ACS SyntheticBiology

Review

pubs.acs.org/synthbio

Transcription Activator-like Effectors: A Toolkit for Synthetic Biology

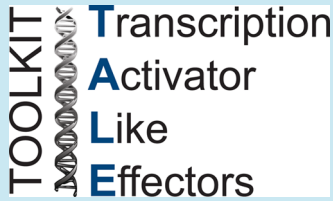
Richard Moore,[†,§,*,∥] Anita Chandrahas,[†,§,∥] and Leonidas Bleris[†,‡,§]

[†]Bioengineering Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas 75080 United States

[‡]Electrical Engineering Department, The University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas 75080 United States

[§]Center for Systems Biology, The University of Texas at Dallas, 800 West Campbell Road, Richardson, Texas 75080 United States

ABSTRACT: Transcription activator-like effectors (TALEs) are proteins secreted by *Xanthomonas* bacteria to aid the infection of plant species. TALEs assist infections by binding to specific DNA sequences and activating the expression of host genes. Recent results show that TALE proteins consist of a central repeat domain, which determines the DNA targeting specificity and can be rapidly synthesized *de novo*. Considering the highly modular nature of TALEs, their versatility, and the ease of constructing these proteins, this technology can have important implications for synthetic biology applications. Here, we review developments in the area with a particular focus on modifications for custom and controllable gene regulation.



Transcription activator-like effectors (TALEs), proteins secreted by *Xanthomonas* plant pathogenic bacteria,[1] have been a major breakthrough in the rapid and systematic targeting of any DNA sequence of choice.[2−5] Efforts in developing TALE technology have led to applications such as activation, repression, deletion, and insertion of a desired DNA sequence in an ever-expanding range of model organisms and cell types.[6−8]

TALEs have a modular DNA-binding domain (DBD) consisting of repetitive sequences of residues; each repeat region consists of 34 amino acids[9,10] (Figure 1). A pair of residues at the 12th and 13th position of each repeat region determines the nucleotide specificity and are referred to as the repeat variable diresidue (RVD).[11] The last repeat region, termed the half-repeat, is typically truncated to 20 amino acids.[3] Combining these repeat regions allows synthesizing sequence-specific synthetic TALEs.[12,13] The C-terminus typically contains a nuclear localization signal (NLS), which directs a TALE to the nucleus, as well as a functional domain that modulates transcription, such as an acidic activation domain (AD).[4,14−16] The endogenous NLS can be replaced by an organism-specific localization signal. For example, an NLS derived from the simian virus 40 large T-antigen can be used in mammalian cells.[5]

The RVDs HD, NG, NI, and NN target C, T, A, and G/A,[3,17,18] respectively. Recent studies suggest that NH has a higher specificity for G than NN and results in stronger TALE binding.[17] The nucleotide preferences of commonly used RVDs appear in Figure 2. This basic code enables DNA targeting where each RVD corresponds to a specific nucleotide.[19] Out of the RVDs that have close to a one-to-one correspondence, HD and NH bind more strongly to DNA and target C and G, respectively.[17] A list of RVDs and their binding preferences for nucleotides appears in Table 1.

There are additional TALE RVDs that can be used for custom degenerate TALE-DNA interactions. For example, NA has high affinity for all four bases of DNA.[17] Additionally, N*, where * is an RVD with a deletion in the 13th residue, can accommodate all letters of DNA including methylated cytosine.[19] Further work has also shown that S* has the ability to bind to any DNA nucleotide.[17]

An important consideration toward building efficient TALEs is to include at least 3−4 strong RVDs in the TALE array while avoiding more than 6 weak RVDs in a row, especially at either end of the repeat region.[19] For TALEs tested as transcriptional activators, the most efficient DBDs range from 15.5 to 19.5 repeats.[4,11]

Identification of new RVDs will likely continue as evidenced by the characterization of TALE protein analogs recently discovered in *Ralstonia solanacearum*.[20,21] These "RipTALs" have a similar repeat variable diresidue domain and several RipTAL RVDs appear to have affinity for specific letters of DNA. In the case of the *Ralstonia* effectors, ND targets C, HN targets A and G, and NP prefers C, A, or G.[20]

To conclude, it is important to mention that TALE proteins prefer a thymine to precede the targeted DNA sequence and may have lower affinity for sequences that lack this 5′ T.[22] Additionally, the Brg11 RipTAL prefers a guanine to precede the target DNA.[21]

CRYSTAL STRUCTURE

Two separate groups have helped elucidate the crystal structure of TALEs. The results show that HD and NN form stronger interactions with DNA by forming hydrogen bonds. In contrast, weaker domains, such as NG and NI form van der Waals interactions with DNA.[23,24]

Deng et al. examined the crystal structure of dHax3, an artificially synthesized TALE with 11.5 repeats comprised of HD, NG, and NS, in the DNA-bound and DNA-free states.[24]

Special Issue: Engineered DNA-Binding Proteins

Received: September 11, 2013
Published: February 5, 2014

   dx.doi.org/10.1021/sb400137b | *ACS Synth. Biol.* 2014, 3, 708−716

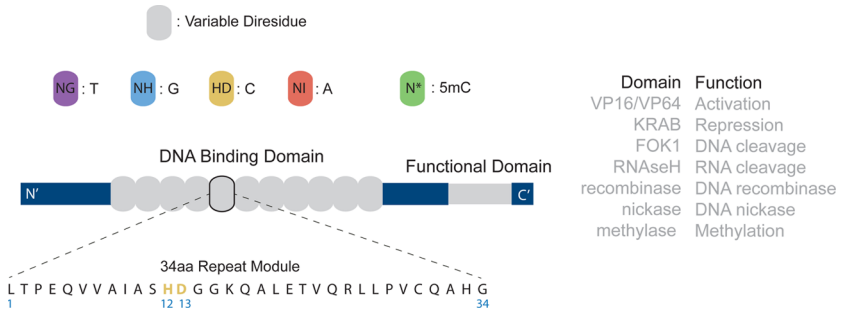


Figure 1. Structural representation of a transcription-activator-like effector. Top left: the corresponding RVD for each letter of DNA, including 5-methylcytosine. Bottom left: the DNA binding domain of the TALE protein includes several repeat modules of 34 residues flanked by an N and C-terminus, which may include mammalian NLS or functional domains. Right: a list of domains for fusion to TALE and their function within cells.

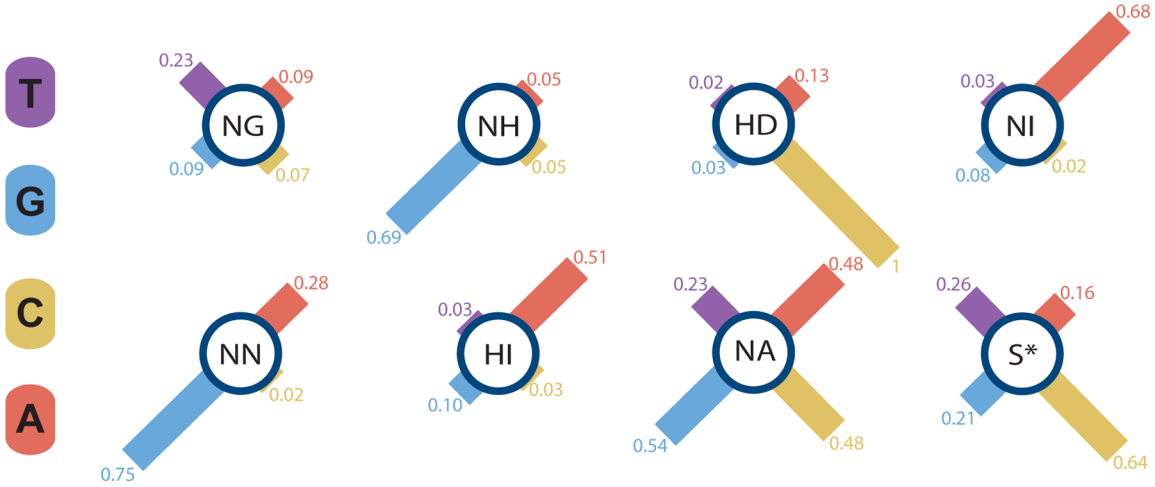


Figure 2. Nucleotide preference of commonly used RVDs. The affinities for all four nucleotides appears for each RVD normalized to the strongest reported affinity of RVD HD for letter C.[17]

The dHax TALE has a right-handed superhelical pitch of 60 Å, which is reduced to 35 Å in the DNA bound state; overall, there is a compression of the superhelical structure in the DNA-bound state that adds to the flexibility of TALEs binding to DNA with minor shifts.

Mak et al. investigated the crystal structure of PthXo1, a TALE protein with 23.5 repeats.[23] The structure of PthXo1 bound to DNA suggests that a proline at the 27th position of each repeat may be important for the consecutive packing of TALE repeats and for the TALE-DNA association. Compared to Deng et al., Mak et al. studied a naturally occurring TALE and analyzed a wider variety of RVDs, including HD, NG, NI, NN, NS, N*, and NG.

Both studies analyzed TALE crystal structure data of the TALE DNA-binding domain. Recent work on the crystal structure of the N and C-terminus of the TALE protein shows that residues 162−288 of the N-terminus have four repeat regions that directly bind to DNA and are structurally similar to the TALE repeats without specificity.[25]

■ ASSEMBLY OF TALE PROTEINS

A number of online tools are available today for designing TALEs to target a specific DNA sequence.[26−29] Accordingly, commercially available kits (summarized in Table 2) allow rapid, custom assembly of TALE repeat regions between the N and C-terminus of the protein.[2,8,29,30] These methods can be used to assemble custom DBDs, which are then cloned into an expression vector containing a functional domain.

Most options for *de novo* synthesis of TALEs or TALE nucleases (TALENs) in the laboratory combine digestion and ligation steps in a Golden Gate reaction with type II restriction enzymes.[2,28] Higher-throughput assembly methods of TALE proteins include Ligation-Independent Cloning (LIC),[31] Fast Ligation-based Automatable Solid-phase High-throughput (FLASH) assembly,[30] and Iterative-Capped Assembly (ICA).[32] FLASH uses a library of 376 plasmids containing 1-, 2-, 3-, or 4-mers to synthesize up to 96 TALEs in less than a day.[30] Alternatively, the iterative capped assembly (ICA) method constructs TALEs by sequentially adding monomers to create custom length TALEs in parallel without relying on an extensive library.[32] Finally, LIC uses larger overhangs (10−30 bp) than Golden Gate based assembly; these overhangs remain stable during transformation and eliminate the need for a prior ligation step. Furthermore, LIC has high fidelity, eliminating the need for a selection procedure under optimal conditions.[31] Table 2 contains a list of TALE assembly options.

Repositories and libraries of TALE proteins may prove important in allowing rapid and high-throughput application of this technology. An initial effort involved the synthesis of a library of 18 740 TALEN pairs that span the human genome.[33] Toward genetic screen applications, we recently developed a methodology to construct versatile transcription activator-like effector libraries.[34] As a proof of principle, we built an 11-mer library that covers all possible combinations of the nucleotides that determine the TALE-DNA binding specificity (i.e., approximately 4 million TALEs). Considering the highly

dx.doi.org/10.1021/sb400137b | *ACS Synth. Biol.* 2014, 3, 708−716

Table 1. RVD Binding Preference

	nucleotides			
RVD	A	G	C	T
NN[a,b,c,d]	medium	medium		
NK[a,b,d]		weak		
NI[a,d]	medium			
NG[a,b,d]				weak
HD[a,c,d]			medium	
NS[b,d]	weak	medium	weak	weak
NG[a,b,d]				weak
N*[a,d]			weak	weak
HN[a,b,d]	weak	medium		
NT[d]	weak	medium		
NP[d]	weak		weak	medium
NH[b,d]		medium		
SN[d]		weak		
SH[d]		weak		
NA[a,d]	weak	strong	weak	weak
IG[a,b]				weak
H*[a,b]	poor	poor	weak	poor
ND[a,b]			weak	
HI[b]	medium			
HG[b]				weak
NC[b]				weak
NQ[b]		weak		
SS[b]		weak		
SN[b]		weak		
S*[b]	medium	medium	strong	medium
NV[b]	weak	medium	poor	poor
HH[b]	poor	poor	poor	poor
YG[b]	poor	poor	poor	poor

[a]Ref 18. [b]Ref 17. [c]Ref 45. [d]Ref 19.

modular nature of TALEs and the versatility and ease of constructing these libraries we foresee broad implications for genomic assays.

Given the application of TALE proteins across the genome, one of the major concerns with this new technology lies in mitigating off-target effects. TALE-NT 2.0 ranks target sites for key characteristics including off-target effects (using Target Finder for TALEs or Paired Target Finder for TALENs).[26] The off-target predictions are based off of RVD nucleotide preference detailed by Moscou and Bogdanove.[18] Furthermore, TALgetter is a recently developed computational tool that incorporates positive and negative efficiencies of each RVD in addition to RVD sequence specificities.[35] TALENoffer uses the TALgetter algorithm to specifically address off-target effects. TALENoffer can rank potential off-target sites and incorporate several parameters including homodimer and heterodimer nuclease domains as well as evaluating both strands of genomic DNA.[36]

FUNCTIONAL DOMAIN BASED REPRESSION AND ACTIVATION

TALE activators and repressors function across many eukaryotic systems.[9,37−39] The majority of the repression examples rely on fusing the TALE with a functional domain known to interfere with the RNA Polymerase II complex. TALEs with either the KRAB domain[40] or the mSin3 interacting domain inhibit mammalian transcription.[17,41] Furthermore, TALE repressors in combination with post-transcriptional repressors such as shRNA show virtually complete repression in mammalian cells.[12]

For activation, fusing the herpes simplex virus derived VP-16 or VP-64 activation domains to a TALE can cause an increase in transcription.[5,42] Weaker activation domains such as the AD of human NF-κB add to the variety of options for gene activation.[41] Furthermore, as shown on endogenous promoters,[43] combinations of TALE activators can be used to introduce synergistic effects.

Several C-termini modifications have shown robust increase in gene expression and may enable stronger activation. Miller et al. used a Δ152 truncation (+152) at the N-terminus based on

Table 2. TALE Assembly Methods

method/time	description
Cellectis Bioresearch[a] 4−12 weeks	Basic: 4 weeks, validation of TALEN with SSA yeast assay.
	First: 4−9 weeks, confirmed efficiency in human cells and mouse/rat cell lines upon request; validation by deep sequencing.
	Premium: 12 weeks, validation in custom cell lines, and validation by deep sequencing.
Life Technologies[a] 2−3 weeks	Create TALEs of either 17.5 or 23.5 repeats.
	Binding sequence must be 19 or 25 bp long and the 5′ end must contain a T.
	Options include a nuclease (Fok1), activator (VP16), repressor (KRAB), or a functional domain of choice (MCS).
GeneCopoeia[a] 2−8 weeks	Creates TALE nucleases (Fok1) or activators (VP64).
	Offers various levels of validation from basic sequencing (2 weeks) to chromosomal level functional validation (8 weeks).
	Also offers "safe-harboring" a selected gene in the AAVS1 chromosomal location via TALEN mediated homologous recombination.
FLASH[b] 1 day	TAL library made of 96-mer repeats.
	Can adopt to high-throughput (automated) or medium throughput (hand) synthesis.
Voytas Kit[c] 5−7 days	Uses type II restriction endonucleases to ligate anywhere from 10.5 to 30.5 repeats.
	Available through Addgene.
ICA[d] 3−4 days	Monomers are added to a growing chain to create the complete DNA binding domain (DBD).
	Allows custom creation of DBD of any length.
LIC[e] 3 days	Library of 2-, 5-, 6-, or 18-mer repeats used to assemble a DBD with medium or high-throughput.
	Under optimal conditions, additional selection steps can be eliminated, reducing the total time.
fairyTALE[f] 1 day	Robotic assisted high-throughput liquid phase assembly.
	Produces TALE activators, repressors, and nucleases that target between 14 and 31 base pairs.

[a]Company Web site. [b]Refs 30 and 90. [c]Ref 2. [d]Ref 32. [e]Ref 31. [f]Ref 91.

dx.doi.org/10.1021/sb400137b | ACS Synth. Biol. 2014, 3, 708−716

a previous study[44] and experimented with truncating the C-terminus. Out of the C-terminal truncations tested (+278, +133, +95, +23), they found that +95 amino acids (AA) yielded nearly ∼70-fold higher activation in comparison to the original TALE.[45] Mussolino et al. looked at N and C-termini truncations using a luciferase assay and found that 153 AA at the N-terminus and 17 AA at the C-terminus gave optimal results.[46] Zhang et al. also studied several truncations and found the TALEs had the most efficient activation with +240 AA and retained ∼80% of activity with +147 AA upstream and +68 AA downstream of the DNA binding domain.[5]

Data suggests that the N-terminus is more important to conserve than the C-terminus. While it appears that the ideal TALE architecture depends on the specific conditions, it can be concluded from these studies that at least 100 residues are needed at the N-terminus and ∼60 residues at the C-terminus.

NUCLEASES

TALE nucleases utilize a C-terminal fusion with the type II restriction enzyme FokI to create a heterodimer which produces a double-stranded break (DSB) in DNA.[47] Nuclease-induced DSBs are repaired by nonhomologous end joining (NHEJ) or homologous directed repair (HDR), where homologous recombination (HR) is the most important type of HDR. NHEJ is an error-prone mechanism that results in a functional gene knockout by creating small insertions or deletions (indels) while HR, in combination with a template donor DNA sequence, results in a gene insertion or direct nucleotide exchange.[48,49]

Introducing heterodimerization for nuclease activity has been a significant step toward minimizing off-target effects, first with zinc finger nucleases[50] and then with TALENs.[47] Using this technique, two TALEN proteins must bind to the plus and minus strand of DNA to create a FokI heterodimer and form a double-stranded break.[47]

TALENs have been useful in creating knockout strains and studying mutations in a variety of organisms such as bacteria,[51] yeast,[2,3,52,53] plants,[54−56] nematodes,[57] fish,[58−61] *Xenopus* embryos,[62] human cell lines,[39,45,63,64] rodents,[65,66] and rat embryonic stem cells.[67] In addition, TALE nucleases have successfully modified human stem cells, allowing editing and gene expression tools for tissue engineering.[68,69]

Several assays allow researchers to assess the cutting efficiencies of TALENs.[28,70] The surveyor method can be used to detect DSBs by PCR amplification. Another method is the traffic light reporter (TLR) assay, which can be used to determine whether a TALEN cuts the target DNA and induces NHEJ or HR. A mutated GFP and a frameshifted RFP provide the initial target DNA for the TALEN. If HR occurs, a functioning GFP protein replaces the mutated GFP; if NHEJ occurs, red fluorescence protein (RFP) is shifted into frame.[70]

Though TALE nucleases have wide applications, the double strand breaks they create are predominately repaired by NHEJ. NHEJ and HR are believed to be competing pathways.[71] Error-prone NHEJ is greatly reduced by utilizing TALE nickases, which only cut a single strand of DNA. TALE-MutH has recently been shown to be an efficient, programmable nickase.[72] Here, a single TALE-MutH protein created the desired single-stranded break (SSB) in DNA, thereby inducing the HR repair mechanism. Other strategies to create TALE nickases may involve the FokI nuclease, where one unit of the heterodimer is catalytically inactive.[73]

RECOMBINASES

Site-specific recombinases (SSRs) can integrate, excise, or invert specified DNA segments. Most SSRs are part of one of two major families: tyrosine (λ) recombinases and serine (resolvase/invertase) recombinases. Tyrosine recombinases use a Holliday junction to break and rejoin single strands in pairs while serine recombinases introduce a DSB before strand exchange.[74]

Mercer et al. created recombinatorial TALE proteins (TALER) by fusing a Gin invertase, a serine recombinase, to edit both mammalian and bacterial cells at specific locations.[75] This study also shows that in *E. coli* longer targets of 26 and 32 bp recombined 100 fold more efficiently than the shorter targets of 14 and 20 base pairs. Translating this assay to mammalian cells showed a ∼20 fold efficiency with a 44 bp target and ∼6 fold efficiency with a 32 bp target, much lower than the observed efficiencies in *E. coli*. In mammalian cells, the combined use of a zinc finger recombinase and TALER to create a heterodimer improved recombinase activity.[75]

METHYLATED DNA TARGETS

The presence of methylation on DNA target regions was initially recognized as a potential limitation of the TALE technology. Subsequently, computational modeling of the TALE structure data suggests that RVDs which have relatively small side-chains or a deletion at residue 13, such as HG and N* respectively, could prevent steric hindrance at methylated cytosine (mC).[76] *In vitro* studies confirm that TALENs with N* better accommodate methylated regions of DNA.[76] However, HG and N* are not highly selective and can accept a purine as well.[23,77]

Further structural analysis from Deng et al. suggests that NG may also be able to recognize mC.[24] In the TALE cipher, NG normally targets T, but experimental data shows that methylated cytosines can be targeted by a TALE with NG.[78] Thus, TALE DBDs engineered to target potentially methylated DNA targets may require the RVD N* in order to accommodate methylated cytosine or cytosine[17] and therefore target a region in a state of dynamic methylation.[78] Alternatively, applications that require specific identification of mC could utilize NG, which recognizes mC but has low preference for cytosine.[17,78]

The ability of TALEs to target methylations may allow introducing modifications on the DNA level. Indeed, a fusion of the LSD1 demethylase to TALE was recently used to demonstrate epigenetic control in mammalian cells. The TALE-LSD1 fusion altered the chromatin at enhancers and modified gene expression.[79]

REGULATION OF TALE FUNCTION

An approach to regulate TALEs involves the use of dimerization domains controlled by a small molecule.[41] In this work, the TALE was fused with the Rheo receptor protein, which can dimerize with a second protein called the Rheo activator that contains a functional domain (e.g., an activation VP-16 domain). While the TALE-Rheo receptor fusion can bind to its DNA target, it lacks an activation domain and thus does not regulate transcription. After the addition of a small molecule, GenoStat, the TALE-Rheo receptor dimerizes with the Rheo activator tethered to the VP-16 activation domain. The small molecule control of TALE expands the ability to bring inducible control to any promoter. The same small

dx.doi.org/10.1021/sb400137b | *ACS Synth. Biol.* 2014, 3, 708−716

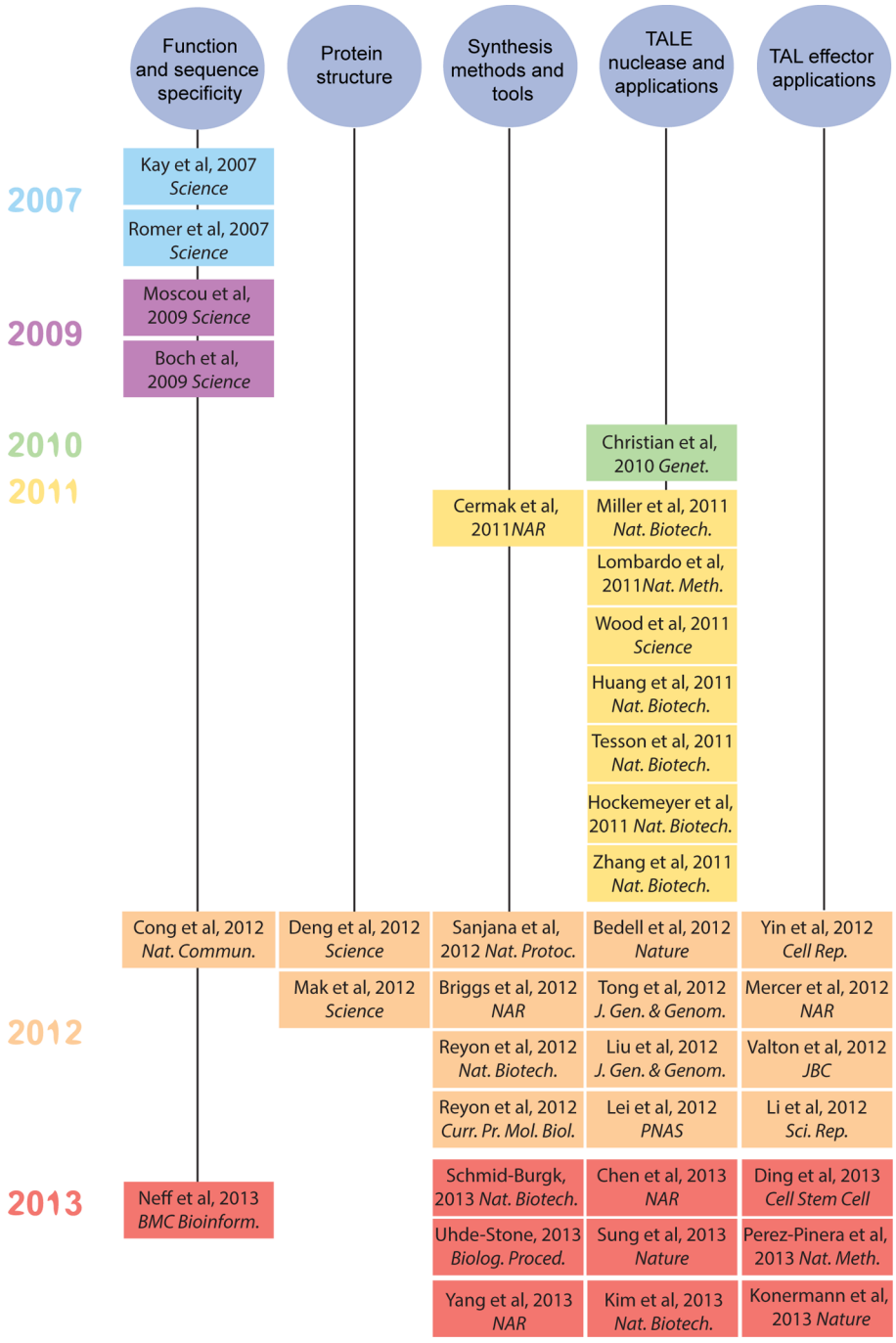


Figure 3. Timeline of papers in TALE technology development.

molecule dimerization methodology was recently adopted to demonstrate inducible control of endogenous genes.[80]

The protein dimerization methodology[41] also allows interfacing the function of TALEs with endogenous signals, such as hypoxia. In particular, by fusing a TALE DBD with ARNT, which forms a heterodimer with the transcription factor HIF-1α under hypoxia or $CoCl_2$ treatment, a TALE can modulate transcription in response to hypoxia. Both the ARNT and Rheo fusions led to activation of the target promoter under dimerization with the application of GenoStat showing a clear dose dependent activation of the target promoter.[41]

Recent work on light-inducible transcriptional activators (LITEs) can be used for spatiotemporal regulation of TALEs.[81] Here, a TALE DBD-CRY2 complex dimerizes with CIB-VP64 when induced with blue light (466 nm). This system shows TALE activity in as little as 30 min (~5-fold activation) after blue-light induction with full saturation at 12 h (~20-fold activation). The system also demonstrated epigenetic modification using epiTALEs where an TALE-CIB1 complex dimerizes with a CRY2-histone effector domain in blue light, resulting to 2−3 fold reduction of the target gene mRNA.

Protein dimerization also allows implementation of Boolean logic where individual vectors express different parts of a TALE protein but only become functional as intein fragments are spliced and combined.[82] Notably, when designing the individual components, the placement of the splice site requires careful consideration of the secondary structures.[82]

dx.doi.org/10.1021/sb400137b | ACS Synth. Biol. 2014, 3, 708−716

Finally, TALEs can be used synergistically to fine-tune gene expression. Maeder et al. targeted DNA-hypersensitive regions for three genes: VEGFA, NTF3, and the microRNA cluster miR-302-367.[43] For the VEGFA gene, 54 TALES were made and tested individually for nine regions. Additionally, repeat lengths varying from 14.5 to 24.5 were tested. In general, TALEs with 16.5−20.5 repeats had the highest efficiency, and the results showed up to a 114 fold increase in activation above basal level. Synergistic effects for all three regions were subsequently tested. Interestingly, for the VEGFA region, the 6 TALEs had a synergistic effect when transfected at 1/6 the amount of each individual TALE but no synergistic effect when transfected together at the higher individual dosage. Of the three genes, miR-302-367 had the most noticeable synergistic effect with approximately 50-fold increase using all five TALEs together rather than each individual effector. Synergistic TALEs may be useful to significantly increase efficiency but may increase off-target effects at the same time.[42,43] Perez-Pinera et al. targeted four genes, IL1RN, KLK3, CEACAM5, and ERBB2, with six to eight synergistic TALEs. Additional data suggests that open chromatin may not be a prerequisite to selecting an efficient TALE target site and silenced genes may be activated by synergistic TALEs. Testing all 63 possible combinations of TALEs for IL1RN, KLK3, and CEACAM5 showed highly tunable gene expression.[42]

DISCUSSION

TALEs hold promise in addressing current challenges in rewiring endogenous and engineered networks to achieve synthetic biology goals.[83−85] A timeline listing TALE research papers of outstanding interest appears in Figure 3. The opportunity to apply custom-made transcriptional activation and repression, targeted cutting and genome modifications, and epigenetic control greatly expands the synthetic biology toolkit and will enable the implementation of more complex genetic circuits. This genome editing and transcriptional control tool is rapidly advancing toward higher precision and efficiency.[86] Furthermore, recent results in optogenetics[81] and small molecule-based dimerizations[41] allow for highly controllable TALE and TALEN spatiotemporal function.

While transcription activator-like effectors show significant promise there are still plenty of opportunities for improving their function. Single-nucleotide precision in transgene integration and modifications remains a challenge for custom genome editing and genetic engineering.[87,88] Another key challenge is increasing the nuclease selectivity for HDR or NHEJ.[87,88] Finally, while recent results offer tools to assess unintended activity,[89] additional work must be performed to ensure tight regulation and quality control, particularly toward therapeutics and industrial scale applications.[87,88]

AUTHOR INFORMATION

Corresponding Author

*Phone: 972-883-5785. E-mail: bleris@utdallas.edu.

Author Contributions

‖R.M. and A.C. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the U.S. National Institutes of Health (NIH) grant GM098984 and National Science Foundation (NSF) grant CBNET-1105524.

DEFINITIONS

Transcription activator-like effectors: modular transcription factors that target DNA on a one-to-one basis with a wide variety of potential applications for this newly developing technology, including genome editing, and genome expression regulation.

Transcription activator-like effector nuclease (TALEN): TALE protein with a nuclease functional domain designed for creating double strand breaks in DNA.

Nuclear localization domain (NLS): domain attached to the C-terminus marking the protein for import into the cell nucleus.

Functional domain: part of the C-terminus of the TALE that gives specific functionality to the TALE (cutting, activation, repression, etc.).

Repeat-variable domain (RVD): Each RVD has 34−35 AA where the 12th and 13th AA is responsible for affinity to a certain base-pair.

DNA binding domain (DBD): portion of the TALE protein containing an array of 12−30 RVDs that targets a DNA sequence.

Homologous recombination (HR): directed method to repair a double-stranded break (DSB) with a template DNA. The template DNA relies on homologous arms of 400−800 bp on either end. HR is a reliable method to create a gene insertion or substitution.

Non-homologous recombination (NHEJ): repair mechanism involving a frameshift that creates small insertions or functional deletions.

REFERENCES

(1) Römer, P., Hahn, S., Jordan, T., Strauß, T., Bonas, U., and Lahaye, T. (2007) Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science 318*, 645−648.

(2) Cermak, T., Doyle, E. L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J. A., Somia, N. V., Bogdanove, A. J., and Voytas, D. F. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res. 39*, e82.

(3) Bogdanove, A. J., and Voytas, D. F. (2011) TAL effectors: Customizable proteins for DNA targeting. *Science 333*, 1843−1846.

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

(5) Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G. M., and Arlotta, P. (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat. Biotechnol. 29*, 149−153.

(6) Mahfouz, M. M., and Li, L. (2011) TALE nucleases and next generation GM crops. *GM Crops 2*, 99−100−103.

(7) Sun, N., Liang, J., Abil, Z., and Zhao, H. (2012) Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Mol. BioSyst. 8*, 1255−1263.

(8) Marx, V. (2012) Genome-editing tools storm ahead. *Nat. Methods 9*, 1055−1059.

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

(10) Muñoz Bodnar, A., Bernal, A., Szurek, B., and López, C. E. (2012) Tell me a tale of TALEs. *Mol. Biotechnol. 53*, 228−235.

(11) Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science 326*, 1509−1512.
(12) Garg, A., Lohmueller, J. J., Silver, P. A., and Armel, T. Z. (2012) Engineering synthetic TAL effectors with orthogonal target sites. *Nucleic Acids Res. 40*, 7584−7595.
(13) Carlson, D. F., Fahrenkrug, S. C., and Hackett, P. B. (2012) Targeting DNA with fingers and TALENs. *Mol. Ther.—Nucleic Acids 1*, e3 DOI: 10.1038/mtna.2011.5.
(14) Schornack, S., Minsavage, G. V., Stall, R. E., Jones, J. B., and Lahaye, T. (2008) Characterization of AvrHah1, a novel AvrBs3-like effector from xanthomonas gardneri with virulence and avirulence activity. *New Phytol. 179*, 546−556.
(15) Gürlebeck, D., Szurek, B., and Bonas, U. (2005) Dimerization of the bacterial effector protein AvrBs3 in the plant cell cytoplasm prior to nuclear import. *Plant J. 42*, 175−187.
(16) Kay, S., Boch, J., and Bonas, U. (2005) Characterization of AvrBs3-like effectors from a brassicaceae pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. *Mol. Plant-Microbe Interact. 18*, 838−848.
(17) Cong, L., Zhou, R., Kuo, Y., Cunniff, M., and Zhang, F. (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nature communications 3*, 968.
(18) Moscou, M. J., and Bogdanove, A. J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science 326*, 1501.
(19) Streubel, J., Blücher, C., Landgraf, A., and Boch, J. (2012) TAL effector RVD specificities and efficiencies. *Nat. Biotechnol. 30*, 593−595.
(20) Li, L., Atef, A., Piatek, A., Ali, Z., Piatek, M., Aouida, M., Sharakou, A., Mahjoub, A., Wang, G., and Khan, S. (2013) Characterization and DNA-binding specificities of ralstonia TAL-like effectors. *Mol. Plant 6*, 1318−1330.
(21) Lange, O., Schreiber, T., Schandry, N., Radeck, J., Braun, K. H., Koszinowski, J., Heuer, H., Strauß, A., and Lahaye, T. (2013) Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytol.*, 773−786.
(22) Lamb, B. M., Mercer, A. C., and Barbas, C. F., 3rd. (2013) Directed evolution of the TALE N-terminal domain for recognition of all 5′ bases. *Nucleic Acids Res. 41*, 9779−9785.
(23) Mak, A. N., Bradley, P., Cernadas, R. A., Bogdanove, A. J., and Stoddard, B. L. (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science 335*, 716−719.
(24) Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J. K., Shi, Y., and Yan, N. (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science 335*, 720−723.
(25) Gao, H., Wu, X., Chai, J., and Han, Z. (2012) Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. *Cell Res. 22*, 1716−1720.
(26) Doyle, E. L., Booher, N. J., Standage, D. S., Voytas, D. F., Brendel, V. P., VanDyk, J. K., and Bogdanove, A. J. (2012) TAL effector-nucleotide targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Res. 40*, W117−W122.
(27) Neff, K. L., Argue, D. P., Ma, A. C., Lee, H. B., Clark, K. J., and Ekker, S. C. (2013) Mojo hand, a TALEN design tool for genome editing applications. *BMC Bioinform. 14*, x DOI: 10.1186/1471-2105-14-1.
(28) Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G., and Zhang, F. (2012) A transcription activator-like effector toolbox for genome engineering. *Nat. Protocols 7*, 171−192.
(29) Li, L., Piatek, M. J., Atef, A., Piatek, A., Wibowo, A., Fang, X., Sabir, J. S. M., Zhu, J. K., and Mahfouz, M. M. (2012) Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. *Plant Mol. Biol. 78*, 407−416.
(30) Reyon, D., Tsai, S. Q., Khayter, C., Foden, J. A., Sander, J. D., and Joung, J. K. (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol. 30*, 460−465.
(31) Schmid-Burgk, J. L., Schmidt, T., Kaiser, V., Höning, K., and Hornung, V. (2013) A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat. Biotechnol. 31*, 76−81.
(32) Briggs, A. W., Rios, X., Chari, R., Yang, L., Zhang, F., Mali, P., and Church, G. M. (2012) Iterative capped assembly: Rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res. 40*, e117.
(33) Kim, Y., Kweon, J., Kim, A., Chon, J. K., Yoo, J. Y., Kim, H. J., Kim, S., Lee, C., Jeong, E., and Chung, E. (2013) A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol. 31*, 251−258.
(34) Li, Y.; Ehrhardt, K.; Zhang, M. Q.; Bleris, L. *Unpublished experiments* 2014.
(35) Grau, J., Wolf, A., Reschke, M., Bonas, U., Posch, S., and Boch, J. (2013) Computational predictions provide insights into the biology of TAL effector target sites. *PLoS Comput. Biol. 9*, e1002962.
(36) Grau, J., Boch, J., and Posch, S. (2013) TALENoffer: Genome-wide TALEN off-target prediction. *Bioinformatics 29*, 2931−2932.
(37) Crocker, J., and Stern, D. L. (2013) TALE-mediated modulation of transcriptional enhancers *in vivo*. *Nat. Methods 10*, 762−767.
(38) Mahfouz, M. M., Li, L., Piatek, M., Fang, X., Mansour, H., Bangarusamy, D. K., and Zhu, J. K. (2012) Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. *Plant Mol. Biol. 78*, 311−321.
(39) Geiβler, R., Scholze, H., Hahn, S., Streubel, J., Bonas, U., Behrens, S. E., and Boch, J. (2011) Transcriptional activators of human genes with programmable DNA-specificity. *PloS One 6*, e19509.
(40) Peng, H., Begg, G. E., Harper, S. L., Friedman, J. R., Speicher, D. W., and Rauscher, F. J., III (2000) Biochemical analysis of the Kruppel-associated box (KRAB) transcriptional repression domain. *J. Biol. Chem. 275*, 18000−18010.
(41) Li, Y., Moore, R., Guinn, M., and Bleris, L. (2012) Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. *Sci. Rep. 2*, 897.
(42) Perez-Pinera, P., Ousterout, D. G., Brunger, J. M., Farin, A. M., Glass, K. A., Guilak, F., Crawford, G. E., Hartemink, A. J., and Gersbach, C. A. (2013) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. *Nat. Methods 10*, 239−242.
(43) Maeder, M. L., Linder, S. J., Reyon, D., Angstman, J. F., Fu, Y., Sander, J. D., and Joung, J. K. (2013) Robust, synergistic regulation of human gene expression using TALE activators. *Nat. Methods 10*, 243−245.
(44) Szurek, B., Rossier, O., Hause, G., and Bonas, U. (2002) Type III-dependent translocation of the xanthomonas AvrBs3 protein into the plant cell. *Mol. Microbiol. 46*, 13−23.
(45) Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., and Hinkley, S. J. (2010) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol. 29*, 143−148.
(46) Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res. 39*, 9283−9293.
(47) Christian, M., Cermak, T., Doyle, E. L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A. J., and Voytas, D. F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics 186*, 757−761.
(48) Moore, F. E., Reyon, D., Sander, J. D., Martinez, S. A., Blackburn, J. S., Khayter, C., Ramirez, C. L., Joung, J. K., and Langenau, D. M. (2012) Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). *PloS One 7*, e37877.
(49) Orlando, S. J., Santiago, Y., DeKelver, R. C., Freyvert, Y., Boydston, E. A., Moehle, E. A., Choi, V. M., Gopalan, S. M., Lou, J. F., and Li, J. (2010) Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res. 38*, e152−e152.

(50) Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y., Rupniewski, I., Beausejour, C. M., Waite, A. J., Wang, N. S., and Kim, K. A. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol. 25*, 778−785.
(51) Politz, M. C., Copeland, M. F., and Pfleger, B. F. (2013) Artificial repressors for controlling gene expression in bacteria. *Chem. Commun. 49*, 4325−4327.
(52) Li, T., Huang, S., Jiang, W. Z., Wright, D., Spalding, M. H., Weeks, D. P., and Yang, B. (2011) TAL nucleases (TALNs): Hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res. 39*, 359.
(53) Li, T., Huang, S., Zhao, X., Wright, D. A., Carpenter, S., Spalding, M. H., Weeks, D. P., and Yang, B. (2011) Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res. 39*, 6315−6325.
(54) Morbitzer, R., Römer, P., Boch, J., and Lahaye, T. (2010) Regulation of selected genome loci using *de novo*-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc. Natl. Acad. Sci. 107*, 21617.
(55) Mahfouz, M. M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., and Zhu, J. K. (2011) *De novo*-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci. 108*, 2623.
(56) Shan, Q., Wang, Y., Chen, K., Liang, Z., Li, J., Zhang, Y., Zhang, K., Liu, J., Voytas, D. F., and Zheng, X. (2013) Rapid and efficient gene modification in rice and brachypodium using TALENs. *Mol. Plant 6*, 1365−1368.
(57) Wood, A. J., Lo, T., Zeitler, B., Pickle, C. S., Ralston, E. J., Lee, A. H., Amora, R., Miller, J. C., Leung, E., and Meng, X. (2011) Targeted genome editing across species using ZFNs and TALENs. *Science 333*, 307−307.
(58) Huang, P., Xiao, A., Zhou, M., Zhu, Z., Lin, S., and Zhang, B. (2011) Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol. 29*, 699−700.
(59) Chen, S., Oikonomou, G., Chiu, C. N., Niles, B. J., Liu, J., Lee, D. A., Antoshechkin, I., and Prober, D. A. (2013) A large-scale *in vivo* analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. *Nucleic Acids Res. 41*, 2769−2778.
(60) Ansai, S., Sakuma, T., Yamamoto, T., Ariga, H., Uemura, N., Takahashi, R., and Kinoshita, M. (2013) Efficient targeted mutagenesis in medaka using custom-designed transcription activator-like effector nucleases (TALENs). *Genetics 193*, 739−749.
(61) Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug, R. G., II, Tan, W., Penheiter, S. G., Ma, A. C., and Leung, A. Y. (2012) *In vivo* genome editing using a high-efficiency TALEN system. *Nature*, DOI: 10.1038/nature11537.
(62) Lei, Y., Guo, X., Liu, Y., Cao, Y., Deng, Y., Chen, X., Cheng, C. H., Dawid, I. B., Chen, Y., and Zhao, H. (2012) Efficient targeted gene disruption in xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). *Proc. Natl. Acad. Sci. 109*, 17484−17489.
(63) Ding, Q., Lee, Y. K., Schaefer, E. A. K., Peters, D. T., Veres, A., Kim, K., Kuperwasser, N., Motola, D. L., Meissner, T. B., and Hendriks, W. T. (2012) A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell 12*, 238−251.
(64) Uhde-Stone, C., Gor, N., Chin, T., Huang, J., and Lu, B. (2013) A do-it-yourself protocol for simple transcription activator-like effector assembly. *Biol. Proc. Online 15*, DOI: 10.1186/1480-9222-15-3.
(65) Tesson, L., Usal, C., Ménoret, S., Leung, E., Niles, B. J., Remy, S., Santiago, Y., Vincent, A. I., Meng, X., and Zhang, L. (2011) Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol. 29*, 695−696.
(66) Wefers, B., Meyer, M., Ortiz, O., de Angelis, M. H., Hansen, J., Wurst, W., and Kühn, R. (2013) Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proc. Natl. Acad. Sci. 110*, 3782−3787.
(67) Tong, C., Huang, G., Ashton, C., Wu, H., Yan, H., and Ying, Q. L. (2012) Rapid and cost-effective gene targeting in rat embryonic stem cells by TALENs. *J. Genet. Genomics 39*, 275−280.
(68) Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., Cost, G. J., Zhang, L., Santiago, Y., and Miller, J. C. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol. 29*, 731−734.
(69) Liu, G. H., Suzuki, K., Qu, J., Sancho-Martinez, I., Yi, F., Li, M., Kumar, S., Nivet, E., Kim, J., and Soligalla, R. D. (2011) Targeted gene correction of laminopathy-associated mutations in patient-specific iPSCs. *Cell Stem Cell 8*, 688−694.
(70) Certo, M. T., Ryu, B. Y., Annis, J. E., Garibov, M., Jarjour, J., Rawlings, D. J., and Scharenberg, A. M. (2011) Tracking genome engineering outcome at individual DNA breakpoints. *Nat. Methods 8*, 671−676.
(71) Hartlerode, A. J., and Scully, R. (2009) Mechanisms of double-strand break repair in somatic mammalian cells. *Biochem. J. 423*, 157.
(72) Gabsalilow, L., Schierling, B., Friedhoff, P., Pingoud, A., and Wende, W. (2013) Site-and strand-specific nicking of DNA by fusion proteins derived from MutH and I-SceI or TALE repeats. *Nucleic Acids Res. 41*, e83.
(73) Ramirez, C. L., Certo, M. T., Mussolino, C., Goodwin, M. J., Cradick, T. J., McCaffrey, A. P., Cathomen, T., Scharenberg, A. M., and Joung, J. K. (2012) Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res. 40*, 5560−5568.
(74) Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006) Mechanisms of site-specific recombination*. *Annu. Rev. Biochem. 75*, 567−605.
(75) Mercer, A. C., Gaj, T., Fuller, R. P., and Barbas, C. F. (2012) Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Res. 40*, 11163−11172.
(76) Valton, J., Dupuy, A., Daboussi, F., Thomas, S., Marechal, A., Macmaster, R., Melliand, K., Juillerat, A., and Duchateau, P. (2012) Overcoming TALE DNA binding domain sensitivity to cytosine methylation. *J. Biol. Chem. 287*, 38427−38432.
(77) Bochtler, M. (2012) Structural basis of the TAL effector−DNA interaction. *Biol. Chem. 393*, 1055−1066.
(78) Deng, D., Yin, P., Yan, C., Pan, X., Gong, X., Qi, S., Xie, T., Mahfouz, M., Zhu, J., and Yan, N. (2012) Recognition of methylated DNA by TAL effectors. *Cell Res. 22*, 1502−1504.
(79) Mendenhall, E. M., Williamson, K. E., Reyon, D., Zou, J. Y., Ram, O., Joung, J. K., and Bernstein, B. E. (2013) Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol. 31*, 1133−1136.
(80) Mercer, A. C., Gaj, T., Sirk, S. J., Lamb, B. M., and Barbas, C. F., III (2013) Regulation of endogenous human gene expression by ligand-inducible TALE transcription factors. *ACS Synth. Bio.*, DOI: 10.1021/sb400114p.
(81) Konermann, S., Brigham, M. D., Trevino, A., Hsu, P. D., Heidenreich, M., Cong, L., Platt, R. J., Scott, D. A., Church, G. M., and Zhang, F. (2013) Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, DOI: 10.1038/nature12466.
(82) Lienert, F., Torella, J. P., Chen, J., Norsworthy, M., Richardson, R. R., and Silver, P. A. (2013) Two-and three-input TALE-based AND logic computation in embryonic stem cells. *Nucleic Acids Res. 41*, 9967−9975.
(83) Pennisi, E. (2012) The tale of the TALEs. *Science 338*, 1408−1411.
(84) Yonekura-Sakakibara, K., Fukushima, A., and Saito, K. (2012) Transcriptome data modeling for targeted plant metabolic engineering. *Curr. Opin. Biotechnol. 24*, 285−290.
(85) Purcell, O., Peccoud, J., and Lu, T. K. (2013) Rule-based design of synthetic transcription factors in eukaryotes. *ACS Synth. Bio.*, DOI: 10.1021/sb400134k.
(86) Gaj, T., Gersbach, C. A., and Barbas, C. F., III (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol. 31*, 397−405.

(87) Owens, J. B., Mauro, D., Stoytchev, I., Bhakta, M. S., Kim, M., Segal, D. J., and Moisyadi, S. (2013) Transcription activator-like effector (TALE)-directed piggyBac transposition in human cells. *Nucleic Acids Res. 41*, 9197−9207.
(88) Yang, L., Guell, M., Byrne, S., Yang, J. L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A. W., and Rios, X. (2013) Optimization of scarless human stem cell genome editing. *Nucleic Acids Res. 41*, 9049−9061.
(89) Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., and Church, G. M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol. 31*, 833−838.
(90) Reyon, D., Khayter, C., Regan, M. R., Joung, J. K., and Sander, J. D. (2012) Engineering designer transcription activator-like effector nucleases (TALENs) by REAL or REAL-Fast assembly. *Curr. Protocols Mol. Biol.*, 12.15. 1−12.15. 14.
(91) Liang, J., Chao, R., Abil, Z., Bao, Z., and Zhao, H. (2013) FairyTALE: A high-throughput TAL effector synthesis platform. *ACS Synth. Bio.*, DOI: 10.1021/sb400109p.