Review

Biotechnology and Bioengineering

Transcription Activator-Like Effector Nucleases (TALENs): A Highly Efficient and Versatile Tool for Genome Editing

Ning Sun,[1] Huimin Zhao[1,2]

[1]Department of Biochemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

[2]Department of Chemical and Biomolecular Engineering, Department of Bioengineering, Department of Chemistry, Center for Biophysics and Computational Biology, and Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801; telephone: 217-333-2631; fax: 217-333-5052; e-mail: zhao5@illinois.edu

ABSTRACT: Transcription activator-like effector (TALE) nucleases (TALENs) have recently emerged as a revolutionary genome editing tool in many different organisms and cell types. The site-specific chromosomal double-strand breaks introduced by TALENs significantly increase the efficiency of genomic modification. The modular nature of the TALE central repeat domains enables researchers to tailor DNA recognition specificity with ease and target essentially any desired DNA sequence. Here, we comprehensively review the development of TALEN technology in terms of scaffold optimization, DNA recognition, and repeat array assembly. In addition, we provide some perspectives on the future development of this technology.

Biotechnol. Bioeng. 2013;110: 1811–1821.

KEYWORDS: double-strand break; genome engineering; gene targeting; homologous recombination; non-homologous end joining

Introduction

Targeted genome engineering or editing enables researchers to modify genomic loci of interest in a precise manner, which has various applications in industry, agriculture, and human therapeutics. In most organisms, efficient genome editing relies on a site-specific DNA double-strand break (DSB). The subsequent DNA repair by homologous recombination (HR) or non-homologous end joining (NHEJ) can generate desired genetic modifications (Fig. 1A) (recently reviewed in Perez-Pinera et al., 2012; Sun et al., 2012a). HR can mediate nucleotide exchange between an endogenous genomic region and an exogenous DNA fragment through flanking homologous sequences, resulting in DNA insertion, deletion, or replacement. The DNA repaired by NHEJ usually contains small deletions and insertions at the break sites. Therefore, the imprecise DSB repair through NHEJ can be applied for gene disruption by the introduction of frame-shift mutations.

To generate a site-specific DSB at the desired genomic region, zinc finger (ZF) nucleases (ZFNs) have been developed (as reviewed in Carroll, 2011; Urnov et al., 2010). ZFNs are artificial proteins constructed by fusing several ZF domains to a sequence-independent cleavage domain of the type IIS restriction endonuclease *Fok*I. Each ZF domain comprises 30 amino acids (aa) and coordinates one zinc atom using two His and two Cys residues. An α-helix in each ZF domain recognizes a specific DNA triplet, so linkage of several ZF domains enables ZFNs to recognize long DNA sequences. Because the *Fok*I nuclease domain functions as a dimer, functional ZFNs are formed when two sets of ZF domains are positioned in close proximity and in the appropriate orientations. Because the DNA binding and cleavage domains are functionally independent, they can be optimized in isolation for efficient DNA cleavage against custom-designed target sites. Even though ZFNs have been used for targeted genome editing in various organisms, two major limitations prevent their wider applications.

Conflict of interest statement: None declared.
Correspondence to: H. Zhao
Contract grant sponsor: National Institutes of Health
Contract grant number: R21 HL089418
Contract grant sponsor: Department of Chemical and Biomolecular Engineering at the University of Illinois at Urbana-Champaign
Received 21 January 2013; Revision received 18 February 2013;
Accepted 26 February 2013
Accepted manuscript online 18 March 2013;
Article first published online 7 April 2013 in Wiley Online Library (http://onlinelibrary.wiley.com/doi/10.1002/bit.24890/abstract)
DOI 10.1002/bit.24890

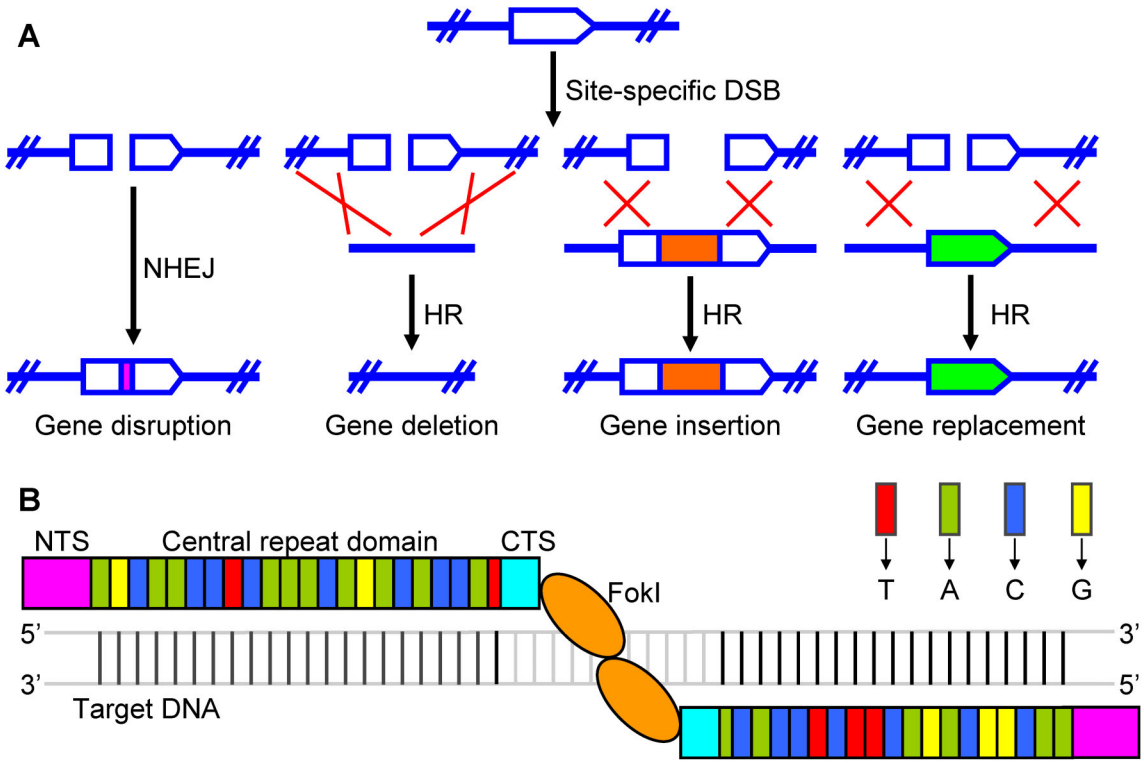

Figure 1. TALEN-mediated genome editing. **A**: The site-specific chromosomal DSB is repaired by either NHEJ or HR, resulting in gene disruption, gene deletion, gene insertion or gene replacement. **B**: Schematic of TALEN architecture. A TALEN is composed of a N-terminal segment (NTS, pink box), a central repeat domain, a C-terminal segment (CTS, cyan box), and a *Fok*I catalytic domain (orange oval). The central repeat domain comprises a series of repeat units that are responsible for specific recognition of thymine (red boxes), adenine (green boxes), cytosine (blue boxes), and guanine (yellow boxes). The formation of a heterodimer by two TALENs in a tail-to-tail orientation at the target site executes a site-specific DNA DSB. The TALE binding sites on the target DNA are shown in black and the spacer is shown in gray.

ZF domains have limited modularity due to the context-dependent DNA-binding effects, making it difficult for ZFNs to target any desired DNA sequence (Ramirez et al., 2008). Moreover, lack of specificity of some ZF domains can generate off-target cleavage, leading to undesired mutations and chromosomal aberrations (Pattanayak et al., 2011; Radecke et al., 2010).

Recently, transcription activator-like effector nucleases (TALENs) have rapidly emerged as an alternative genome editing tool to ZFNs (recently reviewed in Joung and Sander, 2013). Similar to ZFNs, TALENs use the non-specific *Fok*I domain as the DNA cleavage module and function as dimers (Fig. 1B). However, the DNA binding domains of TALENs are composed of a series of tandem repeats as in TALEs of the plant pathogenic bacteria from the genus *Xanthomonas* (as reviewed in Bogdanove and Voytas, 2011; Munoz Bodnar et al., 2013). Each repeat comprises 33–35 aa and recognizes a single nucleotide. The last repeat typically has only 20 aa, and is therefore called a "half-repeat." The DNA recognition specificity is conferred by the highly variable amino acids at positions 12 and 13 (e.g., NI recognizes adenine, HD recognizes cytosine, NG recognizes thymine, and NN recognizes guanine and adenine) (Boch et al., 2009; Moscou and Bogdanove, 2009). Unlike the context-dependent DNA binding of ZFNs, TALENs can be easily and rapidly constructed to target almost any DNA sequence due to the simple protein–DNA code and the modular nature. In addition, TALENs exhibit significantly reduced off-target effects and cytotoxicities compared with ZFNs, making them an efficient genome editing tool (Ding et al., 2013; Mussolino et al., 2011). Within the last 3 years, TALENs have been widely applied to modify endogenous genes in a variety of organisms (Table I). Applications include studying gene functions in model organisms, improving traits in crop plants and livestock, generating disease models, and treating genetic disorders in humans. In this article, we provide a comprehensive review of TALEN technology including the optimization of the scaffold, improvement of the DNA recognition specificity, and assembly of TALE repeat arrays. Due to the ease of design and high efficiency of genome editing, TALENs have opened up many new avenues for basic and applied biological research.

Scaffold Optimization

The original TALEN construction was reported by two independent groups. Li et al. (2011a) fused the full-length

Table I. Applications of TALENs for targeted genome editing in various organisms.

Organisms	Genes	Refs.
Arabidopsis thaliana	*ADH1*	Cermak et al. (2011)
Brachypodium	*BdABA1*, *BdCKX2*, *BdCOI1*, *BdHTA1*, *BdRHT*, *BdSBP*, *BdSMC6*, *BdSPL*	Shan et al. (2013)
Cattle (*Bos taurus*)	*ACAN*, *GDF8*, *GGTA*, *PRNP*	Carlson et al. (2012)
Cricket (*Gryllus imaculatus*)	*Gb′lac2*	Watanabe et al. (2012)
Frog (*Xenopus tropicalis*)	*ets1*, *foxd3*, *grp78/bip*, *hhex*, *noggin*, *ptf1a/p48*, *sox9*, *tyr*, *vpp1*	Ishibashi et al. (2012), Lei et al. (2012)
Fruitfly (*Drosophila melanogaster*)	*CG9797*, *yellow*	Liu et al. (2012)
Hamster (*Cricetulus griseus*)	*FUT8*	Cristea et al. (2013)
Human (*Homo sapiens*)	*ABL1*, *AKT2*, *ALK*, *ANGPTL3*, *APC*, *APOB*, *ATGL*, *ATM*, *AXIN2*, *BAX*, *BCL6*, *BMPR1A*, *BRCA1*, *BRCA2*, *C6orf106*, *CIITA*, *CBX3*, *CBX8*, *CCND1*, *CCR5*, *CDC73*, *CDK4*, *CELSR2*, *CFTR*, *CHD4*, *CHD7*, *CTNNB1*, *CYLD*, *DDB2*, *DDX60*, *DHX58*, *DHX9*, *DICER1*, *EIF2AK2*, *EIF2C2*, *ERCC2*, *EWSR1*, *EXT1*, *EXT2*, *EZH2*, *FANCA*, *FANCC*, *FANCF*, *FANCG*, *FES*, *FGFR1*, *FH*, *FLCN*, *FLT4*, *FOXO1*, *FOXO3*, *GLI1*, *GLUT4*, *HBB*, *HDAC1*, *HDAC2*, *HDAC6*, *HMGA2*, *HOXA13*, *HOXA9*, *HOXC13*, *HPRT1*, *IFI44L*, *IFIT1*, *IFIT2*, *IL2RG*, *JAK2*, *KRAS*, *LINC00116*, *MAOA*, *MAP2K4*, *MB21D1*, *MDM2*, *MET*, *MLH1*, *MSH2*, *MUTYH*, *MYC*, *MYCL1*, *MYCN*, *NAMPT*, *NBN*, *NCOR1*, *NCOR2*, *NDUFA9*, *NLRC5*, *NTF3*, *NUB1*, *OASL*, *OCT4*, *PDGFRA*, *PDGFRB*, *PHF11*, *PHF8*, *PITX3*, *PLA2G4A*, *PLIN1*, *PMS2*, *PPP1R12C*, *PTCH1*, *PTEN*, *QRICH1*, *RARA*, *RBBP5*, *RECQL4*, *RET*, *RIPK4*, *RTP4*, *RUNX1*, *SDHB*, *SDHC*, *SDHD*, *SETDB1*, *SIRT6*, *SMAD2*, *SORT1*, *SS18*, *STAT1*, *STAT6*, *SUZ12*, *TBK1*, *TFE3*, *TP53*, *TRIB1*, *TSC2*, *TTN*, *UNC93B1*, *VHL*, *XPA*, *XPC*	Cermak et al. (2011), Hockemeyer et al. (2011), Kim et al. (2011), Miller et al. (2011), Mussolino et al. (2011), Sanjana et al. (2012), Wang et al. (2012), Reyon et al. (2012b), Sun et al. (2012b), Ding et al. (2013), Schmid-Burgk et al. (2013), Stroud et al. (2013)
Medaka (*Oryzias latipes*)	*DJ-1*	Ansai et al. (2013)
Mouse (*Mus musculus*)	*Pibf1*, *Sepw1*	Sung et al. (2013)
Nematode (*Caenorhabditis elegans*)	*ben-1*	Wood et al. (2011)
Rat (*Rattus norvegicus*)	*BMPR2*, *IgM*	Tesson et al. (2011), Tong et al. (2012)
Rice (*Oryza sativa L.*)	*Os11N3*, *OsBADH2*, *OsCKX2*, *OsDEP1*, *OsSD1*	Li et al. (2012b), Shan et al. (2013)
Silkworm (*Bombyx mori*)	*BmBlos2*	Ma et al. (2012), Sajwan et al. (2013)
Swine (*Sus scrofa*)	*AMELY*, *DMD*, *GDF8*, *GGTA*, *GHRHR*, *IL2Rg*, *p65*, *RAG2*, *RELA*, *SRY*	Carlson et al. (2012)
Tobacco (*Nicotiana tabacum*)	*SurA*, *SurB*	Zhang et al. (2013)
Yeast (*Saccharomyces cerevisiae*)	*ADE2*, *LYS2*, *URA3*	Li et al. (2011b)
Zebrafish (*Danio rerio*)	*aanat2*, *abcc9*, *adora1b*, *adora2aa*, *bmi1*, *cdh5*, *clc*, *crhr1*, *dip2a*, *elmo1*, *epas1b*, *fgf21*, *fh*, *golden*, *gpr103a*, *gria3a*, *hdc*, *hey2*, *hif1ab*, *ikzf1*, *jak3*, *moesina*, *myod*, *nmu*, *npy*, *phf6*, *pmch*, *pmchl*, *ponzr1*, *ppp1cab*, *prok2*, *prokr1*, *prokr2*, *ptpmt1*, *qrfp*, *ryr1a*, *ryr3*, *scl6a*, *tbx6*, *tgfa*, *tnikb*, *vip*	Huang et al. (2011), Sander et al. (2011), Bedell et al. (2012), Cade et al. (2012), Dahlem et al. (2012), Moore et al. (2012), Chen et al. (2013)

natural TALE (AvrXa7 and PthXo1) with the *Fok*I catalytic domain, creating the TALENs bearing a 288 aa N-terminal segment (NTS) and a 295 aa C-terminal segment (CTS). Based on a yeast reporter assay, the optimal spacer length between the two TALEN binding sites was determined to be 16–31 bp (Table II). Alternatively, the TALENs created by Christian et al. (2010) encompasses a 287 aa NTS and a 231 aa CTS. This scaffold allows efficient DNA cleavage against target sites with 13–30 bp spacers (Table II). Since naturally occurring TALEs are transcription activators from a plant bacterial pathogen, their NTSs harbor protein secretion signal peptides while their CTSs contain nuclear localization signal peptides and a transcription activator domain (White et al., 2009). These sequences can impair the catalytic activity when fused with the *Fok*I cleavage domain. To identify the optimum TALEN architecture with highest cleavage efficiency and minimal peptide portion, scaffold optimization has been carried out by several groups. Using a truncated NTS with only 136 aa, Miller et al. (2011) constructed a series of TALENs by trimming the CTS between the central repeat units and the *Fok*I nuclease domain. This study found that different TALEN scaffolds prefer different spacer length between each TALEN binding site. One TALEN scaffold was identified bearing a 63 aa CTS, which could drive efficient gene modification in human cells when separated by 12–20 bp spacers (Table II). This scaffold has since been widely applied for efficient genome editing in various species (Cade et al., 2012; Hockemeyer et al., 2011; Huang et al., 2011; Lei et al., 2012; Moore et al., 2012; Reyon et al., 2012b; Sander et al., 2011; Tesson et al., 2011; Wood et al., 2011). In addition, another TALEN scaffold bearing only a 28 aa CTS with a narrower separation range of 12-13 bp was identified (Table II). Later, Mussolino et al. (2011) and Christian et al. (2012) reported TALENs with

Table II. Engineered TALEN scaffolds with different NTSs and CTSs.

NTS (aa)	CTS (aa)	Spacer (bp)	Reporter system	Refs.
288	295	16–31	β-galactosidase assay in yeast	Li et al. (2011a)
288	285	16[a]	Transient expression assay in tobacco leaves	Mahfouz et al. (2011)
287	231	13–30	β-galactosidase assay in yeast	Christian et al. (2010)
287	63	15[a]	Mutagenesis in medaka embryos	Ansai et al. (2013)
207	63	14–32	β-galactosidase assay in yeast	Sun et al. (2012b)
207	31	10–16	β-galactosidase assay in yeast	Sun et al. (2012b)
153	47	12–21	dsEGFP assay in HEK293	Mussolino et al. (2011)
153	17	12	dsEGFP assay in HEK293	Mussolino et al. (2011)
136	63	12–20	Surveyor nuclease assay in K562	Miller et al. (2011)
136	28	12–13	Surveyor nuclease assay in K562	Miller et al. (2011)
136	18	13–16	β-galactosidase assay in yeast	Christian et al. (2012)

[a]The spacer length was not optimized in the study.

even shorter CTSs that exhibit narrow optimal ranges of spacers (Table II). A systematic study was carried out by Sun et al. (2012b) who constructed 10 different TALEN scaffolds with various NTSs and CTSs. The DNA cleavage activity of each scaffold was assayed against 10 substrates with different spacers in a yeast reporter system. Based on this 10×10 matrix, two TALEN scaffolds with high DNA cleavage efficiency in both yeast and human cells were identified. One bearing a 207 aa NTS and a 31 aa CTS prefers target sites with 10–16 bp spacers while another bearing a 207 aa NTS and a 63 aa CTS has highest efficiency when separated by 14–32 bp spacers (Table II). It is noteworthy that the TALEN with a 50 aa NTS has no catalytic activity against any target sites. Later on, Gao et al. (2012) solved the crystal structure of the TALE NTS and discovered an extended N-terminal DNA binding region composed of the 127 aa immediately preceding the central repeat units. The 127 aa NTS features four continuous repeats. Each repeat contains two α-helices and an intervening loop (Fig. 2A), a structural feature highly similar to that of the central repeat unit. Although the 127 aa NTS does not confer sequence specificity, it is crucial for DNA binding. This feature explains why all the effective TALEN scaffolds have at least 127 aa preceding the central repeat units (Table II).

A second-generation GoldyTALEN scaffold has been demonstrated to improve the genome editing efficiency in zebrafish (Bedell et al., 2012). GoldyTALEN is based on a scaffold previously reported by Miller et al. (2011) which consists of a 136 aa NTS and a 63 aa CTS. However, it has nine different aa substitutions at the NTS and five different aa substitutions at the CTS. Using the GoldyTALEN scaffold and zebrafish delivery system, certain loci were modified with 100% efficiency. Moreover, they provided the first example of HR-based genome editing in zebrafish using single-stranded DNA as a donor. The GoldyTALEN scaffold was also applied for efficient gene knockout in livestock (Carlson et al., 2012).

Because the *Fok*I catalytic domain must dimerize to become active, two TALEN subunits are assembled as heterodimers at the cleavage site. However, cleavage-competent homodimers composed of each subunit may also form and generate off-target cleavage, which can limit safety or efficiency. To address this limitation, obligate heterodimer mutations were introduced at the dimer interface of the *Fok*I cleavage domain which prevented homodimerization based on electrostatic and hydrophobic interactions. The creation of *Fok*I variants that preferentially heterodimerize successfully reduced off-target cleavage of ZFNs and relieved toxicity (Doyon et al., 2011; Miller et al., 2007; Szczepek et al., 2007). A similar principle was applied to TALENs by Cade et al. (2012) for the generation of zebrafish knockout lines. The heterodimeric TALENs show similar or even greater activities than their homodimeric counterparts. Moreover, the TALENs constructed with heterodimeric *Fok*I domains induced smaller numbers of abnormal or dead embryos, indicating reduced toxicity. This obligate heterodimeric TALEN configuration has also been reported by other groups for gene knockout studies (Dahlem et al., 2012; Huang et al., 2011; Lei et al., 2012).

DNA Recognition Specificity

The DNA recognition specificity of TALENs is conferred by the repeat-variable diresidues (RVDs) at positions 12 and 13 of each repeat. More than 20 different RVDs have been identified in TALEs, among which NI, NG, HD, NN, and HG are the most common ones recognizing the nucleotides A, T, C, G/A, and T, respectively (Boch et al., 2009; Moscou and Bogdanove, 2009). Based on crystal structures, TALE binds to target DNA as a right-handed superhelix (Fig. 2B). Each repeat unit forms a left-handed, two-helix bundle that presents an RVD-containing loop to the DNA major groove (Deng et al., 2012a; Mak et al., 2012). The first residue of each RVD (residue 12 of each repeat), either His or Asn, does not contact DNA directly. Instead, the side chain forms a hydrogen bond to the backbone carbonyl oxygen of Ala at position 8 (Ala^8) of each repeat, stabilizing the local conformation of the RVD-containing loop (Fig. 2C–E). Sequence-specific contacts of TALEs to target DNA are made by the second residue of each RVD (residue 13 of each repeat) to the corresponding base on the sense strand. In the

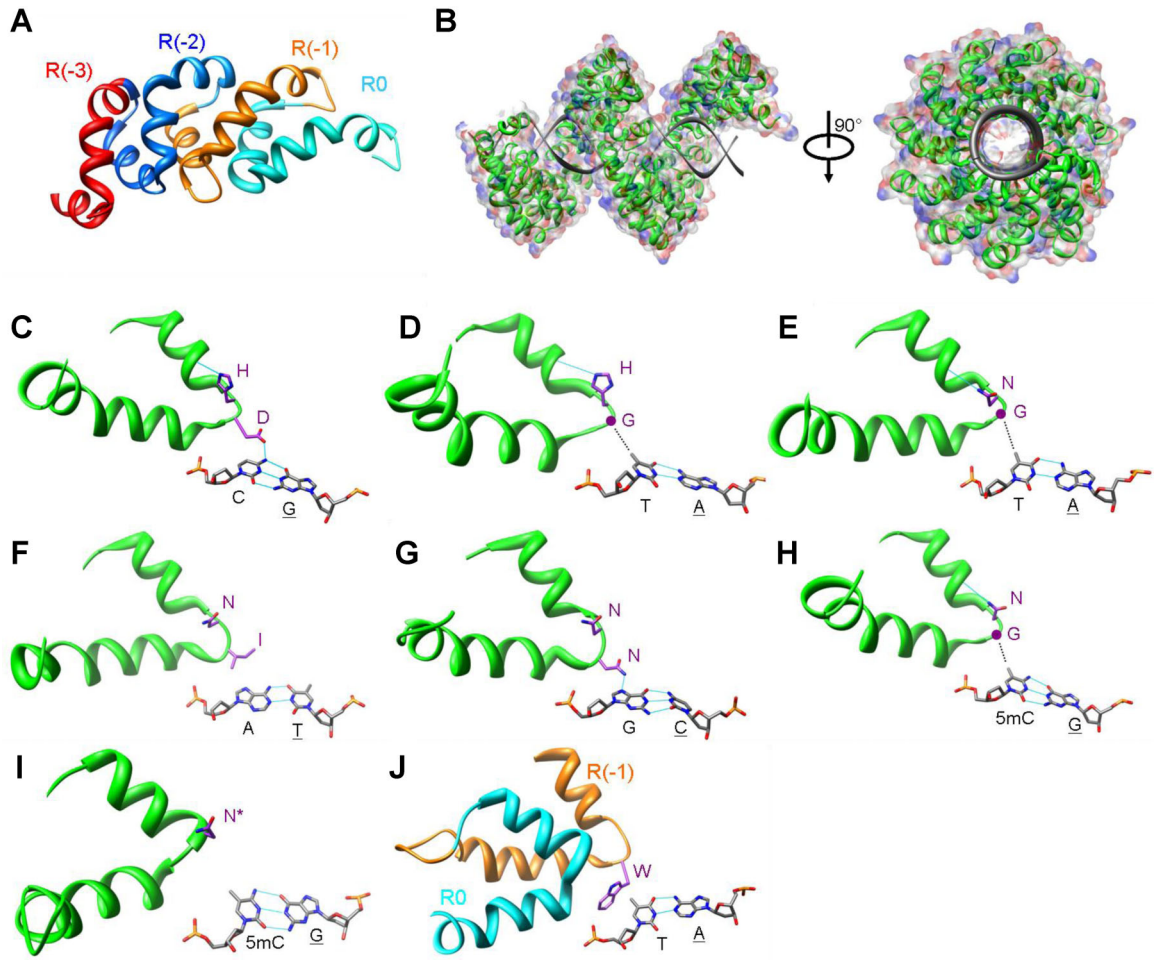

Figure 2. Crystal structures of TALEs. **A**: The TALE NTS features four continuous repeats, which are important for the DNA binding affinity (adapted from Gao et al., 2012, PDB identification code: 4HPZ). **B**: Overall structure of the TALE PthXo1 central repeat domain in complex with its target site (adapted from Mak et al., 2012, protein data bank (PDB) identification code: 3UGM). **C**: Interaction of RVD HD with cytosine (adapted from Gao et al., 2012, PDB identification code: 4HPZ). Hydrogen bonds are indicated by cyan lines. **D**: Interaction of RVD HG with thymine (adapted from Mak et al., 2012, PDB identification code: 3UGM). A non-polar van der Waals interaction is shown in a dotted line. **E**: Interaction of RVD NG with thymine (adapted from Gao et al., 2012, PDB identification code: 4HPZ). **F**: Interaction of RVD NI with adenine (adapted from Mak et al., 2012, PDB identification code: 3UGM). **G**: Interaction of RVD NN with guanine (adapted from Mak et al., 2012, PDB identification code: 3UGM). **H**: Interaction of RVD NG with 5-methyl cytosine (adapted from Deng et al., 2012b, PDB identification code: 4GJR). **I**: Interaction of RVD N* with 5-methyl cytosine based on a structural model. **J**: Interaction of the NTS with the 5′-preceding thymine. (Adapted from Mak et al., 2012, PDB identification code: 3UGM.) The coloring of the two repeats in TALE NTS matches that in A.

HD RVDs specific for C nucleotides, the carboxylate oxygen of Asp[13] forms a hydrogen bond to the amine group of cytosine, which excludes the other bases through physical or electrostatic clash (Fig. 2C). In the case of NG and HG RVDs specific for T nucleotides, the backbone α carbon of the Gly[13] makes a non-polar van der Waals contact with the methyl group of the opposing thymine base, which is less favorable for the other bases (Fig. 2D and E). In the NI RVDs specific for A nucleotides, the aliphatic side chain of Ile[13] makes non-polar van der Waals interactions to C8 and N7 of adenine, which reduces the binding affinity to the other bases (Fig. 2F). NN RVDs are commonly used to recognize G nucleotides. The side chain of Asn[13] residue makes a hydrogen bond with the N7 nitrogen of the opposing guanine base (Fig. 2G). But similar interaction might be made with the N7 nitrogen of adenine, which makes NN RVDs associate with A and G nucleotides with almost identical frequency. Because HD and NN RVDs form hydrogen bonds with DNA bases, the binding affinities of HD to cytosine and NN to guanine are much stronger than the van der Waals contacts of NI to adenine and NG/HG to thymine. It has been suggested to incorporate at least 3–4 strong RVDs for the construction of efficient TALENs (Streubel et al., 2012).

The lack of specific RVDs to recognize guanine limits TALENs' broader applications because non-specific binding can generate off-target cleavage, resulting in unexpected genomic instability and cytotoxicity. Morbitzer et al. (2010) discovered that the NK RVDs can facilitate specific targeting of G nucleotides through *in planta* function analysis. Based

on the SELEX assay, Miller et al. (2011) provided in vitro evidence that RVD NK has a much stronger preference for guanine over adenine, which represents a promising code for the specific recognition of G nucleotides. However, substitution of RVD NN with NK significantly reduced TALEN activity in zebrafish embryo (Huang et al., 2011). Substantially lower activities in NK containing TALEs have also been observed in plants and in mammalian cells (Cong et al., 2012; Streubel et al., 2012). Therefore, RVD NK is not ideal for guanine recognition because the improvement in specificity sacrifices efficiency. Alternatively, NH has been reported as a competent guanine-specific RVD, which has much higher efficiency than RVD NK (Cong et al., 2012; Streubel et al., 2012). Computational modeling analysis showed that the imidazole ring on the His^{13} of the NH RVD has a compact base-stacking interaction with the guanine base, suggesting a possible mechanism for its increased specificity for G nucleotide while maintaining the binding affinity (Cong et al., 2012).

Although successfully used in various cellular contexts, TALE DNA binding domains have been reported to be incapable of targeting methylated DNA (Bultmann et al., 2012). Often considered as the fifth base, 5-methyl cytosine (5mC) is a major epigenetic mark and widely distributed in fungi, plant, and mammalian genomes (Su et al., 2011). In addition, 5mC has been identified in CpG islands of many promoters, which are important regulatory regions for genome modification (Maunakea et al., 2010). Recently, two groups discovered that RVD NG and N^* (an asterisk indicates a deletion at residue 13 in the repeat unit) can accommodate 5mC efficiently in vitro and in vivo (Deng et al., 2012b; Valton et al., 2012). Thymine is structurally similar to 5mC, with the only difference at position 4, which is not involved in binding to TALE repeats. This observation indicates that the NG RVD specific for thymine might be used to recognize 5mC. The protein crystal structure solved by Deng et al. (2012b) shows that lack of side chain of Gly^{13} in NG RVDs provides sufficient space to accommodate the 5-methyl group of 5mC and allows the formation of van der Waals contacts (Fig. 2H). Because RVDs are followed immediately by two conserved Gly residues, N^* is roughly equivalent to NG except for a shortened RVD loop (Fig. 2I). Using N^* to code for 5mC, Valton et al. (2012) demonstrated the first example of TALEN-mediated modification at a methylated locus in human cells. Accommodation of 5mC by TALE repeats through the RVD NG or N^* extends the DNA recognition code and enables researchers to design TALENs to target hyper-methylated DNA regions, which has great potential in epigenetics studies and human therapeutic applications.

All naturally occurring TALE target sites are preceded by a 5′-thymine at position 0, which was previously believed to be essential for TALE function (Boch et al., 2009; Moscou and Bogdanove, 2009). The TALE crystal structure reveals that two degenerate repeats prior to the central repeat domain appear to cooperate to specify the conserved 5′-thymine (Mak et al., 2012). The indole ring of a Trp residue in the repeat R(-1) forms a van der Waals contact with the methyl group of the thymine base, suggesting a possible mechanism for the conserved specificity at position 0 (Fig. 2J). Sun and coworkers reported that TALENs with shorter CTSs (31 aa) show higher efficiency against natural TALE recognition sites preceded by a 5′-T than that against unnatural TALE sites preceded by A, C, or G. However, TALEN variants with longer CTSs (63–117 aa) are capable of cleaving unnatural DNA substrates with similar efficiency compared with that of natural TALE sites (Sun et al., 2012b). Other studies also provided evidence that a thymine at position 0 is not strictly required for TALEN activity (Briggs et al., 2012; Miller et al., 2011; Yu et al., 2011). Notably, there are nine leucine zipper-like heptad repeats closely linked to the C terminus of the TALE central repeat domain (Yang and Gabriel, 1995). These leucine-rich repeats may mediate TALE/DNA interactions and increase DNA binding affinity, making the 5′-T less of a requirement. Detailed structural studies could help solve this uncertainty. The requirement of a preceding 5′-T can be mitigated using certain TALEN scaffolds, allowing greater flexibility in choosing target sites in genome editing endeavors.

Assembly of TALE Repeat Arrays

Because of the high similarity between each TALE repeat unit, it is challenging to construct plasmids encoding long arrays of TALE repeats. To address this limitation, numerous methods have been developed to assemble the highly repetitive TALE central repeat domains rapidly and cost-effectively (Joung and Sander, 2013). Based on a standard cloning strategy, Sander and coworkers described a restriction enzyme and ligation (REAL) method, in which single TALE repeats are joined together using routine restriction digestion and ligation techniques (Sander et al., 2011). Initially, they constructed a library of plasmids encoding various individual TALE repeats by DNA synthesis. In the assembly step, two TALE repeats are first joined together by ligating compatible overhangs generated by digestion with restriction endonucleases. Next, the ligation product encoding two TALE repeats is joined with another TALE repeat dimer in the same manner, resulting in a DNA fragment encoding four TALE repeats. This process continues in an iterative fashion until a TALE repeat array of the desired length is assembled (Fig. 3A). Using a large plasmid library of pre-assembled multiple TALE repeats, REAL can be performed in a more rapid and less labor-intensive fashion, which is referred to as REAL-Fast (Reyon et al., 2012a). With the help of isocaudamer restriction enzymes (e.g., *Nhe*I and *Spe*I), a unit assembly method has been described for building long TALE repeat arrays in the similar hierarchical fashion (Huang et al., 2011).

Utilization of Golden Gate cloning has greatly facilitated and accelerated the synthesis of TALE genes (Cermak et al., 2011; Geissler et al., 2011; Li et al., 2011b, 2012a; Morbitzer et al., 2011; Sanjana et al., 2012; Weber et al., 2011; Zhang

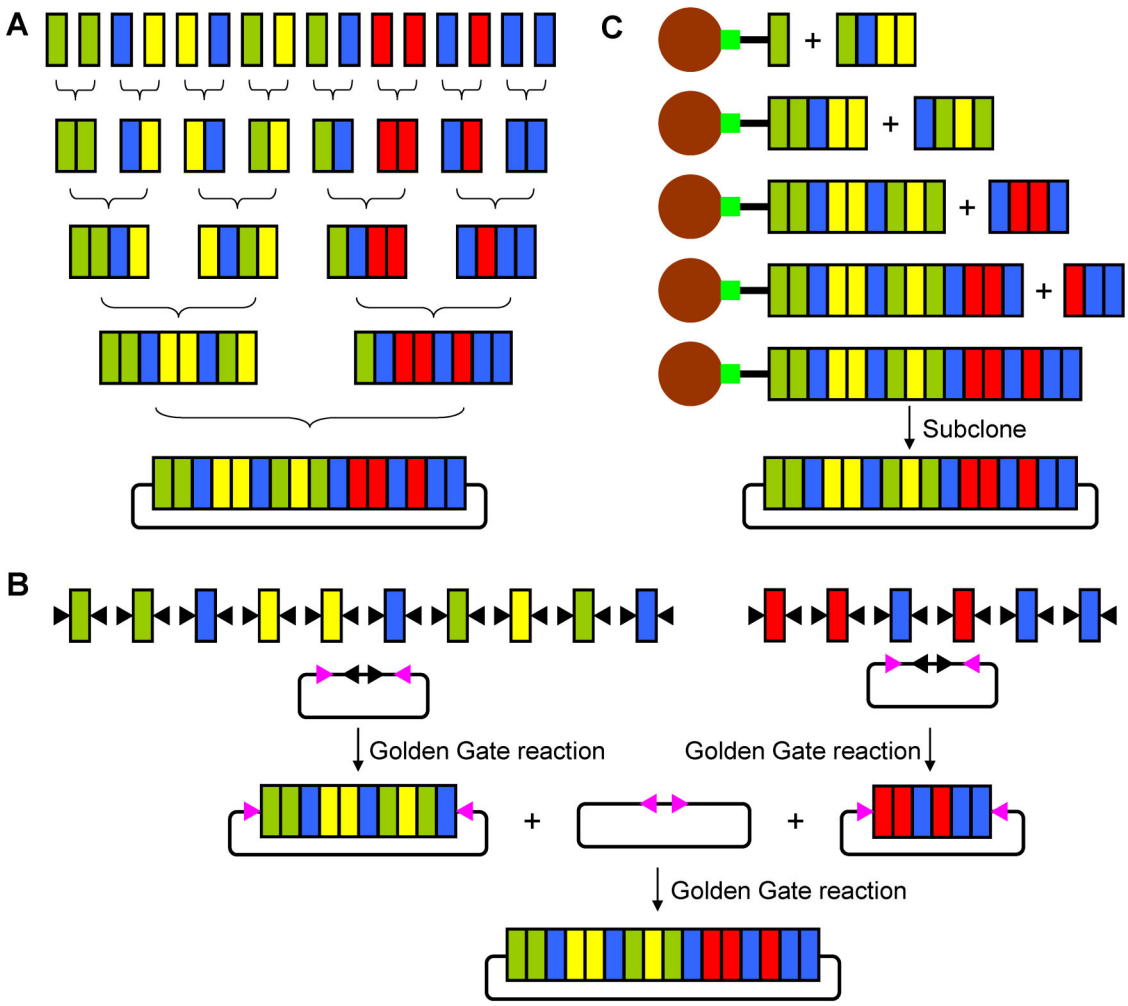

Figure 3. Schematic of the strategies for the assembly of TALE repeat arrays. **A**: The REAL strategy based on hierarchical ligations (Sander et al., 2011). **B**: The Golden Gate cloning-based strategy (Cermak et al., 2011). Each repeat unit is flanked by the recognition sites of a type IIS restriction endonuclease (black triangles). The first round of Golden Gate reaction assembles multiple TALE repeat units in a single step. The second round of Golden Gate reaction relies on the recognition site of a different type IIS restriction enzyme (pink triangles), which assembles the complete TALE genes from pre-assembled repeat multimers. **C**: The FLASH assembly method based on solid-phase ligation (Reyon et al., 2012b). Ligations are carried out iteratively on a streptavidin-coated magnetic bead (brown circle), which contains an immobilized biotinylated DNA double-strand adaptor (green box). The final ligation product is released from the solid phase by restriction digestion and subcloned into the final backbone vector.

et al., 2011). Golden Gate cloning has been developed to overcome the difficulty of assembling the monomers into ordered multimer arrays (Engler et al., 2008). It employs type IIS restriction endonucleases that cut outside of their recognition sequences and produce non-palindromic, 4 bp 5′-overhangs. Since the recognition and cleavage sites are spatially separated in type IIS restriction enzymes, essentially any desired 5′-overhang sequence can be generated. Moreover, because the correct ligation products lack the enzyme recognition site and cannot be recut, the cleavage and ligation can be carried out in the same reaction mixture in a single step. To construct a TALE central repeat domain to recognize a desired target site, each individual repeat must be flanked by the type IIS restriction sites at both 5′- and 3′-ends. Restriction endonuclease digestion removes the flanking restriction sites and generates unique terminal overhangs. The overhangs are designed so that each repeat ligates specifically to another repeat with a compatible overhang. Thus, the 3′-end of first repeat can only ligate to the 5′-end of the second repeat, the 3′-end of the second repeat can only ligate to the 5′-end of the third repeat and so on. Therefore, the position of each repeat within the TALE central repeat domain is defined exclusively by the given overlap. For the first round of Golden Gate reaction, 6–10 repeats are cloned into an intermediate plasmid. The second round of Golden Gate reaction assembles the TALE repeat arrays of each intermediate plasmid into the final backbone plasmid and makes the complete TALE central repeat domain fused with *Fok*I or other functional protein domains (Fig. 3B). To increase the assembly efficiency, the *lacZ* gene

and the toxic *ccdB* gene are introduced for blue/white screening and selection, respectively. With the help of preassembled TALE repeat tetramers and trimers, a single-step Golden Gate strategy has been developed to generate TALENs that recognize 15 bp target sites within 2 days (Ding et al., 2013). Due to its ease of use and public availability, Golden Gate assembly provides a convenient means to construct TALENs for academic laboratories.

For industrial scale synthesis, development of a solid-phase ligation strategy has facilitated the cloning of TALE genes in a high-throughput and cost-effective manner (Briggs et al., 2012; Reyon et al., 2012b; Wang et al., 2012). The solid-phase strategy assembles TALE repeat units on a streptavidin-coated magnetic bead that contains an immobilized biotinylated DNA double-strand adaptor with a restriction endonuclease site on one end. Ligation of TALE repeat units is unidirectional and iterative. In each cycle, newly added TALE repeat units are ligated to the immobilized DNA fragments and subsequent washing steps remove undesired products. The cycle is continued until an array of the desired length is assembled. The final ligation product is then released from the solid phase by restriction digestion and subcloned into the final backbone vector (Fig. 3C). With this strategy, TALE repeat units are assembled on solid-phase rather than in solution, thereby avoiding the need for gel isolation, purification, or analysis of intermediate plasmids. With 376 archived plasmids encoding TALE repeats as tetramers, trimers, dimers, and monomers, the fast ligation-based automatable solid-phase high-throughput (FLASH) system enables assembly of 96 TALE genes in less than one day. With automation, FLASH can make sequence-verified TALE expression plasmids for <$100 each, including the cost of labor (Reyon et al., 2012b). Instead of using a pre-assembled plasmid library, iterative capped assembly (ICA) builds full-length arrays from individual TALE repeat monomers (Briggs et al., 2012). Introduction of capping oligonucleotides eliminates incomplete ligation and monomer self-ligation, which are essential for the production of pure full-length TALE repeat arrays. With automation, ICA enables efficient assembly of TALE genes bearing up to 21 repeats followed by ligation into an expression plasmid within 3 h. Wang et al. (2012) performed solid-phase ligation on a chip, which allows the synthesis of >100 TALE genes bearing 16 or 20 repeats in 3 days.

Recently, a ligation-independent cloning (LIC) technique has been developed for high-throughput assembly of TALE genes (Schmid-Burgk et al., 2013). Compared with Golden Gate cloning, LIC relies on much longer (10–30 bp) non-palindromic overhangs to anneal with the overhangs of other fragments in a highly specific manner. The long overhangs are generated by the controllable 3′-exonuclease activity of T4 DNA polymerase. Because the fragments' long overlaps do not dissociate during transformation, the annealed products can be directly transformed into *Escherichia coli* without prior ligation step and ligated through bacterial ligases. Because of its high fidelity, LIC circumvents agar-based single-colony picking step, which allows growth of cells directly in polyclonal cultures after transformation. Using 64 repeat dimer-containing plasmids, LIC allows generation of correctly assembled TALE genes bearing 18.5 repeat units in three days through a hierarchical, two step assembly process. In addition, a comprehensive 5-mer TALE repeat unit fragment library composed of 3,072 plasmids was created, which enables automated assembly of >600 TALE genes bearing 15.5 repeat units in 1 day.

Conclusions and Future Perspectives

The last 3 years witnessed the tremendous progress of the TALEN technology. The scaffold optimization isolated TALEN variants with high DNA cleavage efficiency, which is essential for targeted genome editing. The characterization of novel RVDs extended the DNA recognition code and helped to minimize off-target cleavage activity of TALENs by increasing guanine recognition specificity. Development of novel strategies for convenient and quick assembly of TALE repeat arrays enabled high-throughput synthesis of TALENs and made TALEN technology accessible and affordable for any academic or industrial lab.

Besides TALENs, there are other tools available for editing genomes. Meganucleases are natural DNA endonucleases with high activity and specificity, but it is difficult to tailor their DNA recognition specificities. It is relatively easier to engineer ZFNs to target custom-designed DNA sequences, but some of them suffer from the requirement of intensive labor for construction and off-target effects (as reviewed in Sun et al., 2012a). Recently, Clustered regularly interspaced short palindromic repeats (CRISPR)-mediated DNA cleavage has been applied for genome editing (Cho et al., 2013; Cong et al., 2013; Hwang et al., 2013; Jiang et al., 2013; Jinek et al., 2013; Mali et al., 2013). This system can be reprogrammed readily using customized RNAs and enable multiplex genome engineering. However, the limited target specificity (14 bp) can cause off-target cleavage and the requirement for a protospacer adjacent motif (PAM) restricts its targeting range. Compared with these tools, TALENs have the advantage of high specificity and modularity, but there are also limitations that remain to be addressed for further improvement. The bulky size of TALENs might limit their broader applications, especially in the cases when efficient gene delivery cannot be achieved. Development of strategies for efficient delivery of TALEN genes into cells would enable TALEN-mediated genome editing in more different organisms and cell types. In eukaryotic cells, DNA is packaged into chromatin. Therefore, chromosomal context and epigenetic modifications play a major role in the DNA accessibility of TALENs. The combination of epigenetic modification tools and TALEN technology could expand the range of target for TALEN-mediated genome modifications, which might be a potential area for future exploration. Unlike ZFNs, the off-

target effects of TALENs have not been comprehensively characterized. Mussolino et al. (2011) carried out a side-by-side comparison between ZFNs and TALENs and found significantly reduced nuclease-associated cytotoxicities of TALENs. Ding et al. (2013) also reported minimal off-target-effects of TALENs using exome sequencing and whole-genome sequencing at low coverage, but they still could not completely rule out TALEN off-target effects. Therefore, careful screening of the complete genome of TALEN-modified cells using deep sequencing analysis would be instructive for safe use of TALENs, especially for human clinical applications.

Other than fusing with *Fok*I to make DNA endonucleases, TALEs have been used to create novel chimeric proteins by fusing with other functional protein domains. TALE-based transcription activators have been constructed to induce transcription of endogenous genes in plants (Morbitzer et al., 2010) and human cells (Bultmann et al., 2012; Garg et al., 2012; Geissler et al., 2011; Li et al., 2012c; Tremblay et al., 2012; Zhang et al., 2011). By fusing with transcription repressor domains, TALEs have been used to generate artificial repressors for sequence-specific gene repression in bacteria (Politz et al., 2012), yeast (Blount et al., 2012), plants (Mahfouz et al., 2012), and human cells (Cong et al., 2012; Li et al., 2012c). In addition, chimeric TALE recombinases (TALERs) have been constructed by fusing a hyperactivated catalytic domain from the DNA invertase Gin with a TALE central repeat domain (Mercer et al., 2012). TALERs with optimized architecture recombine DNA efficiently in bacterial and mammalian cells, providing an alternative approach for targeted genome editing. It would be interesting to combine TALE central repeat arrays with other different functional domains for many different applications in the future. For example, a TALE combined with a ligand-binding domain can be applied for high-throughput drug screening; a TALE combined with a DNA methyltransferase can be used for targeted DNA modification; a TALE combined with a histone deacetylase can be used for specific chromatin modification; a TALE combined with a cytosine deaminase can be applied for endogenous targeted mutagenesis, etc.

Thanks to simplicity in design, convenience in construction and high success rates across species, TALEN technology has received much attention since its invention. Although challenges and obstacles remain, TALEN technology will continue to be an important topic for future research and development and benefit both basic and applied biological sciences.

We thank the National Institutes of Health (R21 HL089418) and Department of Chemical and Biomolecular Engineering at the University of Illinois at Urbana-Champaign for financial support in our previous and current studies on the development of new genome engineering tools. We thank Dr. Julien Valton and Dr. Philippe Duchateau from Cellectis S.A. (Paris, France) for kindly sharing the structural model exhibiting RVD N* interacting with 5mC.

References

Ansai S, Sakuma T, Yamamoto T, Ariga H, Uemura N, Takahashi R, Kinoshita M. 2013. Efficient targeted mutagenesis in medaka using custom-designed transcription activator-like effector nucleases (TALENs). Genetics 193(3):739–749.

Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG 2nd, Tan W, Penheiter SG, Ma AC, Leung AY, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC. 2012. *In vivo* genome editing using a high-efficiency TALEN system. Nature 491:114–118.

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.

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.

Bogdanove AJ, Voytas DF. 2011. TAL effectors: Customizable proteins for DNA targeting. Science 333:1843–1846.

Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, Church GM. 2012. Iterative capped assembly: Rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40:e117.

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.

Cade L, Reyon D, Hwang WY, Tsai SQ, Patel S, Khayter C, Joung JK, Sander JD, Peterson RT, Yeh JR. 2012. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. Nucleic Acids Res 40(16):8001–8010.

Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, Voytas DF, Long CR, Whitelaw CB, Fahrenkrug SC. 2012. Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci USA 109:17382–17387.

Carroll D. 2011. Genome engineering with zinc-finger nucleases. Genetics 188(4):773–782.

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.

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.

Cho SW, Kim S, Kim JM, Kim JS. 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31:230–232.

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 ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF. 2012. Targeting G with TAL effectors: A comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS ONE 7:e45383.

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.

Cristea S, Freyvert Y, Santiago Y, Holmes MC, Urnov FD, Gregory PD, Cost GJ. 2013. In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. Biotechnol Bioeng 110:871–880.

Dahlem TJ, Hoshijima K, Jurynec MJ, Gunther D, Starker CG, Locke AS, Weis AM, Voytas DF, Grunwald DJ. 2012. Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS Genet 8:e1002861.

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

Deng D, Yin P, Yan C, Pan X, Gong X, Qi S, Xie T, Mahfouz M, Zhu JK, Yan N, Shi Y. 2012b. Recognition of methylated DNA by TAL effectors. Cell Res 22:1502–1504.

Ding Q, Lee YK, Schaefer EA, Peters DT, Veres A, Kim K, Kuperwasser N, Motola DL, Meissner TB, Hendriks WT, Trevisan M, Gupta RM, Moisan A, Banks E, Friesen M, Schinzel RT, Xia F, Tang A, Xia Y, Figueroa E, Wann A, Ahfeldt T, Daheron L, Zhang F, Rubin LL, Peng LF, Chung RT, Musunuru K, Cowan CA. 2013. A TALEN genome-editing system for generating human stem cell-based disease models. Cell Stem Cell 12:238–251.

Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC. 2011. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods 8:74–79.

Engler C, Kandzia R, Marillonnet S. 2008. A one pot, one step, precision cloning method with high throughput capability. PLoS ONE 3(11):e3647.

Gao H, Wu X, Chai J, Han Z. 2012. Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. Cell Res 22:1716–1720.

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

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.

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.

Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. 2011. Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29:699–700.

Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, Joung JK. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31:227–229.

Ishibashi S, Cliffe R, Amaya E. 2012. Highly efficient bi-allelic mutation rates using TALENs in *Xenopus tropicalis*. Biol Open 1:1273–1276.

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

Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. 2013. RNA-programmed genome editing in human cells. eLife 2:e00471.

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

Kim H, Um E, Cho SR, Jung C, Kim H, Kim JS. 2011. Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat Methods 8:941–943.

Lei Y, Guo X, Liu Y, Cao Y, Deng Y, Chen X, Cheng CH, Dawid IB, Chen Y, Zhao H. 2012. Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). Proc Natl Acad Sci USA 109:17484–17489.

Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, Yang B. 2011a. TAL nucleases (TALNs): Hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res 39:359–372.

Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, Weeks DP, Yang B. 2011b. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res 39:6315–6325.

Li L, Piatek MJ, Atef A, Piatek A, Wibowo A, Fang X, Sabir JS, Zhu JK, Mahfouz MM. 2012a. Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. Plant Mol Biol 78:407–416.

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

Li Y, Moore R, Guinn M, Bleris L. 2012c. Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. Sci Rep 2:897.

Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM, Jiao R. 2012. Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy. J Genet Genomics 39:209–215.

Ma S, Zhang S, Wang F, Liu Y, Liu Y, Xu H, Liu C, Lin Y, Zhao P, Xia Q. 2012. Highly efficient and specific genome editing in silkworm using custom TALENs. PLoS ONE 7:e45035.

Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu JK. 2011. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci USA 108:2623–2628.

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.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, Dicarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. Science 339(6121):823–826.

Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, Turecki G, Delaney A, Varhol R, Thiessen N, Shchors K, Heine VM, Rowitch DH, Xing X, Fiore C, Schillebeeckx M, Jones SJ, Haussler D, Marra MA, Hirst M, Wang T, Costello JF. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466:253–257.

Mercer AC, Gaj T, Fuller RP, Barbas CF III. 2012. Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic Acids Res 40:11163–11172.

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.

Moore FE, Reyon D, Sander JD, Martinez SA, Blackburn JS, Khayter C, Ramirez CL, Joung JK, Langenau DM. 2012. Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). PLoS ONE 7:e37877.

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.

Morbitzer R, Elsaesser J, Hausner J, Lahaye T. 2011. Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res 39:5790–5799.

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

Munoz Bodnar A, Bernal A, Szurek B, Lopez CE. 2013. Tell me a tale of TALEs. Mol Biotechnol 53:228–2235.

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.

Pattanayak V, Ramirez CL, Joung JK, Liu DR. 2011. Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. Nat Methods 8:765–770.

Perez-Pinera P, Ousterout DG, Gersbach CA. 2012. Advances in targeted genome editing. Curr Opin Chem Biol 16:268–277.

Politz MC, Copeland MF, Pfleger BF. 2012. Artificial repressors for controlling gene expression in bacteria. Chem Commun (Camb), DOI: 10.1039/c2cc37107c

Radecke S, Radecke F, Cathomen T, Schwarz K. 2010. Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: Wanted and unwanted target locus modifications. Mol Ther 18:743–753.

Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA, Cathomen T, Voytas DF, Joung JK. 2008. Unexpected failure rates for modular assembly of engineered zinc fingers. Nat Methods 5:374–375.

Reyon D, Khayter C, Regan MR, Joung JK, Sander JD. 2012a. Engineering designer transcription activator-like effector nucleases (TALENs) by REAL or REAL-Fast assembly. Curr Protoc Mol Biol Chapter 12:Unit 12.15.

Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. 2012b. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30:460–465.

Sajwan S, Takasu Y, Tamura T, Uchino K, Sezutsu H, Zurovec M. 2013. Efficient disruption of endogenous Bombyx gene by TAL effector nucleases. Insect Biochem Mol Biol 43:17–23.

Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JR. 2011. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29:697–698.

Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. 2012. A transcription activator-like effector toolbox for genome engineering. Nat Protoc 7:171–192.

Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V. 2013. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol 31:76–81.

Shan Q, Wang Y, Chen K, Liang Z, Li J, Zhang Y, Zhang K, Liu J, Voytas DF, Zheng X, Zhang Y, Gao C. 2013. Rapid and efficient gene modification in rice and Brachypodium using TALENs. Mol Plant, DOI: 10.1093/mp/sss162

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

Stroud DA, Formosa LE, Wijeyeratne XW, Nguyen TN, Ryan MT. 2013. Gene knockout using transcription activator-like effector nucleases (TALENs) reveals that human NDUFA9 protein is essential for stabilizing the junction between membrane and matrix arms of complex I. J Biol Chem 288:1685–1690.

Su Z, Han L, Zhao Z. 2011. Conservation and divergence of DNA methylation in eukaryotes: New insights from single base-resolution DNA methylomes. Epigenetics 6:134–140.

Sun N, Abil Z, Zhao H. 2012a. Recent advances in targeted genome engineering in mammalian systems. Biotechnol J 7:1074–1087.

Sun N, Liang J, Abil Z, Zhao H. 2012b. Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Biosyst 8:1255–1263.

Sung YH, Baek IJ, Kim DH, Jeon J, Lee J, Lee K, Jeong D, Kim JS, Lee HW. 2013. Knockout mice created by TALEN-mediated gene targeting. Nat Biotechnol 31(1):23–24.

Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. 2007. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol 25:786–793.

Tesson L, Usal C, Menoret S, Leung E, Niles BJ, Remy S, Santiago Y, Vincent AI, Meng X, Zhang L, Gregory PD, Anegon I, Cost GJ. 2011. Knockout rats generated by embryo microinjection of TALENs. Nat Biotechnol 29:695–696.

Tong C, Huang G, Ashton C, Wu H, Yan H, Ying QL. 2012. Rapid and cost-effective gene targeting in rat embryonic stem cells by TALENs. J Genet Genomics 39:275–280.

Tremblay JP, Chapdelaine P, Coulombe Z, Rousseau J. 2012. Transcription activator-like effector proteins induce the expression of the frataxin gene. Hum Gene Ther 23:883–890.

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.

Wang Z, Li J, Huang H, Wang G, Jiang M, Yin S, Sun C, Zhang H, Zhuang F, Xi JJ. 2012. An integrated chip for the high-throughput synthesis of transcription activator-like effectors. Angew Chem Int Ed Engl 51:8505–8508.

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.

Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S. 2011. Assembly of designer TAL effectors by Golden Gate cloning. PLoS ONE 6:e19722.

White FF, Potnis N, Jones JB, Koebnik R. 2009. The type III effectors of Xanthomonas. Mol Plant Pathol 10:749–7766.

Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ. 2011. Targeted genome editing across species using ZFNs and TALENs. Science 333:307.

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

Yu Y, Streubel J, Balzergue S, Champion A, Boch J, Koebnik R, Feng J, Verdier V, Szurek B. 2011. Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae pv. oryzae* depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. Mol Plant Microbe Interact 24:1102–1113.

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.