ACS Synthetic Biology
Research Article
pubs.acs.org/synthbio

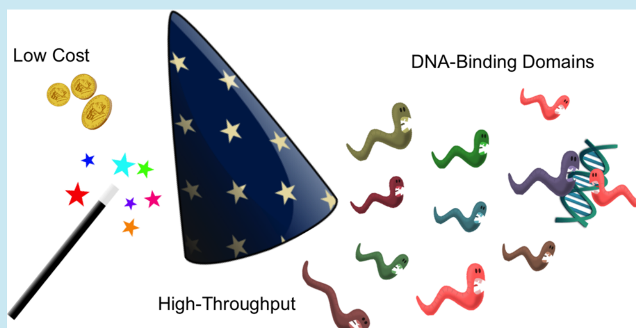
FairyTALE: A High-Throughput TAL Effector Synthesis Platform

Jing Liang,[†] Ran Chao,[†,‡] Zhanar Abil,[§] Zehua Bao,[§] and Huimin Zhao[†,‡,§,*]

[†]Department of Chemical and Biomolecular Engineering, [‡]Institute for Genomic Biology, [§]Department of Biochemistry, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States of America

Supporting Information

ABSTRACT: Recombinant transcription activator-like effectors (TALEs) have been effectively used for genome editing and gene regulation applications. Due to their remarkable modularity, TALEs can be tailored to specifically target almost any user-defined DNA sequences. Here, we introduce fairyTALE, a liquid phase high-throughput TALE synthesis platform capable of producing TALE-nucleases, activators, and repressors that recognize DNA sequences between 14 and 31 bp. It features a highly efficient reaction scheme, a flexible functionalization platform, and fully automated robotic liquid handling that enable the production of hundreds of expression-ready TALEs within a single day with over 98% assembly efficiency at a material cost of just $5 per TALE. As proof of concept, we synthesized and tested 90 TALEs, each recognizing 27 bp, without restrictions on their sequence composition. 96% of these TALEs were found to be functional, while sequencing confirmation revealed that the nonfunctional constructs were all correctly assembled.



KEYWORDS: *synthetic biology, genome engineering, genome editing, TAL effector, TALEN*

Recombinant transcription activator-like effector (TALE) technology has become a versatile platform for genome editing[1−4] and genetic regulation.[5−8] TALE proteins were first discovered in *Xanthomonas*, a genus of plant bacteria, and its DNA recognition code was solved in 2009.[9,10] In the following years, intensive research efforts expanded the host range of TALE from plant to yeast, nematode, zebrafish, frog, cattle, human, and others.[11] DNA recognition by TALE is carried out by the central repeat domain (CRD), which contains multiple repeats of a 33−35 amino acids (aa) motif. Each of the repeats recognizes a single DNA nucleotide, and the recognition specificity is determined by the highly variable aa at position 12 and 13, called the repeat-variable diresidue (RVD). Due to the modular nature of TALE and the one-to-one correspondence between RVD and DNA nucleotide, researchers can specify the DNA binding sequence of a TALE by assembling the appropriate repeats in a tandem array.

A cost-effective, high-throughput TALE synthesis platform is needed to support the next stage of TALE technology development, that is, to expand its use as a tool for the general research community and for large genome-scale engineering applications. Three general strategies have been used to address this challenge: solid-phase,[12,13] ligation-independent cloning (LIC),[14] and Golden Gate (GG).[15,16] While highly scalable, solid-phase platforms suffer from low inherent fidelity due to the lack of specific junctions for every fragment, and high reagent cost because fresh reagents need to be added at every elongation step. LIC and GG both exhibit high inherent fidelity, but the strong antagonistic relationship between maximum number of fragments and assembly efficiency limits the number of fragments that can be assembled at high efficiency. In the LIC and GG high-throughput synthesis platforms reported so far, a maximum of 6 fragments can be simultaneously assembled[16] (typically much fewer[14,15]), beyond which the decrease in efficiency becomes untenable. Limited by the number of fragments, current high-throughput methods suffer from lower flexibility in the number of repeats a TALE can contain, and they also require higher multimer fragments to achieve a targeted number of repeats. Since the size of the substrate library scales exponentially (4^n) with the number of repeats within a fragment and only linearly with the number of fragments, having high multimer fragments leads to large substrate library.

Despite the multitude of assembly technologies, they suffer from a common downstream bottleneck—clonal isolation. In a typical TALE synthesis scheme, the assembled TALEs are transformed into *E. coli* for plasmid amplification. Due to the limited assembly efficiency, *E. coli* clones need to be isolated and screened before expression-ready TALEs can be obtained. The clonal isolation requirement adds significant time and cost to the overall synthesis scheme and encumbers the application of the current technologies for large-scale synthesis. In genome-scale applications such as genetic screening or evolution through knockout (TALEN) or gene activation/repression (transcription factor), thousands of TALEs will be needed. For such applications, the TALEs do not have to be absolutely pure,

Received: August 13, 2013
Published: November 15, 2013

© 2013 American Chemical Society

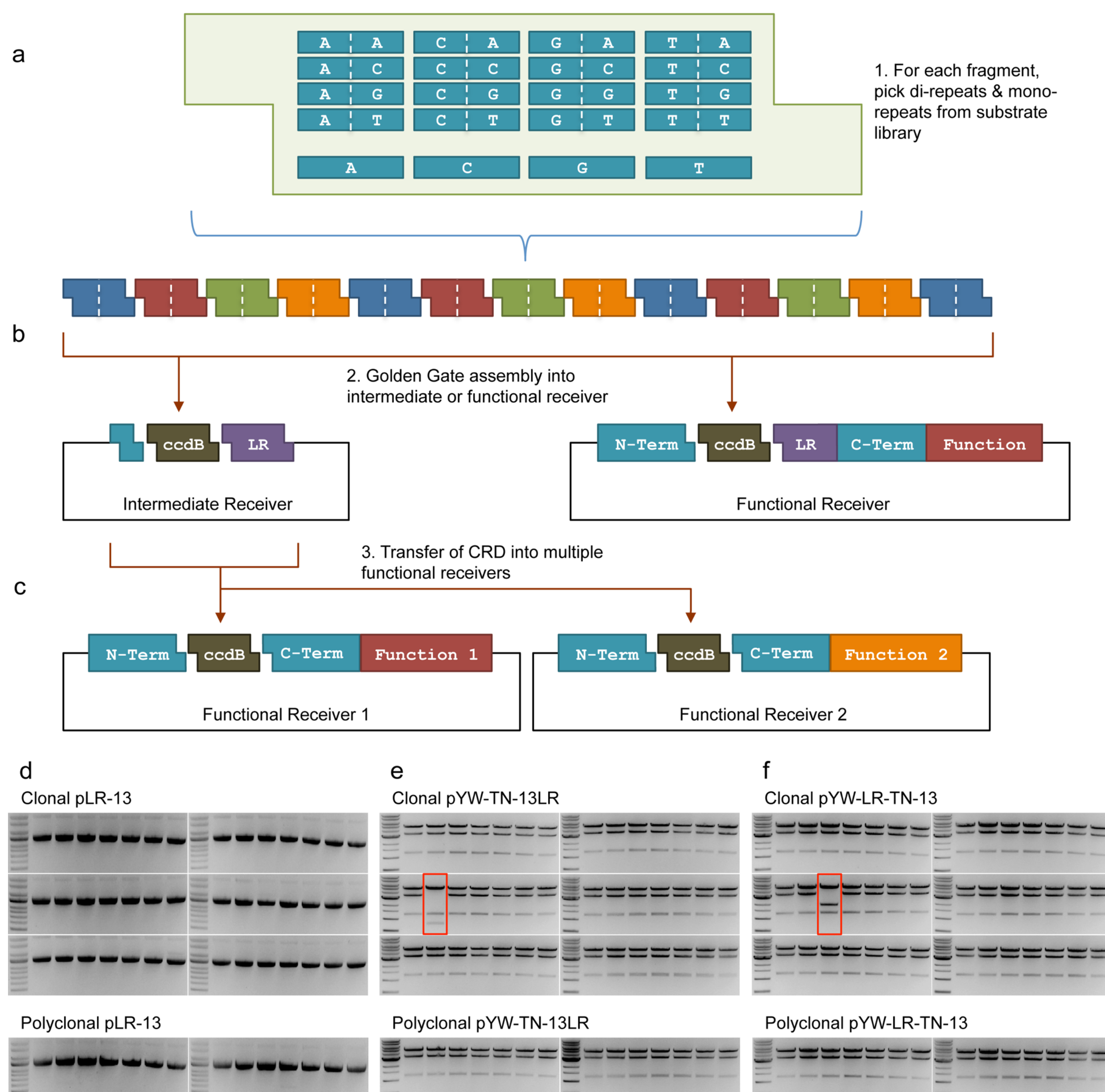


Figure 1. (a) Within a typical group, there are 16 di- and 4 monorepeat plasmids, covering all possible combinations of 2 or 1 RVD. Every group has a unique and defined position in the assembly. (b) A *Bsa*I GG reaction assembles 13 substrate fragments in one step, in the process displaces the *ccdB* toxic gene. Assembly can be made into either an intermediate receiver or functional receiver. Sequence specificity is dictated by the choice of di- and monorepeat from each group. (c) CRD assembled in an intermediate receiver can be transferred using an *Esp*3I GG reaction into multiple functional receivers. (d−f) 42 clonal plasmids from 14 different assemblies are shown in the top three rows, and the corresponding polyclonal plasmids are shown in the bottom row. Incorrect assemblies were highlighted by red boxes. NEB supercoiled ladder is used for (d) and NEB 1 kb ladder is used for (e, f). (d) Typical result of assembly into an intermediate receiver, correct assembly is ∼5 kb. (e) Typical result of a transfer into a functional receiver, digested with *Hind*III and *Stu*I, correct assembly shows a 3.7 kb band. (f) Typical result of a 1-step assembly into a functional receiver, also digested with *Hind*III and *Stu*I.

but they need to be pure enough to be highly functional. Most importantly, they have to be cheap.

Here, we introduce fairyTALE, a liquid-phase GG-based high-throughput TALE synthesis platform capable of transforming a desired target sequence into an expression-ready plasmid in a single day, without the need for clonal isolation, at a material cost of just $5 per TALE. FairyTALE represents a significant improvement over current platforms in terms of efficiency, flexibility, cost, and speed. This new platform should greatly facilitate the expansion of the recombinant TALE technologies for genome-scale applications.

RESULTS

Design of the FairyTALE System. The fairyTALE system assembles TALE CRDs from a library of fragments, which are carried by their respective substrate plasmids, excisable through the Type IIS restriction enzyme − *Bsa*I. The fragments contain either a monorepeat or a direpeat in which two TALE repeats have been preassembled together. The substrate library is comprised of 308 substrate plasmids, divided into 16 groups according to their positions in a CRD assembly. There are 13 groups that each have 20 plasmids—16 direpeats and 4 monorepeat plasmids (Figure 1a). The remaining 3 groups

dx.doi.org/10.1021/sb400109p | ACS Synth. Biol. 2014, 3, 67−73

each contain only 16 direpeat plasmids. Using a GG assembly scheme,[17] the mono- and direpeat fragments are excised from their respective substrate plasmids, spliced together in a predefined order, and ligated onto a receiver plasmid, forming a complete TALE in a single step. The receiver plasmids contain the last half repeat and can also include the N- and C-terminus extension, as well as functional domains such as FokI or VP64 (Figure 1b). FairyTALE can utilize two assembly schemes to generate TALEs with varying numbers of repeats. In the standard assembly scheme, one plasmid from each of the 13 groups is chosen according to the desired target sequence and used for assembly. By choosing either mono- or direpeats at each position, the length of TALEs can be flexibly determined, and TALEs containing between 13.5 and 26.5 repeats can be constructed. In the extended assembly scheme, group 5 in the original 13 is expanded into 3 new groups—5A, B, C—each with only direpeat substrates (Supporting Information (SI) Figure 1). By using substrates from the 15 groups, TALEs containing up to 30.5 repeats can be constructed.

FairyTALE has incorporated the latest RVD usage—NI for adenine, NG for thymine, HD for cytosine, and NH for guanine.[18] All 308 di- and monorepeats have been generated using a robotic liquid handling system from just 28 parts, and they can therefore be updated easily if necessary. Using codon redundancy, unique primer annealing regions have been incorporated into group 5 and group 9 substrate plasmids so that the final product can be fully sequenced.

For maximum functionalization flexibility, the repeats are assembled into pLR intermediate receivers, which contain the last half-repeat as well as a flanking *Esp*3I site for high efficiency transfer of the assembled CRD (Figure 1c) to the functional receivers. This 2-step synthesis route is useful when multiple functionalizations, such as both activator and repressor, of the same CRD are desired. For maximum synthesis speed, assembly can be done via a 1-step synthesis route, where the repeats are assembled directly into functional receivers such as TALE nuclease (TALEN) plasmids. The TALEN plasmids used in this work, pYW-TN and pYW-LR-TN, contain a +207 aa N-terminus extension, a +63 aa C-terminus extension, and a FokI cleavage domain.[19] The C-terminus extension and FokI have been engineered for improved activity in yeast as well as in mammalian cells (SunnyTALEN[22]). All receivers have utilized the *ccdB* counter-selection cassette to eliminate background clones during transformation. The protein sequence of a typical SunnyTALEN can be found in SI Figure 2.

The set of GG junctions used in fairyTALE have been selected for ligation efficiency, which can vary by as much as 30% among the different junction sets tested. All junctions tested are nonpalindromic but differ in their GC contents. In general, we have chosen junctions that have between 50% to 75% GC content. Given the surprising variability among the different sets, we attribute a large part of the exceptional efficiency exhibited by fairyTALE to the specific collection of GG junctions. The exact sequences of these junctions can be found in SI Figure 3.

Assembly Process. To ensure uniqueness in the human genome, a target length of at least 17 bp is required, assuming random sequence composition. However, the human genome is far from random and contains numerous repetitive and low complexity regions. Furthermore, due to the mismatch tolerance of TALE,[20] a longer target sequence will be necessary to ensure uniqueness. For most TALE applications, 26.5 repeats, recognizing a 27 bp target, should offer sufficient specificity. For TALEN applications, a heterodimer is required to make a cut, thus contributing to further stringency. Recent evidence also suggests that longer TALE may not offer better specificity because the specificity of TALE repeats appears to decrease toward the C-terminus end.[20] We have therefore focused our development and characterization on the 13-fragment standard assembly scheme. For proof of concept, we selected 90 sites of 27 bp length from 12 different yeast promoters. The sites have been chosen purely based on their position from the transcriptional start site, with no regard to sequence composition nor any of the TALE design guidelines.[4,21]

A robotic liquid handling system performed the pipetting of substrate plasmids and receiver plasmids. After a one-pot GG reaction, the mixture was transformed into Z-competent *E. coli*. One-third of each transformation was plated while the rest was transferred directly into LB+antibiotic liquid culture for polyclonal plasmid purification. After one day, the colonies on the plate were then picked for monoclonal plasmid purification and assembly efficiency assessment. Using a standard curve, the apparent plasmid yield from the GG reaction can be estimated from the number of colonies on the plate. We defined specific yield as apparent yield divided by theoretical maximum yield, and it determines the transformation efficiency requirement in the later step. Details of specific yield estimation can be found in SI Figure 4.

The performance of the standard assembly scheme was evaluated using both the 2-step and 1-step synthesis routes. For the 2-step route, the specific yield of the intermediate GG reaction was ~2%. Given a theoretical maximum yield of 50 ng, this corresponds to the production of ~1 ng circular plasmid, which is expected to give 1000 colonies at a transformation efficiency of just $1 \times 10^6/\mu g$. This allows us to use a low-cost and highly scalable transformation procedure such as Z-competent transformation.

We randomly selected 64 colonies (3 colonies each from 16 different assembly reactions during large-scale synthesis and another 16 colonies during pilot trial) for monoclonal plasmid purification and analysis by agarose gel electrophoresis. All 64 clones showed the expected 5 kb supercoiled band (Figure 1d), corresponding to an efficiency of 100% (64/64). More importantly, when the monoclonally purified plasmids were compared to the polyclonally purified plasmids, they were virtually indistinguishable. This allowed us to skip plating without compromising the resultant plasmid quality. We fully sequenced 24 of the polyclonal plasmids using two flanking and one internal primer. All the sequencing reactions gave clean reads, indicating that the polyclonal plasmids were of good quality, and all showed the expected sequences. Sample sequencing traces from a polyclonal plasmid preparation can be found in SI Figure 5.

The transfer reaction was carried out in a one-pot GG reaction using the purified polyclonal plasmids as substrates and pYW-TN as the functional receiver. The specific yield of this step was ~23%, and the correct digestion pattern was observed in ~98% (54/55) of the monoclonal plasmids (Figure 1e). Initially, 3 colonies each from 16 different assemblies were assessed, and 1 error was found. To ensure that the error was random, we assessed an additional 7 colonies from the plate containing the original error. The functionalized TALEs (in this case, TALENs) were once again purified from both monoclonal and polyclonal cultures, and the purity of the polyclonal

plasmids were indistinguishable from that of monoclonal plasmids according to restriction digestion.

For the 1-step synthesis route, substrate fragments were directly assembled into functional receivers such as pYW-LR-TN. This reaction scheme achieved a specific yield of ~3% and an efficiency of ~98% (54/55, 3 × 16 + 7). As expected, the polyclonal plasmids were also indistinguishable from that of clonal plasmids in terms of digestion pattern (Figure 1f). All 20 direct assembly polyclonal plasmids we sequenced gave clean reads and showed the expected sequences.

For the extended assembly scheme comprising a similar 15-fragment GG reaction, the specific yield was ~1.6%, the polyclonal plasmid appeared pure on agarose gel, and also generated clean sequencing reads. We sequenced 6 of the polyclonal 15-fragment assemblies, and they all showed the expected sequences. Since few TALE applications would require a 31 bp target sequence, no further characterization on the extended assembly scheme was performed.

Functional Validation. To determine the fidelity of fairyTALE, we assessed the activity of the assembled TALE CRDs on their expected target sequences. The activity was assayed in *Saccharomyces cerevisiae* using a GFP TALEN reporter system.[22] As shown in Figure 2a, two overlapping halves of the GFP gene were separated by a heterodimeric TALEN binding site containing in-frame stop codons. When a cut was made, the GFP reading frame could be restored by homologous recombination. To obtain a good indicator of CRD DNA binding activity, we kept one monomer of the TALEN heterodimer constant, so that the GFP signal reflected the relative DNA binding activity of the variant monomer. The fairyTALE-generated half-TALENs were paired with a previously generated Xa10 half-TALEN[22] in this functional assay, and the GFP signal was measured using flow cytometry.

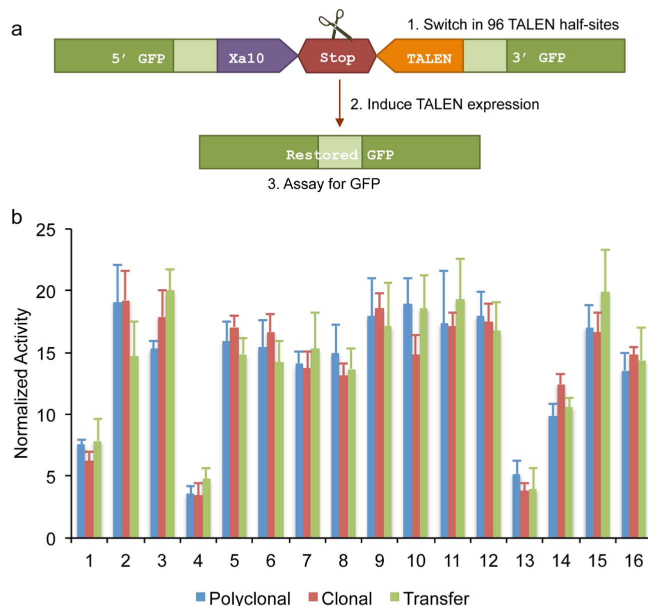


Figure 2. (a) GFP TALEN reporter. The Xa10 binding site is constant, and the TALEN binding site is unique for every TALEN we tested. The light green portion represents the GFP homology region. When a cut is made, homologous recombination restores the disrupted GFP. GFP fluorescence is therefore an indicator of TALEN activity. (b) The TALEN activity of 16 different TALENs, each obtained through three different synthesis route—direct synthesis polyclonal plasmid (polyclonal), direct synthesis clonal plasmid after colony picking (clonal), and 2-step synthesis polyclonal plasmid (transfer). GFP fluorescence has been normalized by the activity of I-CreI positive control. Reporter alone control and uninduced TALEN control showed no detectable activity.

Using the GFP TALEN reporter system, we compared the TALEN activities of 1-step assembly polyclonal plasmids, 1-step assembly monoclonal plasmids after colony picking, and 2-step assembly polyclonal plasmids after transfer. We randomly chose 16 different TALENs for this comparison, and there were no significant differences in activity among the three groups of TALENs (Figure 2b). This demonstrated that the polyclonal plasmids not only digested and sequenced like clonal plasmids but also functioned like clonal plasmids.

Overall, when all the TALENs were assessed, ~96% (87/90) of the TALENs tested showed nuclease activity. A TALEN was considered active if it showed at least half the activity of the I-CreI positive control. We sequenced the three inactive TALENs, and the results showed that all had been assembled correctly. The lack of activity therefore suggested a lack of binding rather than assembly error. The level of activity varied greatly among the TALENs, ranging from ~0.1 to ~20 normalized fluorescence. Under the TALEN scaffold used in our tests, the preceding T appears to be nonessential for TALEN activity. This property is likely an attribute of the scaffold as suggested in previous work[19] and may not be generalizable to all TALEN scaffolds currently used by the scientific community. TALES that target sequences with long stretches of A, T, and very few C appeared to function poorly, in agreement with previous observations.[21] However, having a high number of strong RVD alone is also not sufficient to guarantee high activity. There appears to be as yet undetermined factors that affect TALE binding efficiency. However, we have too few inactive TALEs among our targets to discern any rule, and no obvious trends are observed among the active TALEs. Activity data of the 90 TALENs can be found in SI Figure 6.

■ DISCUSSION

There are few performance indicators available to assess and compare current high-throughput TALE synthesis platforms. We characterized our platform using three parameters that we consider to be key for evaluating the performance of high-throughput synthesis systems: first, specific yield, which represents the efficiency of mass transfer during the assembly reaction and is estimated by actual plasmid yield divided by theoretical maximum yield. High specific yield will allow the use of a smaller amount of starting material and/or less efficient competent cells. Second, efficiency, which represents the effectiveness and accuracy of the assembly reaction, and is estimated by the fraction of correct clones out of all clones. High efficiency will allow the picking of fewer colonies to get a correct clone. Third, cost, which measures the monetary and time input necessary to produce expression-ready plasmids. Monetary cost should include all reagents and consumables. Time cost should include the entire assembly process from a desired sequence to a final expression-ready plasmid.

Analogous to the $1000 genome threshold in genome sequencing,[23] there are performance thresholds in high-throughput TALE synthesis that will allow substantial nonlinear throughput improvement. The most significant contributor to time and cost of a synthesis process is clonal isolation and the subsequent screening. Extremely high assembly efficiency can negate the need for clonal isolation and allow direct polyclonal

plasmid purification. There is no clear guideline for deciding what the minimum efficiency needs to be, but if the polyclonal products digest, sequence, and function just like their clonal counterparts, the efficiency is most likely sufficient. Among all synthesis platforms published thus far, only Schmid-Burgk et al. produced usable polyclonal plasmids at about 85% purity.[14] Limiting dilution can be used in place of plating to obtain a clonal culture, but it increases the total process cost due to duplicated culture and plasmid purification. Likewise, the efficiency of all other high-throughput synthesis systems, including Kim et al.'s trimer Golden Gate assembly platform,[16] is around 85%, and clonal isolation needs to be performed before transformation. In comparison, fairyTALE has an efficiency of ~98%, achieved by using a 13-fragment GG reaction, more than twice as many fragments as previous platforms. The ability to ligate more fragments reduces the size of the substrate library and increases the maximum length of the generated TALEs. Moreover, we have shown that, at ~98% efficiency, the direct polyclonal plasmids obtained using fairyTALE are experimentally equivalent to monoclonal plasmids in all measures.

Although clonal isolation and individual confirmation is still the preferred practice for low-throughput dedicated applications where a few TALEs are used to target a few targets, it is prohibitively expensive for high-throughput applications where thousands of TALEs are used to target thousands of targets. Furthermore, in applications such as genetic screening or engineering using a library of TALENs or TALE transcription factors, absolute purity becomes less important.

High efficiency bacterial transformation can potentially add significant cost to the process. Therefore, the ability to use low efficiency competent cells, which are easier and cheaper to produce and transform, is important. The typical transformation efficiency of homemade Z-competent Top10 is between 10^5 and $10^6/\mu g$ pUC19, and a specific yield threshold that will allow the use of this low-cost transformation method is around 0.5%. FairyTALE meets this threshold and can consistently produce hundreds of transformants in 13-insert GG reactions. Specific yield has been an under-appreciated performance parameter, and no synthesis method published so far has reported it.

The overall cost of synthesis is of paramount importance when it comes to high-throughput synthesis—an extra dollar spent per reaction can very quickly scale to tens of thousands of dollars. At every step of the synthesis process, conscious design considerations are required to minimize the cost. At \$5 per TALE and one-day synthesis speed, fairyTALE can potentially make TALE as accessible as oligonucleotides are today (the breakdown of cost estimate can be found in SI Figure 7). The typical workflow of a one-day synthesis scheme is shown in Figure 3.

The low cost high-throughput synthesis also makes possible a wide range of genome engineering applications that require a large number of TALEs. For example, as suggested by Kim et al.,[16] a collection of human knockout cell lines covering every nonessential gene is highly desirable. Without the need for clonal isolation, fairyTALE can very quickly generate all the necessary TALENs to carry out the knockout. In combination with the SunnyTALEN scaffold,[22] which has improved activity over existing TALENs, knockout efficiency can be increased, therefore reduces the amount of downstream work for cell line development. Knockout collection of many other cell types can be similarly developed.

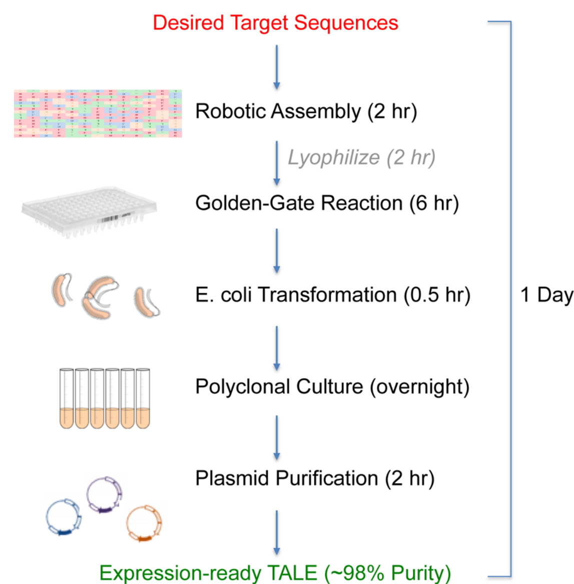


Figure 3. Typical synthesis timeline: fairyTALE can transform desired DNA binding sequences to expression-ready TALE plasmids in one day. Lyophilization is necessary in this case because large volume addition is required to achieve pipetting consistency in our robotic system.

One application where extremely high synthesis efficiency is critical is library synthesis; in which tens of thousands of TALEs are individually assembled and then combined into a library without getting individually verified. The library can be used for genome evolution and/or genetic screening to identify genes of interest associated with certain phenotypes. If all the steps after the initial GG assembly reaction, e.g. transformation and plasmid preparation, can be done as a library, the synthesis cost per TALE can be reduced to just \$2 per TALE. However, because no clonal isolation can be done in such a synthesis scheme, extremely high synthesis efficiency becomes a requirement.

Lastly, in Table 1, we attempt to compare and summarize all TALE synthesis platforms published to date. With superior efficiency and lower cost, fairyTALE has the potential to revolutionize the use of TALEs for genome scale applications.

All material used in this work, including the substrate library and the full set of receivers, will be made available on Addgene.

METHODS

Substrate Plasmid Construction. Each substrate plasmid can be assembled from four parts—front, RVD1, RVD2, back—with 5 junctions in total. The front and back are sequences that bridge adjacent direpeats; they, as well as their distal junctions, are constant within each group of substrate plasmids. We first constructed 16 plasmids, pFB1 through 13 and pFB5A through C, that contain the front and back of each group with an Esp3I-excisable cassette between them. The set of three proximal junctions around the RVDs are constant in all substrate plasmids, and a Type IIS ligation reaction is used to insert RVD1 and RVD2 between the front and back fragment to create direpeats. In a similar manner, monorepeats are generated by inserting RVDM, which has a single RVD with junctions compatible with the front and back fragments. All 308 substrate plasmids were generated using a robotic liquid handling system from 16 pFB plasmids, 4 RVD1 fragments, 4 RVD2 fragments, and 4 RVDM fragments. All RVD fragments

Table 1. Comparison of Existing TALE Synthesis Platform with a Timeline of Development[a]

	Zhang et al.[5]	Li et al.[25]	Cermak et al.[4]	Weber et al.[26]	Reyon et al.[12]	Briggs et al.[13]	Schmid-Burgk et al.[14]		Ding et al.[15]	Kim et al.[16]	this work
pub. date	1/19/11	3/31/11	4/14/11	5/19/11	4/8/12	6/26/12	12/16/12		2/7/13	2/17/13	
no. parts	72	32	50	72	376	12	64	3072	832	424	308
efficiency	NR	NR	NR	NR	NR	85%	85%	82%	NR	82%	>98%
time	NR	NR	5 days	NR	5 days	3 days	3 days	1 day	2 days	3 days	1 day
CRD length	12.5	16.5/24.5	11.5–30.5	17.5	2.5–20.5	2.5–21.5	18.5	15.5	15.5	14.5–18.5	13.5–30.5
type	2-step GG	2-step GG	2-step GG	2-step GG	solid phase	solid phase	2-step LIC	1-step LIC	1-step GG	1-step GG	1-step GG

[a]Specific yield and material cost are not tabulated because no other work has reported them. Synthesis platforms from the time of Reyon et al. onwards are generally considered high-throughput. NR: Not Reported.

are short and can be generated by direct synthesis, thus allowing easy future upgrade or expansion. More details can be found in SI Figure 8.

Receiver Plasmid Construction. A Type IIS ligation strategy was used to construct pLRX-ccdB, a precursor of the pLR series of intermediate receivers, from three parts—two *ccdB* fragments and a LRX fragment. The *ccdB* cassette without chloramphenicol resistance was PCR amplified in two fragments from a Gateway destination vector (Life Technologies, Carlsbad, CA), removing the internal *Bsa*I site within *ccdB*. The LRX fragment, which contains a last repeat precursor, was generated by direct synthesis (Integrated DNA Technologies, Coralville, IA). The series of 4 pLR receivers were generated from pLRX-ccdB by *Xcm*I linearization followed by In-Fusion (Clontech, Mountain View, CA) with the appropriate RVD-containing PCR fragments. pYW-TN and pYW-LR-TN contain a Trp1 marker for yeast selection, a Gal2 promoter for inducible expression, a +207 aa N-terminus extension with NLS added for improved activity in yeast and FLAG tag added for easy expression analysis, an engineered +63 aa C-terminus extension and FokI cleavage domain for improved cleavage activity in yeast as well as in human. In pYW-TN, the *ccdB* cassette is flanked by *Esp*3I and is designed to receive a CRD assembly from pLR. In pYW-LR-TN, the *ccdB* cassette is flanked by *Bsa*I and is generated by a cassette exchange between pYW-TN and pLR using *Esp*3I. pYW-LR-TN will allow direct assembly of substrate plasmids. All *ccdB*-containing plasmids were propagated in *E. coli* DB3.1. More details can be found in SI Figure 9.

13- and 15-Piece Golden Gate Assembly. All substrate plasmids and receiver plasmids were normalized to equal-concentration, diluted with water, and stored in 96-well storage plates. Using a liquid handling robot (CyBio AG, Germany), interfaced with a custom command script, the appropriate substrate and receiver plasmids were pipetted into a 96-well PCR plate. Due to the pipetting precision of our robotic system, 5 μL of each plasmid were added for reproducible transfer, thus resulting in a final volume of >65 μL. The plasmid mixtures were first lyophilized before the Golden Gate reaction. Twenty-five nanograms of each substrate plasmid, 25 ng pLR or 50 ng pYW, were used in a 10 μL Golden Gate reaction −0.5 μL *Bsa*I-HF, 0.5 μL T4 ligase (New England Biolabs, Ipswich, MA) in 1× T4 ligase buffer. The reaction was carried out in a thermocycler using the following program: 37 °C for 5 min, 37 °C for 5 min, 16 °C for 10 min, repeat 2–3 for 20 cycles, 16 °C for 20 min, 37 °C for 30 min, 75 °C for 6 min, 4 °C hold. After the thermocycling, 0.25 μL of plasmid-safe nuclease (Epicenter), and 0.5 μL of 25 mM ATP was added to the reaction and incubated further at 37 °C for 15 min. Four microliters of the resultant reaction were transformed into either Z-competent (Zymo Research, Irvine, CA) Top10 (Life Technologies) or Stellar (Clontech). Competent cells with greater than 5 × 10^5/μg pUC19 transformation efficiency would suffice. The Golden Gate reaction condition reported here has been optimized for high specific yield and high specificity. More information on the reaction optimization can be found in SI Figure 10.

Golden Gate Transfer Reaction. 50 ng of pLR assembly and 50 ng of pYW-TN were used in a 10 μL Golden Gate transfer reaction −0.5 μL *Esp*3I (Thermo Scientific, Waltham, MA) and 0.5 μL T4 ligase in 1× T4 ligase buffer. The reaction is carried out in a thermocycler using a similar program as above but with only 10 cycles. Four microliters of the resultant reaction were transformed into Z-competent Stellar.

Assembly Efficiency Determination. The initial *E. coli* transformations were each split into two parts: 1/3 of it was plated onto LB+antibiotic plate for colony picking and monoclonal DNA isolation; 2/3 of it was directly inoculated into 5 mL LB+antibiotic liquid culture for polyclonal DNA isolation. The plasmids isolated from the monoclonal DNA preparations and that from polyclonal DNA preparations were analyzed by agarose electrophoresis. A more detailed procedure can be found in SI Figure 11.

GFP TALEN Reporter Assay. The reporter assay was carried out in YZE-19, an estrogen-inducible yeast strain,[24] using a 3-plasmid system with a disrupted GFP TALEN reporter, a Xa10 TALEN,[19] and a fairyTALE-generated TALEN. The disrupted GFP reporter precursor, 416K2-GFP-GG, was constructed by PCR and regular ligation. The disruption cassette has a 5′-Xa10 binding site and a *Bsa*I-flanking 3′-ICreI binding site. 90 TALEN reporters were generated by exchanging the 3′-ICreI site with 90 chemically synthesized TALEN binding sites. The sequences of the target sites can be found in SI Figure 12. The Xa10 plasmid was first transformed into YZE-19; subsequently, a fairyTALE TALEN plasmid and its corresponding GFP TALEN reporter plasmid were cotransformed into the same strain. Transformants were allowed to expand in SC-LWU (glucose) media for 2 days with daily one-tenth passage. TALEN expression was induced on the third day using 0.1 μM 17β-estradiol, and GFP expression was assayed using flow cytometry two days later.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional figures as described in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 217-333-2631. Fax: 217-333-5052. E-mail: zhao5@illinois.edu.

Author Contributions

J.L. and H.Z. designed the system; J.L., R.C., Z.A., and Z.B. constructed the parts for the system; R.C. wrote the script to interface with the liquid handling robot; J.L., R.C., and Z.A. performed the assemblies and the experiments; J.L. and H.Z. wrote the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Department of Chemical and Biomolecular Engineering at the University of Illinois at Urbana-Champaign. J.L. acknowledges fellowship support from the Agency for Science, Technology, and Research of Singapore. The authors acknowledge Stella Xinzi Wu for her help with managing the substrate plasmid library.

REFERENCES

(1) Hockemeyer, D., Wang, H., Kiani, S., Lai, C. S., Gao, Q., Cassady, J. P., Cost, G. J., Zhang, L., Santiago, Y., Miller, J. C., Zeitler, B., Cherone, J. M., Meng, X., Hinkley, S. J., Rebar, E. J., Gregory, P. D., Urnov, F. D., and Jaenisch, R. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol. 29*, 731−734.

(2) Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D., and Rebar, E. J. (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol. 29*, 143−148.

(3) 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. U.S.A. 108*, 2623−2628.

(4) 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.

(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) 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.

(7) 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.

(8) 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.

(9) Moscou, M. J., and Bogdanove, A. J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science 326*, 1501.

(10) 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.

(11) Sun, N., Abil, Z., and Zhao, H. (2012) Recent advances in targeted genome engineering in mammalian systems. *Biotechnol. J. 7*, 1074−1087.

(12) 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.

(13) 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.

(14) 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.

(15) Ding, Q., Lee, Y.-K., Schaefer, E. A. K., Peters, D. T., Veres, A., Kim, K., Kuperwasser, N., Motola, D. L., Meissner, T. B., Hendriks, W. T., Trevisan, M., Gupta, R. M., Moisan, A., Banks, E., Friesen, M., Schinzel, R. T., Xia, F., Tang, A., Xia, Y., Figueroa, E., Wann, A., Ahfeldt, T., Daheron, L., Zhang, F., Rubin, L. L., Peng, L. F., Chung, R. T., Musunuru, K., and Cowan, C. A. (2013) A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell 12*, 238−251.

(16) Kim, Y., Kweon, J., Kim, A., Chon, J. K., Yoo, J. Y., Kim, H. J., Kim, S., Lee, C., Jeong, E., Chung, E., Kim, D., Lee, M. S., Go, E. M., Song, H. J., Kim, H., Cho, N., Bang, D., Kim, S., and Kim, J.-S. (2013) A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.*, 251−258.

(17) Engler, C., Kandzia, R., and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One 3*, e3647.

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

(19) 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.

(20) Meckler, J. F., Bhakta, M. S., Kim, M.-S., Ovadia, R., Habrian, C. H., Zykovich, A., Yu, A., Lockwood, S. H., Morbitzer, R., Elsäesser, J., Lahaye, T., Segal, D. J., and Baldwin, E. P. (2013) Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res. 41*, 4118−4128.

(21) Streubel, J., Blücher, C., Landgraf, A., and Boch, J. (2012) TAL effector RVD specificities and efficiencies. *Nat. Biotechnol. 30*, 593−595.

(22) Sun, N., Bao, Z., Xiong, X., and Zhao, H. (2013) SunnyTALEN: A second-generation TALEN system for human genome editing. *Biotechnol. Bioeng.*, DOI: 10.1002/bit.25154.

(23) Service, R. F. (2006) Gene sequencing. The race for the $1000 genome. *Science*, 1544.

(24) Liang, J., Ning, J. C., and Zhao, H. (2013) Coordinated induction of multi-gene pathways in *Saccharomyces cerevisiae*. *Nucleic Acids Res. 41*, e54.

(25) 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.

(26) Weber, E., Gruetzner, R., Werner, S., Engler, C., and Marillonnet, S. (2011) Assembly of designer TAL effectors by golden gate cloning. *PLoS One 6*, e19722.