

# Quantitative assay for TALEN activity at endogenous genomic loci

Yu Hisano<sup>1,\*</sup>, Satoshi Ota<sup>1,\*</sup>, Kazuharu Arakawa<sup>2</sup>, Michiko Muraki<sup>1</sup>, Nobuaki Kono<sup>2</sup>, Kazuki Oshita<sup>2</sup>, Tetsushi Sakuma<sup>3</sup>, Masaru Tomita<sup>2</sup>, Takashi Yamamoto<sup>3</sup>, Yasushi Okada<sup>4</sup> and Atsuo Kawahara<sup>1,†</sup>

<sup>1</sup>Laboratory for Cardiovascular Molecular Dynamics, Quantitative Biology Center, RIKEN, Furuedai 6-2-3, Suita, Osaka, 565-0874, Japan

<sup>2</sup>Institute for Advanced Biosciences, Keio University, Fujisawa, Kanagawa, 252-0882, Japan

<sup>3</sup>Molecular Genetics Laboratory, Department of Mathematical and Life Science, Graduate School of Science, Hiroshima University, Kagamiyama 1-3-1, Higashi-Hiroshima, Hiroshima, 739-8526, Japan

<sup>4</sup>Laboratory for Cell Polarity Regulation, Quantitative Biology Center, RIKEN, Furuedai 6-2-3, Suita, Osaka, 565-0874, Japan

\*These authors contributed equally to this work

†Author for correspondence (a.kawahara@riken.jp)

*Biology Open* 2, 363–367

doi: 10.1242/bio.20133871

Received 18th December 2012

Accepted 7th January 2013

## Summary

Artificially designed nucleases such as zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) can induce a targeted DNA double-strand break at the specific target genomic locus, leading to the frameshift-mediated gene disruption. However, the assays for their activity on the endogenous genomic loci remain limited. Herein, we describe a versatile modified *lacZ* assay to detect frameshifts in the nuclease target site. Short fragments of the genome DNA at the target or putative off-target loci were amplified from the genomic DNA of TALEN-treated or control embryos, and were inserted into the *lacZα* sequence for the conventional blue–white selection. The frequency of the frameshifts in the fragment can be estimated from the

numbers of blue and white colonies. Insertions and/or deletions were easily determined by sequencing the plasmid DNAs recovered from the positive colonies. Our technique should offer broad application to the artificial nucleases for genome editing in various types of model organisms.

© 2013. Published by The Company of Biologists Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial Share Alike License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

Key words: TALEN, LacZ blue–white selection, Quantitative assay, Genome editing

## Introduction

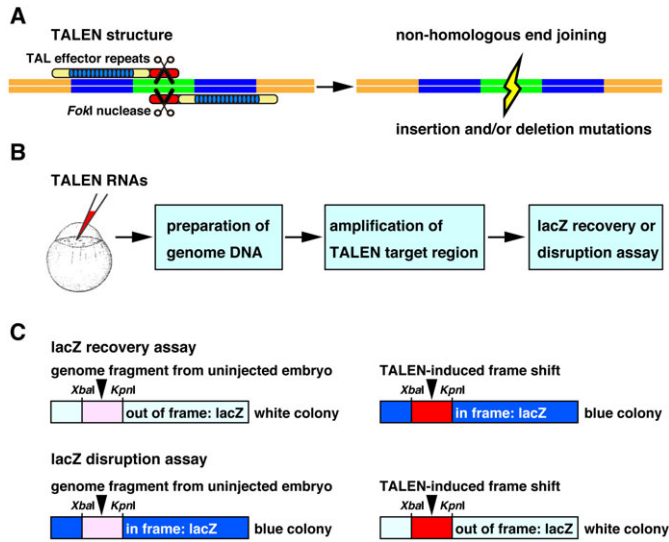
Artificial site-specific nucleases such as transcription activator-like effector nucleases (TALENs) and zinc finger nucleases (ZFNs) induce sequence-specific DNA double-strand breaks (DSBs) that can be repaired by the error-prone non-homologous end joining (NHEJ) system to yield insertion and/or deletion (indel) mutations at targeted genomic loci (Fig. 1A). This technology enables highly efficient targeted gene disruption not only in cultured cells but also in organisms (Bogdanove and Voytas, 2011; Carroll, 2011). However, the designs of these gene disruption systems are still empirical, requiring the use of easy and reliable assays such as the single-strand annealing (SSA) assay, which measures the cleavage activity at the target sequence introduced into the assay plasmid (Kim et al., 2009). Hetero-duplex analyses such as the Cel-I and T7 endonuclease assays can be used for endogenous loci (Kim et al., 2011). However, these assays are not very sensitive and do not confirm the sequence of the mutation. Restriction enzymes are useful for evaluation of TALEN activity when TALEN target region contains a suitable restriction enzyme site, however, that leads the limitation of TALEN design. Sequence analysis of the target loci from hundreds of clones is currently the only method that provides high-quality quantitative results and sequence confirmation (Reyon et al., 2012). Therefore, we developed a new versatile method the “*lacZ* recovery/disruption assay,” for

the quantitative evaluation of genome editing activity at the endogenous loci.

## Results and Discussion

### Principle of the *LacZ* recovery/disruption assay

The *lacZ* recovery/disruption assay is based on the principle of  $\alpha$ -complementation of the  $\beta$ -galactosidase in *Escherichia coli* (*E. coli*) (Fig. 1; supplementary material Fig. S1). pBR-*lacZα* was produced by inserting the *lacZ* promoter and the *lacZα* sequence containing the multi-cloning site (MCS) from pBluescript II (SK+) into the pBR322 plasmid, whose low-copy-number *ori* helps to minimize the number of false blue colonies due to the multiple transformation. The 82 bp fragment between the *KpnI* and *XbaI* sites in the MCS was replaced with a short genome fragment from the target site. Short fragments (100–150 bp) from the target or putative off-target loci were amplified from the genomic DNA of TALEN-injected or control embryos by PCR with locus-specific primers containing the *XbaI* or *KpnI* sites and were inserted into the *lacZα* sequence of the *XbaI*-*KpnI*-cleaved pBR-*lacZα*. TALEN-induced indel mutations will stochastically shift the reading frame. Thus, if the wild-type fragment is out of frame with respect to the *lacZα* sequence, all of the wild-type colonies will be white, whereas 1/3 of the indel mutants will yield an in-frame fusion, leading to the blue colonies (*lacZ* recovery assay). If the wild-type fragment is in frame with the *lacZα*



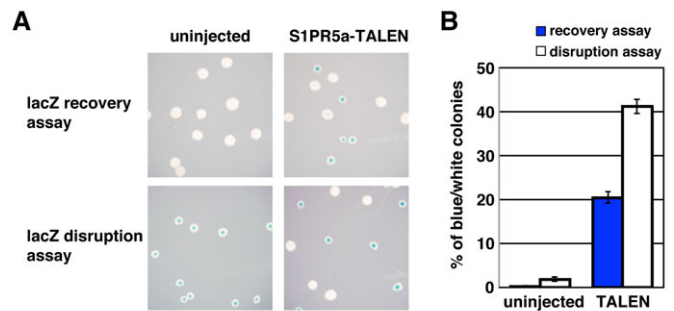
**Fig. 1. Principle of the lacZ recovery/disruption assay for evaluating TALEN activity *in vivo*.** (A) Schematic of TALEN function. TALEN consists of a TAL effector repeat (blue ellipse) that contains the repeat variable di-residues (RVDs) and the *FokI* nuclease catalytic domain (red ellipse). RVDs determine the base-recognition specificity, and TALEN specifically induces double-strand breaks (DSBs) in the targeted genomic locus (spacer region between forward and reverse TALEN target sites), resulting in indel mutations during non-homologous end joining (NHEJ). Blue line: TALEN recognition region; green line: spacer region. (B) Flowchart of the lacZ recovery/disruption assay. TALEN mRNA (400 pg each) was injected into zebrafish embryos at the 1- to 2-cell stage, and the genomic DNA was prepared from uninjected and TALEN-injected embryos. The TALEN target region was amplified by PCR using locus-specific primers for the TALEN target sites. (C) Principles of the lacZ recovery/disruption assay. In the lacZ recovery assay, the primer pairs are designed so that the fragment from the wild-type genome is inserted into the *lacZα* gene out of frame (white colonies). When DSBs are generated at the target genome site by TALENs, 1/3 of the TALEN-induced frameshifts lead to the recovery of lacZ activity because of in-frame fusion (blue colonies). Blue colonies will thus contain sequences with frameshift mutations in the TALEN target site (positive clones). In the lacZ disruption assay, the primer pairs are designed to generate an in-frame fusion of the wild-type genomic fragment with the *lacZα* gene (blue colonies). Two-thirds of the TALEN-mediated frameshifts cause the disruption of lacZ activity because of out-of-frame fusion (white colonies). White colonies will thus contain sequences with frameshift mutations in the TALEN target site (positive clones).

sequence, 2/3 of the indel mutants will yield out-of-frame fusions, leading to white colonies instead of the wild-type blue colonies (lacZ disruption assay). These assays can provide a quantitative measure of the frequency of indel mutations after TALEN injection. Furthermore, we can select mutant clones for sequence confirmation, although only 1/3 or 2/3 of all the indel mutations in the sample can be selected. This limitation is, however, not practically problematic. For gene disruption purposes, the frameshift mutation is the required mutation, which can be selectively identified by the proper design of the primers.

#### Detection of TALEN-induced indel mutations at endogenous loci

To determine whether this assay can detect TALEN-mediated genome editing in zebrafish, we designed TALENs for the receptors of the lipid mediator sphingosine-1-phosphate (S1P) (supplementary material Table S3). In zebrafish, the disruption of *s1p receptor-2* (*s1pr2*) causes the two-heart phenotype, cardia bifida (Kupperman et al., 2000; Kawahara et al., 2009). The

phenotype of *s1pr5a* disruption has not been reported. S1PR2- or S1PR5a-TALEN mRNA (400 pg each) were injected into zebrafish embryos at the 1- to 2-cell stage, and genomic DNA was prepared from uninjected or TALENs-injected embryos at 1 day post fertilization (dpf). TALEN-targeted genomic fragments amplified from genomic DNA were analyzed by the lacZ recovery/disruption assay. In the lacZ recovery assay, the frequency of blue colonies for the S1PR5a-TALEN target site was 0.2% in uninjected embryos. Sequence analysis showed that this background level of blue colonies was primarily due to an error in the primer synthesis (supplementary material Table S2). This frequency of blue colonies serves as the baseline value. The number of blue colonies was significantly increased to 20.5% in S1PR5a-TALEN-injected embryos (Fig. 2A–C). Because only 1/3 of indel mutants can be detected with the lacZ recovery assay, ~60% of genomic fragments in the sample were estimated to



		white colonies	blue colonies	% of blue colonies	% of white colonies	assay	Estimated indel rate [%]
S1PR5a	uninjected	1503	3	0.2 (± 0.1)		recovery	60
	TALEN	758	195	20.5 (± 1.3)			
	uninjected	16	847		1.9 (± 0.5)	disruption	59
	TALEN	371	529		41.2 (± 1.6)		

#### D

##### lacZ recovery assay (S1PR5a)

```

CTCATTTGAGTGGAATAATGCCGATTG-----TTTGGTGGCAATATGGAAGAAC WT
CTCATTTGAGTGGAATAATGCCGATTGTTGATAATTGGTGGCAATATGGAAGAAC +11 (-18/+29)
CTCATTTGAGTGGAATAATGCCGAT-----TTTGGTGGCAATATGGAAGAAC -2
CTCATTTGAGTGGAATAATGCCGATGCGAATA-----TTTGGTGGCAATATGGAAGAAC +8 (-22/+30)
CTCATTTGAGTGGAATAATGCCGATGCGAATA-----TTTGGTGGCAATATGGAAGAAC +2 (-5/+7)
CTCATTTGAGTGGAATAATGCCGAT-----TTTGGTGGCAATATGGAAGAAC -4
CTCATTTGAGTGGAATAATGCCGAT-----TTTGGTGGCAATATGGAAGAAC -13

```

##### lacZ disruption assay (S1PR5a)

```

CTCATTTGAGTGGAATAATGCCGATTG-----TTTGGTGGCAATATGGAAGAAC WT
CTCATTTGAGTGGAATAATGCCGATTGTTGATAATTGGTGGCAATATGGAAGAAC +4 (-2/+6)
CTCATTTGAGTGGAATAATGCCGAT-----TTTGGTGGCAATATGGAAGAAC -2
CTCATTTGAGTGGAATAATGCCAAA-----TTTGGTGGCAATATGGAAGAAC -4 (-7/+3)
CTCATTTGAGTTGTA-----TTTGGTGGCAATATGGAAGAAC -11 (-15/+4)
CTCATTTGAGTGGAATAATGCCGAT-----TTTGGTGGCAATATGGAAGAAC -13

```

**Fig. 2. *In vivo* evaluation of S1PR5a-TALEN.** (A) LacZ blue/white colonies of the lacZ recovery/disruption assay using genomic DNA from uninjected or S1PR5a-TALEN-injected embryos. (B,C) Results of the lacZ recovery/disruption assay for S1PR5a-TALEN. In the lacZ recovery assay, TALEN-induced indel mutations resulted in the recovery of lacZ activity in one frame out of the three possible frames (1/3); the estimated indel rate was calculated by multiplying the percentage of the blue colonies by 3. In the lacZ disruption assay, TALEN-induced indel mutations resulted in the loss of lacZ activity in two frames out of the three possible frames (2/3). Therefore, the estimated indel rate was calculated by multiplying the percentage of white colony by 1.5. “% of blue colonies” and “% of white colonies” are presented as the mean (± s.d., where s.d. =  $\sqrt{\frac{n_{\text{white}} \cdot n_{\text{blue}}}{(n_{\text{white}} + n_{\text{blue}})^3} \times 100}$ ).

(D) DNA sequences of the wild-type and S1PR5a-TALEN-mutated sequences. Deletions are indicated by red dashes, and insertions are indicated by red letters. For the indicated endogenous target, the wild-type sequence is shown in the top line for the respective assay. The numbers of nucleic acids deleted (–) and inserted (+) are indicated to the right. Blue: TALEN recognition sequences. Green: spacer sequences. Red dashes: deleted bases. Red letters: inserted bases.

**Table 1. Potential off-target sites for S1PR5a- or S1PR2-TALEN in zebrafish genome.**

S1PR5a	Score	Left half-site (5' to 3')	Right half-site (5' to 3')	Spacer (bp)		% of blue or white colonies	Assay
Intended	10.78	CTCATTGTAGTGGAAA	GTTCTTCCATATTGCCA	15			
Candidate 1	22	CTCATAATAATGAGAA	CTCATTATAATGAGAA	24	Uninjected	2.1 (±1.1)	Recovery
					TALEN	2.1 (±1.2)	Recovery
Candidate 2	22.8	CTCATTGTAATAAAAAG	TTCATTCTAAAAA	28	Uninjected	2.1 (±0.7)	Recovery
					TALEN	1.6 (±0.7)	Recovery
Candidate 3	23.1	CACAATATAAAAAA	CTCATTTCATGAAAA	25	Uninjected	4.4 (±1.1)	Disruption
					TALEN	4.8 (±1.4)	Disruption
Candidate 4	24.4	CTCATTTTAACGAGAA	CTCATTAATAAGACAA	23	Uninjected	0.3 (±0.3)	Recovery
					TALEN	0.2 (±0.2)	Recovery
S1PR2	Score	Left half-site (5' to 3')	Right half-site (5' to 3')	Spacer (bp)		% of blue or white colonies	Assay
Intended	9.56	CCCTGAACATTTTAT	CCAAGATGATGATGCT	14			
Candidate 1	21.5	CCAAGATGACAATACC	CCCTGAACCTCTTTT	16	Uninjected	0.2 (±0.1)	Recovery
					TALEN	0.2 (±0.2)	Recovery
Candidate 2	21.8	CCCTGAACATTTAAC	TCATGATAATAATCCT	13	Uninjected	1.3 (±0.4)	Recovery
					TALEN	1.1 (±0.4)	Recovery
Candidate 3	22.2	ACAAGTAAACAATACT	CCAAAATGATGATGCT	12	Uninjected	1.0 (±0.7)	Recovery
					TALEN	0.0 (±0.0)	Recovery
Candidate 4	22.4	CCCTTAAAACCTTTAT	CCAAAATATAAAGCT	26	Uninjected	3.5 (±0.8)	Disruption
					TALEN	2.9 (±0.7)	Disruption

have indel mutations caused by S1PR5a-TALEN. Similarly, the lacZ disruption assay yielded 41.2% white colonies after S1PR5a-TALEN injection. Because 2/3 of indel mutants can be detected with the lacZ disruption assay, the indel mutation frequency was estimated to be ~59%. Thus, the recovery and disruption assays gave consistent values, confirming the quantitative reliability of our assay. Further, we found that S1PR2-TALEN injection caused cardia bifida and induced indel mutations at the estimated indel rate, ~27% (supplementary material Figs S2, S3). The sequencing analysis of positive colonies from the lacZ assay confirmed that both S1PR2- and S1PR5a-TALEN can induce various types of indel mutations (Fig. 2D; supplementary material Fig. S2). The quantitative reliability was further confirmed by comparing the estimates from the lacZ assays to those from conventional assays such as massive clonal analysis (supplementary material Fig. S4) and heteroduplex analysis using T7 endonuclease (supplementary material Fig. S5). We observed the germ line transmission of the TALEN-induced indel mutations in F1 embryos obtained from the S1PR2- or S1PR5a-TALEN-injected adult F0 progenitors. Some of the indel mutations detected in the TALEN-injected embryos (1 dpf) were identical to those found in the F1 embryos (supplementary material Fig. S6).

#### Influence of TALEN on off-target genome sites

To determine the influence of S1PR2- or S1PR5a-TALEN on off-target genome sites, we identified the most homologous genomic loci (3–6 bases mismatched) for S1PR2- and S1PR5a-TALEN and performed the lacZ assay using specific primers designed for these loci (Table 1; supplementary material Table S1). The number of positive colonies did not significantly increase in the TALEN-injected embryos relative to the uninjected controls, suggesting that the off-target activities of both S1PR2- and S1PR5a-TALENs were undetectably low, as reported with other TALENs (Dahlem et al., 2012). Consistently, we observed the target-specific phenotype, cardia bifida, in the S1PR2-TALEN-injected embryos (supplementary material Fig. S3), but we

observed no other phenotype. The characterization of the S1PR5a-deficient phenotype will be performed in a future study.

#### Outlook

One attractive application of site-specific nucleases is genome editing in patient-derived induced pluripotent stem (iPS) cells for the development of personalized cell therapy against intractable human diseases (Urnov et al., 2010). The *in vivo* evaluation of site-specific nuclease is indispensable, and the lacZ recovery/disruption assay could be a very efficient method for this analysis based on the following three reasons: (1) TALEN activity for an endogenous target is measured using the genomic DNA of TALEN-injected embryos or TALEN-transfected cultured cells; (2) the efficacy of TALEN activity *in vivo* can be quantitatively evaluated by the ratio of blue/white colonies; and (3) positive colonies with TALEN-induced indel mutations can be selected as the results of the lacZ assay. To facilitate the application of this assay, we developed free software tools for primer selection, which are available at <http://ws.g-language.org/TALEN>. This assay will accelerate the development and application of TALENs and ZFNs in tailored genome editing technology.

#### Materials and Methods

##### Construction of pCS2P-TAF, pCS2P-S1PR5a and pBR-lacZ $\alpha$

To construct pCS2P-TAF, the N- and C-terminal TAL effector domains and the *FokI* catalytic domain were amplified from the pcDNA-TAL-NC plasmid by PCR using the primers TAL-N-*EcoRI* and TAL-C-*EcoRI* (supplementary material Table S1) (Sakuma et al., 2013). After *EcoRI* digestion, the resultant fragments were inserted into the *EcoRI*-cleaved pCS2P vector. The construction of pCS2P-TAF was confirmed by sequencing. We isolated zebrafish *s1pr5a* from a zebrafish cDNA library (AB strain) by PCR using the primers zS1PR5a-S and zS1PR5a-AS (supplementary material Table S1). After *BamHI* and *XbaI* digestion, the resultant DNA fragments were inserted into the *BamHI*-*XbaI*-cleaved pCS2P vector. The GenBank accession number of zebrafish *s1pr5a* is AB743585. We designed S1PR5a-TALEN and S1PR2-TALEN on the basis of *s1pr5a* (AB743585) and *s1pr2* (AF260256), respectively. The DNA fragment containing the lacZ promoter region, the MCS and the  $\beta$ -galactosidase  $\alpha$ -fragment was amplified from the pBluescript II (SK+) vector by PCR using the primers Bg/II-lacZ-S and NheI-lacZ-AS (supplementary material Table S1). After *BglIII* and *NheI* digestion, the resultant fragments were inserted between the *BamHI* and *NheI* sites of pBR322 (pBR-lacZ $\alpha$ ). The construction of pBR-lacZ $\alpha$  was confirmed by sequencing.

### Construction of TALEN plasmids

We constructed TALENs in a two-step assembly system (Addgene, Golden Gate TALEN Kit) as described previously (Cermak et al., 2011). In the first assembly, the pFUS vector (pFUS-A2A, pFUS-A2B or pFUS-B1-B6; 150 ng/μl) and the RVD plasmids (ex. pHD1+pHD2+pHD3+pNG4+pNN5+pNI6; 150 ng/μl) were incubated at 37°C for 30 minutes with *Bsa*I-HF (10 U, New England Biolabs) and 10 mM ATP (Invitrogen) in a 10 μl reaction volume (Sakuma et al., 2013). Subsequently, T7 ligase (1,500 U, Enzymatics) was added, and the reaction mixture was incubated at 25°C for 1 hour. Afterward, 0.5 μl of 10 mM ATP and 0.5 μl of Plasmid-Safe DNase (10 U/μl, Epicentre) were added, and the mixture was incubated at 37°C for 1 hour. After the enzymes were heat-inactivated at 70°C for 30 minutes, the reaction solution was used to transform JM109 cells on an LB plate containing 50 μg/ml spectinomycin, X-gal (5-Bromo-4-Chloro-3-Indolyl-β-D-Galactoside) and IPTG (Isopropyl-β-D-thiogalactopyranoside). The RVD ordering in the assembly vector (pFUS vector) was confirmed by sequencing.

In the second assembly, 100 ng each of the pFUS-A2A, pFUS-A2B and pFUS-B1-B6 plasmids containing the first assembly RVDs; 150 ng of the appropriate last repeat plasmid (pLR-HD, pLR-NI, pLR-NG, or pLR-NN) and 150 ng of pCS2P-TAF were incubated with 0.5 μl of *Bsm*BI (10 U/μl, New England Biolabs) in a 10 μl reaction volume. The ligation reaction and Plasmid-Safe DNase treatment were performed as described for the first assembly. The ligation solution was used to transform JM109 cells on an LB plate containing 50 μg/ml ampicillin, X-gal and IPTG. The TALEN construct was confirmed by sequencing.

### Injection of morpholino and TALEN mRNA

The morpholino for zebrafish *slpr2* (S1PR2-MO, 5'-CCGCAAACAGACGG-CAAGTAGTCAT-3') was used as described previously (Kawahara et al., 2009). Constructed TALEN plasmids were linearised with *Not*I, and mRNA was synthesised using the SP6 mMESSAGE mMACHINE Kit (Ambion) according to the manufacturer's protocol. Morpholinos and synthesised TALEN mRNA were dissolved in the injection buffer (40 mM HEPES (pH 7.4), 240 mM KCl and 0.5% phenol red) and injected into 1- to 2-cell stage embryos.

### Preparation of genomic DNA

Genomic DNA was prepared from uninjected or TALEN-injected embryos at 1 dpf using the Gentra Puregene Tissue Kit (QIAGEN) according to the manufacturer's protocol.

### Primer design for the lacZ assay

In the lacZ disruption assay, the amplicon from the uninjected genome must be inserted into the *Xba*I and *Kpn*I sites of pBR-lacZα in frame. Furthermore, the reading frame used should not contain stop codons (TAG, TAA and TGA) or start codons (ATG and GTG) downstream of the spacer region. Importantly, it is possible that ATG and GTG codons located downstream of the *lacZ* promoter could be used as start codons in *E. coli*. Moreover, it is necessary that neither reading frames used contains a start codon upstream of the spacer region. In the lacZ recovery assay, the amplicon from the uninjected genome must be inserted into the *Xba*I and *Kpn*I sites of pBR-lacZα out of frame. The prerequisites for this assay are as follows: (1) the reading frame connected to the C-terminus of the β-galactosidase α-fragment must not contain either start codons or stop codons downstream of the spacer region; (2) this reading frame must not contain a start codon upstream of the spacer region; and (3) no other reading frame should contain a stop codon upstream of the spacer region. A complementary strand can alternatively be used for these assays.

### Primer design software tool for LacZ recovery/disruption assay

A software tool for designing the primers of this assay is freely available as an online tool at <http://talen.g-language.org>, and the source code is available under GNU General Public License. The software pipeline is mostly written in Perl 5.8.8 and is running on Linux (CentOS 6). The basic strategy for the design of TAL effector target sites for a given sequence of interest follows the methods described by previous works (Cermak et al., 2011; Doyle et al., 2012). Briefly, for every possible binding site pair in a specified range and preceded by T within the given sequence, binding score (lower-the-better) is calculated using the weight matrix developed by Moscou and Bogdanove (Moscou and Bogdanove, 2009), and best-scoring TALEN sites as well as the RVD sequences are displayed. By selecting one of the candidate TALEN sites, the user can further design a set of primers for LacZ recovery or disruption assay in another tool. In this tool, candidate primers are firstly comprehensively identified using Primer3 software (Untergasser et al., 2012), and are subsequently screened under primer design criteria for LacZ assay as described (see previous section), and the screened primer pairs are further checked for off-target sites. In order to achieve high performance in this off-target screening of hundreds of primer pairs, the short read mapping tool BWA 0.6.2 (Li and Durbin, 2010) is utilized, with mismatch setting (-n) of 5. All of the parameters, including the lengths of spacer, TAL effector binding sites, and primers, are customizable from the web interface, and therefore the primer design tool for LacZ assay can also be applied for ZFN assays. Results can be optionally

downloaded in tab-delimited text format, which can be readily imported into standard spreadsheet software.

### LacZ recovery or disruption assay

DNA fragments of 100–150 bp containing the TALEN target site were amplified from 50 ng of genomic DNA by PCR with PrimeSTAR Max DNA polymerase (TaKaRa) in a 50 μl reaction volume. When specific DNA fragments were not sufficient or when non-specific products were amplified, nested PCR was performed. The primers contained *Xba*I and *Kpn*I overhangs to add *Xba*I and *Kpn*I sites to the amplicons (supplementary material Table S2). The PCR conditions were as follows: 94°C for 2 minutes; 35 cycles of 98°C for 10 seconds, 56°C for 10 seconds and 72°C for 20 seconds; and 72°C for 5 minutes. The amplified DNA fragments were purified with a MinElute PCR Purification Kit (QIAGEN) and digested with *Xba*I and *Kpn*I. The pBR-lacZα vector was also digested with *Xba*I and *Kpn*I, following phenol-chloroform-isoamyl alcohol extraction and Thermosensitive Alkaline Phosphatase (Promega) treatment at 37°C for 1 hour. The digested DNA fragments and the vector were electrophoresed in 1.5% and 1% agarose gels in TAE (Tris-acetate-EDTA) buffer, respectively, and extracted from the gels with the MinElute Gel Extraction Kit (QIAGEN). The DNA fragments were then inserted into the *Xba*I-*Kpn*I-cleaved pBR-lacZα with Ligation high (TOYOBO) at 4°C for overnight. After transformation, cells were plated on an LB plate containing 50 μg/ml ampicillin with X-gal and IPTG. After colonies were sufficiently large and colored, the numbers of white and blue colonies were counted.

### T7 endonuclease assay

The TALEN target region was amplified from the genomic DNA by PCR using the gene-specific primers shown in supplementary material Table S1. The amplicons were purified with the MinElute PCR Purification Kit denatured by heating and annealed in hybridisation buffer (10 mM Tris-HCl [pH 8.5], 75 mM KCl, 1.5 mM MgCl<sub>2</sub>) to form heteroduplex DNA. After treatment with T7 endonuclease I (3 U, New England Biolabs) at 37°C for 1 hour, the resulting fragments were subjected to electrophoresis in a 2% agarose gel and were visualized by staining with ethidium bromide.

### Off-target effects

The effects at potential off-target sites for the designed TALENs were identified using Paired Target Finder (<https://tale-nt.cac.cornell.edu/node/add/talef-off-paired>). The minimal and maximal spacer lengths were set to 12 and 31, respectively, and the four most plausible candidates were tested with the lacZ recovery or disruption assay.

### Acknowledgements

The authors thank A. Katsuma for technical assistance; R. Fukuoka, M. Komeno and S. Oohara for zebrafish maintenance; P. Karagiannis for valuable comments; and D.F. Voytas for Golden Gate TALEN Kit. Our work is funded by the Program for Next Generation World-Leading Researchers (NEXT Program), by the Japan Society for the Promotion of Science and by the Sasakawa Scientific Research Grant.

### Competing Interests

The authors have no competing interests to declare.

### References

- Bogdanove, A. J. and Voytas, D. F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science* **333**, 1843–1846.
- Carroll, D. (2011). Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773–782.
- 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.
- Dahlem, T. J., Hoshijima, K., Juryne, M. J., Gunther, D., Starker, C. G., Locke, A. S., Weis, A. M., Voytas, D. F. and Grunwald, D. J. (2012). Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.* **8**, e1002861.
- Doyle, E. L., Booher, N. J., Standage, D. S., Voytas, D. F., Brendel, V. P., Vandyk, J. K. and Bogdanove, A. J. (2012). TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res.* **40** (W1), W117–W122.
- Kawahara, A., Nishi, T., Hisano, Y., Fukui, H., Yamaguchi, A. and Mochizuki, N. (2009). The sphingolipid transporter spns2 functions in migration of zebrafish myocardial precursors. *Science* **323**, 524–527.

- Kim, H. J., Lee, H. J., Kim, H., Cho, S. W. and Kim, J. S. (2009). Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* **19**, 1279-1288.
- Kim, H., Um, E., Cho, S. R., Jung, C., Kim, H. and Kim, J. S. (2011). Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat. Methods* **8**, 941-943.
- Kupperman, E., An, S., Osborne, N., Waldron, S. and Stainier, D. Y. R. (2000). A sphingosine-1-phosphate receptor regulates cell migration during vertebrate heart development. *Nature* **406**, 192-195.
- Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595.
- Moscou, M. J. and Bogdanove, A. J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501.
- 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.
- Sakuma, T., Hosoi, S., Woltjen, K., Suzuki, K., Kashiwagi, K., Wada, H., Ochiai, H., Miyamoto, T., Kawai, N., Sasakura, Y. et al. (2013). Efficient TALEN construction and evaluation methods for human cell and animal applications. *Genes Cells* (in press).
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. and Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115.
- Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. and Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636-646.