Published in final edited form as:

Nat Biotechnol. 2013 January ; 31(1): 76–81. doi:10.1038/nbt.2460.

A ligation-independent cloning technique for high-throughput assembly of transcription activator–like effector genes

Jonathan L Schmid-Burgk¹, Tobias Schmidt¹, Vera Kaiser¹, Klara Höning¹, and Veit Hornung¹

¹Institute for Clinical Chemistry and Clinical Pharmacology, Unit for Clinical Biochemistry, University Hospital, University of Bonn, Bonn, Germany

Abstract

Transcription activator–like (TAL) effector proteins derived from *Xanthomonas* species have emerged as versatile scaffolds for engineering DNA-binding proteins of user-defined specificity and functionality. Here we describe a rapid, simple, ligation-independent cloning (LIC) technique for synthesis of TAL effector genes. Our approach is based on a library of DNA constructs encoding individual TAL effector repeat unit combinations that can be processed to contain long, unique single-stranded DNA overhangs suitable for LIC. Assembly of TAL effector arrays requires only the combinatorial mixing of fluids and has exceptional fidelity. TAL effector nucleases (TALENs) produced by this method had high genome-editing activity at endogenous loci in HEK 293T cells (64% were active). To maximize throughput, we generated a comprehensive 5-mer TAL effector repeat unit fragment library that allows automated assembly of >600 TALEN genes in a single day. Given its simplicity, throughput and fidelity, LIC assembly will permit the generation of TAL effector gene libraries for large-scale functional genomics studies.

INTRODUCTION

TAL effector proteins are proteins secreted by bacteria of the *Xanthomonas* family that can enter the nucleus of an infected host plant cell and bind to genomic DNA in a sequence-specific manner, thereby regulating the expression of adjacent host genes¹. Recently, TAL effector proteins were found to have a highly modular domain architecture that dictates their sequence specificity. They are composed of a variable number of repeat units, each ~34 amino-acid residues in size, of which only two residues—called the repeat variable di-residue (RVD)—confer nucleotide specificity^{2, 3}. This modularity and the recently deciphered RVD code underlying nucleotide specificity^{2, 3} facilitate the generation of synthetic TAL effector proteins that can be linked to various effector domains to interfere with user-defined genomic loci in human cells. For example, overexpression of a C-terminal fusion of a VP16 transcriptional activation domain to an engineered TAL effector

Correspondence and requests for materials should be addressed to V.H. (veit.hornung@uni-bonn.de).

Author Contributions: J.L.S.-B., T.S. and V.H. developed the methodology, designed experiments, analyzed the data and wrote the manuscript. J.L.S.-B., T.S., V.K. and K.H. performed the experiments. V.H. supervised the project.

Competing financial interests: J.L.S.-B., T.S. and V.H. are inventors on a patent application dealing with LIC assembly of TALE genes.

upregulated a gene of interest in a human cell line⁴. As shown first in the field of zinc finger nucleases⁵, TAL effector proteins can also be fused with a nuclease domain to induce double-stranded (ds)DNA breaks at a specific genomic locus. For dsDNA cleavage, two TALENs are engineered to bind adjacent to each other at a genomic locus of interest in such a way as to induce dimerization and activation of the nuclease domains⁶. Subsequently, editing can occur through homologous recombination with a donor DNA or through nonhomologous end-joining repair mechanisms⁵. TALENs have been used to edit endogenous loci in plants⁶, yeast⁷, *Drosophila*⁸, rat⁹, zebrafish^{10, 11}, nematodes¹², human somatic and stem cells^{4, 13}, and crickets¹⁴.

The construction of large synthetic genes from multiple cassettes is challenging using traditional cloning techniques, especially when repetitive cassettes are to be assembled without altering the encoded amino-acid sequence. The most widely used technology to assemble TAL effector genes with a desired domain architecture is based on *Golden Gate* cloning, which employs type IIS restriction endonucleases to cut outside of their recognition sequence and thereby to generate arbitrary, nonpalindromic, 4-bp, single-stranded (ss)DNA overhangs that cannot be recut after ligation to a compatible ssDNA overhang. This approach has greatly facilitated and accelerated the synthesis of TAL effector genes^{7, 15, 16, 17, 18, 19}, although it still requires considerable hands-on time and quality-control measures. At the same time, several critical procedural steps, such as the identification and isolation of correctly assembled TAL effector constructs, limit the applicability of this technology to high-throughput applications. Recently, an additional TAL effector synthesis approach called FLASH was described²⁰. Geared to high-throughput production, FLASH uses a solid-phase strategy to ligate several TAL effector repeat unit fragments in a unidirectional manner. In an iterative process, newly added TAL effector repeat unit fragments are ligated to an immobilized DNA construct, and interlaced washing steps between the ligation events assure correctly ordered ligations. The final ligation product is then released from the solid phase by a restriction digest and is subsequently cloned into an expression-ready vector.

In contrast to conventional cloning techniques, LIC relies on much longer (10–30 bp) nonpalindromic, single-stranded overhangs of dsDNA fragments that are designed to anneal with the overhangs of other fragments in a highly specific manner^{21, 22}. If one of the annealed DNA fragments contains a bacterial origin of replication, *Escherichia coli* can be transformed by an annealing reaction without prior ligation because the fragments' long overlaps do not dissociate during the transformation process and thus can be ligated by bacterial ligases. Recently, we and others published a two-step hierarchical LIC assembly of nine gene-size DNA fragments and a backbone fragment, which relies on the controllable 3' exonuclease activity of T4 DNA polymerase for the sequence-specific generation of single-stranded overhangs²³. If one dNTP is supplied to this proof-reading polymerase, its exonuclease activity outcompetes its polymerase activity exactly until the terminal sequences of the dsDNA fragments in the reaction allow the incorporation of the supplied dNTP (chewback reaction). By designing the terminal fragment sequences to spare one nucleotide, one can thereby generate well-defined overhangs and thus orchestrate LIC-based multiple-fragment assemblies. Notably, upon transformation of multiple fragment

assemblies, the LIC technology usually yields highly reliable assembly results with negligible or no background colonies. Therefore, under optimal conditions, no additional selection procedure, such as blue and white screening or colony PCR, is required to single out positive clones. Indeed, when carefully selecting ssDNA overhangs of high specificity, this technology even allows the expansion of assembled fragments in polyclonal bacterial cultures without prior selection of clonal colonies.

RESULTS

An assembly strategy to synthesize TAL effector genes using LIC

We devised a strategy to assemble TALEN plasmids tailored to the previously published 152/+63-AvrBs3-like TALEN architecture, which is characterized by functionally optimized N- and C-terminal truncations and a C-terminal FokI domain (Fig. 1a)⁴. For the RVDs of the individual repeat units, we used the most common natural RVD-nucleotide cipher for T (NG), A (NI) and C (HD)², and for G we chose the more specific but less frequent RVD NK⁴. To use the unique cloning fidelity of LIC for the rapid synthesis of TAL effector genes, we searched for the optimal combination of base-restricted sequence stretches that could be introduced into the coding sequence lying between two RVDs. For optimal cloning performance, each so-called ID stretch had to be at least ten bases in length and had to not overlap with any other used ID stretch by more than three bases. Also, it had to not alter the encoded amino acid sequence of the repeat unit at any position. By taking advantage of the degenerate nature of the genetic code, we could design four base-restricted ID sequences that were suitable for our approach (ID1–ID4; Supplementary Fig. 1). With four unique IDs it is possible to seamlessly fuse four fragments (e.g., 3 inserts + 1 plasmid backbone) per reaction using LIC technology. To this effect, we decided to use fragments composed of two consecutive TAL effector repeat units as the minimal cloning unit. This approach allows the generation of an 18-RVD TAL effector fragment in a hierarchical, two-step assembly process. To this effect, 64 2-mer fragments containing all 16 possible 2-mers flanked by four different ID combinations (IDs: 1/2, 2/3, 3/4 and 4/1) constitute the source library (Fig. 1b). From this library, three individual 2-mer fragments can be assembled into a level 1 backbone plasmid to generate an intermediate 6-mer fragment containing six consecutive repeat units (Fig. 1c,d). To generate 6-mer fragments with three unique ID combinations for the second LIC assembly, three level 1 backbones are required (Fig. 1d and Supplementary Fig. 2a). Subsequently, three 6-mer fragments are assembled as an 18-mer into a eukaryotic expression-ready level 2 backbone (Supplementary Fig. 2b) that already contains the 152/+63-AvrBs3-like TALEN backbone with a C-terminal half-repeat unit (Fig. 1e,f). The resulting TALEN contains 18.5 TAL effector repeat units and has a target specificity of the sequence 5'-T(N)₁₉-3'.

LIC assembly yields functional TALENs

Testing of these assembly reactions showed that the fidelity of the 6-mer assembly reaction was exceptionally high: shorter fragment assemblies were rarely observed, and in none of the correctly sized 6-mer fragments were assemblies of the wrong order found. This feature allowed us to grow the assembled 6-mer fragments in polyclonal cultures after transformation, thereby avoiding single-clone generation. For example, we generated a

TALEN pair (72L and 72R) that targeted the coding sequence of the human STAT6 gene (Fig. 2a) by assembling three 2-mer fragments into one 6-mer fragment, six times in parallel. The assembled 6-mer fragments were then used to transform bacteria that were grown in polyclonal cultures without further clonal selection procedures (Supplementary Fig. 3a). DNA isolated from these cultures yielded the expected product sizes without any detectable side products. When we used the DNA from these cultures to transform *E. coli* and obtain single clones, we observed that 98.3% (59/60) of all tested colonies contained the correct insert (Supplementary Fig. 3b). Moreover, Sanger sequencing confirmed the correct sequence of 6-mer fragments (data not shown). The polyclonal 6-mer digestions were subsequently used for a T4 DNA polymerase chewback reaction without any prior purification, and the respective three 6-mers were then assembled into the expression-ready level 2 backbone (Supplementary Fig. 3c). Again, we transformed bacteria with the assembly reaction and propagated them without single-colony selection. Control digestions of the obtained 18-mer plasmids showed no discernable extra bands besides the expected 18-mer fragment, and retransformation of the DNA and analysis of the resulting single clones showed 100% correct product sizes (20/20) (Supplementary Fig. 3d). Sequencing furthermore confirmed the fidelity of the assembled 18-mer fragments (9 out of 9 completely sequenced 18-mers were correct). To test the functionality of the hsSTAT6 TALEN pair, we transiently transfected HEK 293T cells with the two TALEN constructs and PCR-amplified the targeted locus 48 h after transfection. Using the mismatch-sensitive T7 endonuclease I (T7EI) assay²⁴, we calculated a 30.6% cutting efficiency of this TALEN pair at the STAT6 locus (Fig. 2b).

Large-scale assembly of TALENs

Without agar-based single-colony picking of the transformants, only 3 days are required to obtain a correctly assembled 18.5-RVD TALEN using the above-described two-level assembly protocol (Supplementary Fig. 4). However, subsequent experiments showed that ~50% of the polyclonally grown 18-mer assemblies contained trace impurities of smaller sizes. This problem could easily be circumvented by cloning the 18-mer assembly transformants using conventional agar plates, which led to 80–90% correctly assembled clones. Nevertheless, to automate LIC-based TAL effector synthesis for high-throughput applications, we wished to devise a strategy that would rely only on liquid-handling steps. We therefore developed a protocol based on limiting-dilution cloning to obtain liquid cultures derived from a single bacterial clone (Supplementary Fig. 5 and Supplementary Note 1). We chose to assemble a set of 64 TALEN constructs in a semi-automated fashion targeted to critical exons within 32 target genes currently under study in our laboratory (Supplementary Table 1). We used the level 2 backbone containing a C-terminal thymine-specific RVD and chose the spacer region to be in the range of 12–19 bp (targetable sites: 5'-T(N)₁₈T-[N]₁₂₋₁₉-A(N)₁₈A-3').

A pipetting robot automatically assembled LIC-ready 2-mer fragments in 384-well plates into 192 6-mer fragments, and these were subsequently transformed into bacteria in 96-well plates (Fig. 3a). Bacteria were grown overnight in polyclonal cultures for subsequent DNA isolation, and plasmid DNA was obtained from all cultures. Control digestions showed that all 6-mer fragments displayed the expected product size of about 600 bp (Supplementary

Fig. 6). The 192 6-mer fragments were then subjected to chewback, assembled into 64 18-mer constructs and used to transform bacteria (Fig. 3a). These were either grown polyclonally or under limiting-dilution conditions, whereas five cultures were inoculated. Ninety-seven percent (62/64) of the polyclonal inoculations displayed the correct 18-mer control digest band at 3.3 kb, but about half of the samples contained impurities of smaller assemblies, which made up roughly 10% of the total plasmid DNA of all samples taken together (Supplementary Fig. 7a). In contrast, a perfectly correct plasmid preparation could be selected for 62 out of 64 assemblies within the first assembly attempt when five liquid cultures per assembly had been inoculated under limiting dilution at 50% growth (Supplementary Fig. 7b). Correct clones of the remaining two assemblies could be obtained by inoculating five additional cultures each. Altogether, this semi-automated assembly procedure allowed us to generate 64 expression-ready TALEN constructs within 3 working days.

Validation of genome editing activity

To study the functional activity of the 32 TALEN pairs, we transfected HEK 293T cells and assessed genome editing activity 48 h later using the T7EI assay. Of the 32 tested, 19 TALEN pairs showed genome-editing activity with calculated mutation frequencies ranging from 2% to 39% (Fig. 3b). Notably, two of the nonfunctional TALEN pairs (MIB2, IKBKE) showed good editing activity in an *in vitro* cleavage assay, indicating that target accessibility within 293T cells is a limiting factor (data not shown). Therefore, to control for possible target-accessibility constraints, we decided to target a genomic locus that we had previously identified to be amenable to TALEN-mediated genome editing (exon 3 of *hsSTAT1*) with another set of TALENs. Using the above-described LIC assembly approach, we generated 10 TALEN pairs that could be functionally tested in 29 reasonable combinations with varying spacer length (Supplementary Fig. 8a,b). Testing these TALENs in HEK 293T cells confirmed that the spacer length is a critical determinant of the genomic targeting efficiency. At the same time, this set of TALENs displayed a higher success rate compared with the randomly chosen 32 target sites. Indeed, within the previously published optimal spacer length of 12–20 bp for the 152/+63-AvrBs3-like TALEN architecture⁴, 12 out of 14 TALEN combinations induced calculated mutation frequencies ranging from 3% to 61% (Supplementary Fig. 8c). When analyzing all 42 TALEN pairs tested, we observed the highest mean genome-editing activity at a spacer length of 14 or 15 bp, with 13 out of 17 TALEN being active (Fig. 4a). When assessing the frequency of the four individual RVD repeat units within these TALEN pairs, we observed no correlation between RVD usage and on-target activity (Fig. 4b).

A 5-mer library for high-throughput TAL effector gene assembly

To further enhance the throughput of our assembly approach and to further reduce the turnover time, we decided to generate a level 1 fragment library containing TAL effector repeat unit fragments large enough to generate functional TALENs within only one LIC assembly step. To this effect, we chose to generate a 3,072-piece plasmid library (3×4^5) containing all possible 5-mer fragments in three ID combinations (Fig. 5a). This size was chosen as a good trade-off between manageability of the library on the one hand and a suitably high sequence specificity of the final TAL effector genes (15.5 RVD repeat units)

on the other. To generate 5-mer fragments with three different ID combinations (IDs: 4/3, 3/2, 2/1), we generated an additional set of 1-mer fragments, an additional level 1 backbone (ID: 2/1) and a new set of level 2 backbones (ID: 4/1). Subsequently, using a robot, we performed 3,072 5-mer LIC assemblies and transformed the resulting constructs into bacteria (Fig. 5b). From these, monoclones were generated and DNA was extracted. For 99.67% (3,062/3,072) of the assembled 5-mers, a correct clone could be selected using analytical restriction digests, and 7 out of 7 randomly selected 5-mer fragments displayed the expected sequence upon Sanger sequencing (Supplementary Fig. 9a). Without further adjustments, DNA of the 5-mer library was then digested and chewed back to prepare for the one-step TAL effector gene assembly (Fig. 5c). To test the assembly process using the 5-mer fragment library, we manually assembled a set of four individual 15.5-RVD TALEN constructs in two independent assembly reactions each. As previously observed for the 18-mer assembly reactions, a high percentage of correctly sized 15-mer fragment clones (82%) was obtained. In addition, two out of two analyzed 15.5-RVD TAL effectorN clones displayed the expected sequence (data not shown). To test the functional activity of the 15.5 RVD TALENs, we generated a set of three TALEN pairs targeting exon 5 of human STAT6 and tested these for genome-editing activity (Fig. 6a). Two out of three TALEN pairs showed genome-editing activity, with calculated mutation frequencies of 30% and 31%, respectively (Fig. 6b). Additional experiments targeting STAT1 showed similar results (data not shown).

To demonstrate the applicability of this library to high-throughput assembly, we accommodated the 5-mer fragment library on a liquid-handling robot and performed 768 15-mer assembly reactions in a single day (Supplementary Table 2). Transformants of the assemblies were grown at limiting-dilution conditions in septuplets to obtain monoclonal 15.5-RVD TALEN constructs (Fig. 5c). For 90% of all assemblies (688/768), at least one inoculated culture grew at the first attempt. Analyzing the clonal 15.5-RVD TALEN constructs using restriction digest revealed that for 89% (619/688) of these assemblies a correctly sized construct of ~3 kb could be obtained (Supplementary Fig. 9b). Notably, subsequent sequencing confirmed the fidelity of the assembly reactions, with 20/20 sequencing reactions displaying the expected sequence. Altogether these results show that LIC assembly allows high-throughput synthesis of TAL effector constructs in a one-step assembly reaction with exceptional fidelity.

DISCUSSION

TAL effector technology represents a breakthrough in the field of genome engineering. The elegance of TAL effector proteins is their modularity, as their domain structure dictates their DNA sequence specificity with a 1:1 repeat unit-to-DNA base relation^{2, 3}. Given that individual repeat units seem to bind their target bases in a relatively context-independent manner, it has become possible to easily apply the TAL effector cipher to engineer designer TAL effectors of any specificity. However, owing to their repetitive domain architecture, synthetic TAL effector genes are difficult to generate using standard cloning procedures. To date, various cloning technologies have been published that make use of either cut-ligation cloning^{7, 15, 16, 17, 18, 19} or conventional ligation cloning combined with a solid-phase system²⁰. All of these methods rely on the specificity of four-nucleotide overhangs and the

use of ligase for the actual assembly process. This poses the risk of imprecise assemblies and also complicates the automation of these techniques.

LIC assembly technologies yield highly reliable cloning results with negligible or no background colonies, even when assembling four fragments at once²³. Based on this observation we developed an LIC-based assembly strategy to synthesize 15.5 and 18.5 RVD TAL effector constructs within 2 or 3 working days. Our technology consists of only simple pipetting steps and does not require any PCR, agarose gel purification or sequencing reaction, and, moreover, it can be practiced with standard laboratory equipment. Another major advantage of LIC technology is its extremely high assembly precision, which can be attributed to the highly specific annealing of long ssDNA overhangs. This results in a very high percentage of correctly assembled constructs. Indeed, up to now, we have never observed incorrect sequences of correctly sized TAL effector constructs. In this respect, a simple restriction digest proved to be sufficient to validate the correctness of an assembly reaction, thereby avoiding sequencing reactions for quality control. Notably, our assembly technology is not restricted to the 15.5 or 18.5 repeat unit TAL effector architecture demonstrated here. With the currently available repeat unit fragments, it is also possible to generate 16.5- and 17.5-RVD TAL effector constructs. Moreover, by adding another set of 1-mer fragments to the library, TAL effector genes with 9.5 to 14.5 could be generated with the same assembly throughput (Supplementary Fig. 10). In addition, the functionality of the TAL effector constructs can be adapted by exchanging the FokI domain for other domains. Also, newly characterized RVD repeat units could easily be incorporated into the LIC assembly platform by generating novel 1-mer or 2-mer fragments.

There have been only a few reports in which the on-target activity at endogenous loci has been studied for a reasonable number of TALENs. One study reported a success rate of 87.5% (84/96) when targeting randomly selected endogenous loci in U2OS cells²⁰. Using the same TALEN platform, another study reported 60% working TALEN pairs (6/10) with an activity of >2% in zebrafish²⁵. The success rate of our LIC-assembled 18.5-RVD TALEN pairs lies within the range of 59% (19/32), for endogenous loci that were randomly selected, to 86% (12/14), for a locus that was preselected for prior TALEN activity. Given that these studies differ in many critical variables, (e.g., the cell type targeted, the technique used to measure on-target activity) and the TALEN architecture itself (e.g., use of the possibly lower-affinity RVD NK²⁶), it is difficult to attribute these differences in targeting activity to a specific experimental or technical parameter. Additional studies are required to dissect the individual contribution of these factors.

The speed, ease of use, high fidelity and amenability to automation of the LIC assembly technology make it well suited to high-throughput cloning. We used our method to generate a plasmid library encompassing all possible 5-mer TAL effector repeat unit arrays, which expedites and simplifies the synthesis of TAL effector constructs. With this resource in hand, only one four-piece assembly reaction followed by a bacterial transformation reaction is required per TAL effector construct, thus effectively reducing the number of hierarchical assembly steps from two to one. Using limiting-dilution cloning, it is possible to obtain ready-to-transfect plasmid DNA of a TAL effector construct within 24 h. As no special laboratory equipment is needed, we expect that every laboratory can easily apply this

technique to generate TAL effector constructs at a fidelity, throughput and simplicity unmatched by any of the currently available techniques (for a comparison with the Golden Gate and FLASH methods, see Supplementary Table 3). At the same time, our one-step LIC assembly platform provides a realistic prospect for the generation of genome-wide TALEN libraries within a reasonable timeframe. We anticipate that such libraries will provide a valuable tool for advancing functional genomics by loss-of-function approaches as did the short interfering RNA and short hairpin RNA libraries in the past decade. Future studies aimed at exploring the determinants of off-target effects and of on-target activity will be important for the design of TALEN libraries with maximum specificity and activity.

Methods

DNA constructs

Two-mer fragments and the coding sequence of the level 2 backbone with the thymine-specific C-terminal half-repeat unit were synthesized by ShineGene (Shanghai, China). All other backbones and the 1-mer fragments were generated by PCR cloning. All sequences can be found in Supplementary Table 4.

Chewback reactions

Prior to the chewback reaction, the 1-mer, 2-mer, 5-mer or 6-mer fragments were cut out of their backbone plasmids using MvaI269I. We mixed 3 μ l of the crude MvaI269I restriction reaction with 7 μ l of a master mix containing the following components: 1 μ l of 10 $\times$ NEB2 buffer (NEB), 0.1 μ l of BSA 10 g/l (NEB), 0.1 μ l of STOP dNTP 100 mM, 0.33 μ l of T4 DNA polymerase 3 U/ μ l (Enzymatics) and 5.47 μ l H₂O. The reaction was incubated at 27 °C for 5 min, briefly put on ice and heated to 75 °C for 20 min.

Assembly reactions

We mixed 2.5 μ l of each dimer and backbone. The assembly reactions were incubated at 55 °C for 30 min and then at 25 °C for 3 h.

Transformation of *E. coli*

We combined 2 μ l assembly mix with 10 μ l of chemocompetent *E. coli* DH10b on ice. The reaction was incubated for 3 min at 37 °C and for 2 min on ice. We added 100 μ l LB medium and the reaction was shaken at 37 °C for 1 h at 900 r.p.m. We added 1 ml LB medium supplemented with 100 μ g/ml Ampicillin or 30 μ g/ml Kanamycin and the reaction was incubated at 37 °C for 12–16 h at 900 r.p.m.

Plasmid preparation

After alkaline lysis, plasmid DNA was purified using silica spin columns according to the manufacturer's protocol (Genetic Biotech or Promega Wizard SV).

Transfection of HEK 293T cells

HEK 293T cells were transfected by 100 ng of each of a pair of TALEN plasmids per well of a 96-well plate at roughly 70% confluency using GeneJuice (Merck Millipore) according to the manufacturer's protocol.

T7 endonuclease I (T7EI) assay

Two days after transfection the medium was removed and the cells were lysed in 150 µl of the following lysis buffer: 0.2 mg/ml proteinase K, 1 mM CaCl₂, 3 mM MgCl₂, 1 mM EDTA, 1% Triton ×100, 10 mM Tris pH 7.5. The reactions were incubated at 65 °C for 10 min and at 95 °C for 15 min. We directly used 1 µl of the lysate in a 25-µl PCR reaction using Phusion DNA polymerase (Finnzymes) according to the manufacturer's protocol. Respective primer sequences can be found in Supplementary Table 5. The annealing temperature was 64 °C, the elongation time was 15 s and the cycle number was 32. After completion, 5 µl of the PCR reaction were mixed with 1.1 µl 10× NEB buffer 2 and 4.4 µl water. The reactions were heated to 95 °C and allowed to cool down to room temperature at a ramp rate of 0.7 °C per minute. Then, 0.5 µl T7 endonuclease I (NEB) were added and the reactions were incubated at 37 °C for 20 min. After adding agarose gel loading buffer, the samples were separated on a 2% agarose gel at 100V and the DNA was visualized using ethidium bromide under UV light. Densitometric quantification of DNA bands was done using ImageJ. Mutation frequencies were calculated using the formula: fractional modification = 1 – (1 – (fraction cleaved))^{0.5} as described in ref. 27.

Reagent availability

All reagents described in this report to generate TALENs using LIC assembly are available to academic researchers upon request (<http://www.hornunglab.de/TALEN>). See Supplementary Methods for detailed protocols.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank T. Cathomen for technical advice with the T7EI assay and M. Hölzel for helpful discussion. V.H. is member of the excellence cluster ImmunoSensation and supported by grants from the German Research Foundation (SFB704 and SFB670) and the European Research Council (ERC-2009-StG 243046).

References

1. Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol.* 2010; 48:419–436. [PubMed: 19400638]
2. Boch J, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science.* 2009; 326:1509–1512. [PubMed: 19933107]
3. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science.* 2009; 326:1501. [PubMed: 19933106]
4. Miller JC, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol.* 2011; 29:143–148. [PubMed: 21179091]

5. Carroll D. Genome engineering with zinc-finger nucleases. *Genetics*. 2011; 188:773–782. [PubMed: 21828278]
6. Christian M, et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*. 2010; 186:757–761. [PubMed: 20660643]
7. Li T, et al. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res*. 2011; 39:6315–6325. [PubMed: 21459844]
8. Liu J, et al. Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. *J Genet Genomics*. 2012; 39:209–215. [PubMed: 22624882]
9. Tesson L, et al. Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol*. 2011; 29:695–696. [PubMed: 21822240]
10. Huang P, et al. Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol*. 2011; 29:699–700. [PubMed: 21822242]
11. Sander JD, et al. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol*. 2011; 29:697–698. [PubMed: 21822241]
12. Wood AJ, et al. Targeted genome editing across species using ZFNs and TALENs. *Science*. 2011; 333:307. [PubMed: 21700836]
13. Hockemeyer D, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol*. 2011; 29:731–734. [PubMed: 21738127]
14. Watanabe T, et al. Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. *Nat Commun*. 2012; 3:1017. [PubMed: 22910363]
15. Zhang F, et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*. 2011; 29:149–153. [PubMed: 21248753]
16. Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S. Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One*. 2011; 6:e19722. [PubMed: 21625552]
17. Morbitzer R, Elsaesser J, Hausner J, Lahaye T. Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Res*. 2011; 39:5790–5799. [PubMed: 21421566]
18. Cermak T, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*. 2011; 39:e82. [PubMed: 21493687]
19. Geissler R, et al. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS One*. 2011; 6:e19509. [PubMed: 21625585]
20. Reyon D, et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*. 2012; 30:460–465. [PubMed: 22484455]
21. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res*. 1990; 18:6069–6074. [PubMed: 2235490]
22. Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res*. 2007; 35:e55. [PubMed: 17389646]
23. Schmid-Burgk JL, et al. Rapid hierarchical assembly of medium-size DNA cassettes. *Nucleic Acids Res*. 2012
24. Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res*. 2009; 19:1279–1288. [PubMed: 19470664]
25. Cade L, et al. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res*. 2012; 40:8001–8010. [PubMed: 22684503]
26. Streubel J, Blucher C, Landgraf A, Boch J. TAL effector RVD specificities and efficiencies. *Nat Biotechnol*. 2012; 30:593–595. [PubMed: 22781676]
27. Miller JC, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*. 2007; 25:778–785. [PubMed: 17603475]

Figure 1. The LIC assembly approach to generating TALEN genes

(a) Architecture of the previously published $\Delta 152/+63$ -AvrBs3-like TALEN that harbors a truncated N terminus ($\Delta 152$), an invariant recognition domain for thymine followed by a stretch of several repeat units, and a truncated and optimized C terminus (+63) fused to a FokI domain. One repeat unit is highlighted indicating the 34 amino acids and the RVDs in position 12 and 13 that were used. At right, a model of a minimal cloning unit consisting of two consecutive repeat units flanked by ID1 and ID2. (b) The library of the 64 possible 2-mer combinations with ID 1/2, 2/3, 3/4 and 4/1. (c,d) First level assembly: three 2-mers

containing the target sequences are picked with alternating ID combinations (e); three 2-mers are assembled to generate a complex of six tandem repeats (6-mer) into a level 1 backbone that encodes for kanamycin resistance (kanaR) (d). (e,f) Second level assembly: assembly of a full TAL effector gene is achieved upon combining three 6-mers (e), and a mammalian expression vector containing the 152/+63-AvrBs3-like TAL effector backbone with a C-terminal FokI nuclease domain (f). Additional features of the mammalian expression vector (level 2 vector) are highlighted (pCMV, cytomegalovirus-promoter, ampR, bacterial resistance gene for ampicillin; NLS, nuclear localization signal; ATG, start codon; STOP, stop codon; T7, promoter sequence for T7 RNA polymerase; XhoI, XbaI, NotI, recognition sequences of the restriction enzymes XhoI, XbaI and NotI, respectively). To prevent undigested fragments from being propagated to the next level of assembly, we switched antibiotic resistance at each level.

Figure 2. LIC assembly and validation of a TALEN targeting STAT6

(a) A schematic view of the human STAT6 gene locus with the targeting site of TALEN pair 72 (yellow letters indicate the intronic sequence portion). (b) The T72L and T72R TALEN pair tested for genome editing activity in HEK 293T cells using the T7EI assay. The left panel provides a sketch of the PCR amplicon used for the assay, whereas the arrows indicate the primers used and the vertical line indicates the expected location of the nuclease activity. The right panel depicts an agarose gel of one representative result. The white numbers indicate quantified mutation frequencies. u, uncut DNA.

Figure 3. Large-scale assembly of TALENs using LIC

(a) The workflow of the 32 TALEN pair assembly. (b) Thirty-two TALEN pairs tested for functional activity in HEK 293T cells using the T7EI assay. Presented are the differences in measured mutation frequency between the TALEN-treated cells and untransfected control cells. BB, backbone.

Figure 4. Evaluation of targeting activity dependence on spacer length and RVD composition
(a) Of all TALEN pairs tested, pairs with a spacer length of 12–19 bp were plotted for genome editing activity versus spacer length. n.t., not tested **(b)** The relative RVD composition of the 42 tested TALEN pairs (left y-axis) alongside their respective targeting efficiencies (right y-axis).

Figure 5. A 5-mer library for LIC assembly of TALENs

The schematic of the assembly strategy of the 5-mer fragment library. **(a)** To obtain all possible 5-mer fragments of one ID combination, 16 2-mer fragments in two different ID combinations and 4 1-mer fragments are assembled to give rise to 1,024 5-mer fragments. **(b)** The 5-mer fragments are then cloned into level 1 backbones available in three different ID combinations to obtain $3 \times 1,024$ 5-mer fragments. **(c)** The arrayed 5-mer fragment library is digested and chewed back in 384-well plates. A liquid-handling robot picks and mixes the appropriate 5-mer fragments with their appropriate level 2 backbone plasmid.

After transformation, bacteria are grown under limiting dilution conditions to obtain monoclonal liquid cultures. A correctly assembled TAL effector construct is identified using a restriction digest and an optional sequencing reaction.

Figure 6. On target activity of 15.5-RVD LIC TALENs

(a) A schematic view of the human STAT6 gene locus with the targeting sites of TALEN pairs T499, T500 and T501 (yellow letters indicate the intronic sequence portion). (b) The TALEN pairs were tested for genome editing activity in HEK 293T cells using the T7EI assay. The panel depicts an agarose gel of one representative result out of two. The white numbers indicate quantified mutation frequencies. *, nonspecific PCR product.