RNA-binding protein that can search and cleave DNA sequences complementary to a bound guide RNA. An improved understanding of protein functional domains allowed researchers to engineer proteins that bind specific sequences of interest and perform tasks, including nicking or cleavage of the target sequence, activation or repression of transcription and alteration or removal of epigenetic marks.

Site-specific double-stranded breaks can improve inefficient homologous recombination in plants

Targeted gene editing in many organisms exploits homologous recombination (HR) (Lloyd et al. 2005; Voytas 2013). In this technique, large fragments (1–5 kb) homologous to the target sequence flank the sequence of interest (donor); the homologous flanking sequences mediate insertion of the donor sequence at the target site. However, plants have lower efficiency of HR than yeast, mice and mammals (reviewed in (Voytas 2013; Weinthal et al. 2010). In plants, the low frequency of actual HR events (10^{-3}–10^{-5} gene targeting events per integration) and the high frequency of random integrations limit the usefulness of HR (reviewed in (Voytas 2013). Such low frequencies require screening of tens of thousands of transformants, or the use of negative selection, to identify rare recombinants (Vasquez et al. 2001).

However, introduction of DSB in the target sequence can increase the frequency of targeted events by 100- to 1,000-fold (Puchta et al. 1993, 1996). Organisms repair DSBs by HDR and non-homologous end joining (NHEJ). In HDR, a highly similar sequence in the genome is used as a template for repair and depending on similarity of the sequence used changes to the repaired cite might be introduced. In NHEJ, the broken DNA ends rejoin, often imprecisely, thereby introducing nucleotide changes at the target site [reviewed in (Waterworth et al. 2011)]. These conserved pathways can be used to change the sequence at the target region, including introduction of small insertions and/or deletions, or the insertion of a gene or tags.

One means to introduce a DSB exploits rare cutting meganucleases such as I-*Sce*I (Puchta et al. 1993, 1996). Moreover, introducing the I-*Sce*I recognition sequence in both the gene targeting vector and the target site (Fauser et al. 2012) further improves the efficiency of HR. However, newly developed genome surveying and editing proteins, such as chimeric nucleases, can be designed to target any sequence for binding and introduction of a DSB, thus eliminating the need to introduce a recognition site. So far there is no major limitation as to which sequence can be targeted, although epigenetic state of the target region might pose difficulties.

Zinc fingers: pioneer engineered DNA-binding proteins

Precision genetic modification requires the ability to target DSBs (and other modifications) by engineering a protein that targets a specific DNA-binding site. The first such engineered DNA-binding proteins included Zinc finger (ZF) proteins. Each ZF motif consists of about 30 amino acids that fold into a ββα-structure, in which the Zn^{2+} ion stabilises the conserved Cys_2His_2 residues (Fig. 1a). When binding to DNA, an α-helix of the ZF inserts into the major groove and each ZF interacts with three base pairs (Pavletich and Pabo 1991). To generate artificial proteins that bind to a defined sequence, researchers identified natural fingers with different triplet specificities and generated artificial fingers targeting GNN, CNN, ANN and TNN triplets [reviewed in (Wilson et al. 2013)]. However, ZFs may bind in a context-dependent and unpredictable fashion, requiring iterative selection to get ZF arrays with the desired specificity (Joung and Sander 2013). Nevertheless, many endogenous loci have been targeted at high frequencies using Zinc Finger Nucleases (ZFNs) in *Arabidopsis* (De Pater et al. 2009; Even-Faitelson et al. 2011; Lloyd et al. 2005), tobacco (Townsend et al. 2009) and maize (Shukla et al. 2009). The reader may refer to recent reviews on this topic (Urnov et al. 2010; Weinthal et al. 2010).

*Fok*I and development of chimeric nucleases

Generation of a site-specific DSB also requires a nuclease such as the *Fok*I restriction endonuclease. *Fok*I is a type IIS restriction endonuclease isolated from *Flavobacterium okeanokoites*. In nature, it recognises the non-palindromic pentadeoxyribonucleotide 5′-GGATG-3′:5′-CATCC-3′ in duplex DNA and cuts the two DNA strands 9 and 13 bases downstream from the target sequence, generating four-base, 5′ overhangs (Sugisaki and Kanazawa 1981). *Fok*I contains two separable domains, a 41-kDa N-terminal DNA-binding domain and a 25-kDa C-terminal domain with non-specific DNA-cleavage activity that requires Mg^{2+} (Li et al. 1992). Cleavage of DNA requires dimerisation of the *Fok*I cleavage domain (Bitinaite et al. 1998; Smith et al. 2000; Vanamee et al. 2001), which may release the sequestered catalytic domain (Bitinaite et al. 1998).

The modular nature of *Fok*I allowed construction of chimeric restriction enzymes with novel sequence-specificities.

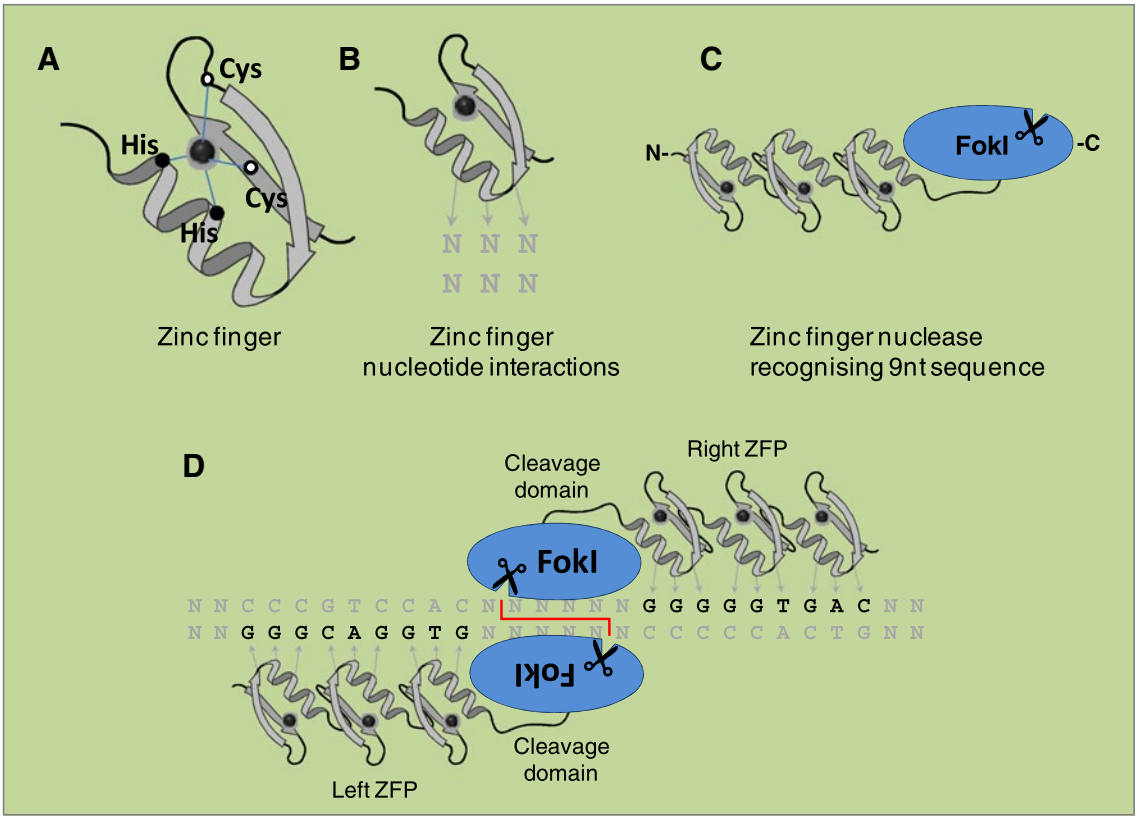

Fig. 1 Zinc finger proteins and zinc finger nucleases in engineering genomes. **a** A typical zinc finger chelating a zinc ion (*black circle* in *centre*). **b** A zinc finger binding to a triplet DNA nucleotide. Multiple zinc fingers can be joined into a functional array targeting any desired sequence for binding. **c** Three zinc finger domains fused to a *Fok*I cleavage domain forming a zinc finger nuclease (ZFN). The recognition site of the ZFN is nine base pairs, where each finger contributes in binding to a precise triplet. **d** Two ZFNs binding on complementary strands, allowing the *Fok*I cleavage domain to dimerise and cleave the double-stranded DNA. Cleavage site is the six base pair spacer between the two ZFN recognition sites. Figure is adapted and modified from Miller et al. (2007) and Urnov et al. (2010)

For example, ZFNs typically consist of three to four ZFs (binding module) and the *Fok*I non-specific nuclease domain (Fig. 1c). The *Fok*I dimerisation requirement has been exploited by generating two independent chimeric proteins targeting adjacent locations that must occur in the correct orientation and appropriate spacing to permit dimer formation for cleavage (Fig. 1d) (Urnov et al. 2010). ZFN pairs containing the *Fok*I nuclease domain function most efficiently when their binding sites are separated by precisely 6 bp (Bibikova et al. 2001).

The use of two chimeric nucleases targeting an adjacent location increases the target binding specificity. For example, one ZFN targeting 9 nt and the other one targeting an adjacent 9 nt sequence increase the target specificity to 18 nt as DNA cleavage requires interaction of *Fok*I cleavage domains of the two ZFNs. With such arrangements it is possible to target a unique region in most genomes and eliminate off-target mutations. However, generation and delivery of two constructs increase the workload and reduce the targeting efficiency. Using two *Fok*I domains attached by a flexible linker to a single DNA-binding protein still allows cleavage of double-stranded DNA, eliminating the need for a second protein (Minczuk et al. 2008). Monomeric nucleases or nickases have also been used in PGM (Beurdeley et al. 2013; Gabsalilow et al. 2013).

Transcription activator-like effectors

To target nucleases and other DNA-modifying enzymes, researchers have harnessed the modular DNA-binding ability of Transcription Activator-like Effectors (TALEs) to generate specific DNA-binding proteins. TALEs were identified from *Xanthomonas* pathogens, which inject TALEs into plants, where they modulate expression of plant genes to promote infection (Fig. 2). A central repeat domain in the TALE mediates DNA recognition and each repeat binds specifically to a single nucleotide. Rearrangement of these repeats allows the user to assemble an engineered protein to recognise a specific target DNA sequence. The binding domain of TALEs, in conjunction with functional domains (fused to the C-terminus) derived from other enzymes such

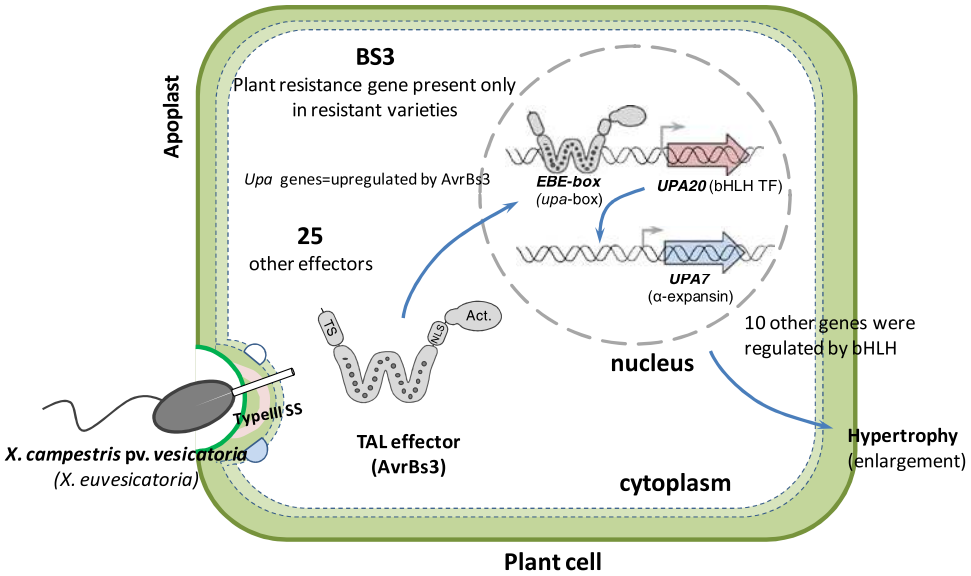

Fig. 2 Discovery of *Xanthomonas* TAL effectors and the proposed mode of action of AvrBs3. AvrBs3 TAL effector protein is secreted into the plant cell via a Type III secretion system. The internal natural nuclear localization signal of AvrBs3 leads to import to the nucleus, where this TALE searches for the base pair sequence recognised by the internal RVD structure of the DNA binding region. Upon binding of the TAL effector to its recognised EBE-box (Effector Binding Element), also known as upa-box, transcription is initiated, leading to physiological effects in the infected plant cell such as hypertrophy. Plant resistance to *Xanthomonas* derives from resistance (*R*) genes having a similar EBE-box and mimicking the natural TALE target site. This leads to enhanced expression of *R* genes upon infection. This figure is adapted and modified from Kay and Bonas (Kay and Bonas 2009)

as nucleases, methylases or repressors, has the potential to be used in genome editing and gene regulation approaches.

Discovery and characterisation of TALEs

The AvrBs3 effector of *Xanthomonas campestris* pv. *vesicatoria* triggers plant defences when the corresponding plant resistance (R) gene *Bs3* is present in pepper [*Capsicum annuum;* (Boch and Bonas 2010; Bonas et al. 1989)]. Comparison of *avrBs3* and its *X. oryzae* pv. *oryzae* homolog *avrXa10* revealed that TALEs have central tandem repeats, which can vary in number and length (Schornack et al. 2013). The sequence of these repeats varies, mainly at amino acids 12 and 13 (Hopkins et al. 1992) (Fig. 3). Fusion of effectors to reporter genes revealed that the members of this large family of homologs in *Xanthomonas* sp. contain an NLS (nuclear localization signal) in the C-terminus, localising them to the plant nucleus (Yang and Gabriel 1995; Van den Ackerveken et al. 1996). Additionally, deletion of C-terminal fragments, which are probably acidic activation domains, from *AvrXa10* and *AvrBs3* indicated that they are essential for triggering plant defences (Fig. 3). Deletions in the N-terminus of *AvrBs3* hindered its transport to the plant cells, indicating that the translocation signal for the bacterial type-III secretion system is located in front of the tandem repeat region (Rossier et al. 1999; Szurek et al. 2002) (Fig. 3).

The clue to the function of these effectors in the plant nucleus came from the identification of a common conserved DNA motif called the *upa*-box or EBE (Effector Binding Element) bound by AvrBs3 located in the promoter region of some pepper genes up-regulated and down-regulated in response to *X. campestris* pv. *vesicatoria* (Kay et al. 2007; Marois et al. 2002; Römer et al. 2007). These EBEs correspond to the repetitive core region of the TALE proteins. Modification of the sequence or number of the 33–34 amino acid repeats causes the TALE to bind different DNA sequences. Thus, the central tandem repeats determine which DNA sequence will be recognised and deletions or rearrangements of these repeats in the protein can modify the binding specificity.

Engineering the TALE DNA-binding code

The variation in the amino acid sequence of the TALE repeats occur mainly at positions 12 and 13 (Hopkins et al. 1992), being designated as RVDs (Repeat Variable Diresidues). Subsequently, It was determined that each RVD recognises a single base in the target sequence, and the binding preferences of individual RVDs were established experimentally and computationally (Boch et al. 2009; Moscou and Bogdanove 2009). The most frequently occurring RVDs in nature are amino acids NI (binding to adenosine), NG (thymidine), NN (guanosine) and

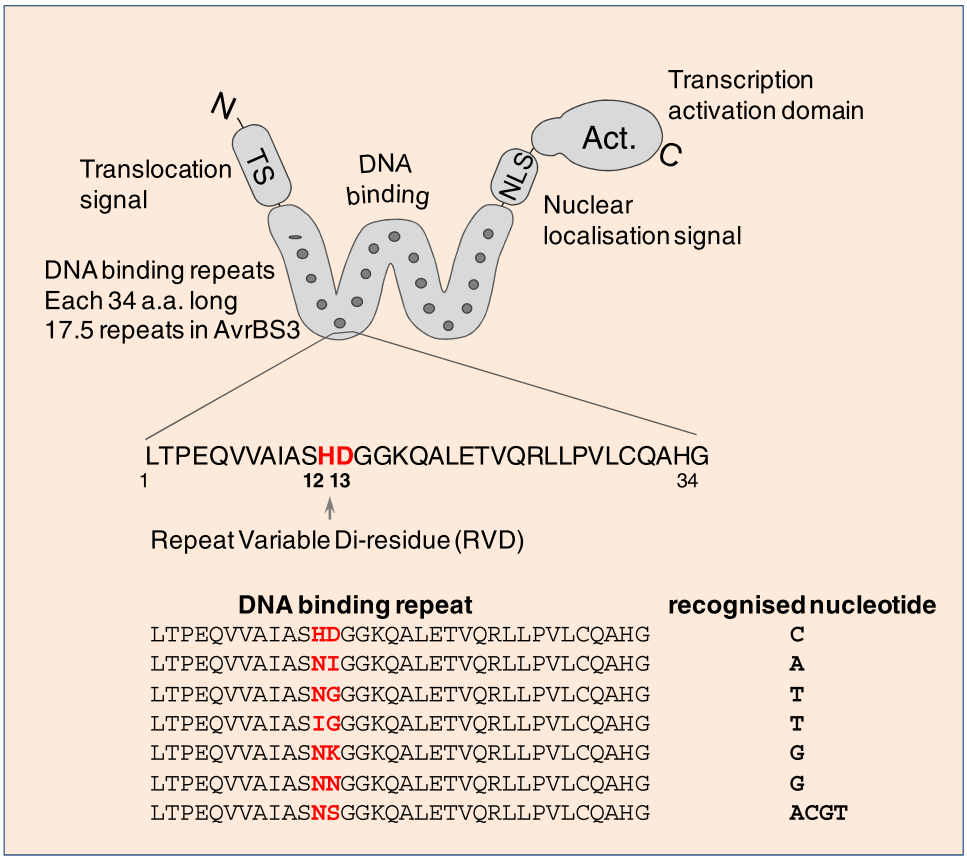

Fig. 3 Typical domain structure of a *Xanthomonas* TALE and its DNA binding specificity. Single base recognition by a 34 a.a. motif differing in the 12th and 13th a.a. (Repeat Variable Di-residue, RVD) in the DNA binding domain. Each RVD combination leads to different nucleotide preferentially bound to the motif. Rearrangement of the individual RVDs leads to a custom recognition sequence allowing binding to almost any target sequence

HD (cystidine). However, the NN RVDs can also bind to adenine and the NH RVDs are highly specific for guanine (Streubel et al. 2012). Secondary structure prediction of complete AvrBs3 TAL effectors revealed a conserved structure of two α-helices flanking a loop domain in each repeat (Boch and Bonas 2010; Schornack et al. 2006). This loop domain harbours the RVD and allows it to access the DNA double helix. Crystal structures of TAL effectors bound to their targets supported the model that TALEs form a superhelical structure wrapping around the DNA at 11 RVDs per helix turn (Deng et al. 2012; Mak et al. 2012), allowing the RVD to face into the middle of the helix and thereby interact with the DNA bases (Schornack et al. 2006). The 12th residue stabilises the structure of the binding domain by interacting with residue number 8, and the 13th residue specifically interacts with the base (Deng et al. 2012; Mak et al. 2012). This protein–DNA interface does not require interaction with any specific host factor (Boch and Bonas 2010).

Using TALEs to activate or repress transcription

TALEs naturally activate plant genes; however, rearranging and assembling the tandem repeats allow researchers to modulate the binding and sequence specificity of TAL effectors. Furthermore, the flexibility of these proteins allows C-terminal fusions of different domains, enabling known DNA modifying enzymes to be precisely bound and function at the target. Although TALEs have a natural activation domain, engineered versions generally include an activation domain from the *herpes simplex* virus VP16, or an artificial tetrameric repeat of the VP16 minimal activation domain, termed VP64 (Boch 2011; Miller et al. 2011). An engineered TALE-TF (TALE-Transcription Factor) was able to increase the transcription of three independent normally silent genes in human HEK293T cells (Perez-Pinera et al. 2013a). Furthermore, several TALE-TFs targeting different regions in the same promoters act synergistically to activate gene expression (Bultmann et al. 2012; Gao et al. 2013; Geissler et al. 2011; Maeder et al. 2013b; Zhang et al. 2011).

TALE-TFs have been effectively used to activate gene transcription in plants. For example, TALE-TFs based on TALEs from *X. campestris* pv. *armoraciae* could upregulate a reporter gene when transiently expressed in *N. benthamina* leaf cells (Boch et al. 2009). TALE-TFs based on AvrBs3 were also used for targeted activation of the endogenous *Arabidopsis* genes *EGL3* and *KNAT1* (Lincoln et al. 1994; Morbitzer et al. 2010; Zhang et al. 2003).

Recently, TALE-TFs were also used to increase *Xanthomonas* resistance in susceptible rice cultivars (Li et al. 2013c). The rice resistance gene *Xa27*, activated by the *AvrBs3* homologous avirulence protein Avr27 from *X. oryzae* pv. *oryzae*, has a susceptible allele termed *xa27*. This allele lacks the EBE and is thus not activated by AvrXa27 (Gu et al. 2005). However, a TALE-TF binding to the *xa27* promoter region could activate transcription and trigger defence mechanisms leading to pathogen resistance (Li et al. 2013c).

An alternative approach to induce plant defences is to incorporate EBEs into the promoter regions of key *R* genes. *R* genes involved in *Xanthomonas* infection are normally recognised by a single TALE that is produced by a specific *Xanthomonas* strain. Consequently, adding different EBEs recognised by different TAL effectors into a single promoter will expand the resistance of a plant species (Römer et al. 2009). For example, introduction of six EBEs in the promoter region of the rice gene *Xa27*, including three EBEs recognised by TALEs from *X. oryzae* pv oryzae PX099^{A} and three from BLS256, broadened rice resistance upon infection by any of the corresponding pathogens (Hummel et al. 2012). However, the possibility to use TALEs for other traits remains to be explored, although their application in other fields, like synthetic biology, is increasing steadily (Blount et al. 2012).

Transcription activator-like effectors fused to known repressor domains can also repress transcription, as shown in mammalian and yeast cells (Blount et al. 2012; Cong et al. 2012), bacteria (Politz et al. 2013) and *Drosophila* (Crocker and Stern 2013). In plants, an AvrBs3-like effector protein (Hax3 from *X. camprestris* pv. *armoraciaceae*) fused to the plant transcription repressor domain SRDX specifically targeting *RD29A* caused a strong reduction in *RD29A* expression (Mahfouz et al. 2012). Therefore, TALEs provide broadly applicable tools to modify gene expression in plants.

Using TALE nucleases (TALENs) for genome editing

When fused to a nuclease domain, TALEs can be used to induce DSBs in the DNA (Christian et al. 2010). TALENs, a pair of TALEs each fused to the catalytic domain of *Fok*I, provide more flexible and efficient tools than ZFNs (Beumer et al. 2013; Chen et al. 2013) and have been used for genome editing in many organisms (Supplementary Table 1). Here, two individual TALENs are targeted to bind to the opposite strands at any desired location. DNA cleavage occurs upon binding of both TALENs and interaction of the *Fok*I domains leading to DSB. Subsequent host DNA repair by NHEJ can produce insertions or deletions at the cleavage site (Puchta and Hohn 2010; Shan et al. 2013a). These DSBs can further alter gene expression by targeting the promoter region, exon, intron or enhancers. In contrast, inserting DNA fragments by HDR can be used to produce more precise alterations (Miller et al. 2011) (Supplementary Table 1).

The feasibility of using TALENs to modify plant genomes was first demonstrated in transient experiments in *N. benthamiana* (Mahfouz and Li 2011). The authors fused a reporter gene downstream of a target construct containing two EBEs flanking a stop codon. Agroinfiltration-mediated delivery of the two TALENs recognising these EBEs led to deletion of the stop codon mediated by DSB resulting in activated reporter gene expression. This confirmed that TALENs can be used to induce deletions/modifications in plants. It was also shown that endogenous loci can be targeted (Cermak et al. 2011; Mahfouz and Li 2011). In *Brachypodium,* it was demonstrated that eight different pairs of TALENs targeting eight endogenous genes were generating indels with an efficiency between 5 to 100 % (Shan et al. 2013a). Furthermore, the repertory of TALE-mediated knock-outs in *Arabidopsis* has been expanded recently (Christian et al. 2013).

Furthermore, TALENs can be used to alter pathogen resistance (Li et al. 2012). In rice, the TAL effector proteins AvrXa7 or PthXo3 (produced by *X. oryzae* pv. *oryzae*) activate the rice susceptibility gene *Os11N3* (also called *OsSweet14*). TALENs targeting the EBE that is bound by these pathogen effectors on the *Os11N3* promoter lead to resistance to *X. oryzae* pv. *oryzae* in the homozygous transgenic rice plants. Of 40 *X. oryzae* pv. *oryzae* strains collected worldwide, 32 are inducing this susceptibility gene, but rice germplasm lacks polymorphisms that prevent *Os11N3* induction. Therefore, the TALE-based approach allows researchers to generate resistance in crops that lack natural variability.

In addition, TALENs have been used for gene insertions by HDR in mammalian (Hockemeyer et al. 2011; Yan et al. 2013; Zu et al. 2013) and plant cells (Zhang et al. 2013). In tobacco, co-delivery of a pair of TALENs and a donor dsDNA with homology to the *ALS* (acetolactate synthase) locus flanking *YFP* (yellow fluorescent protein) generated fusions between the *ALS* genes *SurA* and *SurB* to *YFP*. Depending on the length of the homology region, the integration efficiency ranged from 0.2 to 14 %. Furthermore, it was shown that simultaneous expression of two pairs of TALENs induces large genomic deletions in tobacco, proving the applicability of large gene deletions in plants (Zhang et al. 2013).

Disadvantages of the TALE system

Novel TALEs which are active in vitro sometimes have low or no activity in vivo (Meckler et al. 2013). More precise knowledge of the TAL–DNA interaction will improve downstream applications and target site predictions. It was

demonstrated that TALE binding efficiencies depend on the recognition code and on the combination and position of the RVDs (Meckler et al. 2013; Streubel et al. 2012). Therefore, TALE design should avoid stretches of weak RVDs or repetitions of the same RVD (Streubel et al. 2012). Similarly, RVDs located closer to the N-terminus contribute more to the binding efficiency than RVDs in the C-terminus (Meckler et al. 2013); therefore, between 3 and 4 mismatches can be tolerated in the 3′ end of the binding site (Garg et al. 2012). This increases the probability of off-target effects (Mussolino et al. 2011) or false-positives (Noël et al. 2013). Currently, there are four bio-informatic tools available as web interface and/or standalone software for searching genomes for TAL effector-binding sites (*Target Finder from* TALE-NT 2.0, *Talvez*, *Storyteller*, TALgetter, reviewed in Noël et al. 2013). However, these tools require improvement to accurately predict activity of TALEs in vivo (Chen et al. 2013).

The chromatin context of the target EBE may also affect TALEs activity. 5-methylated cytosine (5mC) and histone acetylation could potentially prevent binding of TALE/TALENs and result in the inability to activate genes (Bultmann et al. 2012) or to induce mutations (Chen et al. 2013). 5mC is found in about 70 % of CpG sites in plants, particularly in CpG islands in promoters and proximal exons (Vanyushin and Ashapkin 2011). Therefore, epigenetic modifiers used along with TALEs (Bultmann et al. 2012) and the repeat N* (instead of HD) may allow TALEs to recognize methylated cytosine (Valton et al. 2012).

The repetitive DNA-binding domains allow TALEs to bind a specific sequence. Although generation of desired repeats and TALE/TALENs as well as cloning into the supplied vectors are straightforward using the Golden Gate kit (Cermak et al. 2011, available from Addgene), in certain cases PCR amplifications are necessary to clone the assembled repeats into user specific vectors. PCR amplification of such repetitive DNAs may fail or generate undesired deletions or artefact products (Supplementary Figure 1) (Hommelsheim, C.M., Frantzeskakis, L., Huang, M. and Ülker, B., manuscript is submitted). Therefore, cloning of TALEs may require special attention.

Continued improvement of the utility of TALEs in PGM will incorporate new insights into TAL–DNA interactions into prediction tools, discover plant genes regulated by natural TALEs and identify novel RVDs with high specificity, to minimise off-target associated problems. Also, accompanying methods to increase target locus accessibility will further improve TALE performance in vivo.

The CRISPR/Cas9 RNA-guided DNA cleavage system

Recent work established the RNA-guided DNA targeting CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated) system for genomic engineering (Fig. 4). CRISPRs were identified in *E. coli* as unusual structures consisting of an array of 29 nt repeats separated by unique 32 nt spacer sequences (Ishino et al. 1987). The spacer sequences derive from extra-chromosomal elements or bacteriophage DNA (Barrangou et al. 2007; Bolotin et al. 2005; Mojica et al. 2005) thus it was demonstrated that this system serves as an innate heritable adaptive bacterial immune system (Barrangou et al. 2007; Deveau et al. 2008; Garneau et al. 2010). Of the three main CRISPR/Cas types (Makarova et al. 2011), only the type-II system has been used to develop RNA-guided gene targeting and genome editing tools.

During the immune reaction the CRISPR locus is transcribed into a long pre-crRNA (CRISPR RNA) which is then processed into a library of short crRNAs by Cas6 in type-I and type-III systems (Brouns et al. 2008; Carte et al. 2008, 2010; Hatoum-Aslan et al. 2011; Haurwitz et al. 2010) or by a housekeeping RNAseIII and the protein Cas9 in type-II systems. In the type-II system, also another RNA called the tracrRNA (*trans*-activating crRNA), encoded on the anti-sense strand upstream of the CRISPR array, functions in processing (Deltcheva et al. 2011). Part of the tracrRNA is complementary to the pre-crRNA repeat regions, thus inducing cleavage by Cas9/RNAseIII complex. After cleavage, the mature crRNA comprises a 20-nt guide sequence derived from the spacer and a 19- to 22-nt sequence derived from the repeat of the CRISPR array. The tracrRNA stays attached to the crRNA, forming a loop structure at the 3′ end (Deltcheva et al. 2011) (Fig. 4).

After processing, each crRNA can be bound by Cas proteins, which scan for complementary DNA sequences. If it finds complementary DNA sequences (like phage or extra-chromosomal DNA), the crRNA forms a RNA:DNA duplex (R-loop) that leads to induction of double-strand breaks by the Cas protein(s). One subtype, the type-IIIB system, interestingly targets and cleaves RNA in vitro and in vivo (Hale et al. 2009, 2012).

Binding and cleaving in the type-I and type-II systems requires the PAM (protospacer adjacent motif), the sequence adjacent to the targeted DNA, either upstream in type-I, or downstream in type-II systems (Fig. 5) (Deveau et al. 2008; Gasiunas et al. 2012; Jinek et al. 2012; Mojica et al. 2009; Sapranauskas et al. 2011; Sinkunas et al. 2013; Westra et al. 2012b). While in type-I and type-III systems, large protein complexes bind and cleave foreign DNA (summarised in (Westra et al. 2012a; Wiedenheft et al. 2012), in the type-II system, binding and cleavage require only the Cas9 endonuclease, the processed tracrRNA and crRNA in vivo and in vitro (Gasiunas et al. 2012; Jinek et al. 2012; Sapranauskas et al. 2011) (Fig. 5).

Cas9 endonucleases have two distinct magnesium-dependent cleavage domains. Upon cleavage, an R-loop is

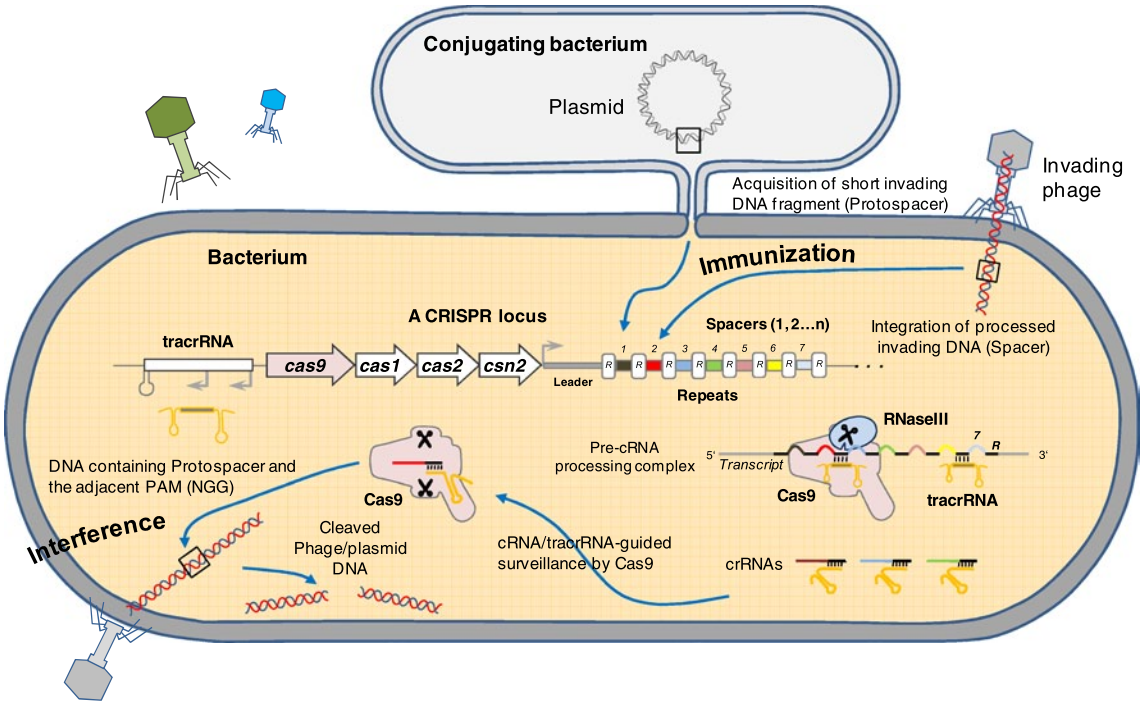

Fig. 4 CRISPR/Cas RNA-mediated DNA cleavage in bacterial immunity. The CRISPR/Cas type-II bacterial immune reaction, which can be divided into immunisation and interference processes. In immunisation, after a phage attack or invasion of an extra-chromosomal element (*grey* bacterial and phage figure), the CRISPR/Cas machinery (clustered regularly interspaced short palindromic repeats/CRISPR associated) recognises foreign DNA. Foreign sequences are cleaved downstream of a PAM (protospacer adjacent motif) sequence (in *Streptococcus pyogenes* 5′-NGG-3′) and integrated into the CRISPR locus, probably by Cas1, Cas2 and Csn2, leading to immunisation. The CRISPR locus is composed of an array of unique spacer sequences derived from foreign DNA elements and identical repeat sequences. In front of each array there is a leader sequence serving as transcription start site and the Cas genes important for the immune reaction. Mostly upstream of the Cas genes, but on the opposite strand, there are also the tracrRNA (*trans*-activating crRNA) coding sequences. Interference starts with transcription of the CRISPR locus into a long pre-crRNA (CRISPR RNA) and of the tracrRNA, which is partially complementary to the repeat sequence on the pre-crRNA, leading to RNA duplex formation. During processing, the precrRNA bound by the tracrRNA is cleaved by an RNAseIII (*blue oval*) and Cas9 into a library of short crRNAs:tracrRNA oligos. Cas9 binds these crRNA:tracrRNA duplexes consisting of 20-nt guiding sequences derived from former spacers of the CRISPR array and of 19- to 22 nt sequences derived from repeat areas. The repeat derived section of the crRNA is still bound by the tracrRNA and stays attached during the immune reaction. After forming a complex with the crRNA:tracrRNA, Cas9 scans the DNA for complementary sequences. Upon detection of a complementary sequence having the PAM downstream of the target site, an R-loop (RNA:DNA base pairing) (see Fig. 5) is formed between the 20-nt guiding sequence and the proto-spacer of the foreign DNA followed by cleavage and destruction of foreign elements by Cas9

formed, letting the RuvC domain cut the single stranded non-complementary DNA strand and the HNH domain cleave the double-stranded complementary strand 3 base pairs upstream of the PAM sequence (Gasiunas et al. 2012; Jinek et al. 2012; Sapranauskas et al. 2011).

The type-II *S. pyogenes* CRISPR/Cas System

In the CRISPR/Cas system of the human pathogen *Streptococcus pyogenes*, the PAM consists of only two nucleotides (5′-NGG-3′; Jinek et al. 2012). Further, this system can be dramatically simplified because a single guide RNA (sgRNA) can mimic the natural loop structure and replace the tracrRNA and crRNA (Jinek et al. 2012) (Fig. 5). Here, the first 20 bp at the 5′ end of the sgRNA function as the guiding sequence, detecting complementary DNA regions. Expressing Cas9 and sgRNA(s) produces genetic modifications via the induction of DSB by Cas9, followed by error-prone NHEJ or more precise HDR. With a codon-optimised version of the *S. pyogenes* Cas9 this system has been used to modify or knock-out genes in many animal systems (Supplemental Table 1) and, recently, in plant cells (Feng et al. 2013; Jiang et al. 2013b; Li et al. 2013b; Mao et al. 2013; Miao et al. 2013; Nekrasov et al. 2013; Shan et al. 2013b; Xie and Yang 2013). Moreover, this system effectively targets heterochromatin in *Drosophila* (Yu et al. 2013) and Cas9 cleavage is not affected by DNA methylation (Hsu et al. 2013). Indeed, in human pluripotent stem cells, the CRISPR/Cas system could target genes that could

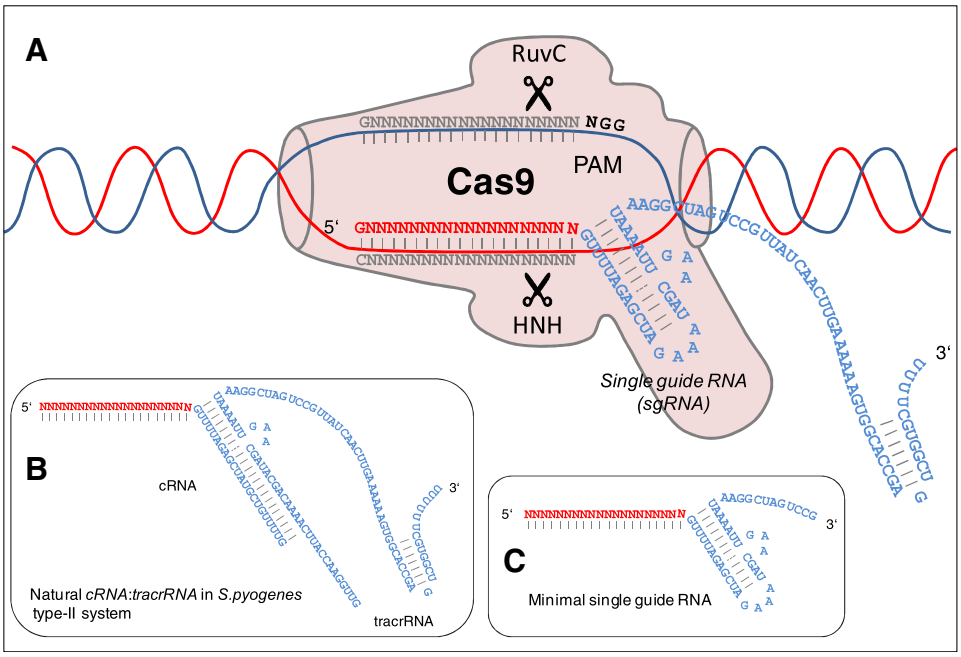

Fig. 5 Cas9 protein scanning to find DNA sequence complementary to the guide RNA. In contrast to the naturally occurring duplex of crRNA:tracrRNA (CRISPR RNA/*trans*-activating CRISPR RNA) bound by Cas9 in *S. pyogenes* (**b**), both RNAs can be replaced by a single guide RNA (sgRNA) that mimics the natural loop structure of crRNA:tracrRNA (**a**, **c**). Combining the necessary regions that allow crRNA and tracrRNA to guide the Cas9 protein produces a minimal scaffold (*blue letters*) for the sgRNA which were shown to be also active (Jinek et al. 2012). An extended, more active version of the sgRNA scaffold (*blue letters*) can also be used (**a**). Binding and cleaving by Cas9 depends on the 20-bp-long target region (*red letters*) and the PAM (protospacer adjacent motif) sequence 5′-NGG-3′ downstream of the target site (*black letters*). When using Cas9 technique-based genome engineering in eukaryotic systems, the sgRNA is expressed by RNA polymerase III U6 promoters, with a Guanine (G in *red letters*) as the first transcribed nucleotide. Therefore, targeting depends on a sequence of 5′-GN_{19}-NGG-3′. Upon binding of the sgRNA/Cas9 complex to complementary DNA regions leading to formation of an R-loop, the RuvC domain of Cas9 cleaves the single-stranded DNA strand and the HNH domain cleaves the RNA:DNA duplex strand (*black* scissors)

not be mutated using TALENs (Ding et al. 2013). Furthermore, the CRISPR/Cas system has a higher mutation rate and less toxicity than ZFN (Fujii et al. 2013).

The CRISPR system can be used to edit multiple sites (multiplex engineering) or delete larger fragments using several sgRNAs because Cas9 requires only the sgRNA as guide to determine the target and the PAM to induce DSB (Chang et al. 2013; Cong et al. 2013; Gratz et al. 2013; Jao et al. 2013; Li et al. 2013b; Mali et al. 2013b; Shan et al. 2013b). In mouse cells, five different loci could be targeted at once using five sgRNAs (Wang et al. 2013). The same mechanisms also enable insertions of DNA fragments, either by single or double cleavage followed by HDR (Chang et al. 2013; Cong et al. 2013; Gratz et al. 2013; Mali et al. 2013b; Wang et al. 2013). Addition of tags or marker gene fusions to a gene of interest has also been shown in mice and human (Ran et al. 2013; Yang et al. 2013) (Supplemental Table 2).

In addition to genomic modifications, a catalytically dead version of Cas9 (dCas9) lacking endonuclease activity can affect gene expression. dCas9 does not cleave dsDNA, but binds to DNA, guided by the sgRNA(s). Upon binding, dCas9 can "hide" transcription start sites or *cis*-elements or block transcript elongation by binding to ORFs of the non-template strand (sense strand), thereby repressing gene expression, as shown in human cells, bacteria and yeast (Bikard et al. 2013; Farzadfard et al. 2013; Qi et al. 2013). Furthermore, dCas9 can be fused to activation or repression domains and modulate expression of one or several genes in parallel (Cheng et al. 2013; Farzadfard et al. 2013; Gilbert et al. 2013; Maeder et al. 2013a; Perez-Pinera et al. 2013b) (Supplemental Table 2).

The sgRNA must be expressed from a fixed transcription start and end site to produce an optimal sgRNA structure. The RNA III polymerase U6 promoter produces a small nucleolar RNA (snoRNA) transcript with an annotated transcription start site, only requiring a guanine as the first nucleotide of the RNA transcript and a termination signal of several thymines in a row (Fig. 6). Expression of sgRNAs from this promoter allows researchers to produce short RNAs lacking secondary terminator structures or a polyadenylation site (Brummelkamp et al. 2002; Miyagishi and Taira 2002; Waibel and Filipowicz 1990; Wakiyama et al. 2005).

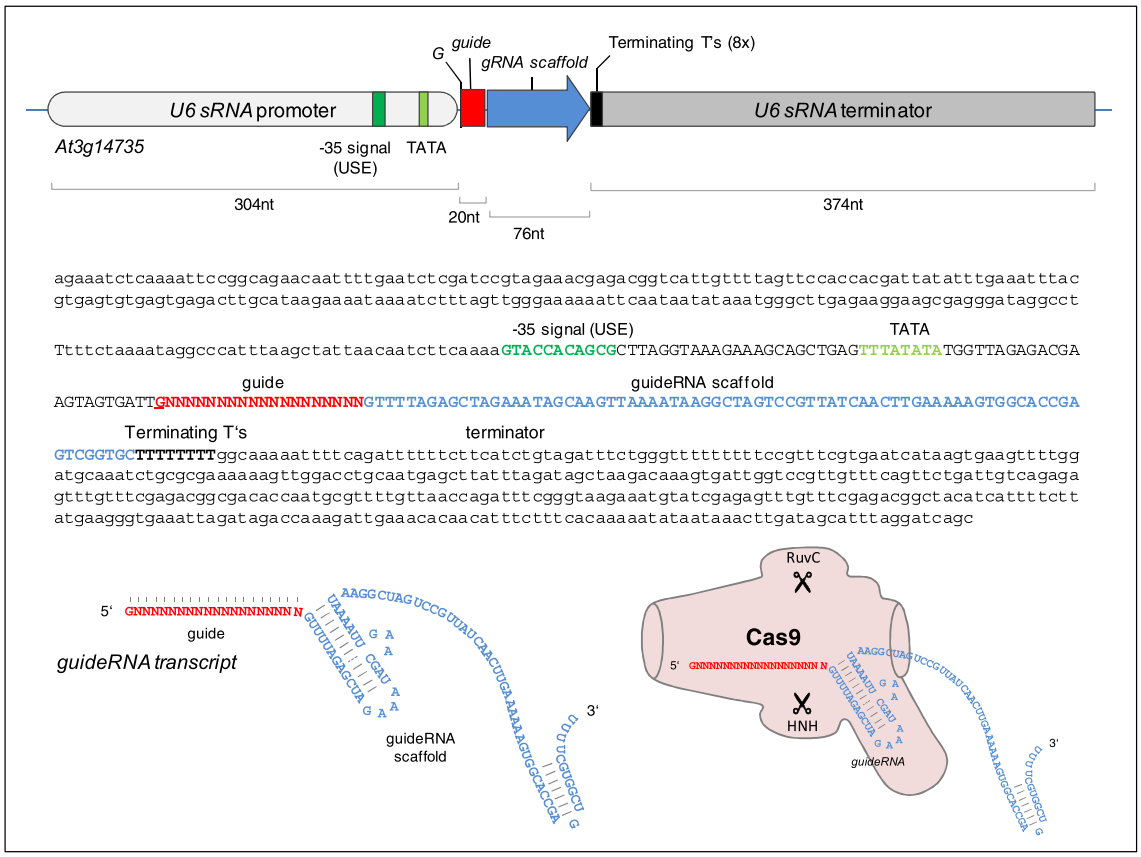

Fig. 6 Single guide RNA expression system with the U6-26 snRNA promoter and termination signal. Guide RNAs can be delivered to plant cells as in vitro synthesised RNA, commercially obtained synthetic RNAs or as DNA expression constructs. Expression of guide RNAs from DNA templates requires carefully selected promoter and terminator sequences. Eukaryotic RNA polymerase III transcribed U6-1/U6-26 snRNAs have accurate transcription start and termination sites. A highly conserved upstream sequence element (USE) and a well-defined TATA box on the promoter region are key determinants for these genes to be transcribed by PolIII. Transcripts under the regulation of the U6-1/U6-26 promoters and terminators all start with a G at the 5′ end and four T's at the 3′ end, and lack polyA tails. Expression constructs that make use of eight T's at the 3′ end of the construct without further extension to the terminator region have been also successfully used as termination signals. Guide RNA transcripts are thought to form a special secondary structure that plays a key role in recognition and binding by Cas9. The first 20 nucleotides at the 5′ of the RNA transcript form the actual guide sequence that pairs with the corresponding complementary DNA sequence which then leads to sequential cleavage of both the guide RNA bound and the unbound DNA strands by the HNH and RuvC nuclease domains

Using the CRISPR/Cas system for genomic engineering

Recently, several groups have used the CRISPR/Cas system for targeted genome editing in crops such as rice and wheat, and model plants such as *A. thaliana* and *N. benthamiana* (Feng et al. 2013; Jiang et al. 2013b; Li et al. 2013b; Mao et al. 2013; Miao et al. 2013; Nekrasov et al. 2013; Shan et al. 2013b; Xie and Yang 2013). CRISPR/Cas systems produce relatively high mutation rates. For example, PEG-based protoplast transformation of rice and wheat with Cas9 and sgRNAs targeting various genes including *Phytoene Desaturase* (*PDS*) resulted in mutagenesis efficiencies of 15–38 % (Shan et al. 2013b) or 3–8 % (Xie and Yang 2013) (Supplementary Table 2). Particle bombardment of callus tissue led to mutation rates around 7–9 % (Shan et al. 2013b). In another study, PEG-based protoplast transformations of *Arabidopsis* and *N. benthamiana* and targeting various endogenous genes including *PDS* produced mutation efficiencies between 1 and 7 % (*Arabidopsis*) or around 38 % (*N. benthamiana*; Li et al. 2013b).

By contrast, agroinfiltration of CRISPR constructs targeting *PDS* in *N. benthamiana* leaves only led to mutation frequencies between 2.7 and 4.8 % (Li et al. 2013b) or 2.1 %, respectively (Nekrasov et al. 2013).

In addition to single modifications, CRISPR/Cas systems can be used for multiplex genome editing in plant cells. Using either a sgRNA with two identical target sites in two loci or two sgRNAs with distinct target sites in one gene can produce mutations at all sites in *Arabidopsis* protoplasts. The single sgRNA produced mutation frequencies of around 2.5 % and the two sgRNAs produced deletions between the two cleavage sites with a frequency of 7.7 %

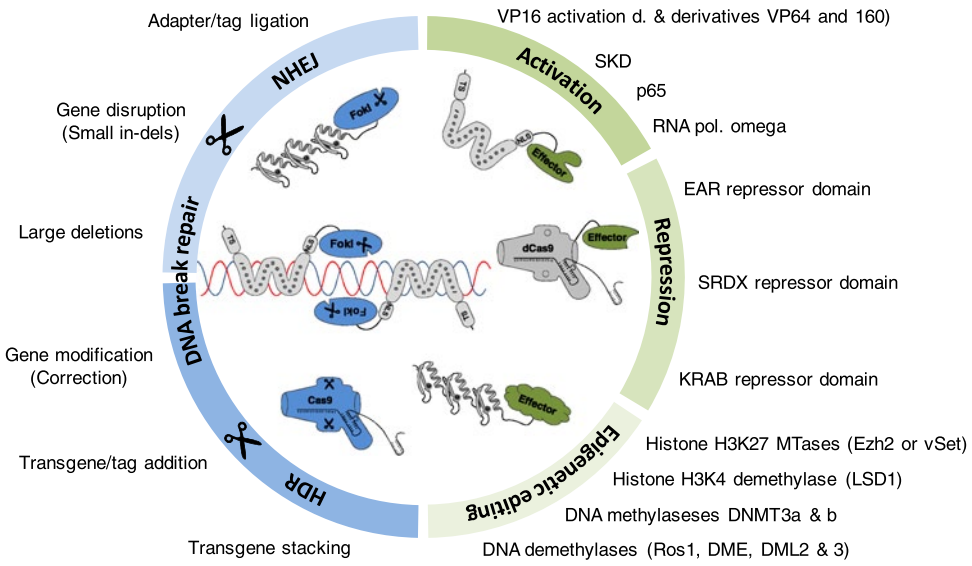

Fig. 7 Overview of possible PGM applications. Different PGM techniques can be used depending on the intended modification (Table 1). The target sequence should be locus-specific to minimise off-target bindings and modifications. *Blue sections* Nucleic acid modification based on cleavage or nicking of the double-stranded DNA. All break-repair dependent modifications rely on cleavage by *Fok*I dimers or Cas9. These modifications use host repair mechanisms, such as non-homologous end joining (NHEJ), which produces small insertions and deletions, or homology-dependent repair (HDR), which can incorporate donor templates into the genome. A pair of target sites in desired proximity can also be used in deleting small or large DNA fragments from genomes. *Green sections* Gene regulation-based activating, repressing or epigenetic editing PGMs, obtained by different domain fusions. Fused activation domains can recruit the host transcription machinery to the target site and initiate or enhance transcription. Fused repression domains can decrease transcription. Fused DNA methylation or histone modification domains can alter epigenetic modifications to modulate transcription and chromatin state. dCas9 is a mutant form of Cas9 where the HNH and the RuvC nuclease domains of the protein are mutated to enable dCAS9 perform other tasks than cleaving the DNA

(Li et al. 2013b). In addition, transgenic *Arabidopsis* lines were generated using two sgRNAs targeting the same gene. Both sites were targeted with a frequency of 74 %; also, 26 % of these had deletions between the two target sites showing that CRISPR/Cas systems can produce deletions in plants (Mao et al. 2013).

In addition to deletions, CRISPR/Cas has been used for other applications in plants. For example, gene replacement by HDR using dsDNA with homology arms was demonstrated in *N. benthamiana* protoplasts and *Arabidopsis* seedlings. Integration of the DNA oligonucleotide occurred in 10–18 % of events (Feng et al. 2013; Li et al. 2013b; Mao et al. 2013). Further, CRISPR/Cas systems can be used to generate stable lines. For example, targeting two *Arabidopsis* genes that lead to a severe phenotype when knocked-out (*BRASSINOSTEROID INSENSITIVE 1* and *GIBBERELLIC ACID INSENSITIVE*) produced T1 progeny showing the expected phenotype of dwarfed growth. Targeting *YOUNG SEEDLING ALBINO* in rice caused colourless albino progeny (Feng et al. 2013). Furthermore, targeting two rice genes with severe phenotypes produced biallelic genome editing in rice with frequencies between 13 % (*LAZY1*) and 50 % (*CHLOROPHYLL A OXYGENASE 1*) (Miao et al. 2013).

The high variation seen for the mutation efficiencies in using the CRISPR/Cas system in plant cells can be partly explained with the biased mutation rate determination systems (either SURVEYOR or amplification/restriction analysis). On the other hand, the type of codon-optimisation as well as the structure of the sgRNA's scaffold seems to be important factors, affecting the mutation rate (Belhaj et al. 2013). This shows that the system is not as easily adaptable as previously assumed because many factors are influencing the Cas9 cleavage efficiencies and thus acquired mutation rates. However, in generating stable lines higher mutation rates seem to be achievable. More stable lines should be generated to gain further insight into the applicability of the CRISPR/Cas system in plant cells.

Disadvantages of the CRISPR/Cas systems

Non-specific binding can cause problems in the CRISPR/Cas system, which tolerates mismatches depending on their position within the sgRNA, allowing Cas9 to cleave off-target sequences. Although in mammalian systems Cas9 causes a high degree of off-target effects (Fu et al. 2013; Mali et al. 2013a; Pattanayak et al. 2013), the limited data in plants indicate fewer off-target effects, which have been detected in rice (Shan et al. 2013b; Xie

and Yang 2013), but not in *Arabidopsis* or *N. benthamiana* protoplasts yet (Li et al. 2013b; Nekrasov et al. 2013).

The targeting nucleotides adjacent to the PAM sequence, called the "seed" sequence, provide crucial specificity. In vitro experiments showed that at least 13 nucleotides upstream of the PAM must match perfectly between RNA and target DNA to induce cleavage, but up to six contiguous mismatches are tolerated in the 5′ end (Jinek et al. 2012). In human cells and bacteria, only the first 12 bp next to the PAM have to match to induce cleavage (Cong et al. 2013; Jiang et al. 2013a). Accordingly, studies using dCas9 in human cells and bacteria showed no significant in vivo off-target effects by RNAseq (Jao et al. 2013; Qi et al. 2013) or microarray analysis (Cheng et al. 2013). Nonetheless, sgRNA/Cas9 can recognize and cleave sequences even if there are mutations in the seed or PAM sequence (Fu et al. 2013; Mali et al. 2013a; Pattanayak et al. 2013). Thereby high concentrations of sgRNA/Cas9 complexes cleave off-target sites at a higher rate. However, these off-target sites seem to be not cleaved when there is a limited supply of sgRNA/Cas9 complexes (Fujii et al. 2013; Hsu et al. 2013; Pattanayak et al. 2013). This suggests a tradeoff between activity and specificity; shorter, less-active sgRNAs can be used to provide higher specificity, as shown in human and mice cells (Fujii et al. 2013; Hsu et al. 2013; Pattanayak et al. 2013).

Reduction of off-target effects may involve modification of the current system or use of alternative CRISPR/Cas systems. A simple extension of the sgRNA targeting sequence does not increase specificity (Ran et al. 2013). However, similar to TALENs, using two distinct sgRNAs but only one Cas9 nicking enzyme can increase specificity by 50- to >1,500-fold compared to using one sgRNA (Ran et al. 2013). In addition, the CRISPR/Cas type-II system of *Neisseria meningitides* has been used for genome editing in human stem cells. This system requires a longer spacer sequence (24 nt) and a longer PAM sequence (5′-NNNN-GATT-3′ or 5′-NNNNGCTT-3′), potentially increasing its specificity (Hou et al. 2013).

Comparison of PGM techniques

PGM techniques can be used in many diverse applications. An overview of possible PGM applications are illustrated in Fig. 7. The Cas9 and TALE systems can be more precise and effective than ZF PGM systems. Especially the use of two TALENs for editing increases target specificity and reduces off-target cleavages. In the TALE/TALEN system, each target site requires a specific TALE/TALEN, thus limiting editing of multiple sites at once. By contrast, the Cas9 system uses a single protein with only a new designed short sgRNA for each site, thus facilitating multiplex applications. See Table 1 for a more detailed comparison of the PGM techniques.

Future prospects for PGM

Non-transgenic genetic modifications

Transient delivery of PGM proteins or their coding mRNAs might eliminate the need to transform plants with the DNA coding sequences of the PGM enzymes. PGM proteins or their coding mRNAs could be transiently transfected to protoplasts, pollen, egg cells, cells in culture, or plant tissues. For example, microinjection of mRNAs encoding ZFNs or TALENs into cricket embryos effectively targeted a transgene or endogenous gene for specific mutations (Watanabe et al. 2012). Similar applications of mRNA injections have also been used in zebrafish, mice and rat (Hwang et al. 2013; Jao et al. 2013; Li et al. 2013a). Viral delivery of RNA transcripts may also enable non-transgenic precision genome editing. Tobacco rattle virus, an RNA virus, was used to deliver ZFNs to tissues and cells of intact plants and produced genome edited tobacco and petunia plants (Marton et al. 2010). Alternatively, transgenic lines carrying coding sequences for PGM proteins under the control of an inducible promoter can be crossed with elite lines. After the desired modification occurs, the transgenes can be crossed out of the genetic background, leaving a non-transgenic line carrying the desired modification.

If no DNA integrates into the genome, the resulting plants would not be considered transgenic. The results would be analogous to chemical mutagenesis used to induce random mutations. Mutagenesis requires selection or screening for a desired phenotype or specific mutations; with PGM, the desired mutation can be rapidly introduced and identified. However, PGM-based reverse genetics requires substantial understanding of the effects of specific mutations.

Other limits of PGMs

Precision genetic modification techniques offer high precision, but require a strong understanding of the system to design the modifications of interest. However, the functions of many genes and the connections in complex regulatory networks remain unclear. Therefore, we cannot always predict how to modulate a given pathway to produce a specific phenotype. Thus, we cannot alter many traits regulated by quantitative trait loci. Classical plant breeding will still be necessary to select gene combinations that lead to desired phenotypes. Similarly, the lack of sequence data from many crop species limits sequence-based targeting, alterations and prediction/analysis of off-target effects. However, PGMs can improve our understanding of these complex systems, supplementing traditional breeding.

Table 1 Comparison of precision gene targeting and genome editing systems

	ZFNs	TALENs	Cas9
Protein size (kDa)	40	105	160
Components	Two ZFNs addressing two adjacent locations on DNA	Two TALENs addressing two adjacent locations on DNA[a]	Cas9 and sgRNA
Assembly	Difficult and slow	Was difficult and slow but now improved	Easy and rapid
Cost[b]	High	Relatively high	Cheap
Commercial source	Yes	Yes	Yes
Public resources	Zinc finger consortium, Addgene	Addgene (https://www.addgene.org/)	Addgene
Target size (bp)	18–24[d] (2 × 9 or 2 × 12)	11–23[c] 22–48[c] (2 × 11–24)	20[c]
Target required to have	GNN triplets or no specific requirements[e]	T at the 5′ (preferred)	N_{20}NGG GN_{19}NGG
Off-target	High but can be reduced	Less but can be eliminated	High but depends on the sgRNA structure[f] Double nicking was also found to be effective[g]
Multiplexing[h]	Difficult	Difficult	Easy
Double-strand breaks	Yes	Yes	Yes
Nicking	Yes (one *Fok*I dimer is mutated)	Yes (Gabsalilow et al. 2013)	Yes[i]
DSB induced mutations	Yes	Yes	Yes
DSB induced large fragment insertions/deletions	Yes	Yes	Yes
DSB induced gene corrections using oligos	Yes	Yes	Yes
Tagging genes with epitopes and assayable markers	Yes	Yes	Yes
Activation of gene expression	Yes	Yes	Only with dCas9
Repression of gene expression	Yes	Yes	Only with dCas9
Epigenetic editing[j]	Yes	Yes	Not yet

[a] A single chimeric TALEN containing two *Fok*I cleavage domains separated by a short flexible linker was also shown to be functional (Minczuk et al. 2008)

[b] Currently couple of thousand Euros for ZFN and TALENs but expected to be couple of hundred Euros for Cas9 and gRNA

[c] Naturally occurring protein

[d] Engineered version

[e] The zinc fingers were found to behave in a context-dependent and often unpredictable fashion requiring iterative selection to get ZF arrays with desired specificity (Joung and Sander 2013)

[f] Hsu et al. (2013), Pattanayak et al. (2013)

[g] Paired nicking induced by aCas9 nickase mutant and paired single guide RNAs can reduce off-target activity by 50- to 1,500-fold in cell lines and facilitate gene knockouts in mouse zygotes (Ran et al. 2013)

[h] Multiplexing (targeting several locations in a given genome at once)

[i] Requires mutation in either HNH or RuvC nuclease domain of the Cas9

[j] A term used to indicate that epigenetic state of the target site can be manipulated by altering chromatin structure by using enzymes modifying histones or DNA

Agronomic traits from other germplasms

Loss of genetic variability because of extensive monoculturing of a limited number of crop varieties and difficulties and cost-intensive maintenance of genetic resources limits crop breeding. Also, most of the disease, pest and abiotic stress resistance genes occur in wild-type progenitors and old landraces. However, discovering and transferring these and other characters to elite crop varieties requires sequence information, bioinformatic analysis and associated phenotypic data. With decreasing cost of genome sequencing and availability of better bioinformatics tools,

we will be able to tap into this valuable information. PGM techniques should allow researchers to make allelic changes in desired genes in modern lines using sequence information from old landraces or re-introduce the missing gene by HDR. Even sequences and genes from extinct plant species can be resurrected using PGM techniques.

Engineering organellar genomes using PGMs

Transformation and engineering of chloroplast and mitochondria organelle genomes will improve our understanding of organelle gene functions and enable engineering of metabolic pathways such as increasing photosynthetic output of C3 plants or altering the amount and content of organelle-produced metabolites. Other advantages include robust protein expression, high transgene copy number and no detected gene silencing (Chong-Pérez and Angenon 2013; Iamtham and Day 2000). Organelle genome engineering may also limit gene flow of transgenes, by maternal inheritance (i.e. no transgene inheritance through pollen) or by engineering of novel cytoplasmic male sterility lines.

Mitochondrial genomes have been targeted using PGM in humans (Minczuk et al. 2006, 2008), using an enzymatically active chimeric ZF protein to bind a mutant mtDNA sequence and induce methylation with the catalytic domain of human DNMT3a. Methylation was chosen deliberately as a marker activity because it is practically absent from mtDNA (Maekawa et al. 2004). The engineered ZF-methylase produced sequence-specific modification of mtDNA in the vicinity of the target site (Minczuk et al. 2006). So far, PGM engineering of chloroplast genomes remains to be tested.

The main challenge in engineering mitochondrial and chloroplast genomes is that each organelle contains hundreds of genomes (Thomas and Rose 1983). Therefore, unless selected for, each cell would be heteroplasmic, with WT and engineered genomes. Organellar modification requires the development of effective technologies to retain the modified sequence. Nuclease-based cleavage of unmodified organelle DNA and protection of modified DNA offer one potential means to enrich for homoplasmic organelles.

New resources for basic and applied research

The availability of genomic resources and tools, such as mutant collections, can limit research, even in *Arabidopsis*. For example, T-DNA insertion mutants exist for 95 % of *Arabidopsis* genes, but many genes lack insertions and many of the available insertions are in introns, UTRs, promoter or terminator regions. Furthermore, since the T-DNAs contain marker genes and their regulatory sequences, they can influence flanking genes and potentially other loci (Ülker et al. 2008). Last, generation of point mutations and small or large deletions remain limiting and difficult to identify with other technologies. Current PGM techniques give 1–30 % targeting efficiency; this can enable generation of null mutants, allelic series of genes, deletions of a few to thousands of base pairs and addition of defined length DNA fragments to specific sites.

Furthermore, the ease and cost effective generation of lines by PGM should facilitate research in non-model plant species and crops. For example, non-*Agrobacterium* based delivery of the PGM proteins and mRNAs should enable researchers to engineer the genomes of hard to transform plants and generate useful allelic series of genes in many other crops. PGM may also prove useful to modify the regulatory sequences of agriculturally and economically important genes. Such modifications will enable important basic science and gene function analyses and could also be useful in improving crop performance and increasing yield.

Precision genetic modification may also be useful to tag genes and defined regions in genomes with epitopes or fluorescent markers, enabling researcher to monitor the expression, localization and regulation of nucleic acids and proteins in their native locations. Such tags could also be used to trace DNA from genetically modified crops or elite lines, thus facilitating breeding and estimation of gene flow.

The ease of generating mutations or tagging genes in complex genomes should also accelerate identification of genes for important traits in crop improvement. For example, targeted editing of each candidate gene in a genetically defined region could identify the responsible gene and thus accelerate crop improvement. Targeted mutations will also allow construction of double mutants in closely linked loci, and targeted deletions will allow identification of true null phenotypes.

Important developments needed

All PGM techniques display off-target effects to varying degrees. Therefore, the specificity of these techniques requires improvement. Adaptation of the heterodimer *Fok*I nuclease structure needed to address two adjacent locations significantly reduced the off-target effects of ZFNs and TALENs. In the Cas9 system, using other CRISPR/Cas systems requiring longer sgRNAs or developing Cas systems addressing adjacent locations with two sgRNAs may address off-target effects.

As PGM technology develops, companies will produce kits for facile cloning, effective delivery and rapid screening of plants. Currently, Addgene produces cloning and limited vector systems for TALE/TALENs. Also, customised ZFNs (Sigma) and TALE/TALENs (Cellectis,

ThermoFischer) can be ordered. Companies will likely produce CRISPR/Cas9 based kits and customised guideRNA(s). Researchers, both in academia and industry, will also likely develop PGM-based genomics resources. For example, fully characterised transgenic lines carrying Cas9 under an inducible promoter and transient expression constructs and systems, as well as methods for delivery of protein/mRNAs directly to plant cells, would be extremely useful for non-transgenic genome editing.

Most of the data obtained from PGM techniques currently derive from transient assays or plants early in development. Better assessment of PGM techniques requires data from stable modifications and all developmental stages. Improved delivery and expression methods will also improve the effectiveness and safety of PGM technologies.

Targeting RNA for modifications

Unlike *P. pyogenes*, the *P. furiosus* CRISPR/Cas complex targets complementary invading RNAs for cleavage (Hale et al. 2009, 2012). Also, the code of sequence-specific RNA recognition by Pumilio and FBF homology repeat proteins has been uncovered (Filipovska et al. 2011, 2012) and determination of the binding specificity of the RNA-editing Pentatricopeptide Repeat proteins is underway (Ban et al. 2013; Barkan et al. 2012; Ke et al. 2013; Laluk et al. 2011; Yagi et al. 2013; Yin et al. 2013). These developments will enable numerous applications to eliminate, regulate or modulate RNA. The resulting techniques could be used, for example, in developing plant resistance to RNA viruses.

Conclusions

Emerging developments in PGM techniques are timely, occurring with growing concerns about the future of agriculture, human population increases and climate change. These precise, controlled, new technologies may provide better, faster, cheaper and safer alternatives to current genetic modification techniques and supplement classical crop breeding methods requiring extensive selection, backcrossing, time and effort. The diversity, versatility and simplicity of PGM technologies would allow broad innovation in crop improvement and the promised safety and precision of PGMs should defer many public critics of genetic modification technologies. Perhaps naming the crops that will be generated using these PGM technologies as PGMO (Precision Genetically Modified) instead of just GMO (Genetically Modified) will help differentiate them from previous technologies and speed their acceptance.

Acknowledgments Research in the Ülker lab is supported in part by a grant from the German Research Foundation, DFG (UE 160/1-1) and basic funding from IZMB, University of Bonn. We thank members of the Ülker lab particularly Carl Hommelsheim, Désirée V. Waidmann and Tobias Berson for reading the manuscript, contributions to the literature survey and for many helpful discussions.

References

Blount BA, Weenink T, Vasylechko S, Ellis T (2012) Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology. PLoS ONE 7:e33279–e33279

Ban T, Ke J, Chen R, Gu X, Tan MH, Zhou XE, Kang Y, Melcher K, Zhu JK, Xu HE (2013) Structure of a PLS-class pentatricopeptide repeat protein provides insights into mechanism of RNA recognition. J Biol Chem 288:31540–31548

Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet 8:e1002910

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. Science 315:1709–1712

Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 9:39

Beumer KJ, Trautman JK, Christian M, Dahlem TJ, Lake CM, Hawley RS, Grunwald DJ, Voytas DF, Carroll D (2013) Comparing ZFNs and TALENs for gene targeting in Drosophila. G3 (Bethesda) 3:1717–1725

Beurdeley M, Bietz F, Li J, Thomas S, Stoddard T, Juillerat A, Zhang F, Voytas DF, Duchateau P, Silva GH (2013) Compact designer TALENs for efficient genome engineering. Nat Commun 4:1762

Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim Y-G, Chandrasegaran S (2001) Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. Mol Cell Biol 21:289–297

Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res 41:7429–7437

Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I (1998) *Fok*I dimerization is required for DNA cleavage. Proc Natl Acad Sci USA 95:10570–10575

Boch J (2011) TALEs of genome targeting. Nat Biotechnol 29:135–136

Boch J, Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48:419–436

Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512

Bolotin A, Quinquis B, Sorokin A, Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 151:2551–2561

Bonas U, Stall RE, Staskawicz B (1989) Genetic and structural characterization of the avirulence gene avrBs3 from *Xanthomonas campestris* pv. *vesicatoria*. Mol Gen Genet 218:127–136

Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV, van der

Oost J (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321:960–964

Brummelkamp TR, Bernards R, Agami R (2002) A system for stable expression of short interfering RNAs in mammalian cells. Science 296:550–553

Bultmann S, Morbitzer R, Schmidt CS, Thanisch K, Spada F, Elsaesser J, Lahaye T, Leonhardt H (2012) Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res 40:5368–5377

Carte J, Wang R, Li H, Terns RM, Terns MP (2008) Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. Genes Dev 22:3489–3496

Carte J, Pfister NT, Compton MM, Terns RM, Terns MP (2010) Binding and cleavage of CRISPR RNA by Cas6. RNA 16:2181–2188

Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82

Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X, Xiong J-W, Xi JJ (2013) Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Res 23:465–472

Chen S, Oikonomou G, Chiu CN, Niles BJ, Liu J, Lee DA, Antoshechkin I, Prober DA (2013) A large-scale in vivo analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. Nucleic Acids Res 41:2769–2778

Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, Rangarajan S, Shivalila CS, Dadon DB, Jaenisch R (2013) Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res 23:1163–1171

Chong-Pérez B, Angenon G (2013) Strategies for generating marker-free transgenic plants. InTech

Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186:757–761

Christian M, Qi Y, Zhang Y, Voytas DF (2013) Targeted mutagenesis of arabidopsis thaliana using engineered TAL effector nucleases (TALENs). G3 (Bethesda) (early online)

Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3:968

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 339:819–823

Crocker J, Stern DL (2013) TALE-mediated modulation of transcriptional enhancers in vivo. Nat Methods 10:762–767

De Pater S, Neuteboom LW, Pinas JE, Hooykaas PJJ, Van Der Zaal BJ (2009) ZFN-induced mutagenesis and gene-targeting in Arabidopsis through *Agrobacterium*-mediated floral dip transformation. Plant Biotechnol J 7:821–835

Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Za Pirzada, Eckert MR, Jr Vogel, Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471:602–607

Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, Shi Y, Yan N (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. Science 335:720–723

Deveau H, Barrangou R, Garneau JE, Labonte J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. J Bacteriol 190:1390–1400

Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K (2013) Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. Cell Stem Cell 12:393–394

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

Farzadfard F, Perli SD, Lu TK (2013) Tunable and multi-functional eukaryotic transcription factors based on CRISPR/Cas. ACS Synth Biol 2:604–613

Fauser F, Roth N, Pacher M, Ilg G, Sanchez-Fernandez R, Biesgen C, Puchta H (2012) In planta gene targeting. Proc Natl Acad Sci USA 109:7535–7540

Feng Z, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, Zhu JK (2013) Efficient genome editing in plants using a CRISPR/Cas system. Cell Res 23:1229–1232

Filipovska A, Rackham O (2012) Modular recognition of nucleic acids by PUF, TALE and PPR proteins. Mol BioSyst 8: 699–708

Filipovska A, Razif MF, Nygard KK, Rackham O (2011) A universal code for RNA recognition by PUF proteins. Nat Chem Biol 7:425–427

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 31:822–826

Fujii W, Kawasaki K, Sugiura K, Naito K (2013) Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. Nucleic Acids Res 41:e187

Gabsalilow L, Schierling B, Friedhoff P, Pingoud A, Wende W (2013) Site- and strand-specific nicking of DNA by fusion proteins derived from MutH and I-*Sce*I or TALE repeats. Nucleic Acids Res 41:e83

Gao X, Yang J, Tsang JCH, Ooi J, Wu D, Liu P (2013) Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. Stem Cell Rep 1:183–197

Garg A, Lohmueller JJ, Silver PA, Armel TZ (2012) Engineering synthetic TAL effectors with orthogonal target sites. Nucleic Acids Res 40:7584–7595

Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadan AH, Moineau S (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468:67–71

Gasiunas G, Barrangou R, Horvath P, Siksnys V (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci USA 109:E2579–E2586

Geissler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens SE, Boch J (2011) Transcriptional activators of human genes with programmable DNA-specificity. PLoS ONE 6:e19509

Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA, Lim WA, Weissman JS, Qi LS (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154:442–451

Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM (2013) Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. Genetics 194:1029–1035

Griggs D, Stafford-Smith M, Gaffney O, Rockstrom J, Ohman MC, Shyamsundar P, Steffen W, Glaser G, Kanie N, Noble I (2013) Policy: sustainable development goals for people and planet. Nature 495:305–307

Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, Yang F, Chu Z, Wang G-L, White FF, Yin Z (2005) R gene expression induced by a type-III effector triggers disease resistance in rice. Nature 435:1122–1125

Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, Terns MP (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell 139:945–956

Hale CR, Majumdar S, Elmore J, Pfister N, Compton M, Olson S, Resch AM, Glover CVC, Graveley BR, Terns RM, Terns MP (2012) Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. Mol Cell 45:292–302

Hatoum-Aslan A, Maniv I, Marraffini LA (2011) Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. Proc Natl Acad Sci USA 108:21218–21222

Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Ja Doudna (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. Science 329:1355–1358

Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2011) Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29:731–734

Hopkins CM, White FF, Choi SH, Guo A, Leach JE (1992) A family of avirulence genes from *Xanthomonas oryzae* pv. *oryzae*. Mol Plant Microbe Interact 5:451–459

Hou Z, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, Thomson JA (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. Proc Natl Acad Sci USA (early online)

Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31:827–832

Hummel AW, Doyle EL, Bogdanove AJ (2012) Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. New Phytol 195:883–893

Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, Joung JK, Peterson RT, Yeh JR (2013) Heritable and precise zebrafish genome editing using a CRISPR-Cas system. PLoS ONE 8:e68708

Iamtham S, Day A (2000) Removal of antibiotic resistance genes from transgenic tobacco plastids. Nat Biotechnol 18:1172–1176

Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. J Bacteriol 169:5429–5433

Jao LE, Wente SR, Chen W (2013) Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc Natl Acad Sci USA 110:13904–13909

Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013a) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31:233–239

Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP (2013b) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res 41:e188

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 337:816–821

Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14:49–55

Kay S, Bonas U (2009) How Xanthomonas type III effectors manipulate the host plant. Curr Opin Microbiol 12:37–43

Kay S, Hahn S, Marois E, Hause G, Bonas U (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318:648–651

Ke J, Chen RZ, Ban T, Zhou XE, Gu X, Tan MH, Chen C, Kang Y, Brunzelle JS, Zhu JK, Melcher K, Xu HE (2013) Structural basis for RNA recognition by a dimeric PPR-protein complex. Nat Struct Mol Biol 20:1377–1382

Laluk K, Abuqamar S, Mengiste T (2011) The Arabidopsis mitochondria-localized pentatricopeptide repeat protein PGN functions in defense against necrotrophic fungi and abiotic stress tolerance. Plant Physiol 156:2053–2068

Li L, Wu LP, Chandrasegaran S (1992) Functional domains in Fok I restriction endonuclease. Proc Natl Acad Sci USA 89:4275–4279

Li T, Liu B, Spalding MH, Weeks DP, Yang B (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol 30:390–392

Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M, Li Y, Gao N, Wang L, Lu X, Zhao Y (2013a) Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. Nat Biotechnol 31:681–683

Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J (2013b) Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. Nat Biotechnol 31:688–691

Li T, Huang S, Zhou J, Yang B (2013c) Designer TAL effectors induce disease susceptibility and resistance to *Xanthomonas oryzae* pv. *oryzae* in rice. Mol Plant (early online)

Lincoln C, Long J, Yamaguchi J, Serikawa K, Hake S (1994) A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. Plant Cell 6:1859–1876

Lloyd A, Plaisier CL, Carroll D, Drews GN (2005) Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. Proc Natl Acad Sci USA 102:2232–2237

Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK (2013a) CRISPR RNA-guided activation of endogenous human genes. Nat Methods 10:977–979

Maeder ML, Linder SJ, Reyon D, Angstman JF, Fu Y, Sander JD, Joung JK (2013b) Robust, synergistic regulation of human gene expression using TALE activators. Nat Methods 10:243–245

Maekawa M, Taniguchi T, Higashi H, Sugimura H, Sugano K, Kanno T (2004) Methylation of mitochondrial DNA is not a useful marker for cancer detection. Clin Chem 50:1480–1481

Mahfouz MM, Li L (2011) TALE nucleases and next generation GM crops. GM Crops 2:99–103

Mahfouz MM, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, Zhu JK (2012) Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. Plant Mol Biol 78:311–321

Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 335:716–719

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV (2011) Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 9:467–477

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

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 339:823–826

Mao Y, Zhang H, Xu N, Zhang B, Gao F, Zhu JK (2013) Application of the CRISPR-Cas system for efficient genome engineering in plants. Mol Plant (early online)

Marois E, Van den Ackerveken G, Bonas U (2002) The xanthomonas type III effector protein AvrBs3 modulates plant gene

expression and induces cell hypertrophy in the susceptible host. Mol Plant Microbe Interact 15:637–646

Marton I, Zuker A, Shklarman E, Zeevi V, Tovkach A, Roffe S, Ovadis M, Tzfira T, Vainstein A (2010) Nontransgenic genome modification in plant cells. Plant Physiol 154:1079–1087

Meckler JF, Bhakta MS, Kim MS, Ovadia R, Habrian CH, Zykovich A, Yu A, Lockwood SH, Morbitzer R, Elsaesser J, Lahaye T, Segal DJ, Baldwin EP (2013) Quantitative analysis of TALE–DNA interactions suggests polarity effects. Nucleic Acids Res 41:4118–4128

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu L-J (2013) Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res 23:1233–1236

Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, Gregory PD, Pabo CO, Rebar EJ (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25:778–785

Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148

Minczuk M, Papworth MA, Kolasinska P, Murphy MP, Klug A (2006) Sequence-specific modification of mitochondrial DNA using a chimeric zinc finger methylase. Proc Natl Acad Sci USA 103:19689–19694

Minczuk M, Papworth MA, Miller JC, Murphy MP, Klug A (2008) Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. Nucleic Acids Res 36:3926–3938

Miyagishi M, Taira K (2002) U6 promoter-driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. Nat Biotechnol 20:497–500

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 60:174–182

Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 155:733–740

Morbitzer R, Romer P, Boch J, Lahaye T (2010) Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci USA 107:21617–21622

Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501

Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res 39:9283–9293

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

Noël LD, Denancé N, Szurek B (2013) Predicting promoters targeted by TAL effectors in plant genomes: from dream to reality. Front Plant Sci 4:333

Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol 31:839–843

Pavletich NP, Pabo CO (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 252:809–817

Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW, Guilak F, Crawford GE, Reddy TE, Gersbach CA (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods 10:973–976

Perez-Pinera P, Ousterout DG, Brunger JM, Farin AM, Glass Ka, Guilak F, Crawford GE, Hartemink AJ, Gersbach Ca (2013b) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. Nat Methods 10:239–242

Politz MC, Copeland MF, Pfleger BF (2013) Artificial repressors for controlling gene expression in bacteria. Chem Commun 49:4325–4327

Puchta H, Hohn B (2010) Breaking news: plants mutate right on target. Proc Natl Acad Sci USA 107:11657–11658

Puchta H, Dujon B, Hohn B (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res 21:5034–5040

Puchta H, Dujon B, Hohn B (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proc Natl Acad Sci USA 93:5055–5060

Qi 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

Ran FA, Hsu PD, Lin CY, 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 154:1380–1389

Römer P, Hahn S, Jordan T, Strauss T, Bonas U, Lahaye T (2007) Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. Science 318:645–648

Römer P, Recht S, Lahaye T (2009) A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. Proc Natl Acad Sci USA 106:20526–20531

Rossier O, Wengelnik K, Hahn K, Bonas U (1999) The Xanthomonas Hrp type III system secretes proteins from plant and mammalian bacterial pathogens. Proc Natl Acad Sci USA 96:9368–9373

Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. Nucleic Acids Res 39:9275–9282

Schornack S, Meyer A, Romer P, Jordan T, Lahaye T (2006) Gene-for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins. J Plant Physiol 163:256–272

Schornack S, Moscou MJ, Ward ER, Horvath DM (2013) Engineering plant disease resistance based on TAL effectors. Annu Rev Phytopathol 51:383–406

Shan Q, Wang Y, Chen K, Liang Z, Li J, Zhang Y, Zhang K, Liu J, Voytas DF, Zheng X, Zhang Y, Gao C (2013a) Rapid and efficient gene modification in rice and brachypodium using TALENs. Mol Plant 6:1365–1368

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013b) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31:686–688

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, Mitchell JC, Arnold NL, Gopalan S, Meng X, Choi VM, Rock JM, Wu YY, Katibah GE, Zhifang G, McCaskill D, Simpson MA, Blakeslee B, Greenwalt SA, Butler HJ, Hinkley SJ, Zhang L, Rebar EJ, Gregory PD, Urnov FD (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature 459:437–441

Sinkunas T, Gasiunas G, Waghmare SP, Dickman MJ, Barrangou R, Horvath P, Siksnys V (2013) In vitro reconstitution of cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. EMBO J 32:385–394

Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D (2000) Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res 28:3361–3369

Streubel J, Blucher C, Landgraf A, Boch J (2012) TAL effector RVD specificities and efficiencies. Nat Biotechnol 30:593–595

Sugisaki H, Kanazawa S (1981) New restriction endonucleases from *Flavobacterium okeanokoites* (*Fok*I) and *Micrococcus luteus* (*Mlu*I). Gene 16:73–78

Szurek B, Rossier O, Hause G, Bonas U (2002) Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. Mol Microbiol 46:13–23

Thomas MR, Rose RJ (1983) Plastid number and plastid structural changes associated with tobacco mesophyll protoplast culture and plant regeneration. Planta 158:329–338

Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459:442–445

Ülker B, Peiter E, Dixon DP, Moffat C, Capper R, Bouche N, Edwards R, Sanders D, Knight H, Knight MR (2008) Getting the most out of publicly available T-DNA insertion lines. Plant J 56:665–677

Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11:636–646

Valton J, Dupuy A, Daboussi F, Thomas S, Marechal A, Macmaster R, Melliand K, Juillerat A, Duchateau P (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem 287:38427–38432

Van den Ackerveken G, Marois E, Bonas U (1996) Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. Cell 87(7):1307–1316

Vanamee ES, Santagata S, Aggarwal AK (2001) *Fok*I requires two specific DNA sites for cleavage. J Mol Biol 309:69–78

Vanyushin BF, Ashapkin VV (2011) DNA methylation in higher plants: past, present and future. Biochim Biophys Acta 1809:360–368

Vasquez KM, Marburger K, Intody Z, Wilson JH (2001) Manipulating the mammalian genome by homologous recombination. Proc Natl Acad Sci USA 98:8403–8410

Voytas DF (2013) Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol 64:327–350

Waibel F, Filipowicz W (1990) U6 snRNA genes of Arabidopsis are transcribed by RNA polymerase III but contain the same two upstream promoter elements as RNA polymerase II-transcribed U-snRNA genes. Nucleic Acids Res 18:3451–3458

Wakiyama M, Matsumoto T, Yokoyama S (2005) Drosophila U6 promoter-driven short hairpin RNAs effectively induce RNA interference in Schneider 2 cells. Biochem Biophys Res Commun 331:1163–1170

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153:910–918

Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N, Nakamura T, Bando T, Ohuchi H, Yamamoto T, Noji S, Mito T (2012) Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. Nat Commun 3:1017

Waterworth WM, Drury GE, Bray CM, West CE (2011) Repairing breaks in the plant genome: the importance of keeping it together. New Phytol 192:805–822

Weinthal D, Tovkach A, Zeevi V, Tzfira T (2010) Genome editing in plant cells by zinc finger nucleases. Trends Plant Sci 15:308–321

Westra ER, Swarts DC, Staals RHJ, Jore MM, Brouns SJJ, van der Oost J (2012a) The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. Annu Rev Genet 46:311–339

Westra ER, van Erp PBG, Künne T, Wong SP, Staals RHJ, Seegers CLC, Bollen S, Jore MM, Semenova E, Severinov K, de Vos WM, Dame RT, de Vries R, Brouns SJJ, van der Oost J (2012b) CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. Mol Cell 46:595–605

Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. Nature 482:331–338

Wilson KA, McEwen AE, Pruett-Miller SM, Zhang J, Kildebeck EJ, Porteus MH (2013) Expanding the repertoire of target sites for zinc finger nuclease-mediated genome modification. Mol Ther Nucleic Acids 2:e88

Xie K, Yang Y (2013) RNA-guided genome editing in plants using a CRISPR-Cas system. Mol Plant 6:1975–1983

Yang Y, Gabriel DW (1995) Xanthomonas avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. Mol Plant Microbe Interact 8(4):627–631

Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T (2013) Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. PLoS ONE 8:e57286

Yan W, Smith C, Cheng L (2013) Expanded activity of dimer nucleases by combining ZFN and TALEN for genome editing. Sci Rep 3:2376

Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, Mali P, Aach J, Kim-Kiselak C, Briggs AW, Rios X, Huang PY, Daley G, Church G (2013) Optimization of scarless human stem cell genome editing. Nucleic Acids Res 41:9049–9061

Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang HW, Wang J, Zhu JK, Shi Y, Yan N (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. Nature 504:168–171

Yu Z, Ren M, Wang Z, Zhang B, Rong YS, Jiao R, Gao G (2013) Highly efficient genome modifications mediated by CRISPR/Cas9 in Drosophila. Genetics 195:289–291

Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. Development 130:4859–4869

Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29:149–153

Zhang Y, Zhang F, Li X, Baller JA, Qi Y, Starker CG, Bogdanove AJ, Voytas DF (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. Plant Physiol 161:20–27

Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, Hu Y, Luo Z, Huang P, Wu Q, Zhu Z, Zhang B, Lin S (2013) TALEN-mediated precise genome modification by homologous recombination in zebrafish. Nat Methods 10:329–331