Supplementary Information

Programmable Sequence-Specific Transcriptional Regulation of Mammalian Genome Using Designer TAL Effectors

Feng Zhang[1,2,3,5,7*,±], Le Cong[2,3,4*], Simona Lodato[5,6], Sriram Kosuri[2,3], George Church[2,3,#,±], and Paola Arlotta[5,6,#,±]

[1] Society of Fellows
[2] Department of Genetics, Harvard Medical School
[3] Wyss Institute for Biologically Inspired Engineering
[4] Program in Biological and Biomedical Sciences, Harvard Medical School
[5] Department of Stem Cell and Regenerative Biology Harvard University
[6] Center for Regenerative Medicine and Department of Neurosurgery, Massachusetts General Hospital, Harvard Medical School
[7] Present Address: Broad Institute of MIT and Harvard, Cambridge, MA
McGovern Institute for Brain Research, MIT, Cambridge, MA
Department of Brain and Cognitive Sciences, MIT Cambridge, MA

* These authors contributed equally to this work.

\# These authors contributed equally to this work.

± To whom correspondence should be addressed.
E-mail: zhang_f@mit.edu, gmc@harvard.edu, paola_arlotta@hms.harvard.edu

Supplementary Information

Methods

Design and Synthesis Strategy for Designer TALEs:

Designer TALEs with customized DNA binding domains were constructed using hierarchical ligation as described below.

Step 1: Optimization of DNA sequence for each repeat monomer to minimize repetitiveness of the final product.

Repetitive DNA sequences are difficult to manipulate for a number of reasons, including susceptibility to recombination and difficulty for PCR amplification. To reduce the repetitiveness of designer TALEs, we first optimized the DNA sequence of the four monomers (NI, HD, NN, NG) to minimize homology while preserving the amino acid sequence. We used the 34aa repeat monomer from the *Xanthomonas sp*. hax3 gene and generated 4 monomers (**Supplementary Table 1**). The new monomers were synthesized (DNA2.0, Menlo Park, CA) and cloned into individual plasmids to be used as amplification templates.

Step 2: Design of a ligation strategy that utilizes orthogonal sticky ends to specify the position of each monomer in the ligated tandem repeat.

In order to assemble the individual monomers in a specific order, we altered the DNA sequence at the junction between each pair of monomers, similar to the Golden Gate cloning method[1, 2]. The junction is Gly-Leu and has a total of 24 possible codon pairs (4 codons for Gly and 6 codons for Leu) with 4 variable bases. We initially chose 12 ligation linkers and found that the ability to ligate 12 pieces together into the specified order was very inefficient. Therefore, we tested multi-piece ligation reactions containing 2 to 12 pieces and found that 4-piece ligation was most efficient. This led us to revise our assembly strategy using hierarchical ligation where three 4-mer tandem repeats were assembled first, and the pre-assembled 4-mer tandem repeats were subsequently ligated to form the final 12-mer tandem repeats. To minimize the formation of incorrectly ligated products, we tested a series of ligases and found that T7 ligase gave the highest efficiency for multi-piece ligation. For specifying the order of the 4 monomers, we used the following linkers to construct each 4-mer tandem repeat:

```
AA :                  G   L          G   L          G   L
DNA:   5′ ............GGA CTC........GGC CTC........GGA TTA............ 3′
       3′ ............CCT GAG........CCG GAG........CCT AAT............ 5′
          |---repeat 1---|---repeat 2---|---repeat 3---|---repeat 4---|
```

The 4-mers were used as assembly blocks for constructing the full 12-mer repeat. The completely assembled 12-mer was cloned into the appropriate destination plasmid containing the N- and C-termini of hax3 as well as one 0.5 repeat. The ligated tandem repeat along with the N- and C-termini form a fully functional designer TALE. The sequences for the destination plasmids are listed in **Supplementary Sequences**.

Step 3: Construction of designer TALEs

We used PCR amplification to generate a set of monomers (HD, NI, NN, NG) for each position. We designed a set of 24 primers (**Supplementary Table 2**) to attach the right ligation junction

onto each monomer. Using type IIs enzymes (e.g. *Bsm*BI and BsaI) we can process the ends of each repeat monomer to expose the sticky-end ligation junctions. An example is illustrated below:

```
AA:                  L  T  P  E  Q  V  V  .....  C  Q  A  H  G  L
DNA:  5′ cgtctcGACTCACCCCAGAGCAGGTCGTG.......TGCCAAGCGCACGGCCTCAgagacc
      3′ gcagagCTGAGTGGGGTCTCGTCCAGCAC.......ACGGTTCGCGTGCCGGAGTctctgg
          BsmBI Site                                        BsaI Site

                                    |
                                    |   Digest with BsaI
                                   V

AA:                     L  T  P  E  Q  V  V  .....  C  Q  A  H  G  L
DNA:              5′ ACTCACCCCAGAGCAGGTCGTG.......TGCCAAGCGCACGG
                  3′     TGGGGTCTCGTCCAGCAC.......ACGGTTCGCGTGCCGGAG
```

Step 4: Ligate monomers into specific tandem repeats (Fig. 1b). We constructed 12mer tandem repeats in two steps. 4mer tandem repeats were first assembled in 10ul ligation reactions consisting of 25ng for each monomer. We specifically chose the T7 DNA ligase based on its 1000x higher activity on sticky ends than blunt ends. The correct size ligation product (~440bp) for each 4mer tandem repeat is then purified and PCR amplified. The 4mer PCR products were then processed with the appropriate type IIs enzyme and then ligated again to form 12mer tandem repeats. The correctly ligated 12mer product is then PCR amplified again and processed with Type IIs enzyme for ligation into the backbone plasmid. The final assembled dTALE is verified by sequencing.

* * * * *

The original set of primers (**Supplementary Table 2**) was designed for optimizing the dTALE assembly procedure. Each monomer had two different restriction sites flanking the 5' and 3' ends. We tested a number of ligation conditions by varying the number of pieces being ligated simultaneously and found that 4-piece ligation works most efficiently. To avoid the need for double digests using two different restriction enzymes at different temperatures, we revised the original dTALE construction protocol as well as primer design (**Supplementary Table 3**) to streamline the assembly process. A step-by-step protocol for the simplified dTALE construction method is presented below:

1. A library consisting of 48 monomers (4 monomers for each position in the final assembled 12-mer tandem repeat) is generated using PCR. Plasmids containing each type of monomer repeat (monomer sequence listed in **Supplementary Table 1**) are used as the template for amplification. PCR reactions are set up for each monomer, according to the following table of template and primer pairing (primers shown in **Supplementary Table 3**).

Primer: / Template:												
	F1/R1 NI	F2/R2 NI	F3/R3 NI	F4/R4 NI	F5/R5 NI	F6/R6 NI	F7/R7 NI	F8/R8 NI	F9/R9 NI	F10/R10 NI	F11/R11 NI	F12/R12 NI
	F1/R1 HD	F2/R2 HD	F3/R3 HD	F4/R4 HD	F5/R5 HD	F6/R6 HD	F7/R7 HD	F8/R8 HD	F9/R9 HD	F10/R10 HD	F11/R11 HD	F12/R12 HD
	F1/R1 NG	F2/R2 NG	F3/R3 NG	F4/R4 NG	F5/R5 NG	F6/R6 NG	F7/R7 NG	F8/R8 NG	F9/R9 NG	F10/R10 NG	F11/R11 NG	F12/R12 NG
	F1/R1 NN	F2/R2 NN	F3/R3 NN	F4/R4 NG	F5/R5 NN	F6/R6 NN	F7/R7 NN	F8/R8 NN	F9/R9 NN	F10/R10 NN	F11/R11 NN	F12/R12 NN

2. For monomer PCR, high-fidelity polymerase (e.g. Herculase II) (Stratagene) is used to minimize mutation and achieve the highest product yield. Monomers are amplified in 100μl PCR reactions following appropriate protocols of polymerase manufacturers.

3. After completion of the PCR reaction, each monomer is purified using the 96 QIAquick PCR Purification Kit (Qiagen) and the product eluted in 70μl of ddH_2O.

4. Each monomer is digested using *Bsa*I (New England BioLabs) at 37°C for 1 hour in a 100μl reactions as follows:

70μl	purified PCR Product
5μl	BsaI (50 units)
5μl	10X Buffer #4
1μl	100X BSA
19μl	ddH_2O

5. After digestion, digested monomers are purified using the 96 QIAquick PCR Purification Kit (Qiagen) and eluted in 70μl of ddH_2O.

6. The concentration of each monomer is adjusted to 25ng/μl for monomers 1, 4, 5, 8, 9, 12; and 20ng/μl for monomers 2,3,6,7,10,11.

7. For each dTALE to be assembled, individual 4-mer tandem repeats are first constructed by simultaneously ligating four repeat monomers together at equal molar ratio (25ng for monomers 1 and 4, 20ng for monomers 2 and 3 in 10ul total ligation mix, using 300units of T7 ligase from Enzymatics and 10X ligation buffer). The ligation is incubated at room temperature for 30 minutes.

8. 5μl of the 4-mer ligation reactions are run on a 2% E-Gel EX (Invitrogen) and the correct size products are amplified by gel-stab PCR[3]. Specifically, a 10 μL pipette tip is used to puncture the gel at the location of the desired product. The stab is mixed up and down in 10 μL of water, and the water is heated to 65 °C for 2 min. 2.5 μL of the gel-isolated product diluted in water is then amplified in a 50 μL PCR reaction using Herculase II polymerase (Stratagene).

9. The amplified 4-mer tandem repeats are purified using the QIAquick PCR Purification kit and eluted in 40μl of ddH_2O.

10. Purified 4-mer tandem repeats as well as the appropriate dTALE backbone are digested using *Bsm*BI at 55°C for 1 hour.

40ul	purified PCR product		500ng	dTALE backbone vector
5ul	10X Buffer #3	and	5ul	10X Buffer #3
5ul	BsmBI (50 units)		5ul	*Bsm*BI (50 units)
				(bring volume to 50ul with ddH_2O)

11. Purify digested 4-mer tandem repeats using QIAquick PCR Purification Kit (Qiagen). Gel purify the digested backbone.

12. Fully assembled dTALEs are generated by simultaneously ligating the three 4-mer tandem repeats with the backbone vector at equal molar ratio (1ng for each 4-mer tandem repeat and 28ng for backbone vector; in a 10ul ligation reaction using 1500U T7 ligase from Enzymatics). A negative control reaction should be set up with 28ng of the backbone vector alone. All ligation reactions are incubated in a thermal cycler using the following parameters: 37^{o}C for 1 min followed by 25^{o}C for 5min, for 30 cycles.

13. 2ul of each dTALE ligation reaction is transformed into XL-10 Gold chemically competent cells (Stratagene).

14. Plasmid DNA for assembled dTALE are prepared from ligation transformants and analyzed via restriction digest and DNA sequencing. Transformation of the negative control ligation should not yield any transformants.

Supplementary Table 1 | List of repeat monomer sequences. Forward and reverse priming sites are highlighted in blue and red respectively.

NI

```
 L  T  P  E  Q  V  V  A  I  A  S  N  I  G  G  K  Q  A  L
CTGACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGGGAAACAGGCACTC
 E  T  V  Q  R  L  L  P  V  L  C  Q  A  H  G
GAGACTGTCCAGCGCCTGCTTCCCGTGCTGTGCCAAGCGCACGGA
```

HD

```
 L  T  P  E  Q  V  V  A  I  A  S  H  D  G  G  K  Q  A  L
TTGACCCCAGAGCAGGTCGTGGCGATCGCAAGCCACGACGGAGGAAAGCAAGCCTTG
 E  T  V  Q  R  L  L  P  V  L  C  Q  A  H  G
GAAACAGTACAGAGGCTGTTGCCTGTGCTGTGCCAAGCGCACGGG
```

NN

```
 L  T  P  E  Q  V  V  A  I  A  S  N  N  G  G  K  Q  A  L
CTTACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGGAAAACAGGCTTTG
 E  T  V  Q  R  L  L  P  V  L  C  Q  A  H  G
GAAACGGTGCAGAGGCTCCTTCCAGTGCTGTGCCAAGCGCACGGG
```

NG

```
 L  T  P  E  Q  V  V  A  I  A  S  N  G  G  G  K  Q  A  L
CTGACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTG
 E  T  V  Q  R  L  L  P  V  L  C  Q  A  H  G
GAAACCGTACAACGATTGCTGCCGGTGCTGTGCCAAGCGCACGGC
```

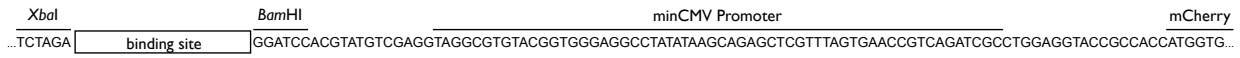
Supplementary Table 2 | Primers used for the amplification and assembly of artificial TALEs reported in this manuscript. Unique linkers used to specify the ligation ordering are highlighted in blue; sequences written from 5' to 3'. *Bsm*BI sites are highlighted in yellow and *Bsa*I sites are highlighted in gray.

F1	AGATGCCGTCCTAGCGcgtctcCTGACCCCAGAGCAGGTCGTGG
F2	AGATGCCGTCCTAGCGcgtctcG**ACTC**ACCCCAGAGCAGGTCGTG
F3	AGATGCCGTCCTAGCGcgtctcG**CCTC**ACCCCAGAGCAGGTCGTG
F4	AGATGCCGTCCTAGCGcgtctcG**ATTA**ACCCCAGAGCAGGTCGTG
F5	AGATGCCGTCCTAGCGcgtctcG**CTTA**ACCCCAGAGCAGGTCGTG
F6	AGATGCCGTCCTAGCGcgtctcG**ACTT**ACCCCAGAGCAGGTCGTG
F7	AGATGCCGTCCTAGCGcgtctcG**CCTT**ACCCCAGAGCAGGTCGTG
F8	AGATGCCGTCCTAGCGcgtctcG**ACTA**ACCCCAGAGCAGGTCGTG
F9	AGATGCCGTCCTAGCGcgtctcG**GCTC**ACCCCAGAGCAGGTCGTG
F10	AGATGCCGTCCTAGCGcgtctcG**GCTA**ACCCCAGAGCAGGTCGTG
F11	AGATGCCGTCCTAGCGcgtctcG**CCTA**ACCCCAGAGCAGGTCGTG
F12	AGATGCCGTCCTAGCGcgtctcG**GTTA**ACCCCAGAGCAGGTCGTG
R1	GTATCTTTCCTGTGCCCAggtctcT**GAGT**CCGTGCGCTTGGCAC
R2	GTATCTTTCCTGTGCCCAggtctcT**GAGG**CCGTGCGCTTGGCAC
R3	GTATCTTTCCTGTGCCCAggtctcT**TAAT**CCGTGCGCTTGGCAC
R4	GTATCTTTCCTGTGCCCAggtctcT**TAAG**CCGTGCGCTTGGCAC
R5	GTATCTTTCCTGTGCCCAggtctcT**AAGT**CCGTGCGCTTGGCAC
R6	GTATCTTTCCTGTGCCCAggtctcT**AAGG**CCGTGCGCTTGGCAC
R7	GTATCTTTCCTGTGCCCAggtctcT**TAGT**CCGTGCGCTTGGCAC
R8	GTATCTTTCCTGTGCCCAggtctcT**GAGC**CCGTGCGCTTGGCAC
R9	GTATCTTTCCTGTGCCCAggtctcT**TAGC**CCGTGCGCTTGGCAC
R10	GTATCTTTCCTGTGCCCAggtctcT**TAGG**CCGTGCGCTTGGCAC
R11	GTATCTTTCCTGTGCCCAggtctcT**TAAC**CCGTGCGCTTGGCAC
R12	GTATCTTTCCTGTGCCCAggtctcT**TAAA**CCGTGCGCTTGGCAC
F-assem	ATATAGATGCCGTCCTAGCGC
R-assem	AAGTATCTTTCCTGTGCCCAG

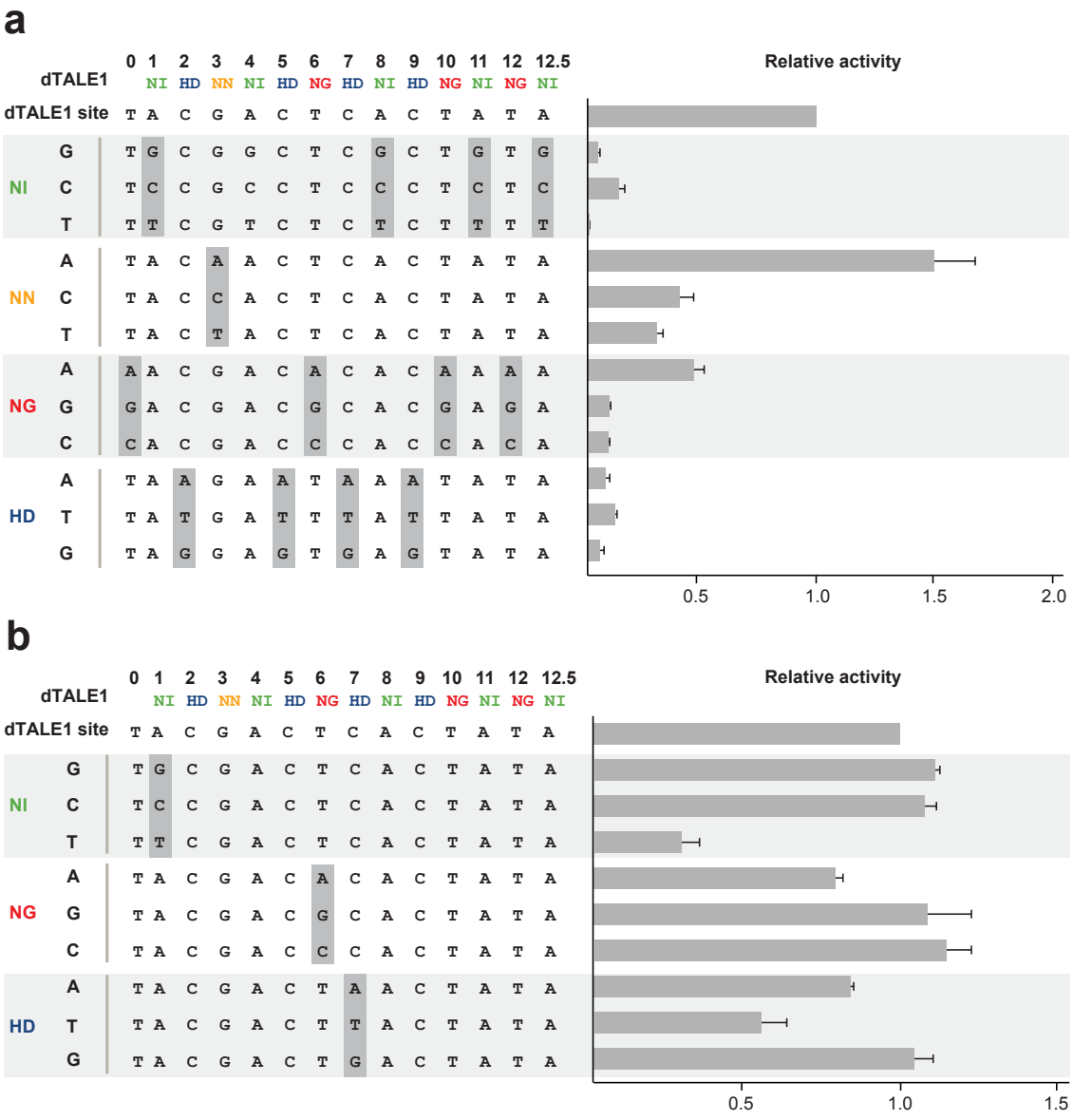
Supplementary Table 3 | Primers used in the simplified step-by-step dTALE construction method. Unique linkers used to specify the ligation ordering are highlighted in blue; sequences written from 5' to 3'. *Bsm*BI sites are highlighted in yellow and *Bsa*I sites are highlighted in gray.

F1	ATATAGATGCCGTCCTAGCGcgtctcCTGACCCCAGAGCAGGTCGTGG
F2	TGCTCTTTATTCGTTGCGTCggtctcGACTCACCCCAGAGCAGGTCGTG
F3	TGCTCTTTATTCGTTGCGTCggtctcGCCTCACCCCAGAGCAGGTCGTG
F4	TGCTCTTTATTCGTTGCGTCggtctcGATTAACCCCAGAGCAGGTCGTG
F5	ATATAGATGCCGTCCTAGCGcgtctcGCTTAACCCCAGAGCAGGTCGTG
F6	TGCTCTTTATTCGTTGCGTCggtctcGACTCACCCCAGAGCAGGTCGTG
F7	TGCTCTTTATTCGTTGCGTCggtctcGCCTCACCCCAGAGCAGGTCGTG
F8	TGCTCTTTATTCGTTGCGTCggtctcGATTAACCCCAGAGCAGGTCGTG
F9	ATATAGATGCCGTCCTAGCGcgtctcGGCTCACCCCAGAGCAGGTCGTG
F10	TGCTCTTTATTCGTTGCGTCggtctcGACTCACCCCAGAGCAGGTCGTG
F11	TGCTCTTTATTCGTTGCGTCggtctcGCCTCACCCCAGAGCAGGTCGTG
F12	TGCTCTTTATTCGTTGCGTCggtctcGATTAACCCCAGAGCAGGTCGTG
R1	TCTTATCGGTGCTTCGTTCTggtctcTGAGTCCGTGCGCTTGGCAC
R2	TCTTATCGGTGCTTCGTTCTggtctcTGAGGCCGTGCGCTTGGCAC
R3	TCTTATCGGTGCTTCGTTCTggtctcTTAATCCGTGCGCTTGGCAC
R4	AAGTATCTTTCCTGTGCCCAcgtctcTTAAGCCGTGCGCTTGGCAC
R5	TCTTATCGGTGCTTCGTTCTggtctcTGAGTCCGTGCGCTTGGCAC
R6	TCTTATCGGTGCTTCGTTCTggtctcTGAGGCCGTGCGCTTGGCAC
R7	TCTTATCGGTGCTTCGTTCTggtctcTTAATCCGTGCGCTTGGCAC
R8	AAGTATCTTTCCTGTGCCCAcgtctcTGAGCCCGTGCGCTTGGCAC
R9	TCTTATCGGTGCTTCGTTCTggtctcTGAGTCCGTGCGCTTGGCAC
R10	TCTTATCGGTGCTTCGTTCTggtctcTGAGGCCGTGCGCTTGGCAC
R11	TCTTATCGGTGCTTCGTTCTggtctcTTAATCCGTGCGCTTGGCAC
R12	AAGTATCTTTCCTGTGCCCAcgtctcTGAGTCCGTGCGCTTGGCAC
F-assem	ATATAGATGCCGTCCTAGCG
R-assem	AAGTATCTTTCCTGTGCCCA

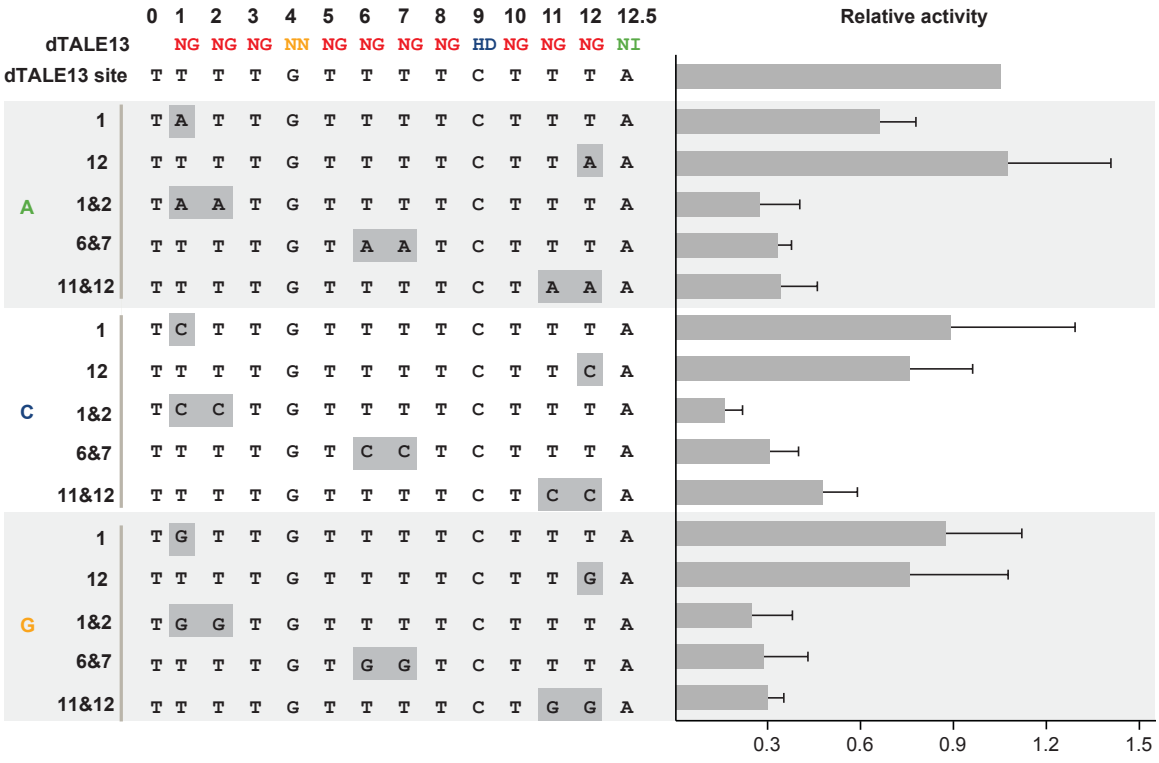
Supplementary Figure 1 | Design of mCherry reporter plasmid. Target binding site of a dTALE was cloned into the mCherry reporter plasmid between the *Xba*I and *Bam*HI restriction sites. Hence, the dTALE binding site is placed -96bp upstream of the transcription start site of a full-length mCherry gene, with a minimal CMV promoter in the middle.



Supplementary Figure 2 | Test of the DNA binding specificity of dTALE using reporters with varying numbers of mismatches. a, Each diresidue in dTALE1 was tested against its non-preferred DNA bases to determine the binding specificity for each diresidue. **b,** Single base pair mismatches were used to test the binding specificity of dTALE1. The relative activity of dTALE1 for each mutant reporter compared to the intended reporter is shown on the right. All error bars indicate s.e.m, n=3. The fold induction was determined via flow cytometry analysis of mCherry expression in transfected 293FT cells, and calculated as the ratio of the total mCherry fluorescence intensity of cells transfected with and without the specified dTALE.



Supplementary Figure 3 | Tests of the DNA binding specificity of dTALE to mismatched sequences. DNA binding specificity of dTALE was tested using dTALE13 and a series of reporters bearing systematically designed mutations in the binding site for dTALE13. The design of the reporter series is shown on the left, with three different groups: A, C, and G. Within each group, the base T at one or two positions of the dTALE-binding site (designated by numbers to the left of each target sequence) were altered to A, C, or G (the mutated bases are highlighted). The relative activity of dTALE13 for each mutant reporter compared to the original reporter is shown on the right. All error bars indicate s.e.m, n=3. The fold induction was determined via flow cytometry analysis of mCherry expression in transfected 293FT cells, and calculated as the ratio of the total mCherry fluorescence intensity of cells transfected with and without the specified dTALE.



Supplementary Sequences

Type IIs sites are colored in blue. NLS is colored in red. 2A-GFP is colored in green. The variable diresidue is highlighted yellow.

>dTALE-Backbone(NI in 0.5 repeat)-NLS-VP64-2A-EGFP

ATGTCGCGGACCCGGCTCCCTTCCCCACCCGCACCCAGCCCAGCGTTTTCGGCCGACTCGTTCTCAGACCTGCTTAGGCAGTT
CGACCCCTCACTGTTTAACACATCGTTGTTCGACTCCCTTCCTCCGTTTGGGGCGCACCATACGGAGGCGGCCACCGGGGAGT
GGGATGAGGTGCAGTCGGGATTGAGAGCTGCGGATGCACCACCCCCAACCATGCGGGTGGCCGTCACCGCTGCCCGACCGC
CGAGGGCGAAGCCCGCACCAAGGCGGAGGGCAGCGCAACCGTCCGACGCAAGCCCCGCAGCGCAAGTAGATTTGAGAACTT
TGGGATATTCACAGCAGCAGCAGGAAAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGGG
TCATGGGTTTACACATGCCCACATCGTAGCCTTGTCGCAGCACCCTGCAGCCCTTGGCACGGTCGCCGTCAAGTACCAGGACA
TGATTGCGGCGTTGCCGGAAGCCACACATGAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCGAGCGCTTGAGG
CCCTGTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGCTGGACACGGGCCAGTTGCTGAAGATCGCGAAGCGGG
GAGGAGTCACGGCGGTCGAGGCGGTGCACGCGTGGCGCAATGCGCTCACGGGAGCACCCCTCAACCTGACCgagacgGTACAT
GAAACGCATGGCACGGcgtctcAACTCACGCCTGAGCAGGTAGTGGCTATTGCATCCAATATCGGGGGCAGACCCGCACTGGAG
TCAATCGTGGCCCAGCTTTCGAGGCCGGACCCCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTGCCTCG
GCGGACGACCCGCCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTGCATTGATTAAGCGGACCAACAGAAGGATTCC
CGAGAGGACATCACATCGAGTGGCAGATCACGCGCAAGTGGTCCGCGTGCTCGGATTCTTCCAGTGTCACTCCCACCCCGCA
CAAGCGTTCGATGACGCCATGACTCAATTTGGTATGTCGAGACACGGACTGCTGCAGCTCTTTCGTAGAGTCGGTGTCACAGA
ACTCGAGGCCCGCTCGGGCACACTGCCTCCCGCCTCCCAGCGGTGGGACAGGATTCTCCAAGCGAGCGGTATGAAACGCGC
GAAGCCTTCACCTACGTCAACTCAGACACCTGACCAGGCGAGCCTTCATGCGTTCGCAGACTCGCTGGAGAGGGATTTGGACG
CGCCCTCGCCCATGCATGAAGGGGACCAAACTCGCGCGTCAgctagccccaagaagaagagaaaggtggaggccagcggttccGGACGGGCTG
ACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCC
TTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATTAACtctagaggcagtggagag
ggcagaggaagtctgctaacatgcggtgacgtcgaggagaatcctggcccaGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTG
GTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG
ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT
GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAC
CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATAT
CATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC
GACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCC
TGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGTAA

>dTALE-Backbone(NG in 0.5 repeat)-NLS-VP64-2A-EGFP

ATGTCGCGGACCCGGCTCCCTTCCCCACCCGCACCCAGCCCAGCGTTTTCGGCCGACTCGTTCTCAGACCTGCTTAGGCAGTT
CGACCCCTCACTGTTTAACACATCGTTGTTCGACTCCCTTCCTCCGTTTGGGGCGCACCATACGGAGGCGGCCACCGGGGAGT
GGGATGAGGTGCAGTCGGGATTGAGAGCTGCGGATGCACCACCCCCAACCATGCGGGTGGCCGTCACCGCTGCCCGACCGC
CGAGGGCGAAGCCCGCACCAAGGCGGAGGGCAGCGCAACCGTCCGACGCAAGCCCCGCAGCGCAAGTAGATTTGAGAACTT
TGGGATATTCACAGCAGCAGCAGGAAAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGGG
TCATGGGTTTACACATGCCCACATCGTAGCCTTGTCGCAGCACCCTGCAGCCCTTGGCACGGTCGCCGTCAAGTACCAGGACA
TGATTGCGGCGTTGCCGGAAGCCACACATGAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCGAGCGCTTGAGG
CCCTGTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGCTGGACACGGGCCAGTTGCTGAAGATCGCGAAGCGGG
GAGGAGTCACGGCGGTCGAGGCGGTGCACGCGTGGCGCAATGCGCTCACGGGAGCACCCCTCAACCTGACCgagacgGTACAT
GAAACGCATGGCACGGcgtctcAACTCACGCCTGAGCAGGTAGTGGCTATTGCATCCAATGGCGGGGGCAGACCCGCACTGGAG
TCAATCGTGGCCCAGCTTTCGAGGCCGGACCCCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTGCCTCG
GCGGACGACCCGCCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTGCATTGATTAAGCGGACCAACAGAAGGATTCC
CGAGAGGACATCACATCGAGTGGCAGATCACGCGCAAGTGGTCCGCGTGCTCGGATTCTTCCAGTGTCACTCCCACCCCGCA
CAAGCGTTCGATGACGCCATGACTCAATTTGGTATGTCGAGACACGGACTGCTGCAGCTCTTTCGTAGAGTCGGTGTCACAGA
ACTCGAGGCCCGCTCGGGCACACTGCCTCCCGCCTCCCAGCGGTGGGACAGGATTCTCCAAGCGAGCGGTATGAAACGCGC
GAAGCCTTCACCTACGTCAACTCAGACACCTGACCAGGCGAGCCTTCATGCGTTCGCAGACTCGCTGGAGAGGGATTTGGACG
CGCCCTCGCCCATGCATGAAGGGGACCAAACTCGCGCGTCAgctagccccaagaagaagagaaaggtggaggccagcggttccGGACGGGCTG
ACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCC
TTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATTAACtctagaggcagtggagag
ggcagaggaagtctgctaacatgcggtgacgtcgaggagaatcctggcccaGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTG
GTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG
ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT
GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAC
CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATAT
CATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC
GACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCC
TGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGTAA

>dTALE-Backbone(HD in 0.5 repeat)-NLS-VP64-2A-EGFP

ATGTCGCGGACCCGGCTCCCTTCCCCACCCGCACCCAGCCCAGCGTTTTCGGCCGACTCGTTCTCAGACCTGCTTAGGCAGTT
CGACCCCTCACTGTTTAACACATCGTTGTTCGACTCCCTTCCTCCGTTTGGGGCGCACCATACGGAGGCGGCCACCGGGGAGT
GGGATGAGGTGCAGTCGGGATTGAGAGCTGCGGATGCACCACCCCCAACCATGCGGGTGGCCGTCACCGCTGCCCGACCGC
CGAGGGCGAAGCCCGCACCAAGGCGGAGGGCAGCGCAACCGTCCGACGCAAGCCCCGCAGCGCAAGTAGATTTGAGAACTT
TGGGATATTCACAGCAGCAGCAGGAAAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGGG
TCATGGGTTTACACATGCCCACATCGTAGCCTTGTCGCAGCACCCTGCAGCCCTTGGCACGGTCGCCGTCAAGTACCAGGACA
TGATTGCGGCGTTGCCGGAAGCCACACATGAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCGAGCGCTTGAGG
CCCTGTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGCTGGACACGGGCCAGTTGCTGAAGATCGCGAAGCGGG
GAGGAGTCACGGCGGTCGAGGCGGTGCACGCGTGGCGCAATGCGCTCACGGGAGCACCCCTCAACCTGACCgagacgGTACAT
GAAACGCATGGCACGGcgtctcAACTCACGCCTGAGCAGGTAGTGGCTATTGCATCCCATGACGGGGGCAGACCCGCACTGGAG
TCAATCGTGGCCCAGCTTTCGAGGCCGGACCCCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTGCCTCG
GCGGACGACCCGCCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTGCATTGATTAAGCGGACCAACAGAAGGATTCC
CGAGAGGACATCACATCGAGTGGCAGATCACGCGCAAGTGGTCCGCGTGCTCGGATTCTTCCAGTGTCACTCCCACCCCGCA
CAAGCGTTCGATGACGCCATGACTCAATTTGGTATGTCGAGACACGGACTGCTGCAGCTCTTTCGTAGAGTCGGTGTCACAGA
ACTCGAGGCCCGCTCGGGCACACTGCCTCCCGCCTCCCAGCGGTGGGACAGGATTCTCCAAGCGAGCGGTATGAAACGCGC
GAAGCCTTCACCTACGTCAACTCAGACACCTGACCAGGCGAGCCTTCATGCGTTCGCAGACTCGCTGGAGAGGGATTTGGACG
CGCCCTCGCCCATGCATGAAGGGGACCAAACTCGCGCGTCAgctagccccaagaagaagagaaaggtggaggccagcggttccGGACGGGCTG
ACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCC
TTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATTAACtctagaggcagtggagag
ggcagaggaagtctgctaacatgcggtgacgtcgaggagaatcctggcccaGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTG
GTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG
ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT
GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAC
CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATAT
CATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC
GACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCC
TGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGTAA

>dTALE-Backbone(NN in 0.5 repeat)-NLS-VP64-2A-EGFP

ATGTCGCGGACCCGGCTCCCTTCCCCACCCGCACCCAGCCCAGCGTTTTCGGCCGACTCGTTCTCAGACCTGCTTAGGCAGTT
CGACCCCTCACTGTTTAACACATCGTTGTTCGACTCCCTTCCTCCGTTTGGGGCGCACCATACGGAGGCGGCCACCGGGGAGT
GGGATGAGGTGCAGTCGGGATTGAGAGCTGCGGATGCACCACCCCCAACCATGCGGGTGGCCGTCACCGCTGCCCGACCGC
CGAGGGCGAAGCCCGCACCAAGGCGGAGGGCAGCGCAACCGTCCGACGCAAGCCCCGCAGCGCAAGTAGATTTGAGAACTT
TGGGATATTCACAGCAGCAGCAGGAAAAGATCAAGCCCAAAGTGAGGTCGACAGTCGCGCAGCATCACGAAGCGCTGGTGGG
TCATGGGTTTACACATGCCCACATCGTAGCCTTGTCGCAGCACCCTGCAGCCCTTGGCACGGTCGCCGTCAAGTACCAGGACA
TGATTGCGGCGTTGCCGGAAGCCACACATGAGGCGATCGTCGGTGTGGGGAAACAGTGGAGCGGAGCCCGAGCGCTTGAGG
CCCTGTTGACGGTCGCGGGAGAGCTGAGAGGGCCTCCCCTTCAGCTGGACACGGGCCAGTTGCTGAAGATCGCGAAGCGGG
GAGGAGTCACGGCGGTCGAGGCGGTGCACGCGTGGCGCAATGCGCTCACGGGAGCACCCCTCAACCTGACCgagacgGTACAT
GAAACGCATGGCACGGcgtctcAACTCACGCCTGAGCAGGTAGTGGCTATTGCATCCAATAACGGGGGCAGACCCGCACTGGAG
TCAATCGTGGCCCAGCTTTCGAGGCCGGACCCCGCGCTGGCCGCACTCACTAATGATCATCTTGTAGCGCTGGCCTGCCTCG
GCGGACGACCCGCCTTGGATGCGGTGAAGAAGGGGCTCCCGCACGCGCCTGCATTGATTAAGCGGACCAACAGAAGGATTCC
CGAGAGGACATCACATCGAGTGGCAGATCACGCGCAAGTGGTCCGCGTGCTCGGATTCTTCCAGTGTCACTCCCACCCCGCA
CAAGCGTTCGATGACGCCATGACTCAATTTGGTATGTCGAGACACGGACTGCTGCAGCTCTTTCGTAGAGTCGGTGTCACAGA
ACTCGAGGCCCGCTCGGGCACACTGCCTCCCGCCTCCCAGCGGTGGGACAGGATTCTCCAAGCGAGCGGTATGAAACGCGC
GAAGCCTTCACCTACGTCAACTCAGACACCTGACCAGGCGAGCCTTCATGCGTTCGCAGACTCGCTGGAGAGGGATTTGGACG
CGCCCTCGCCCATGCATGAAGGGGACCAAACTCGCGCGTCAgctagccccaagaagaagagaaaggtggaggccagcggttccGGACGGGCTG
ACGCATTGGACGATTTTGATCTGGATATGCTGGGAAGTGACGCCCTCGATGATTTTGACCTTGACATGCTTGGTTCGGATGCCC
TTGATGACTTTGACCTCGACATGCTCGGCAGTGACGCCCTTGATGATTTCGACCTGGACATGCTGATTAACtctagaggcagtggagag
ggcagaggaagtctgctaacatgcggtgacgtcgaggagaatcctggcccaGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTG
GTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG
ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGT
GCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAC
CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATAT
CATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC
GACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCC
TGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGA
CGAGCTGTACAAGTAA

References

1. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4**, e5553 (2009).
2. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, e3647 (2008).
3. Kosuri, S. et al. Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat Biotechnol* **28**, 1295-1299 (2010).