A transcription activator-like effector toolbox for genome engineering

Neville E Sanjana[1,2,4], Le Cong[1–4], Yang Zhou[1,2,4], Margaret M Cunniff[1,2], Guoping Feng[1,2] & Feng Zhang[1,2]

[1]Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts, USA. [2]Department of Brain and Cognitive Sciences, McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [3]Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, Massachusetts, USA. [4]These authors contributed equally to this work. Correspondence should be addressed to F.Z. (zhang_f@mit.edu).

Published online 5 January 2012; doi:10.1038/nprot.2011.431

Transcription activator-like effectors (TALEs) are a class of naturally occurring DNA-binding proteins found in the plant pathogen *Xanthomonas sp*. The DNA-binding domain of each TALE consists of tandem 34–amino acid repeat modules that can be rearranged according to a simple cipher to target new DNA sequences. Customized TALEs can be used for a wide variety of genome engineering applications, including transcriptional modulation and genome editing. Here we describe a toolbox for rapid construction of custom TALE transcription factors (TALE-TFs) and nucleases (TALENs) using a hierarchical ligation procedure. This toolbox facilitates affordable and rapid construction of custom TALE-TFs and TALENs within 1 week and can be easily scaled up to construct TALEs for multiple targets in parallel. We also provide details for testing the activity in mammalian cells of custom TALE-TFs and TALENs using quantitative reverse-transcription PCR and Surveyor nuclease, respectively. The TALE toolbox described here will enable a broad range of biological applications.

INTRODUCTION

Systematic reverse-engineering of the functional architecture of the mammalian genome requires the ability to perform precise perturbations on gene sequences and transcription levels. Tools capable of facilitating targeted genome editing and transcription modulation are essential for elucidating the genetic and epigenetic basis of diverse biological functions and diseases. The recent discovery of the TALE code[1,2] has enabled the generation of custom TALE DNA-binding domains with programmable specificity[3–12]. When coupled to effector domains, customized TALEs provide a promising platform for achieving a wide variety of targeted genome manipulations[3–5,8,11,13,14]. Previously, we reported efficient construction of TALEs with customized DNA-binding domains for activating endogenous genes in the mammalian genome[3]. Here we describe an improved protocol for rapid construction of customized TALEs and methods to apply these TALEs to achieve endogenous transcriptional activation[3–5,8] and site-specific genome editing[4,7,9,11–15]. Investigators should be able to use this protocol to construct TALEs for targets of their choice in less than 1 week.

Transcription activator-like effectors

TALEs are natural bacterial effector proteins used by *Xanthomonas sp.* to modulate gene transcription in host plants to facilitate bacterial colonization[16,17]. The central region of the protein contains tandem repeats of 34-aa sequences (termed monomers) that are required for DNA recognition and binding[18–21] (**Fig. 1a**). Naturally occurring TALEs have been found to have a variable number of monomers, ranging from 1.5 to 33.5 (ref. 16). Although the sequence of each monomer is highly conserved, they differ primarily in two positions termed the repeat variable diresidues (RVDs, 12th and 13th positions). Recent reports have found that the identity of these two residues determines the nucleotide-binding specificity of each TALE repeat and that a simple cipher specifies the target base of each RVD (NI = A, HD = C, NG = T, NN = G or A)[1,2]. Thus, each monomer targets one nucleotide and the linear sequence of monomers in a TALE specifies the target DNA sequence in the 5′ to 3′ orientation. The natural TALE-binding sites within plant genomes always begin with a thymine[1,2], which is presumably specified by a cryptic signal within the nonrepetitive N terminus of TALEs. The tandem repeat DNA-binding domain always ends with a half-length repeat (0.5 repeat, **Fig. 1a**). Therefore, the length of the DNA sequence being targeted is equal to the number of full repeat monomers plus two.

Comparison with other genome manipulation methods

For targeted gene insertion and knockout, there are several techniques that have been used widely in the past, such as homologous gene targeting[22–24], transposases[25,26], site-specific recombinases[27], meganucleases[28] and integrating viral vectors[29,30]. However, most of these tools target a preferred DNA sequence and cannot be easily engineered to function at noncanonical DNA target sites. The most promising, programmable DNA-binding domain has been the artificial zinc-finger (ZF) technology, which enables arrays of ZF modules to be assembled into a tandem array and target new DNA-binding sites in the genome. Each finger module in a ZF array targets three DNA bases[31,32]. In comparison, TALE DNA-binding monomers target single nucleotides and are much more modular than ZF modules. For instance, when two independent ZF modules are assembled into a new array, the resulting target site cannot be easily predicted based on the known binding sites for the individual finger modules. Perhaps the biggest caveat of ZFs is that most of the intellectual property surrounding the ZF technology platform is proprietary and expensive (>$10,000 per target site). A public effort for ZF technology development also exists through the Zinc Finger Consortium, but the publicly available ZF modules can only target a subset of the 64 possible trinucleotide combinations[33–35]. TALEs theoretically can target any sequence and have already been used in many organisms with impressive success (**Table 1**). Although TALEs seem superior in many ways, ZFs have a much longer track record in DNA-targeting applications[32], including their use in human clinical trials[36]. Despite their relatively recent development, early results with TALEs have been promising and it seems that they can be applied in the same way as ZFs for many

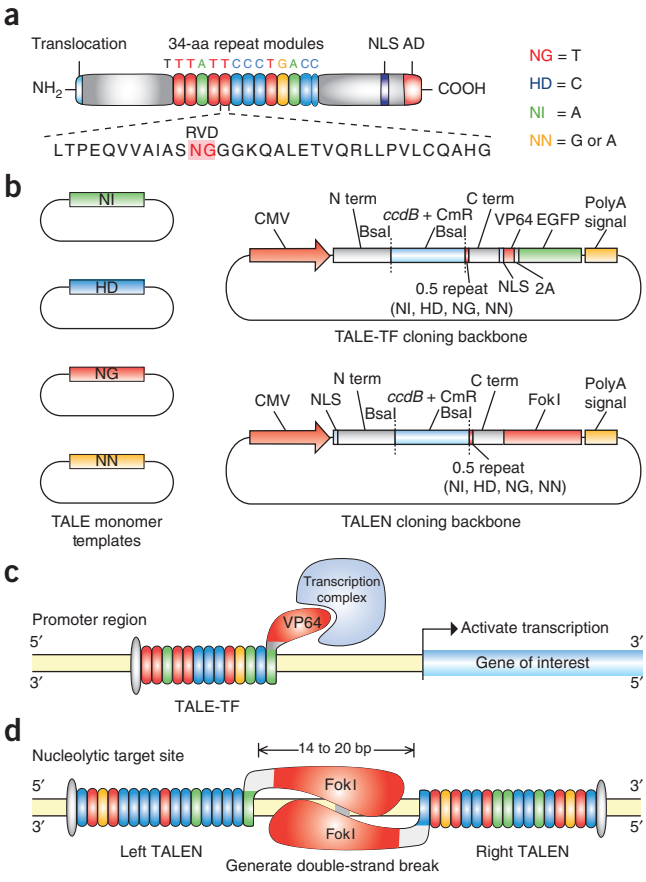


Figure 1 | A TALE toolbox for genome engineering. (**a**) Natural structure of TALEs derived from *Xanthomonas sp*. Each DNA-binding module consists of 34 amino acids, where the RVDs in the 12th and 13th amino acid positions of each repeat specify the DNA base being targeted according to the cipher NG = T, HD = C, NI = A, and NN = G or A. The DNA-binding modules are flanked by nonrepetitive N and C termini, which carry the translocation, nuclear localization (NLS) and transcription activation (AD) domains. A cryptic signal within the N terminus specifies a thymine as the first base of the target site. (**b**) The TALE toolbox allows rapid and inexpensive construction of custom TALE-TFs and TALENs. The kit consists of 12 plasmids in total: four monomer plasmids to be used as templates for PCR amplification, four TALE-TF and four TALEN cloning backbones corresponding to four different bases targeted by the 0.5 repeat. CMV, cytomegalovirus promoter; N term, nonrepetitive N terminus from the Hax3 TALE; C term, nonrepetitive C terminus from the Hax3 TALE; BsaI, type IIs restriction sites used for the insertion of custom TALE DNA-binding domains; *ccdB* + CmR, negative selection cassette containing the *ccdB* negative selection gene and chloramphenicol resistance gene; NLS, nuclear localization signal; VP64, synthetic transcriptional activator derived from VP16 protein of herpes simplex virus; 2A, 2A self-cleavage linker; EGFP, enhanced green fluorescent protein; polyA signal, polyadenylation signal; FokI, catalytic domain from the FokI endonuclease. (**c**) TALEs can be used to generate custom TALE-TFs and modulate the transcription of endogenous genes from the genome. The TALE DNA-binding domain is fused to the synthetic VP64 transcriptional activator, which recruits RNA polymerase and other factors needed to initiate transcription. (**d**) TALENs can be used to generate site-specific double-strand breaks to facilitate genome editing through nonhomologous repair or homology directed repair. Two TALENs target a pair of binding sites flanking a 16-bp spacer. The left and right TALENs recognize the top and bottom strands of the target sites, respectively. Each TALE DNA-binding domain is fused to the catalytic domain of FokI endonuclease; when FokI dimerizes, it cuts the DNA in the region between the left and right TALEN-binding sites.

DNA-targeting applications (e.g., transcriptional modulator[3–5,8], nuclease[4,7,9,11–15], recombinase[37–39], transposase[40,41]).

Constructing customized TALE-TFs and TALENs

Because of the repetitive nature of TALEs, construction of the DNA-binding monomers can be difficult. Previously, we and other groups have used a hierarchical ligation strategy to overcome the difficulty of assembling the monomers into ordered multimer arrays, taking advantage of degeneracy in the codons surrounding the monomer junction and Type IIs restriction enzymes[3,6–10]. In this protocol, we use the same basic strategy that we previously used[3] to construct TALE-TFs to modulate transcription of endogenous human genes. We have further improved the TALE assembly system with a few optimizations, including maximizing the dissimilarity of ligation adaptors to minimize misligations and combining separate digest and ligation steps into single Golden Gate[42–44] reactions. Briefly, we first amplify each nucleotide-specific monomer sequence with ligation adaptors that uniquely specify the monomer position within the TALE tandem repeats. Once this monomer library is produced, it can conveniently be reused for the assembly of many TALEs. For each TALE desired, the appropriate monomers are first ligated into hexamers, which are then amplified via PCR. Then, a second Golden Gate digestion-ligation with the appropriate TALE cloning backbone (**Fig. 1b**) yields a fully assembled, sequence-specific TALE. The backbone contains a *ccdB* negative selection cassette flanked by the TALE N and C termini, which is replaced by the tandem repeat DNA-binding domain when the TALE has been successfully constructed. *ccdB* selects against cells transformed with an empty backbone, thereby yielding clones with tandem repeats inserted[7].

Assemblies of monomeric DNA-binding domains can be inserted into the appropriate TALE-TF or TALEN cloning backbones to construct customized TALE-TFs and TALENs. TALE-TFs are constructed by replacing the natural activation domain within the TALE C terminus with the synthetic transcription activation domain VP64 (ref. 3; **Fig. 1c**). By targeting a binding site upstream of the transcription start site, TALE-TFs recruit the transcription complex in a site-specific manner and initiate gene transcription. TALENs are constructed by fusing a C-terminal truncation (+63 aa) of the TALE DNA-binding domain[4] with the nonspecific FokI endonuclease catalytic domain (**Fig. 1d**). The +63-aa C-terminal truncation has also been shown to function as the minimal C terminus sufficient for transcriptional modulation[3]. TALENs form dimers through binding to two target sequences separated by ~17 bases. Between the pair of binding sites, the FokI catalytic domains dimerize and function as molecular scissors by introducing double-strand breaks (DSBs; **Fig. 1d**). Normally, DSBs are repaired by the nonhomologous end-joining[45] pathway (NHEJ), resulting in small deletions and functional gene knockout. Alternatively, TALEN-mediated DSBs can stimulate homologous recombination, enabling site-specific insertion of an exogenous donor DNA template[4,13].

We also present a short procedure for verifying correct TALE assembly by using colony PCR to verify the correct insert length followed by DNA sequencing. With our cloning procedure, we routinely achieve high efficiency (correct length) and high accuracy (correct sequence). The cloning procedure is modular in several ways: we can construct TALEs to target DNA sequences of different lengths, and the protocol is the same for producing either

TABLE 1 | Applications of custom TALEs on endogenous genome targets.

	Species	Genomic loci	References
TALE-TF	*Arabidopsis thaliana*	*egl3*	5
		knat1	
	Homo sapiens	*KLF4*	3
		SOX2	
		NTF3	4
		PUMA	8
		IFNA1	
		IFNB1	
TALEN	*Saccharomyces cerevisiae*	*ura3*	9
		lys2	
		ade2	
	H. sapiens	*CCR5*	4
		NTF3	
		PPP1R12C (AAVS1)	13
		OCT4 (POU5F1)	
		PITX3	
	Caenorhabditis elegans	*ben-1*	11
	Danio rerio	*hey2*	58,59
		gria3a	
		tnikb	
	Rattus norvegicus	*Igm*	60

TALE-TFs or TALENs. The backbone vectors can be modified with different promoters to achieve cell type–specific expression.

Our protocol includes functional assays for evaluating TALE-TF and TALEN activity in human cells. This step is important because we have observed some variability in TALE activity on the endogenous genome, possibly because of epigenetic repression and/or inaccessible chromatin at certain loci. For TALE-TFs, we perform quantitative reverse-transcription PCR (qRT-PCR) to quantify changes in gene expression. For TALENs, we use the Surveyor mutation detection assay (i.e., the base-mismatch cleaving endonuclease *Cel2*) to quantify NHEJ. Although these assays are standard and have already been described elsewhere[46,47], we feel that the functional characterization is integral to TALE production and therefore have presented it here with the assembly procedure. Other functional assays, such as plasmid-based reporter constructs[3,7], restriction sites destroyed by NHEJ[48] or other enzymes that detect DNA mismatch[49], may also be used to validate TALE activity.

Our protocol (**Fig. 2**) begins with the generation of a monomer library, which takes 1 d and can be reused for building many TALEs. Using the monomer library, several TALEs can be constructed in a single day with an additional 2 d for transformation and sequence verification. To assess TALE function on the endogenous genome, we take ~3 d to go from mammalian cell transfection to qRT-PCR or Surveyor results.

Comparison with other TALE assembly procedures

A number of TALE assembly procedures have described the use of Golden Gate cloning to construct customized TALE DNA-binding domains[3,6–10]. These methods rely on the use of a large collection of plasmids (typically over 50 plasmids) encoding repeat monomers and intermediate cloning vectors. Our PCR-based approach requires substantially less initial plasmid preparation, as our monomer library can be amplified on one 96-well PCR plate, and it facilitates more rapid construction of custom TALEs. Plasmid-based amplification has a much lower mutation/error rate but, in our experience, the combination of a high-fidelity polymerase and the short length of the monomer template (~100 nt) results in accurate assembly. For building similar-length TALEs to those presented in this protocol, the plasmid-based approaches also require an additional transformation and colony selection that extends the time needed to build TALEs. Thus, these alternative assembly protocols require a greater time investment both up-front (for monomer library preparation) and on a recurring basis (for each new TALE). For laboratories seeking to produce TALEs quickly, our protocol requires only a few hours to prepare a complete monomer library and less than 1 d to proceed from monomers to the final transformation into bacteria.

Targeting limitations

There are a few key limitations with the TALE technology. Although the RVD cipher is known, it is still not well understood as to why different TALEs designed according to the same cipher act on their target sites in the native genome with different levels of activity. It is possible that there are yet-unknown sequence dependencies for efficient binding or site-specific constraints (e.g., chromatin states) that are responsible for differences in functional activity. Therefore, we suggest constructing at least two or three TALE-TFs or TALEN pairs for each target locus. In addition, it is possible that engineered TALEs can have off-target effects—i.e., binding unintended genomic loci—which can be difficult to detect without additional functional assays at these loci. Given the relatively early state of TALE technology development, these issues remain to be addressed in a conclusive manner.

Experimental design

TALE-TF target site selection. The programmable nature of TALEs allows for a virtually arbitrary selection of target DNA-binding sites. As previously reported, the N terminus of the TALE requires that the target site begin with a thymine nucleotide. For TALE-TFs, we have been successfully targeting 14- to 20-bp sequences within 200 bp of the transcription start site (**Fig. 1c**). It can be advantageous to select a longer sequence to reduce off-target activation, as it is known from reporter activation assays that TALEs interact less efficiently with targets containing more than one mismatching base. In our assembly protocol, we describe ligation of 18 monomers into a backbone containing a nucleotide-specific final 0.5 monomer; combined with the initial thymine requirement, this yields a total sequence specificity of 20 nt. Specifically, the TALE-TF-binding site takes the form 5′-TN[19]-3′. When selecting TALE-TF-targeting sites for modulating endogenous gene transcription, we recommend selecting multiple target sites within the proximal promoter region (targeting either the sense or antisense strand), as epigenetic and local chromatin dynamics might impede TALE binding.

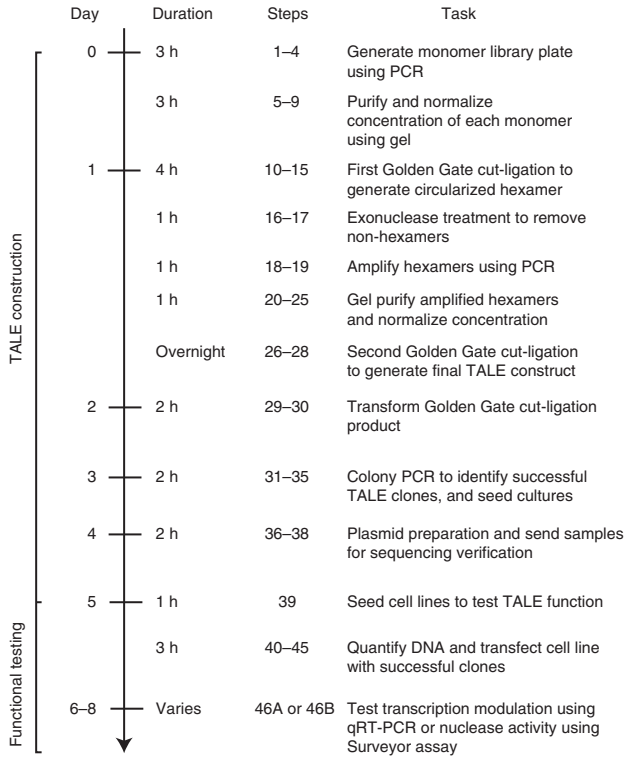


Figure 2 | Timeline for the construction of TALE-TFs and TALENs. Steps for the construction and functional testing of TALE-TFs and TALENs are outlined. TALEs can be constructed and sequence verified in 5 d following a series of ligation and amplification steps. During the construction phase, samples can be stored at −20 °C at the end of each step and continued at a later date. After TALE construction, functional validation via qRT-PCR (for TALE-TFs) and Surveyor nuclease assay (for TALENs) can be completed in 2–3 d.

Larger TALEs might be beneficial for TALE-TFs targeting genes with less unique regions upstream of their transcription start site.

TALEN target site selection. Because TALENs function as dimers, a pair of TALENs, referred to as the left and right TALENs, need to be designed to target a given site in the genome. The left and right TALENs target sequences on opposite strands of DNA (**Fig. 1d**). As with TALE-TF, we design each TALEN to target a 20-bp sequence. TALENs are engineered as a fusion of the TALE DNA-binding domain and a monomeric FokI catalytic domain. To facilitate FokI dimerization, the left and right TALEN target sites are chosen with a spacing of approximately 14–20 bases. Therefore, for a pair of TALENs, each targeting 20-bp sequences, the complete target site should have the form 5′-TN^{19}N$^{14\text{-}20}$N^{19}A-3′, where the left TALEN targets 5′-TN^{19}-3′ and the right TALEN targets the antisense strand of 5′-N^{19}A-3′ (N = A, G, T or C). TALENs should have fewer off-target effects because of the dimerization requirement for the FokI nuclease, although no significant off-target effects have been observed in limited sequencing verifications[13]. Because DSB formation only occurs if the spacer between the left and right TALEN-binding sites (**Fig. 1d**) is approximately 14–20 bases, nuclease activity is restricted to genomic sites with both the specific sequences of the left TALEN and the right TALEN with this small range of spacing distances between those sites. These constraints should greatly reduce potential off-target effects.

TALE monomer design. To ensure that all synthesized TALEs are transcribed at a similar level, all of the monomers have been optimized to share identical DNA sequences except in the variable diresidues, and they are codon-optimized for expression in human cells (**Supplementary Data 1**). This should minimize any difference in translation due to codon availability.

Construction strategy. Synthesis of monomeric TALE DNA-binding domains in a precise order is challenging because of their highly repetitive nature. Previously[3], we took advantage of codon redundancy at the junctions between neighboring monomers and devised a hierarchical ligation strategy to construct ordered assemblies of multiple monomers. In this protocol, we describe a similar strategy, but with several important improvements that make the procedure easier, more flexible and more reliable (**Fig. 3**).

In our initial protocol[3], the digestion and ligation steps were carried out separately with an intervening DNA purification step. This improved protocol adopts the powerful Golden Gate cloning technique[42–44], requiring less hands-on time and resulting in a more efficient reaction. The Golden Gate procedure involves combining the restriction enzyme and ligase together in a single reaction with a mutually compatible buffer. The reaction is cycled between optimal temperatures for digestion and ligation. Golden Gate digestion-ligation capitalizes on Type IIs restriction enzymes, for which the recognition sequence is spatially separated from where the cut is made. During a Golden Gate reaction, the correctly ligated products no longer contain restriction enzyme recognition sites and cannot be further digested. In this manner, Golden Gate drives the reaction toward the correct ligation product, as the number of cycles of digestion and ligation increases.

For the hierarchical ligation steps, we have optimized our previous cloning strategy for faster TALE production. The improved design takes advantage of a circularization step that allows only properly assembled hexameric intermediates to be preserved (**Fig. 3**). Correctly ligated hexamers consist of six monomers ligated together in a closed circle, and incomplete ligation products are left as linear DNA. After this ligation step, an exonuclease degrades all noncircular DNA, leaving intact only the complete circular hexamers. Without circularization and exonuclease treatment, the correct ligation product would need to be gel purified before proceeding. The combination of Golden Gate digestion-ligation and circularization reduces the overall hands-on time required for TALE assembly.

Primer design for monomer library preparation. Each monomer in the tandem repeat must have its position uniquely specified. The monomer primers are designed to add ligation adaptors that enforce this positioning. Our protocol uses a hierarchical ligation strategy: For the 18-mer tandem repeat, we first ligate monomers into hexamers. Then, we ligate three hexamers together to form the 18-mer. By breaking down the assembly into two steps, we do not need unique ligation junctions for each monomer in the 18-mer. Instead, the same set of ligation junctions internal to each hexamer are reused in all three hexamers (first ligation step), whereas unique (external) ligation junctions are used to flank each hexamer (second ligation step). As shown in **Figure 4**, the internal primers used to amplify the monomers within each hexamer are the same, but the external primers differ between the

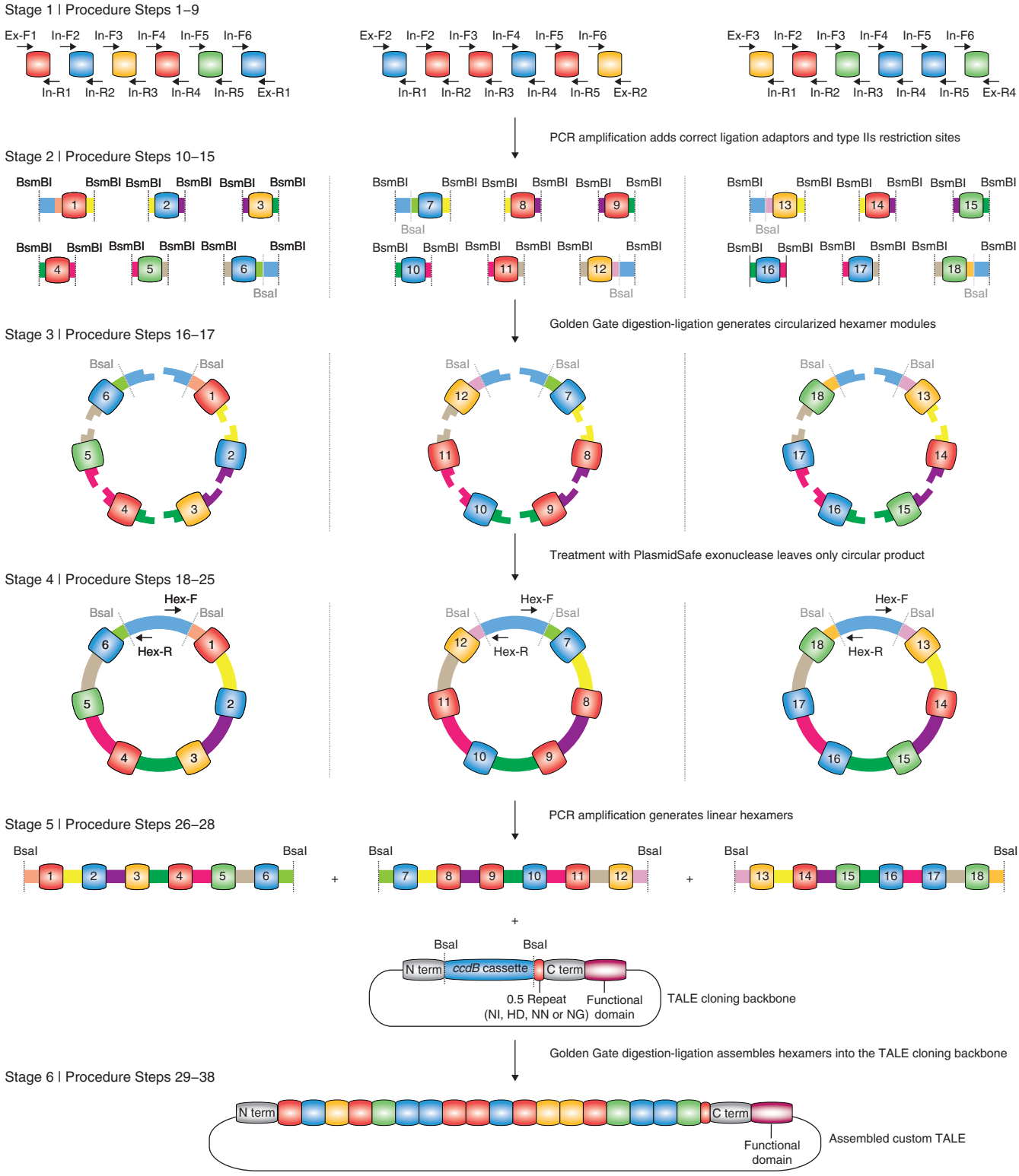


Figure 3 | Construction of TALE DNA-binding domains using hierarchical ligation assembly. Schematic of the construction process for a custom TALE containing an 18-mer tandem repeat DNA-binding domain. Stage 1: specific primers are used to amplify each monomer and add the appropriate ligation adaptors (Procedure Steps 1–9). Stage 2: hexameric tandem repeats (1–6, 7–12 and 13–18) are assembled first using Golden Gate digestion-ligation. The 5′ ends of monomers 1, 7 and 13 and the 3′ ends of monomers 6, 12 and 18 are designed so that each tandem hexamer assembles into an intact circle (Procedure Steps 10–15). Stage 3: the Golden Gate reaction is treated with an exonuclease to remove all linear DNA, leaving only the properly assembled tandem hexamer (Procedure Steps 16 and 17). Stage 4: each tandem hexamer is amplified individually using PCR and purified (Procedure Steps 18–25). Stage 5: tandem hexamers corresponding to 1–6, 7–12 and 13–18 are ligated into the appropriate TALE-TF or TALEN cloning backbone using Golden Gate cut-ligation (Procedure Steps 26–28). Stage 6: the assembled TALE-TF or TALEN is transformed into competent cells, and successful clones are isolated and sequence verified (Procedure Steps 29–38).

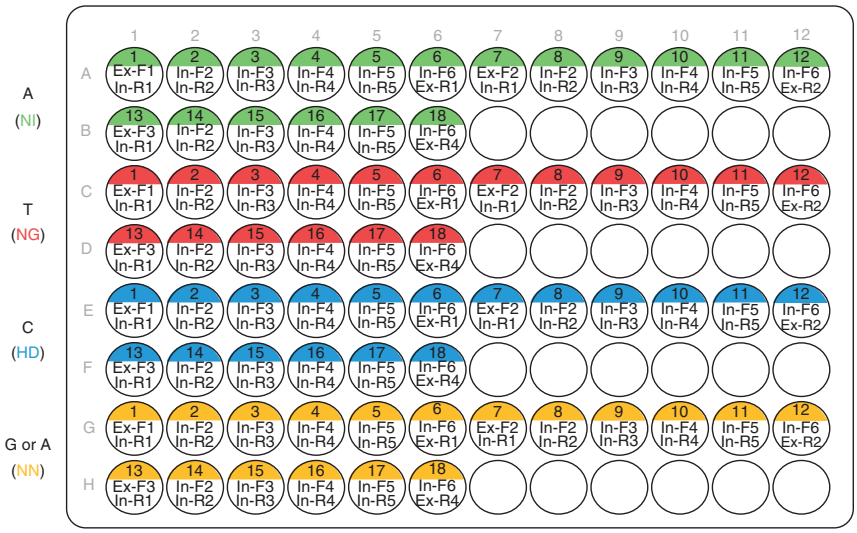


Figure 4 | PCR plate setup used to generate a plate of monomers for constructing custom 18-mer TALE DNA-binding domains. One 96-well plate can be used to carry out 72 reactions (18 for each monomer template). The position of each monomer and the primers used for the position is indicated in the well. Color coding in the well indicates the monomer used as the PCR template. Typically, two to four plates of 100-µl PCRs are pooled together and purified to generate a monomer library of sufficient quantity for production of many TALEs. During TALE construction, the corresponding monomer for each DNA base in the 18-bp target sequence can be easily picked from the plate.

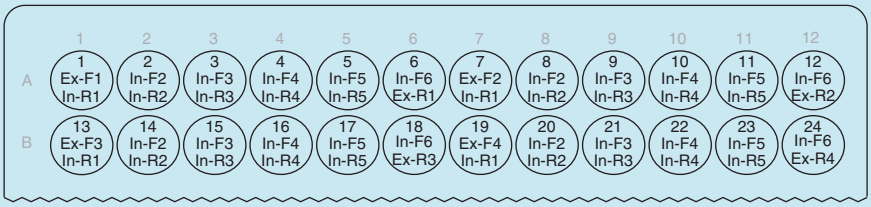
hexamers. By reusing the same internal primers between different hexamers, our protocol minimizes the number of primers necessary for monomer amplification.

Controls. As a negative control for Golden Gate assembly, we recommend performing a separate reaction with only the TALE-TF or TALEN backbone. Transformation of this negative control should result in few or no colonies because of the omission of the tandem repeats and resulting religation of the toxic *ccdB* insert. After completing the TALE cloning, we use colony PCR or restriction digests to screen for correct length clones. For the final verification of proper assembly, we sequence the entire length of the tandem repeats. Owing to limits in Sanger sequencing read length, other TALE assembly protocols have difficulty sequencing the entire tandem repeat region[7,9,10]. The similarity of the monomers within the region makes primer annealing to specific monomers impossible. We have overcome this problem by slightly modifying the codon usage at the 5′ end of monomer 7 to create a unique annealing site, so that a TALE with an 18-mer DNA-binding array can be verified through a combination of three staggered sequencing reads. Specifically, during the monomer amplification, the codons for the first five amino acids in monomer 7 are mutated

Box 1 | Building TALEs that target DNA sequences of different lengths

In the main protocol, we present a hierarchical ligation strategy for the construction of TALEs that contain 18 full monomer repeats; however, the general approach can be easily adapted to construct TALEs of any length. These TALEs containing 18 full repeat monomers bind to 20-bp DNA sequences, where the first and last bases are specified by the N terminus and the 0.5 repeat, respectively (**Fig. 1a**). We chose this length because, empirically, we have observed that 20-bp sequences tend to be unique within the human genome. Nevertheless, for different species (e.g., with larger or more repetitive genomes) or for repetitive regions within the human genome, it can be advantageous to construct longer or shorter TALEs. For certain genomic loci, it might also be difficult to identify TALEN target sites that satisfy the spacing constraints when the binding sites for both left and right TALENs are restricted to 20-bp sequences.

Our protocol is easily modified for the construction of TALEs containing up to 24 full monomer repeats by changing the order in which particular primers are used during the preparation of the monomer library plate (as described in Procedure Steps 1–9). All other steps remain essentially the same. A plate of monomer amplification primers (similar to **Fig. 4**) can be prepared for building TALEs with 24 full monomer repeats, which bind to 26-bp DNA sequences, as illustrated below. In this case, a fourth circular hexamer, corresponding to monomers 19 through 24, is also built and treated identically as the other three circular hexamers (1–6, 7–12 and 13–18).



For building shorter TALEs, only a single change to monomer amplification is needed: the final monomer should be amplified with the Ex-R4 reverse primer. For example, to build TALEs with 17 monomers instead of 18, the monomer templates (NI, NG, NN, HD) should be amplified with the forward/reverse primer combination In-F5/Ex-R4. Note that during gel purification (Procedure Step 20) the desired PCR amplicon is a pentamer containing monomers 13–17 and it will run faster than the hexamers (1–6, 7–12). After purification, ensure that the pentameric and hexameric intermediates are used at an equimolar ratio in the final Golden Gate digestion-ligation.

via PCR to use different but synonymous codons, creating a unique priming site without changing the encoded TALE protein. This modification allows each hexamer in the 18-mer to be sequenced with a separate sequencing read and requires only a standard read length of ~700 bp for complete sequence verification. For TALEs containing more than 18 full monomer repeats, we introduced a third unique priming site for sequencing at the 3′ end of the 18th monomer using a similar approach. For the construction of TALEs containing up to 24 full monomers with the entire tandem repeat region easily sequenced, see **Box 1**.

Design of functional validation assays. For TALE-TFs, qRT-PCR quantitatively measures the increase in transcription driven by the TALE-TF. For TALENs, the Surveyor assay provides a functional validation of TALEN cutting and quantifies the cutting efficiency of a particular pair of TALENs. These assays should be performed in the same cell type as intended for the TALE application, as TALE efficacy can vary between cell types, presumably because of differences in chromatin state or epigenetic modifications.

For qRT-PCR, we use commercially available probes to measure increased transcription of the TALE-TF–targeted gene. For most genes in the human or mouse genomes, specific probes can be purchased (e.g., TaqMan gene expression probes from Applied Biosystems). There are a wide variety of qRT-PCR protocols, and although we describe one of them here others can be substituted. For example, a more economical option is to design custom, transcript-specific primers (e.g., with NCBI Primer-BLAST) and use a standard fluorescent dye to detect amplified double-stranded DNA (e.g., SYBR Green).

For Surveyor, we follow the recommendations given by the assay manufacturer when designing specific primers for genomic PCR. We typically design primers that are ~30 nt long and with melting temperatures of ~65 °C. The primers should flank the TALEN target site and generate an amplicon of approximately 300–800 bp with the TALEN target site near the middle. During the design, we also check to make sure the primers are specific over the intended genome using NCBI Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Before using the primers for Surveyor, the primers and specific PCR cycling parameters should be tested to ensure that amplification results in a single clean band. In difficult cases in which a single-band product cannot be achieved, it is acceptable to gel-extract the correct-length band before proceeding with heteroduplex reannealing and Surveyor nuclease digestion.

MATERIALS

REAGENTS

TALE construction

- TALE monomer template plasmids (Addgene): pNI_v2, pNG_v2, pNN_v2, pHD_v2
- TALE transcriptional activator (TALE-TF) plasmids (Addgene): pTALE-TF_v2 (NI), pTALE-TF_v2 (NG), pTALE-TF_v2 (NN), pTALE-TF_v2 (HD)
- TALE nuclease (TALEN) backbone plasmids (Addgene): pTALEN_v2 (NI), pTALEN_v2 (NG), pTALEN_v2 (NN), pTALEN_v2 (HD). These plasmids can be obtained individually or bundled together as a single kit from the Zhang Lab plasmid collection at Addgene (http://www.addgene.org/TALE_Toolbox). See **Supplementary Data 1** for plasmid sequences.
- PCR primers for TALE construction (**Table 2**, Integrated DNA Technologies, custom DNA oligonucleotides)
- Herculase II fusion polymerase (Agilent Technologies, cat. no. 600679) ▲ **CRITICAL** Standard Taq polymerase, which lacks 3′-5′ exonuclease proofreading activity, has lower fidelity and can lead to errors in the final assembled TALE. Herculase II is a high-fidelity polymerase (equivalent fidelity to Pfu) that produces high yields of PCR product with minimal optimization. Other high-fidelity polymerases may be substituted.
- Herculase II reaction buffer (5×; Agilent Technologies, included with polymerase)
- Taq-B polymerase (Enzymatics, cat. no. P725L)
- Taq-B buffer (10×; Enzymatics, included with polymerase)
- dNTP solution mix (25 mM (each); Enzymatics, cat. no. N205L)
- MinElute gel extraction kit (Qiagen, cat. no. 28606) ▲ **CRITICAL** MinElute columns should be stored at 4 °C until use.
- QIAprep spin miniprep kit (Qiagen, cat. no. 27106)
- QIAquick 96 PCR purification (Qiagen, cat. no. 28181)
- UltraPure DNaseRNase-free distilled water (Invitrogen, cat. no. 10977-023)
- UltraPure TBE buffer (10×; Invitrogen, cat. no. 15581-028)
- SeaKem LE agarose (Lonza, cat. no. 50004)
- SYBR Safe DNA stain (10,000×; Invitrogen, cat. no. S33102)
- Low-DNA mass ladder (Invitrogen, cat. no. 10068-013)
- 1-kb Plus DNA ladder (Invitrogen, cat. no. 10787-018)
- TrackIt CyanOrange loading buffer (Invitrogen, cat. no. 10482-028)
- Restriction enzymes: BsmBI (Esp3I) (Fermentas/ThermoScientific, cat. no. ER0451), BsaI-HF (New England Biolabs, cat. no. R3535L), AfeI (New England Biolabs, cat. no. R0652S)
- Fermentas Tango Buffer and 10× NEBuffer 4 (included with enzymes)
- Bovine serum albumin (100×; New England Biolabs, included with BsaI-HF)
- DL-dithiothreitol (DTT; Fermentas/ThermoScientific, cat. no. R0862)
- T7 DNA ligase (3,000 U μl^{-1}; Enzymatics, cat. no. L602L) ▲ **CRITICAL** Do not substitute the more commonly used T4 ligase. T7 ligase has 1,000-fold higher activity on the sticky ends than on the blunt ends and higher overall activity than commercially available concentrated T4 ligases.
- Adenosine 5′-triphosphate (10 mM; New England Biolabs, cat. no. P0756S)
- PlasmidSafe ATP-dependent DNase (Epicentre, cat. no. E3101K)
- One Shot Stbl3 chemically competent *Escherichia coli* (*E. coli*) (Invitrogen, cat. no. C7373-03)
- SOC medium (New England Biolabs, cat. no. B9020S)
- LB medium (Sigma, cat. no. L3022)
- LB agar medium (Sigma, cat. no. L2897)
- Ampicillin, sterile filtered (100 mg ml^{-1}; Sigma, cat. no. A5354)

TALEN and TALE-TF functional validation in mammalian cells

- HEK293FT cells (Invitrogen, cat. no. R700-07)
- Dulbecco's minimum Eagle's medium (DMEM, 1×, high glucose; Invitrogen, cat. no. 10313-039)
- Dulbecco's phosphate-buffered saline (DPBS, 1×; Invitrogen, cat. no. 14190-250)
- Fetal bovine serum, qualified and heat inactivated (Invitrogen, cat. no. 10438-034)
- Opti-MEM I reduced-serum medium (FBS; Invitrogen, cat. no. 11058-021)
- GlutaMAX-I (100×; Invitrogen, cat. no. 35050079)
- Penicillin-streptomycin (100×; Invitrogen, cat. no. 15140-163)
- Trypsin, 0.05% (wt/vol) (1×) with EDTA·4Na (Invitrogen, cat. no. 25300-062)
- Lipofectamine 2000 transfection reagent (Invitrogen, cat. no. 11668027)
- QuickExtract DNA extraction solution (Epicentre, cat. no. QE09050)
- Herculase II fusion polymerase ▲ **CRITICAL** As Surveyor assay is sensitive to single-base mismatches, it is important to use only a high-fidelity polymerase. Other high-fidelity polymerases can be substituted; refer to the Surveyor manual for PCR buffer compatibility details.
- Herculase II reaction buffer (5×)
- Surveyor mutation detection kit for standard gel electrophoresis (Transgenomic, cat. no. 706025) ▲ **CRITICAL** The Surveyor assay includes the Cel2 base-mismatch nuclease. Alternatives include the Cel1, T7, mung

TABLE 2 | Primer sequences for TALE construction.

Name	Sequence	Purpose
Ex-F1	5′-TGCGTCcgtctcCGAACCTTAAACCGGCCAACATACCggtctcCTGACCCCAGAGCAGGTCGTG-3′	Monomer amplification (primers Ex-F1 through In-R5)
Ex-F2	5′-TGCGTCcgtctcCGAACCTTAAACCGGCCAACATACCggtctcGACTTACACCCGAACAAGTCGTGGCA ATTGCGAGC-3′	
Ex-F3	5′-TGCGTCcgtctcCGAACCTTAAACCGGCCAACATACCggtctcGCGGCCTCACCCCAGAGCAGGTCG-3′	
Ex-F4	5′-TGCGTCcgtctcCGAACCTTAAACCGGCCAACATACCggtctcGTGGGCTCACCCCAGAGCAGGTCG-3′	
Ex-R1	5′-GCTGACcgtctcCGTTCAGTCTGTCTTTCCCCTTTCCggtctcTAAGTCCGTGCGCTTGGCAC-3′	
Ex-R2	5′-GCTGACcgtctcCGTTCAGTCTGTCTTTCCCCTTTCCggtctcAGCCGTGCGCTTGGCACAG-3′	
Ex-R3	5′-GCTGACcgtctcCGTTCAGTCTGTCTTTCCCCTTTCCggtctcTCCCATGGGCCTGACATAACACAGGCAG CAACCTCTG-3′	
Ex-R4	5′-GCTGACcgtctcCGTTCAGTCTGTCTTTCCCCTTTCCggtctcTGAGTCCGTGCGCTTGGCAC-3′	
In-F2	5′-CTTGTTATGGACGAGTTGCCcgtctcGTACGCCAGAGCAGGTCGTGGC-3′	
In-F3	5′-CCAAAGATTCAACCGTCCTGcgtctcGAACCCCAGAGCAGGTCGTG-3′	
In-F4	5′-TATTCATGCTTGGACGGACTcgtctcGGTTGACCCCAGAGCAGGTCGTG-3′	
In-F5	5′-GTCCTAGTGAGGAATACCGGcgtctcGCCTGACCCCAGAGCAGGTCGTG-3′	
In-F6	5′-TTCCTTGATACCGTAGCTCGcgtctcGGACACCAGAGCAGGTCGTGGC-3′	
In-R1	5′-TCTTATCGGTGCTTCGTTCTcgtctcCCGTAAGTCCGTGCGCTTGGCAC-3′	
In-R2	5′-CGTTTCTTTCCGGTCGTTAGcgtctcTGGTTAGTCCGTGCGCTTGGCAC-3′	
In-R3	5′-TGAGCCTTATGATTTCCCGTcgtctcTCAACCCGTGCGCTTGGCACAG-3′	
In-R4	5′-AGTCTGTCTTTCCCCTTTCCcgtctcTCAGGCCGTGCGCTTGGCACAG-3′	
In-R5	5′-CCGAAGAATCGCAGATCCTAcgtctcTTGTCAGTCCGTGCGCTTGGCAC-3′	
Hex-F	5′-CTTAAACCGGCCAACATACC-3′	Hexamer amplification
Hex-R	5′-AGTCTGTCTTTCCCCTTTCC-3′	
TALE-Seq-F1 (aka colony PCR forward)	5′-CCAGTTGCTGAAGATCGCGAAGC-3′	Sequencing forward primer used to check monomers 1–6; also used as colony PCR forward primer
TALE-Seq-F2	5′-ACTTACACCCGAACAAGTCG-3′	Sequencing forward primer used to check monomers 7–12
TALE-Seq-R1 (aka colony PCR reverse)	5′-TGCCACTCGATGTGATGTCCTC-3′	Sequencing primer used to check monomers 13–18 for TALEs with less than 18 full monomer repeats, and used to check monomers 19–24 for TALEs with more than 18 monomers (use TALE-Seq-R2 to check monomers 13–18 in this case); also used as colony PCR reverse primer
TALE-Seq-R2	5′-CCCATGGGCCTGACATAA-3′	Sequencing reverse primer used to check monomers 13–18 in TALEs with more than 18 full monomer repeats

bean and S1 nucleases[50,51]. Of these, Cel1 has been applied extensively for mutation detection[52–54] and established protocols are available for its purification[52,54].
- Primers for Surveyor assay of TALEN cutting efficiency (Integrated DNA Technologies, custom DNA oligonucleotides; see Experimental design for further information on primer design)
- RNeasy mini kit (Qiagen, cat. no. 74104)
- QIAshredder (Qiagen, cat. no. 79654)
- RNAse ZAP (Applied Biosystems, cat. no. AM9780)
- iScript cDNA synthesis kit (Bio-Rad, cat. no. 170-8890)
- TaqMan universal master mix (Applied Biosystems, cat. no. 4364341)
- TaqMan gene expression assay probes for the TALE-TF-targeted gene (Applied Biosystems, http://bioinfo.appliedbiosystems.com/genomic-products/gene-expression.html)

EQUIPMENT
- 96-well thermocycler with programmable temperature stepping functionality (Applied Biosystems Veriti, cat. no. 4375786) ▲ **CRITICAL** Programmable temperature stepping is needed for the TALEN (Surveyor) functional assay. Other steps only require a PCR-capable thermocycler.
- qPCR system (96 well; StepOnePlus real-time PCR system, Applied Biosystems, cat. no. 4376600)
- Optical plates (96 well; MicroAmp, Applied Biosystems, cat. no. N801-0560)
- PCR plates (96 well; Axygen, cat. no. PCR-96-FS-C)
- Strip PCR tubes (8 well; Applied Biosystems, cat. no. N801-0580)
- QIAvac 96 vacuum manifold (Qiagen, cat. no. 19504)
- Gel electrophoresis system (PowerPac basic power supply, Bio-Rad, cat. no. 164-5050, and Sub-Cell GT System gel tray, Bio-Rad, cat. no. 170-4401)
- Digital gel imaging system (GelDoc EZ, Bio-Rad, cat. no. 170-8270, and blue sample tray, Bio-Rad, cat. no. 170-8273)
- Blue light transilluminator and orange filter goggles (SafeImager 2.0, Invitrogen, cat. no. G6600)
- Sterile 20-µl pipette tips for colony picking
- Gel quantification software (Bio-Rad, ImageLab, included with GelDoc EZ, or open-source ImageJ from the National Institutes of Health, available at http://rsbweb.nih.gov/ij/)
- TALE reference sequence generator (Zhang Lab, http://taleffectors.com/tools/)
- Petri dishes (60 mm × 15 mm; BD Biosciences, cat. no. 351007)
- Incubator for bacteria plates (Quincy Lab, cat. no. 12-140E)
- Shaking incubator for bacteria suspension culture (Infors HT Ecotron)
- Cell culture-treated polystyrene plates (6 well; Corning, cat. no. 3506)
- UV spectrophotometer (NanoDrop 2000c, Thermo Scientific)
- Kimwipes (Kimberly-Clark)

REAGENT SETUP

Tris-borate EDTA (TBE) electrophoresis solution Dilute TBE buffer in distilled water to 1× working solution for casting agarose gels and for use as a buffer for gel electrophoresis. Buffer can be stored at room temperature (18 – 22 °C) for at least 1 year.

BSA, 10× Dilute 100× BSA (supplied with BsaI-HF) to 10× concentration and store it at −20 °C for at least 1 year in 20-µl aliquots.

ATP, 10 mM Divide 10 mM ATP into 50-µl aliquots and store at −20 °C for up to 1 year; avoid repeated freeze-thaw cycles.

DTT, 10 mM Prepare 10 mM DTT solution in distilled water and store in 20-µl aliquots at −70 °C for up to 2 years; for each reaction, use a new aliquot, as DTT is easily oxidized.

D10 culture medium For culture of HEK293FT cells, prepare D10 culture medium by supplementing DMEM with 1× GlutaMAX and 10% (vol/vol) FBS. As indicated in the protocol, this medium can also be supplemented with 1× penicillin-streptomycin. D10 medium can be made in advance and stored at 4 °C for up to 1 month.

PROCEDURE

Amplification and normalization of monomer library with ligation adaptors for 18-mer TALE DNA-binding domain construction ● TIMING 6 h

1| *Prepare diluted forward and reverse monomer primer mixes.* In a 96-well PCR plate, prepare primer mixes for amplifying a TALE monomer library (**Fig. 3**, stage 1). Mix forward and reverse primers for each of the 18 positions according to the first two rows (A and B) of **Figure 4** and achieve a final concentration of 10 µM for each primer. If you are using multichannel pipettes, arrange the oligonucleotide primers in the order indicated in **Figure 4** to allow for easy pipetting. Typically, prepare 50-µl mixes for each primer pair (40 µl of ddH_2O, 5 µl of 100 µM forward primer, 5 µl of 100 µM reverse primer).

2| Set up two 96-well monomer library plates according to the organization shown in **Figure 4**; each plate will contain a total of 72 PCRs (18 positions for each monomer × 4 types of monomers). Although it is acceptable to have smaller-volume PCRs, we typically make the monomer set in larger quantities, as one monomer library plate can be used repeatedly for the construction of many TALEs. Each PCR should be made up as follows to a total volume of 200 µl, and then split between the two 96-well plates so that each well contains a 100-µl PCR:

Component	Amount (µl)	Final concentration
Monomer template plasmid (5 ng µl^{-1})	2	50 pg µl^{-1}
dNTP, 100 mM (25 mM each)	2	1 mM
Herculase II PCR buffer, 5×	40	1×
Primer mix, 20 µM (10 µM forward primer and 10 µM reverse primers from Step 1)	4	200 nM
Herculase II Fusion polymerase	2	
Distilled water	150	
Total	200 (for 2 reactions)	

3| Perform PCR on the reactions from Step 2 using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	95 °C, 2 min		
2–31	95 °C, 20 s	60 °C, 20 s	72 °C, 10 s
32			72 °C, 3 min

4| After the reaction has completed, use gel electrophoresis to verify that monomer amplification was successful. Cast a 2% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 1× SYBR Safe dye. The gel should have enough lanes to run out 2 µl of each PCR product from Step 3. Run the gel at 15 V cm^{-1} for 20 min. It is not necessary to check all 72 reactions at this step; it is sufficient to check all 18 reactions for one type of monomer template. Successful amplification should show an ~100-bp product. Monomers positioned at the ends of each hexamer (monomers 1, 6, 7, 12, 13 and 18) should be slightly longer than the other monomers because of the length difference of the longer external primers.
? TROUBLESHOOTING

5| Pool both of the 100-µl PCR plates into a single deep-well plate. Purify the combined reactions using the QIAquick 96 PCR purification kit according to the manufacturer's directions. Elute the DNA from each well using 100 µl of Buffer EB (included with the kit), prewarmed to 55 °C. Alternatively, PCR products can also be purified using individual columns found in standard PCR cleanup kits.
▲ CRITICAL STEP Before eluting the DNA, let the 96-well column plate air-dry, preferably at 37 °C, for 30 min on a clean Kimwipe so that all residual ethanol has enough time to evaporate.

6| *Normalization of monomer concentration*. Cast a 2% (wt/vol) agarose gel. The gel should have enough lanes to run out 2 µl of each purified PCR product from Step 5. Include in one lane 10 µl of the quantitative DNA ladder. Run the gel at 20 V cm^{-1} for 20 min.

7| Image the gel using a quantitative gel imaging system. Monomers 1, 6, 7, 12, 13 and 18 are ~170 bp in size, whereas the other monomers are ~150 bp in size (**Fig. 5**, lanes 1–6). Make sure the exposure is short enough so that none of the bands are saturated.

8| Quantify the integrated intensity of each PCR product band using ImageJ or other gel quantification software. Use the quantitative ladder with known DNA mass (5, 10, 20, 40, 100 ng) to generate a linear fit and quantify the concentration of each purified PCR product.
? TROUBLESHOOTING

9| Adjust the plate of purified PCR products by adding Buffer EB so that each monomer has the same molar concentration. As monomers 1, 6, 7, 12, 13 and 18 are longer than the other monomers, it is necessary to adjust them to a slightly

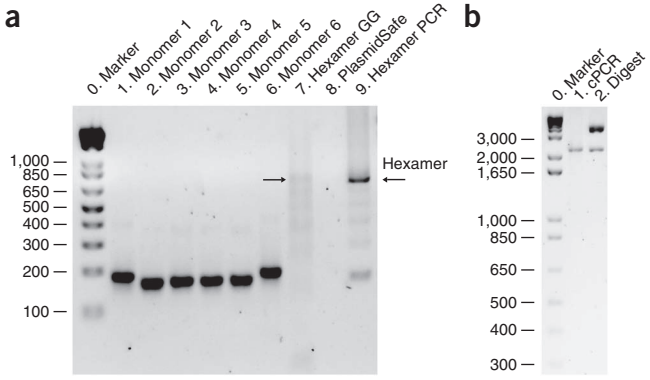


Figure 5 | Example gel results from the TALE construction procedure. (**a**) Lanes 1–6: products from the monomer PCR (Stage 1 in **Fig. 3**) after purification and gel normalization (Procedure Steps 8 and 9). The molar concentrations of samples shown on this gel have been normalized so that equal moles of monomers are mixed for downstream steps. Monomers 1 and 6 are slightly longer than monomers 2–5 because of the addition of sequences used for circularization. Lane 7: result of the hexamer Golden Gate cut-ligation (Procedure Step 15). A series of bands with size ~700 bp and lower can be seen. Successful hexamer Golden Gate assembly should show a band ~700 bp (as indicated by arrow). Lane 8: hexamer assembly after PlasmidSafe exonuclease treatment (Procedure Step 17). Typically, the amount of circular DNA remaining is difficult to visualize by gel. Lane 9: result of hexamer amplification (Procedure Step 20). A band of ~700 bp should be clearly visible. The hexamer gel band should be gel purified to remove shorter DNA fragments. (**b**) Properly assembled TALE-TFs and TALENs can be verified using bacterial colony PCR (2,175-bp band, lane 1; Procedure Step 35) and restriction digestion with AfeI (2,118-bp band for correctly assembled 18-mer in either backbone; other bands for TALE-TF are 165, 3,435, 3,544 bp; other bands for TALEN are 165, 2,803, 3,236 bp; the digest shown is for TALE-TF backbone vector, lane 2, see Procedure Step 35).

higher concentration. For example, we adjust monomers 1, 6, 7, 12, 13 and 18 to 18 ng μl^{-1} and the other monomers to 15 ng μl^{-1}.
▲ **CRITICAL STEP** For subsequent digestion and ligation reactions, it is important that all monomers are at equimolar concentrations.
■ **PAUSE POINT** Amplified monomers can be stored at −20 °C for several months and can be reused for assembling additional TALEs.

Construction of custom 20-bp-targeting TALEs ● TIMING 1.5 d (5 h hands-on time)
10| *Select target sequence(s)*. Typical TALE recognition sequences are identified in the 5′ to 3′ direction and begin with a 5′ thymine. The procedure below describes the construction of TALEs that bind a 20-bp target sequence (5′-$T_0N_1N_2N_3N_4N_5N_6N_7N_8N_9N_{10}N_{11}N_{12}N_{13}N_{14}N_{15}N_{16}N_{17}N_{18}N_{19}$-3′, where N = A, G, T or C), where the first base (typically a thymine) and the last base are specified by sequences within the TALE backbone vector. The middle 18 bp are specified by the RVDs within the middle tandem repeat of 18 monomers according to the cipher NI = A, HD = C, NG = T and NN = G or A. For targeting shorter or longer sequences, see **Box 1**.

11| *Divide target sequences into hexamers*. Divide N_1–N_{18} into subsequences of length 6 ($N_1N_2N_3N_4N_5N_6$, $N_7N_8N_9N_{10}N_{11}N_{12}$ and $N_{13}N_{14}N_{15}N_{16}N_{17}N_{18}$). For example, a TALE targeting 5′-TGAAGCACTTACTTTAGAAA-3′ can be divided into hexamers as (T) GAAGCA CTTACT TTAGAA (A), where the initial thymine and final adenine (in parentheses) are encoded by the appropriate backbone. In this example, the three hexamers will be: hexamer 1 = NN-NI-NI-NN-HD-NI, hexamer 2 = HD-NG-NG-NI-HD-NG and hexamer 3 = NG-NG-NI-NN-NI-NI. Because of the adenine in the final position, we will use one of the NI backbones: pTALE-TF_v2(NI) or pTALEN_v2(NI).

12| *Assembling hexamers using Golden Gate digestion-ligation* (**Fig. 3**, stage 2). Prepare one reaction tube for each hexamer. Using the monomer plate schematic (**Fig. 4**), pipette 1 µl of each normalized monomer into the corresponding hexamer reaction tube. Repeat this for all hexamers. For example, for the target from Step 10, set up tube 1 (1 µl from each of G1, A2, A3, G4, E5 and A6), tube 2 (1 µl from each of E7, C8, C9, A10, E11 and C12) and tube 3 (1 µl from each of D1, D2, B3, H4, B5 and B6). To construct a TALE with 18 full repeats, three separate hexamer tubes are used.
▲ **CRITICAL STEP** Pay close attention when pipetting the monomers; it is very easy to accidentally pipette from the wrong well during this step.

13| To perform a simultaneous digestion-ligation (Golden Gate) reaction to assemble each hexamer (**Fig. 3**, stage 2), add the following reagents to each hexamer tube:

Component	Amount (µl)	Final concentration
Esp3I (BsmBI), 10 U μl^{-1}	0.75	0.375 U μl^{-1}
Tango buffer, 10×	1	1×
DTT, 10 mM	1	1 mM
T7 ligase, 3,000 U μl^{-1}	0.25	75 U μl^{-1}
ATP, 10 mM	1	1 mM
	4	
Six monomers	6 × 1	
Total	10	

▲ **CRITICAL STEP** DTT is easily oxidized in air. It should be freshly made or thawed from aliquots stored at −70 °C and used immediately.

14| Place each hexamer tube in a thermocycler to carry out the Golden Gate reactions using the following cycling conditions for ~3 h:

Cycle number	Digest	Ligate
1–15	37 °C, 5 min	20 °C, 5 min
Hold at 4 °C		

■ **PAUSE POINT** This reaction can be left to run overnight.

15| Run out the ligation product on a gel to check for ~700-bp bands corresponding to the hexamer products (**Fig. 5a**, lane 7). Cast a 2% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 2× SYBR Safe dye. The additional dye helps to visualize faint bands. The gel should have enough lanes to run out each Golden Gate reaction from Step 14; load 3 μl of each ligation product in separate lanes. Include 1 μg of the 1-kb Plus DNA ladder in one lane. Run the gel at 15 V cm^{-1} until there is separation of the 650-bp ladder band from neighboring bands.
? TROUBLESHOOTING

16| *Exonuclease treatment to degrade noncircular ligation products* (**Fig. 3**, stage 3). During the Golden Gate reaction, only fully ligated hexamers should be able to circularize. PlasmidSafe exonuclease selectively degrades noncircular (incomplete) ligation products. Add the following reagents to each hexamer reaction tube:

Component	Amount (μl)	Final concentration
PlasmidSafe DNAse, 10 U μl^{-1}	1	0.66 U μl^{-1}
PlasmidSafe reaction buffer, 10×	1	1×
ATP, 10 mM	1	1 mM
	3	
Golden Gate reaction from Step 14	7	
Total	10	

17| Incubate each hexamer reaction tube with PlasmidSafe at 37 °C for 30 min; follow by inactivation at 70 °C for 30 min.
■ **PAUSE POINT** After completion, the reaction can be frozen and continued later. The circular DNA should be stable for at least 1 week.

18| *Hexamer PCR* (**Fig. 3**, stage 4). Amplify each PlasmidSafe-treated hexamer in a 50-μl PCR using high-fidelity Herculase II polymerase and the hexamer forward and reverse primers (Hex-F and Hex-R; **Table 2**). Add the following reagents to each PCR:

Component	Amount (μl)	Final concentration
dNTP, 100 mM (25 mM each)	0.5	1 mM
Herculase II reaction buffer, 5×	10	1×
Hex-F and Hex-R primers, 10 μM each	1	200 nM
Herculase II Fusion DNA polymerase	0.5	1×
Distilled water	37	
	49	
PlasmidSafe-treated hexamer from Step 17	1	
Total	50	

19| Perform PCR on the reactions in Step 18 using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	95 °C, 2 min		
2–36	95 °C, 20 s	60 °C, 20 s	72 °C, 30 s
37			72 °C, 3 min

20| *Gel purification of amplified hexamers.* Because of the highly repetitive template, it is necessary to purify the amplified hexamer product from the other amplicons. Cast a 2% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 1× SYBR Safe dye. The gel should have enough lanes to run out each PCR product from Step 19, and the comb size should be big enough to load 40–50 µl of PCR product. Include 1 µg of the 1-kb Plus DNA ladder in one lane. Run the gel at 15 V cm^{-1} until there is separation of the 650-bp ladder band from neighboring bands. Use a clean razor blade to excise each hexamer band, which should be nearly aligned with the 650-bp band from the ladder (**Fig. 5**, lane 9).
! CAUTION Wear appropriate personal protective equipment, including a face mask, to minimize risks associated with prolonged light or mutagenic DNA dye exposure.
▲ CRITICAL STEP Avoid any cross-contamination by ethanol sterilization of work surfaces, razor blades, etc. during the gel extraction and between each individual band excision.
? TROUBLESHOOTING

21| Purify the hexamer gel bands from Step 20 using the MinElute gel extraction kit according to the manufacturer's directions. Elute the DNA from each reaction using 20 µl of Buffer EB prewarmed to 55 °C.

22| *Gel normalization of purified hexamer concentrations*. Cast a 2% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 1× SYBR Safe dye. The gel should have enough lanes to run out 2 µl of each purified hexamer from Step 21. Include 10 µl of the quantitative DNA ladder in one lane. Run the gel at 15 V cm^{-1} until all lanes of the quantitative ladder are clearly separated. Each hexamer lane should contain only a single (purified) band.

23| Image the gel using a quantitative gel imaging system. Each lane should have only the ~700-bp hexamer product. Make sure the exposure is short enough so that none of the bands are saturated.

24| Quantify the integrated intensity of each hexamer band using ImageJ or other gel quantification software. Use the quantitative ladder with known DNA mass (5, 10, 20, 40, 100 ng) to generate a linear fit and quantify the concentration of each purified hexamer.
? TROUBLESHOOTING

25| Adjust the concentration of each hexamer to 20 ng μl^{-1} by adding Buffer EB.

26| *Golden Gate assembly of hexamers into TALE backbone* (**Fig. 3**, stage 5). Combine the hexamers and the appropriate TALE backbone vector (transcription factor or nuclease) in a Golden Gate digestion-ligation. For example, we will use a TALE backbone with NI as the 0.5 repeat for the target sequence in Step 10 as N_{19} = A. For this ligation, a 1:1 molar ratio of insert to vector works well. Set up one reaction tube for each TALE. In addition, prepare a negative control ligation by including the TALE backbone vector without any hexamers.

Component	TALE (µl)	Negative control (µl)	Final concentration
TALE backbone vector (100 ng μl^{-1})	1	1	10 ng μl^{-1}
BsaI-HF (20 U μl^{-1})	0.75	0.75	1.5 U μl^{-1}
NEBuffer 4, 10×	1	1	1×
BSA, 10×	1	1	1×
ATP, 10 mM	1	1	1 mM
T7 ligase (3,000 U μl^{-1})	0.25	0.25	75 U μl^{-1}
	5	5	
Three purified hexamers (20 ng μl^{-1})	3 (1 each)		2 ng μl^{-1} each
Distilled water	2	5	
Total	10	10	

▲ CRITICAL STEP As a negative control, set up a separate reaction substituting an equal volume of water in place of the purified hexamers (i.e., including only the TALEN or TALE-TF backbone).

27| Place the tubes from Step 26 in a thermocycler to carry out the Golden Gate reactions using the following cycling conditions for ~4 h:

Cycle number	Digest	Ligate	Inactivate
1–20	37 °C, 5 min	20 °C, 5 min	
21			80 °C, 20 min

■ **PAUSE POINT** Ligation products can be frozen at −20 °C and stored for at least 1 month for transformation into bacteria at a later time.

28| Although it is not necessary, it is possible to run out the ligation product on a gel to check for ~1.8-kbp band corresponding to the properly assembled 18-mer tandem repeat. To check the ligation product, cast a 2% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 2× SYBR Safe dye. The additional dye helps to visualize faint bands. Load 5 µl of the ligation product from Step 27. Include 1 µg of the 1-kb Plus DNA ladder in one lane. Run the gel at 15 V cm^{-1} until there is clear separation of the 1,650- and 2,000-bp ladder bands. Alternatively, proceed directly to transformation (Step 29) without running a gel; transformation is very sensitive and, even when a clear band cannot be visualized on the gel, there is often enough plasmid for transformation of high-competency cells.
? TROUBLESHOOTING

Verifying the correct TALE repeat assembly ● TIMING 3 d (4 h hands-on time)
29| *Transformation*. Transform the ligation products from Step 27 into a competent *E. coli* strain; in our lab, we use Stbl3 for routine transformation. Transformation can be done according to the protocol supplied with the cells. Briefly, add 5 µl of the ligation product to 50 µl of ice-cold chemically competent Stbl3 cells, incubate on ice for 5 min, incubate at 42 °C for 45 s, return immediately to ice for 5 min, add 250 µl of SOC medium, incubate at 37 °C for 1 h on a shaking incubator (250 r.p.m.), plate 100 µl of the transformation on an LB plate containing 100 µg ml^{-1} ampicillin and incubate overnight at 37 °C.

30| Inspect all plates from Step 29 for bacterial colony growth. Typically, we see few colonies on the negative control plates (only backbone in the Golden Gate digestion-ligation) and tens to hundreds of colonies on the complete TALE ligation plates.
? TROUBLESHOOTING

31| For each TALE plate, pick eight colonies to check the assembly fidelity. Use a sterile 20-µl pipette tip to touch a single colony, streak onto a single square on a prewarmed, new, gridded LB-ampicillin plate to save the colony, and then swirl the tip in 100 µl of distilled water to dissolve the colony for colony PCR. Repeat this procedure for all colonies to be checked, streaking each new colony into a separate square on the gridded LB-ampicillin plate. After finishing, incubate the gridded plate at 37 °C for at least 4 h to grow the colony streaks.

32| *Colony PCR*. By using the colonies selected in Step 31 as templates, set up colony PCR to verify that the correctly assembled tandem 18-mer repeat has been ligated into the TALE backbone. We have found that the colony PCR is sensitive to excessive template concentration, and therefore we typically use 1 µl of the 100-µl colony suspension from Step 31. For colony PCR, use primers TALE-Seq-F1 and TALE-Seq-R1 for amplification (**Table 2**). Set up the following colony PCR:

Component	Amount (µl)	Final concentration
Colony suspension from Step 31	1	
dNTP, 100 mM (25 mM each)	0.25	1 mM
Taq-B polymerase buffer, 10×	2.5	1×
TALE-Seq-F1 and TALE-Seq-R1 primers, 10 µM each	0.25	100 nM
Taq-B polymerase (5 U $µl^{-1}$)	0.1	0.02 U $µl^{-1}$
Distilled water	20.9	
Total	25	

33| Perform colony PCR on the reactions in Step 32 using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	94 °C, 3 min		
2–31	94 °C, 30 s	60 °C, 30 s	68 °C, 2 min
32			68 °C, 5 min

34| To check the colony PCR result, cast a 1% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 1× SYBR Safe dye. The gel should have enough lanes to run out 10 µl of each PCR product from Step 33. Include 1 µg of the 1-kb Plus DNA ladder in one lane. Run the gel at 15 V cm^{-1} until there is clear separation of the 1,650- and 2,000-bp ladder bands.

35| Image the gel and identify which colonies have the correct insert size. For an insert of 18 monomers (three hexamers ligated into the TALE backbone vector), the product should be a single band of size 2,175 bp (**Fig. 5b**, lane 1). Incorrect ligation products will show bands of different sizes. In place of colony PCR, plasmid DNA from prepared clones can be digested with AfeI. In both backbones (TALE-TF and TALEN), AfeI cuts four times. For both backbones, one fragment contains the entire tandem repeat region and should be 2,118 bp in size for a correctly assembled 18-mer. For the TALE-TF backbone, the correct clone will produce four bands with sizes 165, 2,118, 3,435 and 3,544 bp (**Fig. 5b**, lane 2). The 3,435- and 3,544-bp bands are difficult to separate on a 1% (wt/vol) agarose gel, and therefore a correct clone will show three bands with the middle 2,118-bp band indicating an intact tandem 18-mer repeat (**Fig. 5b**, lane 2). For the TALEN backbone, the correct clone will produce four bands with sizes 165, 2,118, 2,803 and 3,236 bp.
? TROUBLESHOOTING

36| *Miniprep and sequencing*. For each clone with the correct band size, inoculate a colony from the gridded plate into 3 ml of LB medium with 100 µg ml^{-1} ampicillin and incubate it at 37 °C in a shaking incubator overnight.

37| Isolate plasmid DNA from overnight cultures using a QIAprep Spin miniprep kit according to the manufacturer's instructions.

38| Verify the sequence of each clone by sequencing the tandem repeat region using sequencing primers (**Table 2**) TALE-Seq-F1 (forward primer annealing just before the first monomer), TALE-Seq-F2 (forward primer annealing at the beginning of the seventh monomer) and TALE-Seq-R1 (reverse primer annealing after the final 0.5 monomer). For most TALEs, reads from all three primers are necessary to unambiguously verify the entire sequence. Reference sequences for each custom TALE can be generated using our free online software (http://taleffectors.com/tools/). After entering the target site sequence, our software generates a TALE-TF or TALEN reference sequence in either FASTA format or as an annotated GenBank vector map (*.gb file) that can be viewed using standard plasmid editor software (e.g., everyVECTOR, VectorNTI or LaserGene SeqBuilder). Detailed instructions can be found on our website.
? TROUBLESHOOTING

Transfection of TALE-TF and TALEN into HEK293FT cells ● TIMING 2 d (1 h hands-on time)
39| Plate HEK293FT cells onto six-well plates in D10 culture medium without antibiotics ~24 h before transfection at a seeding density of around 1 × 10^6 cells per well and a seeding volume of 2 ml. Scale up and down the culture according to the manufacturer's manual provided with the 293FT cells if needed.

40| *Prepare DNA for transfection*. Quantify the DNA concentration of the TALE plasmids used for transfection using reliable methods (such as UV spectrophotometry or gel quantification).
▲ CRITICAL STEP The DNA concentration of the TALE plasmids should be quantified to guarantee that an accurate amount of TALE DNA will be used during the transfection.

41| Prepare the DNA–Opti-MEM mix as follows using option A if you are testing transcriptional modulation, or option B if you are testing nuclease activity.
(A) DNA-Opti-MEM mix for testing transcriptional modulation
(i) Mix 4 µg of TALE-TF plasmid DNA with 250 µl of Opti-MEM medium. Include controls (e.g., RFP plasmid or mock transfection) to monitor transfection efficiency and cell health, respectively.

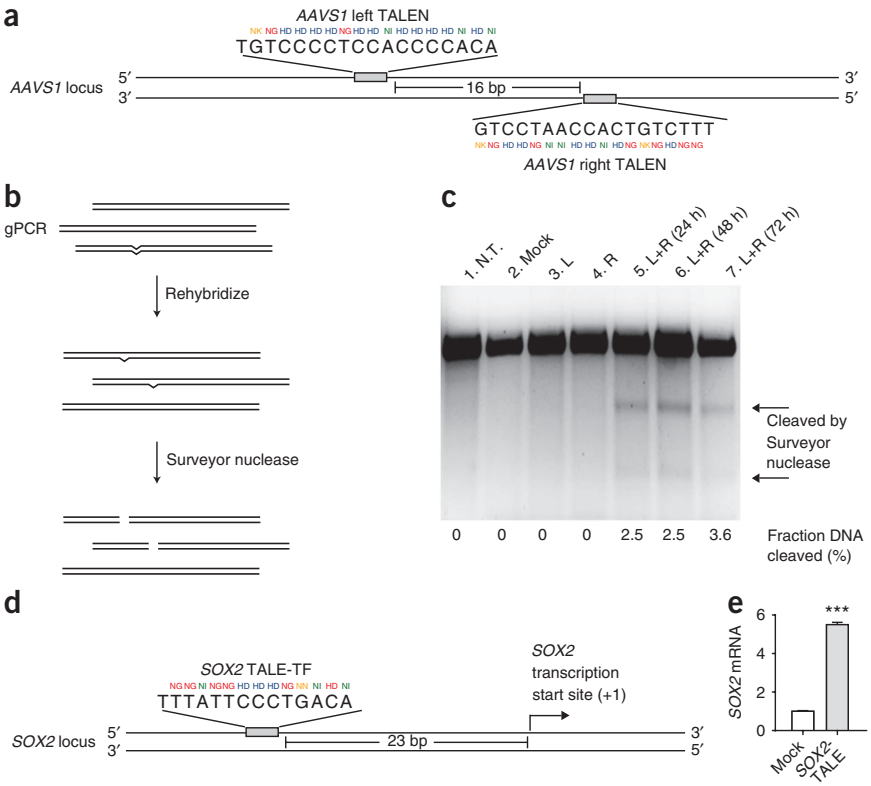


Figure 6 | Examples of TALE-TF and TALEN activity in 293FT cells. (**a**) This schematic shows a pair of TALENs designed to target the *AAVS1* locus in the human genome. The TALENs target a pair of binding sites flanking a 16-bp spacer. The left and right TALENs recognize the top and bottom strands of the target sites, respectively, and each recognition site begins with a T. The nucleotide sequences of the target sites are shown, with the corresponding TALEN RVD specifying the DNA base being targeted shown above. Each TALE DNA-binding domain is fused to the catalytic domain of FokI endonuclease; when FokI dimerizes, it cuts the DNA in the region between the left and right TALEN-binding sites. (**b**) Schematic of the Surveyor nuclease assay used to determine TALEN cleavage efficiency. First, genomic PCR (gPCR) is used to amplify the TALEN target region from a heterogeneous population of TALEN-modified and TALEN-unmodified cells, and the gPCR products are reannealed slowly to generate heteroduplexes. The reannealed heteroduplexes are cleaved by Surveyor nuclease, whereas homoduplexes are left intact. TALEN cleavage efficiency is calculated based on the fraction of cleaved DNA. (**c**) Gel showing the Surveyor nuclease result from the *AAVS1* TALEN pair. Lanes 1–4: controls from un-transfected (NT) cells and cells transfected with a plasmid carrying GFP (Mock), *AAVS1* left TALEN only (L), and *AAVS1* right TALEN only (R). Lanes 5–7: cells transfected with *AAVS1* left and right TALENs (L + R) for 24, 48 and 72 h. The two lower bands indicated by the arrows are Surveyor-cleaved DNA products. (**d**) This schematic shows a TALE-TF designed to target the *SOX2* locus in the human genome. The *SOX2* TALE-TF recognizes the sense strand of the *SOX2* proximal promoter, and the recognition site begins with T. The nucleotide sequence of the target site is shown, with the corresponding TALEN repeat variable diresidue (RVD) specifying each DNA base being targeted shown above. The TALE DNA-binding domain is fused to the synthetic VP64 transcriptional activator, which recruits RNA polymerase and other factors needed to initiate transcription. (**e**) 293FT cells transfected with the *SOX2* TALE-TF exhibited a fivefold increase in the amount of *SOX2* mRNA compared with mock-transfected cells. Error bars indicate s.e.m.; $n = 3$. *** indicates $P < 0.005$. Panel **e** was modified with permission from ref. 3.

(B) DNA-Opti-MEM mix for testing nuclease activity

(i) Mix 2 µg of the left and 2 µg of the right TALEN (**Fig. 6a**) plasmid DNA with 250 µl of Opti-MEM medium. Control transfections should be done by omitting one or both of the TALENs. Also include controls (e.g., an RFP plasmid or mock transfection) to monitor transfection efficiency and cell health, respectively. For all transfections, make sure the total amount of DNA transfected is the same across conditions—when omitting one or both TALENs, supplement with empty vector DNA to maintain the same total DNA amount.

42| Prepare the Lipofectamine–Opti-MEM solution by diluting 10 µl of Lipofectamine 2000 with 250 µl of Opti-MEM. Mix the solution thoroughly by tapping the tube and incubating for 5 min at room temperature.

43| Add the Lipofectamine–Opti-MEM solution to the DNA–Opti-MEM solution to form the DNA-Lipofectamine complex. Mix well by gently pipetting up and down. Incubate for 20 min at room temperature.

▲ **CRITICAL STEP** Make sure the complex is thoroughly mixed. Insufficient mixing results in lower transfection efficiency.

■ **PAUSE POINT** The transfection complex will remain stable for 6 h at room temperature.

44| Add 500 µl of the DNA-Lipofectamine complex to each well of the six-well plate from Step 39 directly. Mix gently by rocking the plates back and forth.

45| Incubate cells at 37 °C with 5% CO_2 for 24 h. At this point, determine the transfection efficiency by estimating the fraction of fluorescent cells in the positive control transfection (e.g., RFP plasmid) using a fluorescence microscope.

▲ **CRITICAL STEP** If incubation beyond 48 h is needed, change the culture medium with fresh D10 supplemented with antibiotics on a daily basis. This will not affect the transfection efficiency.

? TROUBLESHOOTING

TALE functional characterization

46| To measure TALEN cutting efficiency using Surveyor nuclease follow option A, or to measure TALE-TF transcriptional activation using qRT-PCR, follow option B.

(A) Measuring TALEN cutting efficiency using Surveyor nuclease ● TIMING 6 h (3 h hands-on time)

(i) Remove culture medium from each well from Step 45 and add 100 µl of QuickExtract DNA extraction solution to each well and pipette thoroughly to lyse cells. Transfer the lysate to a PCR tube.

(ii) Extract DNA from the lysate from Step 46A(i) using the following cycling conditions:

Cycle number	Condition
1	68 °C, 15 min
2	95 °C, 8 min

(iii) *PCR amplification of the region surrounding TALEN target site*. Prepare the following PCR using the genomic DNA from Step 46A(ii):

Component	Amount (µl)	Final concentration
gDNA from Step 46A(ii)	0.5	
dNTP, 100 mM (25 mM each)	0.5	1 mM
Herculase II reaction buffer, 5×	10	1×
Target-specific Surveyor forward and reverse primers, 10 µM each (see EXPERIMENTAL DESIGN)	1	200 nM
Herculase II Fusion DNA polymerase	0.5	1×
Distilled water	37.5	
Total	50	

▲ **CRITICAL STEP** The Surveyor procedure (Steps 46A(iii–xv)) is carried out according to the manufacturer's protocol and is described in greater detail in the Surveyor manual. We provide brief details here, as mutation detection by mismatch endonuclease is not a common procedure for most laboratories.

▲ **CRITICAL STEP** When performing the Surveyor assay for the first time, we suggest carrying out the positive control reaction included with the Surveyor nuclease kit.

(iv) Perform PCR using the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	95 °C, 3 min		
2–36	95 °C, 30 s	55 °C, 15 s	72 °C, 30 s
37			72 °C, 5 min

(v) Check the PCR result by running 5 µl of PCR product on a 2% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 1× SYBR Safe dye. Include 10 µl of the quantitative DNA ladder in one lane. Run the gel at 15 V cm^{-1} until all bands are clearly separated. For all templates, it is important to make sure that there is only a single band corresponding to the intended product for the primer pair. The size of this band should be the same as calculated from the distance between the two primer annealing sites in the genome.

▲ **CRITICAL STEP** If multiple amplicons are generated from the PCR, redesign primers and reoptimize the PCR conditions to avoid off-target amplification.

? TROUBLESHOOTING

(vi) Image the gel using a quantitative gel imaging system. Make sure the exposure is short enough so that none of the bands are saturated. Quantify the integrated intensity of each PCR product using ImageJ or other gel quantification software. Use the quantitative ladder with known DNA mass (5, 10, 20, 40, 100 ng) to generate a linear fit. Adjust the DNA concentration of the PCR product by diluting it with 1× Herculase II reaction buffer so that it is in the range of 25–80 ng μl^{-1}.

(vii) *DNA heteroduplex formation*. At this point, the amplified PCR product includes a mixture of both modified and unmodified genomic DNA (TALEN-modified DNA will have a few bases of sequence deletion near the TALEN cut site because of NHEJ exonuclease activity). For Surveyor mismatch detection, this mixture of products must first be melted and reannealed such that heteroduplexes are formed. DNA heteroduplexes contain strands of DNA that are slightly different but annealed (imperfectly) together. Given the presence of both unmodified and modified DNA in a sample, a heteroduplex may include one strand of unmodified DNA and one strand of TALEN-modified DNA. Heteroduplexes can also be formed from reannealing of two different TALEN-modified products, as NHEJ exonuclease activity can produce different mutations. To cross-hybridize wild-type and TALEN-modified PCR products into hetero- and homoduplexes, all strands are melted and then slowly reannealed (**Fig. 6b**). Place 300 ng of the PCR product from Step 46A(vi) in a thermocycler tube and bring it to a total volume of 20 µl with 1× Herculase II reaction buffer.

(viii) Perform cross-hybridization on the diluted PCR amplicon from Step 46A(vii) using the following cycling conditions:

Cycle number	Condition
1	95 °C, 10 min
2	95–85 °C, −2 °C s^{-1}
3	85 °C, 1 min
4	85–75 °C, −0.3 °C s^{-1}
5	75 °C, 1 min
6	75–65 °C, −0.3 °C s^{-1}
7	65 °C, 1 min
8	65–55 °C, −0.3 °C s^{-1}
9	55 °C, 1 min
10	55–45 °C, −0.3 °C s^{-1}
11	45 °C, 1 min
12	45–35 °C, −0.3 °C s^{-1}
13	35 °C, 1 min
14	35–25 °C, −0.3 °C s^{-1}
15	25 °C, 1 min

(ix) *Surveyor Nuclease S digestion*. To treat the cross-hybridized homo- and heteroduplexes using Surveyor Nuclease S to determine TALEN cleavage efficiency (**Fig. 6b**), add the following components together on ice and mix by pipetting gently:

Component	Amount (µl)	Final concentration
$MgCl_2$ solution, 0.15 M	2	15 mM
Surveyor nuclease S	1	1×
Surveyor enhancer S	1	1×
	4	
Reannealed duplexes from Step 46A(viii)	16	
Total	20	

(x) Incubate the reaction from Step 46A(ix) at 42 °C for 1 h.

(xi) Add 2 µl of the Stop Solution from the Surveyor kit.
■ **PAUSE POINT** The digestion product can be stored at −20 °C for analysis at a later time.
(xii) Cast a 2% (wt/vol) agarose gel in 1× TBE electrophoresis buffer with 1× SYBR Safe dye. When casting the gel, it is preferable to use a thin comb size (<1 mm) for the sharpest possible bands. The gel should have enough lanes to run out 20 µl of each digestion product band from Step 46A(xi). Include 1 µg of the 1-kb Plus DNA ladder in one lane. Run the gel at 5 V cm^{-1} until the Orange G loading dye has migrated two-thirds of the way down the gel.
(xiii) Image the gel using a quantitative gel imaging system. Make sure the exposure is short enough so that none of the bands are saturated. Each lane from samples transfected with both left and right TALENs should have a larger band corresponding to the uncut genomic amplicon (the same size as in Step 46A(v)) and smaller bands corresponding to the DNA fragments resulting from the cleavage of the genomic amplicon by Surveyor nuclease. Controls (no transfection, control plasmid transfection or transfection omitting one of the TALENs) should only have the larger band corresponding to the uncut genomic amplicon.
? TROUBLESHOOTING
(xiv) Quantify the integrated intensity of each band using ImageJ or other gel quantification software. For each lane, calculate the fraction of the PCR product cleaved (f_{cut}) using the following formula: $f_{cut} = a / (a + b)$, where a = the integrated intensity of both of the cleavage product bands and b = the integrated intensity of uncleaved PCR product band. A sample Surveyor gel for TALENs targeting human *AAVS1* is shown in **Figure 6c**.
(xv) Estimate the percentage of TALEN-mediated gene modification using the following formula[47]:

$$100 \times (1 - (1 - f_{cut})^{1/2})$$

This calculation can be derived from the binomial probability distribution given a few conditions: that strand reassortment during the duplex formation is random, that there is a negligible probability of the identical mutations reannealing during duplex formation and that the Surveyor nuclease digestion is complete.

(B) Measuring TALE-TF transcriptional activation using qRT-PCR ● TIMING 5 h (3 h hands-on time)

(i) *RNA extraction*. Aspirate the medium in each well of the six-well plates from Step 45 at 72 h after transfection.
▲ **CRITICAL STEP** Use proper RNA handling techniques to prevent RNA degradation, including cleaning bench surfaces and pipettes with RNaseZAP. Use RNase-free consumables and reagents.
(ii) Wash the cells in each well twice with 1 ml of DPBS.
(iii) Harvest approximately 1 × 10^6 cells for subsequent total RNA extraction by trypsinizing the cells with 500 µl trypsin with EDTA. Incubate for 1–2 min to let the cells detach from the bottom of the wells.
▲ **CRITICAL STEP** Do not leave the cells in trypsin for longer than a few minutes.
(iv) Neutralize the trypsin by adding 2 ml of D10 medium.
(v) In a 15-ml centrifuge tube, centrifuge the cell suspension at 300*g* for 5 min at 4 °C. Carefully aspirate all of the supernatant.
▲ **CRITICAL STEP** Incomplete removal of the supernatant can result in inhibition of cell lysis.
■ **PAUSE POINT** Cells can be frozen at −80 °C for 24 h.
(vi) Extract and purify RNA from the cells in Step 46B(v) using the RNeasy mini kit and QIAshredder following the manufacturer's directions. Elute the RNA from each column using 30 µl of nuclease-free water.
(vii) Measure the RNA concentration using a UV spectrophotometer.
(viii) *cDNA reverse transcription*. Generate cDNA using the iScript cDNA synthesis kit according to the manufacturer's directions. For matched negative controls, perform the reverse transcription without the reverse-transcriptase enzyme.
(ix) *Quantitative PCR*. Thaw on ice the appropriate TaqMan probe for the target gene and for an endogenous control gene.
▲ **CRITICAL STEP** Protect the probes from light and do not allow the thawed probes to stay on ice for an extended time.
(x) By following the TaqMan Universal PCR Master Mix manufacturer's directions, prepare four technical replicate qPCRs for each sample in optical thermocycler strip tubes or 96-well plates. Set up negative controls for nonspecific amplification as indicated in the directions: namely, RNA template processed without reverse transcriptase ('no RT') and a no-template control.
(xi) Briefly centrifuge the samples to remove any bubbles and amplify them in a TaqMan-compatible qRT-PCR machine with the following cycling parameters.

Cycle number	Denature	Anneal and extend
1	95 °C, 10 min	
2–41	92 °C, 15 s	60 °C, 1 min

(xii) Analyze data and calculate the level of gene activation using the $\Delta\Delta C_t$ method[46,55]. TALE-TF results from qRT-PCR assay of *SOX2* activation in HEK293 cells are shown in **Figure 6d,e**.

▲ **CRITICAL STEP** The $\Delta\Delta C_t$ method assumes that amplification efficiency is 100% (i.e., the number of amplicons doubles after each cycle). For new probes (such as custom TaqMan probes), amplification from a template dilution series (spanning at least five orders of magnitude) should be performed to characterize amplification efficiency. For standard TaqMan gene expression assay probes, this is not necessary, as they are designed to have 100 ± 10% amplification efficiency.

? TROUBLESHOOTING

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 3**.

TABLE 3 | Troubleshooting table.

Step	Problem	Possible reason	Solution
4	Uneven amplification across monomers	Not using Herculase II Fusion polymerase	Optimize annealing temperature and Mg^{2+} and DMSO concentrations
8	Low DNA concentration after elution	Residual ethanol on purification column	Air-dry columns before elution at 37 °C for a longer period of time
		Incorrect vacuum pressure during DNA binding	Adjust vacuum pressure according to the manufacturer's suggestions
15	No visible hexamer band (~700 bp)	Equimolar amounts of monomers were not added	Gel-normalize the monomer concentration
		Degraded DTT or ATP	Use fresh stocks of DTT and ATP, which degrade easily
	No visible hexamer band (~700 bp) but smaller bands present	Wrong monomer(s) added during pipetting	Re-select monomers
		Monomer concentration is too low	Increase the number of Golden Gate digestion-ligation cycles and/or increase the concentration of monomers to >20 ng μl^{-1}; there is no detrimental effect to using more monomers in an equimolar ratio
20	No visible hexamer band (~700 bp)	Unsuccessful Golden Gate digestion-ligation	Verify on a gel that the Golden Gate digestion-ligation product from Step 15 is visible; increase the monomer concentration
24	Low concentration for purified hexamers	Unsuccessful gel extraction	Ensure that there is no residual ethanol during elution or increase PCR reaction volume
28	No visible 18-mer band (~1.8 kbp)	Unsuccessful Golden Gate digestion-ligation	Increase hexamer concentration in Golden Gate digestion-ligation in Step 26 or proceed directly to transformation in Step 29
30	More than a few colonies on negative control plate	Compromised TALE backbone	Perform a restriction digest of the backbone to verify integrity
35	Colony PCR bands are smeared	Too much template	Dilute colony suspension 10× to 100×
38	Monomers assembled in incorrect order	Misligation	Misligation occurs at a very low frequency; analyze two additional clones

(continued)

TABLE 3 | Troubleshooting table (continued).

45	Low transfection efficiency	Low DNA quality	Prepare DNA using high-quality plasmid preparation
		Suboptimal ratio of DNA to Lipofectamine 2000	Titrate the ratio of DNA to Lipofectamine 2000 to determine optimal transfection conditions
46A(v)	Multiple amplicons	Nonspecific primers	Design new primers and verify specificity using PrimerBLAST; use touchdown PCR
	No amplification	Suboptimal PCR condition	Optimize annealing temperature and Mg^{2+} and DMSO concentrations
46A(xiii)	No cleavage bands visible	TALEN is unable to cleave the target site	Design new TALEN pairs targeting nearby sequences
46B(xii)	No increase in transcription in target mRNA	TALE-TF is unable to access the target site	Design new TALE-TFs targeting nearby sequences

● TIMING

Steps 1–9, Monomer library amplification and normalization: 6 h
Steps 10–28, TALE hierarchical ligation assembly: 1.5 d (5 h hands-on time)
Steps 29–38, TALE transformation and sequence verification: 3 d (4 h hands-on time)
Steps 39–45, Transfection of TALE-TF and TALEN into HEK293FT cells: 2 d (1 h hands-on time)
Steps 46A and 46B, TALE functional characterization with qRT-PCR or Surveyor: 5–6 h (3 h hands-on time)

ANTICIPATED RESULTS

TALE-TFs and TALENs can facilitate site-specific transcriptional modulation[3–5,8] and genome editing[4,7,9,11–15] (**Table 1**). TALENs can be readily designed to introduce double-stranded breaks at specific genomic loci with high efficiency. In our experience, a pair of TALENs designed to target the human *AAVS1* locus is able to achieve up to 3.6% cutting efficiency in 293FT cells, as determined by Surveyor nuclease assay (**Fig. 6a–c**). TALE-TFs can also robustly increase the mRNA levels of endogenous genes. For example, a TALE-TF designed to target the proximal promoter region of *SOX2* in human cells is able to elevate the level of endogenous *SOX2* gene expression by up to fivefold[3] (**Fig. 6d,e**). The ability for TALE-TFs and TALENs to act at endogenous genomic loci is dependent on the chromatin state, as well as yet-to-be-determined mechanisms regulating TALE DNA binding[56,57]. For these reasons, we typically build several TALE-TFs or TALEN pairs for each genomic locus we aim to target. These TALE-TFs and TALENs are designed to bind to neighboring regions around a specific target site, as some binding sites might be more accessible than others. The reason why some TALEs exhibit significantly lower levels of activity remains unknown, although it is likely to be due to position- or cell-state-specific epigenetic modifications preventing access to the binding site. Because of differences in epigenetic states between different cells, it is possible that TALEs that fail to work in a particular cell type might work in a different cell type.

Note: Supplementary information is available via the HTML version of this article.

ACKNOWLEDGMENTS We thank the entire Zhang laboratory for their support. L.C. is supported by a Howard Hughes Medical Institute International Student Research Fellowship. Y.Z. is supported by a Simons Foundation Fellowship. M.M.C. is supported by a Massachusetts Institute of Technology Undergraduate Research Opportunities scholarship. F.Z. is supported by a US National Institutes of Health Transformative R01 and by the McKnight and Simons Foundations, Robert Metcalfe and Michael Boylan.

AUTHOR CONTRIBUTIONS N.E.S., L.C., Y.Z. and F.Z. wrote the manuscript. M.M.C. designed the online TALE sequence verification software. F.Z. and G.F. supervised the research.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

Published online at http://www.natureprotocols.com/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
2. Moscou, M.J. & Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
3. Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol.* **29**, 149–153 (2011).
4. Miller, J.C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
5. Morbitzer, R., Romer, P., Boch, J. & Lahaye, T. 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 (2010).

6. Weber, E., Gruetzner, R., Werner, S., Engler, C. & Marillonnet, S. Assembly of designer TAL effectors by golden gate cloning. *PLoS ONE* **6**, e19722 (2011).
7. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).
8. Geissler, R. *et al.* Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE* **6**, e19509 (2011).
9. Li, T. *et al.* Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res.* **39**, 6315–6325 (2011).
10. Morbitzer, R., Elsaesser, J., Hausner, J. & Lahaye, T. Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res.* **39**, 5790–5799 (2011).
11. Wood, A.J. *et al.* Targeted genome editing across species using ZFNs and TALENs. *Science* **333**, 307 (2011).
12. Christian, M. *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757–761 (2010).
13. Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, 731–734 (2011).
14. Li, T. *et al.* TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372 (2011).
15. Mahfouz, M.M. *et al. 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 (2011).
16. Boch, J. & Bonas, U. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* **48**, 419–436 (2010).
17. Bogdanove, A.J., Schornack, S. & Lahaye, T. TAL effectors: finding plant genes for disease and defense. *Curr. Opin. Plant Biol.* **13**, 394–401 (2010).
18. Romer, P. *et al.* Plant pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. *Science* **318**, 645–648 (2007).
19. Kay, S., Hahn, S., Marois, E., Hause, G. & Bonas, U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648–651 (2007).
20. Kay, S., Hahn, S., Marois, E., Wieduwild, R. & Bonas, U. Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3Deltarep16. *Plant J.* **59**, 859–871 (2009).
21. Romer, P. *et al.* Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper *Bs3* alleles. *Plant Physiol.* **150**, 1697–1712 (2009).
22. Hinnen, A., Hicks, J.B. & Fink, G.R. Transformation of yeast. *Proc. Natl. Acad. Sci. USA* **75**, 1929–1933 (1978).
23. Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. & Stahl, F.W. The double-strand-break repair model for recombination. *Cell* **33**, 25–35 (1983).
24. Thomas, K.R., Folger, K.R. & Capecchi, M.R. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419–428 (1986).
25. Ivics, Z., Hackett, P.B., Plasterk, R.H. & Izsvak, Z. Molecular reconstruction of *Sleeping Beauty*, a *Tc1*-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510 (1997).
26. Kawakami, K., Shima, A. & Kawakami, N. Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. *Proc. Natl. Acad. Sci. USA* **97**, 11403–11408 (2000).
27. Akagi, K. *et al.* Cre-mediated somatic site-specific recombination in mice. *Nucleic Acids Res.* **25**, 1766–1773 (1997).
28. Epinat, J.C. *et al.* A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.* **31**, 2952–2962 (2003).
29. Lois, C., Hong, E.J., Pease, S., Brown, E.J. & Baltimore, D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295**, 868–872 (2002).
30. Khan, I.F., Hirata, R.K. & Russell, D.W. AAV-mediated gene targeting methods for human cells. *Nat. Protoc.* **6**, 482–501 (2011).
31. Pavletich, N.P. & Pabo, C.O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252**, 809–817 (1991).
32. Klug, A. The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. *Q. Rev. Biophys.* **43**, 1–21 (2010).
33. Maeder, M.L., Thibodeau-Beganny, S., Sander, J.D., Voytas, D.F. & Joung, J.K. Oligomerized pool engineering (OPEN): an 'open-source' protocol for making customized zinc-finger arrays. *Nat. Protoc.* **4**, 1471–1501 (2009).
34. Kim, J.S., Lee, H.J. & Carroll, D. Genome editing with modularly assembled zinc-finger nucleases. *Nat. Methods* **7**, 91; author reply 91–92 (2010).
35. Sander, J.D. *et al.* Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat. Methods* **8**, 67–69 (2011).
36. Perez, E.E. *et al.* Establishment of HIV-1 resistance in $CD4^+$ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008).
37. Keenholtz, R.A., Rowland, S.J., Boocock, M.R., Stark, W.M. & Rice, P.A. Structural basis for catalytic activation of a serine recombinase. *Structure* **19**, 799–809 (2011).
38. Gersbach, C.A., Gaj, T., Gordley, R.M., Mercer, A.C. & Barbas, C.F. III. Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. *Nucleic Acids Res.* **39**, 7868–7878 (2011).
39. Gaj, T., Mercer, A.C., Gersbach, C.A., Gordley, R.M. & Barbas, C.F. III. Structure-guided reprogramming of serine recombinase DNA sequence specificity. *Proc. Natl. Acad. Sci. USA* **108**, 498–503 (2011).
40. Urnov, F.D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
41. Wilson, M.H., Kaminski, J.M. & George, A.L. Jr. Functional zinc finger/*sleeping beauty* transposase chimeras exhibit attenuated overproduction inhibition. *FEBS Lett.* **579**, 6205–6209 (2005).
42. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **3**, e3647 (2008).
43. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS ONE* **4**, e5553 (2009).
44. Weber, E., Engler, C., Gruetzner, R., Werner, S. & Marillonnet, S. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE* **6**, e16765 (2011).
45. Huertas, P. DNA resection in eukaryotes: deciding how to fix the break. *Nat. Struct. Mol. Biol.* **17**, 11–16 (2010).
46. Nolan, T., Hands, R.E. & Bustin, S.A. Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**, 1559–1582 (2006).
47. Guschin, D.Y. *et al.* A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.* **649**, 247–256 (2010).
48. Zhang, F. *et al.* High frequency targeted mutagenesis in *Arabidopsis thaliana* using zinc finger nucleases. *Proc. Natl. Acad. Sci. USA* **107**, 12028–12033 (2010).
49. Buzdin, A.A. in *Nucleic Acids Hybridization* (eds. Buzdin, A., Lukyanov, S.) 211–239 (Springer, 2007).
50. Till, B.J., Burtner, C., Comai, L. & Henikoff, S. Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Res.* **32**, 2632–2641 (2004).
51. Babon, J.J., McKenzie, M. & Cotton, R.G. The use of resolvases T4 endonuclease VII and T7 endonuclease I in mutation detection. *Mol. Biotechnol.* **23**, 73–81 (2003).
52. Yang, B. *et al.* Purification, cloning, and characterization of the CEL I nuclease. *Biochemistry* **39**, 3533–3541 (2000).
53. Kulinski, J., Besack, D., Oleykowski, C.A., Godwin, A.K. & Yeung, A.T. CEL I enzymatic mutation detection assay. *Biotechniques* **29**, 44–46, 48 (2000).
54. Oleykowski, C.A., Bronson Mullins, C.R., Godwin, A.K. & Yeung, A.T. Mutation detection using a novel plant endonuclease. *Nucleic Acids Res.* **26**, 4597–4602 (1998).
55. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45 (2001).
56. Murakami, M.T. *et al.* The repeat domain of the type III effector protein PthA shows a TPR-like structure and undergoes conformational changes upon DNA interaction. *Proteins* **78**, 3386–3395 (2010).
57. Scholze, H. & Boch, J. TAL effectors are remote controls for gene activation. *Curr. Opin. Microbiol.* **14**, 47–53 (2011).
58. Huang, P. *et al.* Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol.* **29**, 699–700 (2011).
59. Sander, J.D. *et al.* Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat. Biotechnol.* **29**, 697–698 (2011).
60. Tesson, L. *et al.* Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol.* **29**, 695–696 (2011).