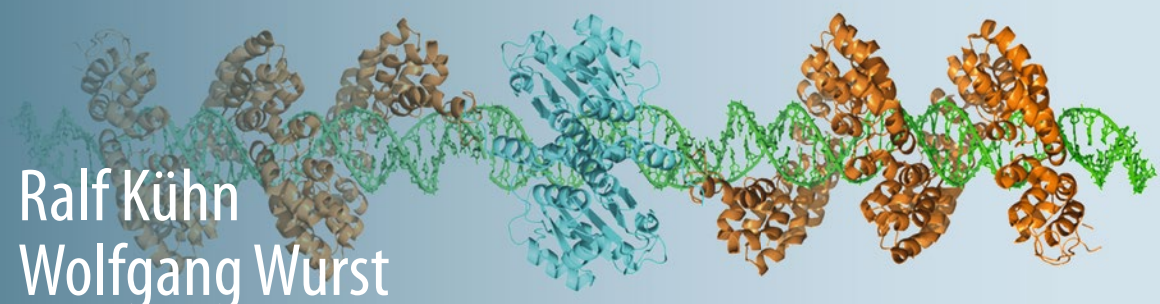
Methods in
Molecular Biology 1338

Springer Protocols



Ralf Kühn
Wolfgang Wurst
Benedikt Wefers *Editors*

TALENs

Methods and Protocols

Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
http://www.springer.com/series/7651

TALENs

Methods and Protocols

Edited by

Ralf Kühn

Max-Delbrück-Center for Molecular Medicine, Berlin, Germany;
Berlin Institute of Health, Berlin, Germany

Wolfgang Wurst

Institute of Developmental Genetics, Helmholtz Zentrum München,
German Research Center for Environmental Health, Neuherberg, Germany;
Technische Universität München, Freising-Weihenstephan, Germany;
Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE), Munich, Germany;
Munich Cluster for Systems Neurology (SyNergy),
Ludwig-Maximilians-Universität München, Munich, Germany

Benedikt Wefers

Institute of Developmental Genetics, Helmholtz Zentrum München,
German Research Center for Environmental Health, Neuherberg, Germany;
Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE),
Munich, Germany

Humana Press

Editors
Ralf Kühn
Max-Delbrück-Center for Molecular Medicine
Berlin, Germany

Berlin Institute of Health, Berlin, Germany

Benedikt Wefers
Institute of Developmental Genetics
Helmholtz Zentrum München
German Research Center for Environmental Health
Neuherberg, Germany

Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE)
Munich, Germany

Wolfgang Wurst
Institute of Developmental Genetics
Helmholtz Zentrum München
German Research Center for Environmental Health
Neuherberg, Germany

Technische Universität München
Freising-Weihenstephan, Germany

Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE)
Munich, Germany

Munich Cluster for Systems Neurology (SyNergy)
Ludwig-Maximilians-Universität München
Munich, Germany

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-2931-3 ISBN 978-1-4939-2932-0 (eBook)
DOI 10.1007/978-1-4939-2932-0

Library of Congress Control Number: 2015949823

Springer New York Heidelberg Dordrecht London

Printed on acid-free paper

Humana Press is a brand of Springer
Springer Science+Business Media LLC New York is part of Springer Science+Business Media (www.springer.com)

Preface

Genetic engineering to create targeted mutations in cells and organisms is a key technology for biomedical research and biotechnology. The wide use of mice as mammalian genetic model is based on gene targeting in embryonic stem cells that undergo spontaneous homologous recombination with externally provided gene targeting vectors. Besides mouse stem cells, most other cell lines and organisms were for a long time refractory to HR and the available tools enabled mainly random genome integrations. These limitations are overcome with the generation of designer nucleases. The first generation of sequence-specific nucleases, the zinc-finger nucleases (ZFN), provided proof-of-principle that the induction of targeted double-strand breaks is a powerful tool for gene editing by the stimulation of DNA repair. Nevertheless, this system was not convenient since ZFNs cannot be easily programmed for the recognition of new target sequences. In 2009 a novel, modular DNA recognition code from the transcription activator-like (TAL) proteins of Xanthomonas was discovered and, based on the extensive experience with ZFNs, could be rapidly adapted for use as sequence-specific TAL effector nucleases (TALEN) by fusion with the FokI nuclease domain. Within the last years TAL proteins and TALENs were further characterized and successfully applied in various species which were previously inaccessible for gene editing. In line with the successful format of Methods in Molecular Biology, this volume provides a comprehensive collection of step-by-step protocols to support the use of this new technology across various model organisms. While most of the chapters describe species-specific methods to generate new mutants, the content is completed by chapters on natural TAL effectors, TAL element DNA binding principles, TALEN target site prediction, and methods for the efficient construction of TALEN coding regions. In addition to the application of TALEN as sequence-specific nucleases, this volume includes a chapter on TAL-based gene activators or inhibitors and the visualization of chromatin dynamics in live cells. More recently a new generation of nuclease technology was provided by the clustered, regularly interspaced, short palindromic repeat (CRISPR)/Cas bacterial defense system, using short RNAs for DNA sequence recognition. We expect that in future both TALEN and CRISPR/Cas will be successfully used for genetic engineering in a wide range of cell and animal model systems. For this purpose we also included one chapter on CRISPR/Cas mutagenesis in mammalian cells.

Berlin, Germany — *Ralf Kühn*
Munich, Germany — *Wolfgang Wurst*
Munich, Germany — *Benedikt Wefers*

Contents

Contributors

Zach N. Adelman • *305 Fralin Life Science Institute, Virginia Tech, Blacksburg, VA, USA; Department of Entomology, Virginia Tech, Blacksburg, VA, USA*

Ignacio Anegon • *Transgenic Rats Nantes IBiSA-Centre National de Recherche Scientifique, Nantes, France; ITUN, CHU Nantes, Nantes, France; INSERM UMR 1064, Center for Research in Transplantation and Immunology, Nantes, France*

Azadeh Aryan • *305 Fralin Life Science Institute, Virginia Tech, Blacksburg, VA, USA; Department of Entomology, Virginia Tech, Blacksburg, VA, USA*

Sanjay Basu • *305 Fralin Life Science Institute, Virginia Tech, Blacksburg, VA, USA; Department of Entomology, Virginia Tech, Blacksburg, VA, USA*

Jens Boch • *Department of Genetics, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany*

Adam J. Bogdanove • *Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, USA*

Nicholas J. Booher • *Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, USA*

Christina Brandl • *Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; Technische Universität München, Freising-Weihenstephan, Germany*

Erika Brunet • *Museum National d'Histoire Naturelle, Paris, France; CNRS, UMR7196, Paris, France; Inserm, U1154, Paris, France*

Yixu Chen • *State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China*

Christopher H.K. Cheng • *School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong, China; School of Biomedical Sciences Core Laboratory, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China*

Anne De Cian • *INSERM U1154, CNRS UMR7196, Museum National d'Histoire Naturelle, Paris, France*

Jean-Paul Concordet • *INSERM U1154, CNRS UMR7196, Museum National d'Histoire Naturelle, Paris, France*

Thomas S. Ebert • *Institute of Molecular Medicine, University Hospital, University of Bonn, Bonn, Germany*

Moritz M. Gaidt • *Institute of Molecular Medicine, University Hospital, University of Bonn, Bonn, Germany*

Charles A. Gersbach • *Department of Biomedical Engineering, Duke University, Durham, NC, USA; Center for Genomic and Computational Biology, Duke University, Durham, NC, USA; Department of Orthopaedic Surgery, Duke University Medical Center, Durham, NC, USA*

Hind Ghezraoui • *Museum National d'Histoire Naturelle, Paris, France; CNRS, UMR7196, Paris, France; Inserm, U1154, Paris, France*

Carine Giovannangeli • *INSERM U1154, CNRS UMR7196, Museum National d'Histoire Naturelle, Paris, France*

Marian Goldsmith • *Department of Biological Sciences, University of Rhode Island, Kingston, RI, USA*

Shengjie Guo • *Biodynamic Optical Imaging Center (BIOPIC), State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing, China*
Yawen Guo • *State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China*
Mary Etna Haac • *305 Fralin Life Science Institute, Virginia Tech, Blacksburg, VA, USA; Department of Entomology, Virginia Tech, Blacksburg, VA, USA*
Toshinori Hayashi • *School of Life Science, Faculty of Medicine, Tottori University, Yonago, Tottori, Japan*
Veit Hornung • *Institute of Molecular Medicine, University Hospital, University of Bonn, Bonn, Germany*
Peng Huang • *Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, College of Life Sciences, Peking University, Beijing, P. R. China*
Renjie Jiao • *State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China; Guangzhou Hoffmann Institute of Immunology, School of Basic Sciences, Guangzhou Medical University, Guangzhou, China*
Ralf Kühn • *Max-Delbrück-Center for Molecular Medicine, Berlin, Germany; Berlin Institute of Health, Berlin, Germany*
Changqing Li • *State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China*
Shuo Lin • *Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, CA, USA*
Jiyong Liu • *State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China; Guangzhou Hoffmann Institute of Immunology, School of Basic Sciences, Guangzhou Medical University, Guangzhou, China*
Yun Liu • *School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong, China; School of Biomedical Sciences Core Laboratory, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China*
Séverine Ménoret • *Transgenic Rats Nantes IBiSA-Centre National de Recherche Scientifique, Nantes, France; ITUN, CHU Nantes, Nantes, France; INSERM UMR 1064, Center for Research in Transplantation and Immunology, Paris, France*
Taro Mito • *Department of Life Systems, Institute of Technology and Science, The University of Tokushima, Tokushima, Japan*
Yusuke Miyanari • *Okazaki Institute for integrative bioscience, Okazaki, Japan*
Kevin M. Myles • *305 Fralin Life Science Institute, Virginia Tech, Blacksburg, VA, USA; Department of Entomology, Virginia Tech, Blacksburg, VA, USA*
Sumihare Noji • *Center for Collaboration among Agriculture, Industry and Commerce, The University of Tokushima, Tokushima, Japan*
Oskar Ortiz • *Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany*
David G. Ousterout • *Department of Biomedical Engineering, Duke University, Durham, NC, USA*
Marion Piganeau • *Museum National d'Histoire Naturelle, Paris, France; CNRS, UMR7196, Paris, France; Inserm, U1154, Paris, France*

Séverine Remy • *Transgenic Rats Nantes IBiSA-Centre National de Recherche Scientifique, Nantes, France; ITUN, CHU Nantes, Nantes, France; INSERM UMR 1064, Center for Research in Transplantation and Immunology, Nantes, France*
Benjamin Renouf • *Museum National d'Histoire Naturelle, Paris, France; CNRS, UMR7196, Paris, France; Inserm, U1154, Paris, France*
Annekatrin Richter • *Department of Genetics, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany*
Tetsushi Sakuma • *Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Hiroshima, Japan*
Chloé Savignard • *Transgenic Rats Nantes IBiSA-Centre National de Recherche Scientifique, Nantes, France; ITUN, CHU Nantes, Nantes, France; INSERM UMR 1064, Center for Research in Transplantation and Immunology, Nantes, France*
Jonathan L. Schmid-Burgk • *Institute of Molecular Medicine, University Hospital, University of Bonn, Bonn, Germany*
Tobias Schmidt • *Institute of Molecular Medicine, University Hospital, University of Bonn, Bonn, Germany*
Jana Streubel • *Department of Genetics, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany*
Yoko Takasu • *National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan*
Takashi Takeuchi • *School of Life Science, Faculty of Medicine, Tottori University, Yonago, Tottori, Japan*
Toshiki Tamura • *National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan*
Laurent Tesson • *Transgenic Rats Nantes IBiSA-Centre National de Recherche Scientifique, Nantes, France; ITUN, CHU Nantes, Nantes, France; INSERM UMR 1064, Center for Research in Transplantation and Immunology, Nantes, France*
Pratiksha I. Thakore • *Department of Biomedical Engineering, Duke University, Durham, NC, USA*
Reynald Thinard • *Transgenic Rats Nantes IBiSA-Centre National de Recherche Scientifique, Nantes, France; ITUN, CHU Nantes, Nantes, France; INSERM UMR 1064, Center for Research in Transplantation and Immunology, Nantes, France*
Xiangjun Tong • *Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, College of Life Sciences, Peking University, Beijing, P. R. China*
Claire Usal • *Transgenic Rats Nantes IBiSA-Centre National de Recherche Scientifique, Nantes, France; ITUN, CHU Nantes, Nantes, France; INSERM UMR 1064, Center for Research in Transplantation and Immunology, Nantes, France*
Takahito Watanabe • *Department of Life Systems, Institute of Technology and Science, The University of Tokushima, Tokushima, Japan; Center for Collaboration among Agriculture, Industry and Commerce, The University of Tokushima, Tokushima, Japan*
Benedikt Wefers • *Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE), Munich, Germany*
Wensheng Wei • *Biodynamic Optical Imaging Center (BIOPIC), State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing, China*
Frank White • *Department of Plant Pathology, University of Florida, Gainesville, FL, USA*

Wolfgang Wurst • *Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; Technische Universität München, Freising-Weihenstephan, Germany; Deutsches Zentrum für Neurodegenerative Erkrankungen e. V. (DZNE), Munich, Germany; Munich Cluster for Systems Neurology (SyNergy), Ludwig-Maximilians-Universität München, Munich, Germany*

An Xiao • *Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, College of Life Sciences, Peking University, Beijing, P. R. China*

Takashi Yamamoto • *Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Hiroshima, Japan*

Junjiao Yang • *Biodynamic Optical Imaging Center (BIOPIC), State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing, China*

Pengfei Yuan • *Biodynamic Optical Imaging Center (BIOPIC), State Key Laboratory of Protein and Plant Gene Research, College of Life Sciences, Peking University, Beijing, China*

Bo Zhang • *Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, College of Life Sciences, Peking University, Beijing, P. R. China; College of Life Sciences, Peking University, Beijing, P. R. China*

Hui Zhao • *School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong, China; School of Biomedical Sciences Core Laboratory, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China*

Michal Zurovec • *Biology Centre, Czech Academy of Sciences, Ceske Budejovice, Czech Republic*

Chapter 1

Xanthomonas and the TAL Effectors: Nature's Molecular Biologist

Frank White

Abstract

Agrobacterium, due to the transfer of T-DNA to the host genome, is known as nature's genetic engineer. Once again, bacteria have led the way to newfound riches in biotechnology. *Xanthomonas* has emerged as nature's molecular biologist as the functional domains of the sequence-specific DNA transcription factors known as TAL effectors were characterized and associated with the cognate disease susceptibility and resistance genes of plants.

Key words AvrBs3, TAL effector, PthA, Bs3, Xa27, Os8N3/OsSWEET11, UPA20

1 In the Beginning There Was AvrBs3

The term TAL effector was introduced in 2006 to cover the large family of type III effector genes related to *avrBs3* from what was known at the time as *Xanthomonas campestris* pathovar *vesicatoria* [1, 2]. The gene *avrBs3* was one of an amazing, and seemingly endless, string of so-called Avr genes cloned from various plant pathogenic bacteria and corresponding to dominant resistance (R) genes of the respective hosts [2]. Identification of family members from pathogenic *Xanthomonas* strains that were associated with symptoms of pathogenicity and not associated with a cognate dominant R gene in the respective host led to the coining of the term *pth*ogenicity or *pth* for some members, including the founding *pth* gene *pthA* [3]. Subsequently, general references to the *avrBs3* family often identified the genes as *avrBs3/pthA* family members. The need for a new term arose, in part, to accommodate a growing number of effectors, in general, with unknown functions or phenotypes, avoid the confusion of the term "avirulence" in plant and animal bacterial pathogenicity terminology, and simplify the nomenclature. Due to the potent

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_1,

acidic transcription activator-like (TAL) domain in the C-terminal end, the name was born.

At the time of discovery, AvrBs3-related sequences were also reported to be found in other species of *Xanthomonas* [2]. The number of TAL effector genes with and without known phenotypic effects in other species rapidly grew, notably *X. oryzae* pv. *oryzae* and *X. campestris* pv. *malvacearum*, the causal agents of bacterial blight of rice and angular leaf spot of cotton, respectively [4, 5]. The gene *avrb6*, so named due to the presence of a cognate recessive resistance gene *b6* in cotton, confers water-soaking phenotype to strains of *X. campestris* pv. *malvacearum* with the gene on cotton leaves [5]. A series of genes, including *avrXa7*, *pthXo1*, *pthXo2*, and *pthXo3*, were found to be required for full virulence, or aggressiveness in some parlance, in *X. oryzae* pv. *oryzae* on rice [6–8]. AvrXa7 both elicits resistance in host plants with the R gene *Xa7* and is the major TAL effector virulence factor in some strains. One oddball was added to the collection, a truncated homology originally cloned as a truncated version known as *avrBsP* [9]. The gene, as cloned, ended within the repetitive domain-coding region and elicited a resistance response in some tomato lines. In addition to imperfect repeats within the repetitive region, several genes were cloned with 105 bp repeats and, at least, one gene was identified with a hybrid 102/105 repeat domain [10, 11]. More homologs were identified in strains of *Ralstonia* with the advent of routine genome sequencing [12], and, more recently, by structural similarity, in *Burkholderia rhizoxinica* [13].

2 Interesting from Any Perspective

The structure of the *avrBs3* gene product was impressive from the start. The gene contained near-perfect direct repeats within the middle of the protein of 102 base pairs, and the repetitive region of AvrBs3 was soon identified as a source of functional variation in regard to R gene recognition [1, 14]. The repetitive domain, consisting of varying unit repeat number and sequence polymorphisms, is also the source of the gross variations in the gene family among the various species of *Xanthomonas* [4, 15]. In the case of genes with phenotypic consequences, swapping repetitive domains among the family members, for the most part, results in genes that maintain the biological function of the gene, either resistance or virulence, from which the repetitive domain originated [7, 16, 17]. Alternations in the repetitive domain through intergenic and intragenic rearrangements give rise to new apparent R gene specificities and altered virulence function [14, 16, 18]. In the case of the dual-functioning TAL effector AvrXa7, spontaneous and selected repetitive region alterations resulted in loss of avirulence activity while maintaining virulence activity, indicating that selective pressure

against R gene deployment could potentially select for repeat variants that maintained virulence [18].

Clues as to the function of TAL effectors were first sought by comparison to known proteins and effectors. However, the only similarity outside of closely related genes in *Xanthomonas* and *Ralstonia* was tetratricopeptide repeat (TPR) proteins, which share a 34-amino acid repeat structure [19]. However, TPR proteins are involved in diverse functions, the repeats are highly polymorphic, and little insight into possible functions was garnered. The concept that type III substrate effectors were internalized into the host cells was first developed for animal bacterial pathogen effectors, and the concept, including AvrBs3, was soon extended for effectors of similar pathways in plant pathogens [20].

3 It Looks Like a Duck and Talks Like a Duck

The road to transcription factor-related hypotheses for TAL effectors started at about the same time upon recognition of functional nuclear localization signal in the C-terminal coding region of TAL effectors [21]. The functional significance of the NLS sequences was elegantly demonstrated for AvrBs3 with regard to avirulence activity in pepper and virulence activities of Avrb6 and PthA [18, 22]. Furthermore, avirulence activity of AvrBs3 could be restored by the substitution of NLS of the Simian virus 40 (SV40) T-antigen [18]. The involvement of nuclear localization was also subsequently shown for the avirulence activity of AvrXa10, AvrXa27, and AvrXa7 and the virulence activity of AvrXa7 [7, 23, 24]. TAL effectors also contain a potent eukaryotic-like acidic transcription activator (AAD) at the C-terminal end of the proteins, and the domain was required for the avirulence activity of AvrXa10 and AvrXa7 [7, 23]. Similar to nuclear localization, the AAD from the herpes virus VP16 protein could substitute for the requirement of the endogenous domain for AvrXa10 and AvrXa7 for avirulence activity [7, 25]. The involvement of the AAD was extended to AvrBs3 with regard to avirulence [26]. AvrBs3 was also recognized to be associated with hypertrophy and swelling in infected tissue, indicating that like AvrXa7, PthA, and Avrb6, the TAL effector served a function in bacterial virulence [27]. However, the AAD from VP16 could not substitute for the endogenous domain for virulence activities in either AvrXa7 or AvrBs3, indicating, at the time, a qualitative or quantitative difference in AAD function in the two functions [7, 27]. Given what is known regarding TAL effector function now, a reasonable hypothesis is that AvrXa7 and AvrBs3 are relatively weak transcription factors, and VP16 would likely function in virulence on a more robust TAL effector like PthXo1.

The features of the TAL effectors led to the hypothesis that the effectors affected or interacted in some way with the host

transcription machinery. One prediction was that the proteins would bind DNA, and, indeed, AvrXa7 was shown to be a double-stranded DNA-binding protein with preference for A/T-rich DNA [7]. However, no specific promoter or DNA-binding site was identified in the host. The second observation was the identification of AvrBs3-dependent transcripts from pepper that were cycloheximide resistant and dependent on the salient features of the TAL effector [27]. Cycloheximide resistance is thought to be indicative of an immediate effect and not due to de novo synthesis of endogenous transcription factors. At the same time, specific TAL effector-dependent host genes had not been identified with accompanying evidence for a causal relationship to the associated phenotypes.

4 Potential Host Target Genes Emerge

The first cloned gene related to the activity of a specific TAL effector was *Xa27* of rice [25]. Related ORFs were present in susceptible and resistant near-isogenic rice lines. However, only the allele in the resistant lines was expressed, and expression was only detected upon inoculation with strains of *Xanthomonas* with *avrXa27*. Transfer of the gene, including the promoter region, to an otherwise susceptible cultivar conferred *avrXa27*-dependent resistance. Three genes in rice were identified as candidate TAL effector targets in virulence. Principal among the genes was the one associated with the major TAL effector for virulence, PthXo1, which directed the expression of a host gene named *Os8N3* [1, 28]. Again, expression of the gene *Os8N3* was only detected in leaf tissue upon inoculations with bacterial strains carrying a specific TAL effector gene, in this case, *pthXo1*. Remarkably, *Os8N3* was mapped in rice to the position of a recessive resistance gene to bacterial blight of rice on chromosome 8 known as *xa13* [1, 29]. Alleles of *xa13* had rearrangements in the promoter region of the gene and were not expressed in a PthXo1-dependent manner [1]. PthXo1 is the major virulence TAL effector for strain PXO99 and, in the absence of *Os8N3* expression, the strain is virtually nonpathogenic on plants homozygous for *xa13*. Similarly, suppression of *Os8N3* expression by RNAi also rendered the plants resistant to infection. Expression of *Os8N3* renders the host susceptible to infection and, hence, *Os8N3* and related host genes are referred to as disease-susceptibility (S) genes. S genes follow somewhat the same rules as resistance (R) genes with the exception of the outcome.

5 It Is a Duck: TAL Effector-Specific DNA Binding

The association of AvrXa27 and PthXo1 with gene expression of specific host S and R genes still lacked evidence for the direct interaction of the TAL effectors with the respective genes. Direct binding evidence came with the identification of genes *Bs3* and *Upa20* of pepper as targets of AvrBs3 [30, 31]. In addition to the TAL effector host gene expression, AvrBs3 was shown to promote the expression of reporter genes under the control of the respective promoters when expressed transiently in plant tissue, and the protein was bound at higher levels to the respective promoters in gel mobilization analyses in comparison to effector proteins of varying specificity. Curiously, *Bs3* and *Upa20* encoded proteins unrelated to the products of *Xa27* or *Os8N3*. UPA20 is a helix-loop-helix transcription factor, which induces cell hypertrophy upon transient expression in *Nicotiana benthamiana* cells and is proposed to be the direct functional target of AvrBs3 in virulence [30]. *Os8N3* encodes a member of membrane-bound sugar transporter family [32]. Other TAL effector targeted S gene products include OsTFX1, a bZip transcription factor in rice that is induced by PthXo6, OsSultr3;6 that is a sulfate transporter induced by Tal2g of *X. oryzae* pv. *oryzicola*, and CsLOB1 that is a member of the lateral organ boundaries family from citrus [28, 33, 34]. The latter is the target of PthA-related TAL effectors involved in pustule formation in bacterial citrus canker [34, 35]. The products of the so-called suicide genes that are induced by AvrBs3, AvrXa27, AvrBs4, and AvrXa10 are equally diverse [24, 31, 36, 37].

The demonstration that TAL effectors bind DNA in a sequence-specific manner depending on the repetitive domain repeat structure and binding prediction based on promoter/repetitive domain amino acid residues provided a unifying model for TAL effector function [38, 39]. In addition to providing a boon to biotechnology, TAL effector technology has also facilitated our understanding of TAL effector function in bacterial disease. Despite the diversity of gene targets for avirulence and virulence, most natural effectors to date have been shown to target the predicted TATAA region in the respective promoters of the candidate S genes [34, 38, 40, 41]. Given that a few do not, for example PthXo1, TalC, and Tal2g, the data indicate that, while not a strict TATAA-binding factor, the TATAA box represents the most likely promoter sequence for a TAL effector to recognize in a complex genome [33, 34, 40–42]. Natural TAL effectors also do not necessarily represent an optimal repetitive domain for promoter binding to the respective intended host gene. At least four TAL effectors, AvrXa7, Tal2g, PthA4, and PthAw, are associated, individually, with the induction of multiple host genes due presumably to wobble in the binding specificity

and/or similarity in promoter sequences [33, 34, 40]. Artificial designer TAL effectors (dTALes) provide a rapid method to assess the functional consequences of host gene induction and discriminate between multiple candidate host S genes [33, 34, 43]. PthA4 and PthAw are both associated with the induction of *CsLOB1* and a second gene, *CsN3-1*, which is another member of the sugar transporter family. However, only dTALes targeting CsLOB1 are associated with pustule formation and enhanced bacterial leaf populations [34]. Similarly, Tal2g is associated with the induction of two rice genes in bacterial leaf streak. In addition to the sulfate transporter gene, *OsSultr3;6*, the locus *Os06g46500*, which encodes a predicted monocopper oxidase, is also up-regulated [33]. Again, only dTALes targeting *OsSultr3;6*, and not *Os06g46500*, compensated in terms of the disease phenotype for the loss of Tal2g function. The dTALes have also been used to identify potential S genes that have not been observed associated with field isolates of the pathogens [43]. Many questions remain regarding TAL effector function and the physiology affected by the accompanying host response. Undoubtedly, TAL effector biotechnology will contribute and facilitate obtaining answers to the questions.

References

1. Yang B, Sugio A, White FF (2006) *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. Proc Natl Acad Sci U S A 103:10503–10508
2. Bonas U, Stall RE, Staskawicz BJ (1989) Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. vesicatoria. Mol Gen Genet 218:127–136
3. Swarup S, Yang Y, Kingsley MT, Gabriel DW (1992) A *Xanthomonas citri* pathogenicity gene, *pthA*, pleiotropically encodes gratuitous avirulence on nonhosts. Mol Plant Microbe Interact 5:204–213
4. Hopkins CM, White FF, Choi S et al (1992) Identification of a family of avirulence genes from *Xanthomonas oryzae* pv. oryzae. Mol Plant Microbe Interact 5:451–459
5. De Feyter R, Yang Y, Gabriel DW (1993) Gene-for-genes interactions between cotton R genes and *Xanthomonas campestris* pv. malvacearum avr genes. Mol Plant Microbe Interact 6:225–237
6. Yang B, White FF (2004) Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. Mol Plant Microbe Interact 17:1192–1200
7. Yang B, Zhu W, Johnson LB, White FF (2000) The virulence factor AvrXa7 of *Xanthomonas oryzae* pv. oryzae is a type III secretion pathway-dependent, nuclear-localized, double-stranded DNA binding protein. Proc Natl Acad Sci U S A 97:9807–9812
8. Bai J, Choi SH, Ponciano G et al (2000) *Xanthomonas oryzae* pv. oryzae avirulence genes contribute differently and specifically to pathogen aggressiveness. Mol Plant Microbe Interact 13:1322–1329
9. Canteros B, Minsavage G, Bonas U et al (1991) A gene from *Xanthomonas campestris* pv. vesicatoria that determines avirulence in tomato is related to avrBs3. Mol Plant Microbe Interact 4:628–632
10. Kay S, Boch J, Bonas U (2005) Characterization of AvrBs3-like effectors from a Brassicaceae pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. Mol Plant Microbe Interact 18:838–848
11. Schornack S, Minsavage GV, Stall RE et al (2008) Characterization of AvrHah1, a novel AvrBs3-like effector from *Xanthomonas gardneri* with virulence and avirulence activity. New Phytol 179:546–556
12. Salanoubat M, Genin S, Artiguenave F et al (2002) Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature 415:497–502
13. Juillerat A, Bertonati C, Dubois G et al (2014) BurrH: a new modular DNA binding protein for genome engineering. Sci Rep 4:3831

14. Herbers K, Conrads-Strauch J, Bonas U (1992) Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. Nature 356:172–174
15. Yang Y, Yuan Q, Gabriel DW (1996) Water-soaking function(s) of XcmH1005 are redundantly encoded by members of the *Xanthomonas avr/pth* gene family. Mol Plant Microbe Interact 9:105–113
16. Yang Y, Gabriel DW (1995) Intragenic recombination of a single plant pathogen gene provides a mechanism for the evolution of new host specificities. J Bacteriol 177:4963–4968
17. Yang Y, De Feyter R, Gabriel DW (1994) Host-specific symptoms and increased release of *Xanthomonas citri* and *X. campestris* pv. malvacearum from leaves are determined by the 102-bp tandem repeats of pthA and avrb6, respectively. Mol Plant Microbe Interact 7:345–355
18. Yang B, Sugio A, White FF (2005) Avoidance of host recognition by alterations in the repetitive and C-terminal regions of AvrXa7, a type III effector of *Xanthomonas oryzae* pv. oryzae. Mol Plant Microbe Interact 18:142–149
19. D'Andrea LD, Regan L (2003) TPR proteins: the versatile helix. Trends Biochem Sci 28: 655–662
20. Van den Ackerveken G, Marois E, Bonas U (1996) Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. Cell 87:1307–1316
21. Yang Y, Gabriel DW (1995) *Xanthomonas* avirulence/pathogenicity gene family encodes functional plant nuclear targeting signals. Mol Plant Microbe Interact 8:627–631
22. Gabriel DW, Yuan Q, Yang Y, Chakrabarty PK (1996) Role of nuclear localizing signal sequences in three disease phenotypes determined by the *Xanthomonas avr/pth* gene family. In: Stacey G, Mullin B, Gresshoff PM (eds) Biology of plant-microbe interactions, vol 8. International Society for Molecular Plant-Microbe Interactions, St Paul, pp 197–202
23. Zhu W, Yang B, Chittoor JM et al (1998) AvrXA10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. Mol Plant Microbe Interact 11:824–832
24. Gu K, Yang B, Tian D et al (2005) R gene expression induced by a type-III effector triggers disease resistance in rice. Nature 435:1122–1125
25. Zhu W, Yang B, Wills N et al (1999) The C terminus of AvrXa10 can be replaced by the transcriptional activation domain of VP16 from the herpes simplex virus. Plant Cell 11:1665–1674
26. Szurek B, Marois E, Bonas U, Van Den AG (2001) Eukaryotic features of the *Xanthomonas* type III effector AvrBs3: protein domains involved in transcriptional activation and the interaction with nuclear import receptors from pepper. Plant J 26:523–534
27. Marois E, van den Ackerveken G, Bonas U (2002) The *Xanthomonas* type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. Mol Plant Microbe Interact 15:637–646
28. Sugio A, Yang B, Zhu T, White FF (2007) Two type III effector genes of *Xanthomonas oryzae* pv. oryzae control the induction of the host genes OsTFIIAγ1 and OsTFX1 during bacterial blight of rice. Proc Natl Acad Sci U S A 104:10720–10725
29. Chu Z, Yuan M, Yao J et al (2006) Promoter mutations of an essential gene for pollen development result in disease resistance in rice. Genes Dev 20:1250–1255
30. Kay S, Hahn S, Marois E et al (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318:648–651
31. Romer P, Hahn S, Jordan T et al (2007) Plant pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. Science 318:645–648
32. Chen L, Hou B, Lalonde S et al (2010) A novel class of sugar transporters essential for intercellular transport and nutrition of pathogens. Nature 468:527–532
33. Cernadas RA, Doyle EL, Nino-Liu DO et al (2014) Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. PLoS Pathog 10:e1003972
34. Hu Y, Zhang J, Jia H, Sosso D et al (2014) Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. Proc Natl Acad Sci U S A 111:E521–E529
35. Al-Saadi A, Reddy JD, Duan YP et al (2007) All five host-range variants of *Xanthomonas citri* carry one *pthA* homolog with 17.5 repeats that determines pathogenicity on citrus, but none determine host-range variation. Mol Plant Microbe Interact 20:934–943
36. Tian D, Wang J, Zeng X et al (2014) The rice TAL effector-dependent resistance protein XA10 triggers cell death and calcium depletion in the endoplasmic reticulum. Plant Cell 26:497–515

37. Strauss T, van Poecke RM, Strauss A et al (2012) RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. Proc Natl Acad Sci U S A 109:19480–19485
38. Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326: 1509–1512
39. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326(5959):1501
40. Antony G, Zhou J, Huang S et al (2010) *xa13* recessive resistance to bacterial blight is defeated by the induction of disease susceptibility gene *Os11N3*. Plant Cell 22: 3864–3874
41. Romer P, Recht S, Strauss T et al (2010) Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. oryzae. New Phytol 187:1048–1057
42. Yu Y, Streubel J, Balzergue S et al (2011) Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. oryzae depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. Mol Plant Microbe Interact 24:1102–1113
43. Streubel J, Pesce C, Hutin M et al (2013) Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. oryzae. New Phytol 200:808–819

Chapter 2

TAL Effector DNA-Binding Principles and Specificity

Annekatrin Richter, Jana Streubel, and Jens Boch

Abstract

Transcription activator-like effectors (TALEs) are proteins with a unique DNA-binding domain that confers both a predictable and programmable specificity. The DNA-binding domain consists typically of 34-amino acid near-identical repeats. The repeats form a right-handed superhelical structure that wraps around the DNA double helix and exposes the variable amino acids at position 13 of each repeat to the sense strand DNA bases. Each repeat binds one base in a highly specific, non-overlapping, and comma-free fashion. Although TALE specificities are encoded in a simple way, sophisticated rules can be taken into account to build highly efficient DNA-binding modules for biotechnological use.

Key words TALE, TALEN, CRISPR, Genome engineering, Genome editing, *Xanthomonas*

1 Introduction

The targeted binding of proteins to specific DNA sequences can be accomplished by a wide variety of protein folds. Typically, the protein-DNA interaction is mediated by several amino acids that together determine the specific recognition of DNA sequences. Accordingly, DNA-binding specificities of proteins are usually extremely difficult to modify. In contrast, a DNA-binding specificity that can be designed would enable wide-reaching possibilities for biotechnological applications. Recently, such a programmable DNA-binding specificity was identified in transcription activator-like effectors (TALEs), which are virulence proteins of plant-pathogenic *Xanthomonas* spp. bacteria [1]. After injection of TALEs into the cytoplasm of the plant cell via the bacterial type III secretion system, the proteins enter the plant cell nucleus where they bind to specific DNA sequences in promoter regions to activate transcription of downstream target genes [2–4].

Typically, TALEs are composed of three domains (Fig. 1). The N-terminal domain contains the type III secretion and translocation signal and is also important to initiate DNA binding [5, 6]. The C-terminal domain of TALEs contains nuclear localization signals

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_2,

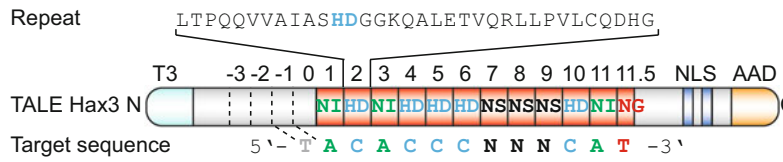


Fig. 1 Organization and DNA recognition of the natural TALE Hax3. The N-terminal type III secretion signal (T3, *light blue*) of natural TALEs is responsible for translocation by the type III secretion system of *Xanthomonas* spp. into the plant cell. TALEs contain nuclear localization signals (NLS, *dark blue*), an acidic activation domain (AAD, *orange*), and a central DNA-binding domain that consist of highly conserved, tandem 34 amino acid repeats (*red*). Each repeat recognizes one nucleotide of the target DNA sense strand in a consecutive manner and the amino acids 12 and 13 termed repeat variable diresidue (RVD) determine the base specificity. Recognition of an additional 5′ thymine is mediated by the degenerated repeat −1 in the N-terminal domain

and an acidic activation domain. Both features are prerequisites to activate plant gene expression [4, 7]. The central domain is composed of repeats that comprise a novel type of DNA-binding domain and are primarily responsible for specifying the TALE target sequence [2]. The repeat region is composed of tandem repeats of a near-identical 33–35 amino acid motif (Fig. 1). The most C-terminal repeat only contains 20 amino acids and is, therefore, termed "half" repeat. Among *Xanthomonas* TALEs the repeat region is highly conserved and mainly differs in the repeat number (from 1.5 to more than 30 repeats) [1]. The most significant differences in the amino acid sequences of repeats are found in positions 12 and 13, termed repeat variable diresidue (RVD) [1, 8, 9]. Interestingly, the number of TALE repeats determines the number of nucleotides in natural DNA target sequences and individual RVDs match to specific DNA bases [9, 10]. This pointed to a simple and novel one-to-one binding mode in which one repeat binds to one nucleotide in the target DNA sequence and the RVD specifies the base that is bound [9, 10]. In addition, an initial thymine precedes almost all target sequences. These recognition principles allow to predict DNA-binding specificities of TALEs with a given RVD sequence [11–13], and the straightforward construction of artificial TALEs with designer DNA-binding specificities [14–26]. The fusion of TALEs with different executer domains led to the design of TALE repressors, epigenetic modifiers, TALE nucleases (TALEN), TALE recombinases, and TALE transposases [27–29].

2 RVD Specificities

According to the TALE code the DNA-binding specificity of TALEs is determined by the RVDs of all repeats in a TALE and a preceding thymine [9, 10, 30]. Each RVD recognizes only one DNA base

without significant neighbor-dependent effects [9, 10, 21]. Initially, several rules have been defined for positional RVD requirements based on preferences in natural TALEs [15], but these have been disproven by more systematic studies [21, 31]. Until now 25 natural RVD types are known. Among these, the most commonly used RVDs HD (histidine, aspartic acid), NI (asparagine, isoleucine), and NG (asparagine, glycine) are highly specific for the nucleobases cytosine, adenine, and thymine, respectively (Fig. 2) [10]. For the base guanine, several alternative RVDs are used, NN (asparagine, asparagine) which recognizes guanine and to a lesser degree also adenine, or the guanine-specific RVDs NH (asparagine, histidine) and NK (asparagine, lysine) (Fig. 2) [32, 33]. For a flexible recognition, some RVDs recognize even more than two bases and the RVDs NS (asparagine, serine) and NA (asparagine, alanine) can be used as almost nonspecific universal ones (Fig. 2) [32, 33]. By now, all 400 theoretically possible RVD amino acid combinations have been analyzed [34]. Besides the naturally occurring RVDs, a few novel functional ones were discovered (Fig. 2). Notably, predominantly histidine, lysine, asparagine, and arginine at position 12 resulted in functional RVDs, somewhat similar to what is found in nature [34]. Amino acids accepted at position 13 reflect the specificities known from natural RVDs (e.g., KI, NI, RI specify for adenine; Fig. 2) [34].

3D structures uncovered the structural basis of the RVD-DNA interaction of TALE-DNA complexes [35–38]. In general, each repeat consists of two α-helices that expose the RVD loop to the DNA [35, 37]. Only the repeat variable residue 2 (RVR2, amino acid position 13) directly interacts with the corresponding base in the sense strand and thereby determines base specificity which matches well to the observed RVD specificities [35, 37]. Computational simulations implied that RVR2 might interact not only with its target base, but also with the preceding base [39], but a neighbor-dependent specificity of RVDs has not been detected experimentally, so far. In contrast to RVR2, the RVR1 (amino acid position 12) indirectly contributes to DNA binding. RVR1 stabilizes the RVD loop by interaction with amino acid 8 of the same repeat [35–37, 39]. Thereby, RVR1 can have an indirect influence on efficiency and apparent specificity of the RVR2-DNA interaction. In addition to the RVD-dependent DNA sequence-specific interactions, each TALE repeat binds to the sugar-phosphate backbone of the DNA sense strand via amino acids 14–17 (usually GGKQ) exerting a general DNA affinity [35–37].

According to the 3D structure the RVR2-base interaction is governed by at least two parameters: (1) Binding of RVR2 to the DNA base: This is mainly facilitated by amino acids interacting with target bases via hydrogen bonds (e.g., D-to-cytosine or N-to-purine rings) or van der Waals interactions [35, 37]. (2) Discrimination of nontarget bases: Several RVR2-base combinations lead to sterical clashes (e.g., D to thymine or N to pyrimidines) [39].

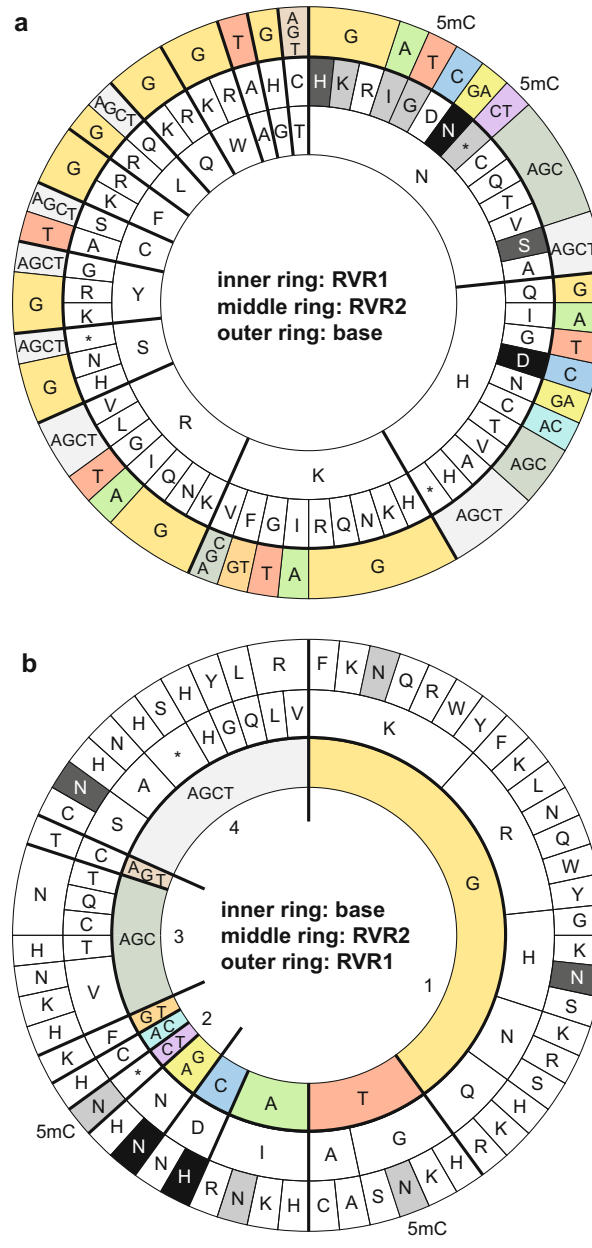


Fig. 2 RVD base specificities. The TALE code is shown in two circular presentations either grouped by RVD (**a**) or by base (**b**). (**a**) *Inner ring*, RVR1/position 12; *middle ring*, RVR2/position 13; *outer ring*, base (in color). (**b**) *Center*, number of different bases specified by RVD; *inner ring*, base (in color); *middle ring*, RVR2/position 13; *outer ring*, RVR1/position 12. RVD amino acids are abbreviated in single-letter code. Frequently used RVDs are shaded (*black*, strong RVD; *dark grey*, intermediate RVD; *light grey*, weak RVD). Please note that the RVD NN is strong for G, but weak for A, and that the RVD NS is intermediate for G and A, but weak for C and T. RVD specificities and efficiencies are assembled from these refs. 32–34. NG and N* also recognize 5-methyl-cytosine (5mC)

In summary, the RVD specificity is executed by "best matches" according to binding strength, sterical conformation, and charge distribution.

In some repeats the RVR2 is missing. In RVD nomenclature this is represented by an asterisk (*), e.g., H*, N*, and S*. Because of steric reasons these RVDs prefer pyrimidines in the target sequence to purines (Fig. 2). In addition, N* allows the positioning of 5-methylated cytosine in the target sequence [40]. The RVD NG also tolerates 5-methylated cytosine because of the structural similarity to the normally favored thymine nucleobase [40, 41]. This finding is relevant for the application of TALE derivatives in a genomic context because it allows the binding of potentially methylated DNA target sequences that are not tolerated by the cytosine-specific RVD type HD [40, 41]. In summary, specific TALE-DNA recognition is encoded by a direct and specific "one amino acid to one nucleotide" correlation (Fig. 2). This simple code makes TALE DNA-protein interactions unique.

3 RVD Efficiencies

The 3D structure and the observation that RVR2 defines the target base in the DNA suggest that this is the only important feature that has to be taken into account to build artificial TALEs with designer DNA-binding specificities. Surprisingly, some studies showed that TALEs and TALEN differ in their efficiency of target gene activation and modification, respectively, depending on their RVD composition [33, 42, 43] (Fig. 2). TALEs composed of exclusively so-called weak RVDs (e.g., NI, NG, NK) did not induce target gene expression [33]. This phenomenon was independent of whether the same RVD was repeated in tandem, or whether different weak RVDs were mixed [33]. The presence of at least three so-called strong RVDs (HD and NN) restored full functionality to the TALE. This suggests that a minimum number of strong RVDs is required for full activity. The efficiencies of the G-binding RVDs NN, NH, and NK were classified as strong, intermediate, and weak, respectively [33], which was subsequently corroborated using TALENs [42]. Semiquantitative DNA-binding assays using electromobility-shift assays (EMSA) suggested that the DNA-binding affinity of TALEs with strong NN and weak NK RVDs can differ 1000-fold [43]. TALE protein-DNA complex computation indicated that the non-base-specific interactions of the amino acids 16 (K) and 17 (Q) in a repeat to the sugar-phosphate backbone form an *oxyanion clip* and are significantly stronger than the specific RVD-base interactions [39]. In contrast, computational and experimental approaches showed that the distance to the DNA base T increases in stretches of tandem NG RVDs and thereby these weaken the overall DNA interaction capacity [35, 44].

Apparently, small RVDs (e.g., NG) compensate for steric imperfections of larger RVD-base interactions. In addition, strong RVDs might pull neighboring weak RVDs closer to the DNA leading to an increased DNA binding of NG. Together, this indicates that the RVD composition and succession are important for the overall TALE activity.

Three aspects contribute to RVD efficiency. (1) RVDs with similar specificity, but different RVR2: For example, the RVDs NN, NH, and NK recognize guanine, but with a variable efficiency [32, 33, 42]. In this case RVD specificity and efficiency are caused by RVR2-base interactions. (2) RVDs with similar specificity, but variable RVR1: In contrast to the high variability of RVR2 amino acids, only few different RVR1 amino acids are accepted, with a majority of asparagine (N) and histidine (H) [1, 34]. Because RVR1 mediates intra-repeat stability [35, 37] it can indirectly influence the RVD-DNA interaction. Accordingly, RVDs with similar specificity can significantly differ in their efficiency (e.g., HD is stronger than ND; [32]). (3) RVDs that recognize different bases with different efficiencies: The RVD NN recognizes both purine bases, but the interaction with guanine is preferred to adenine, because its negative charge distribution is more suitable for the interaction [36, 39]. Accordingly, NN is a strong RVD for guanine, but a weak one for adenine [10, 32, 33, 43]. NS is a universal RVD that recognizes all four DNA bases [10]. Nevertheless, the RVR2 serine favors the interaction with purines over pyrimidines [39]. Accordingly, the RVD NS has a high efficiency for adenine and guanine and a low one for thymine and cytosine.

In summary, specificity and efficiency are separate features of RVDs. Both need to be taken into account when constructing artificial TALEs. Conveniently, only few (i.e., 3 out of 18) strong RVDs need to be present in a TALE for full activity and most designed TALEs automatically follow this rule [33, 42]. In addition, the possibility to adjust the level of specificity for each target site position makes TALEs and TALENs a more flexible tool than the CRISPR/Cas system.

4 TALE Repeat Number

Naturally occurring *Xanthomonas* TALEs contain 1.5–33.5 repeats, but it is unknown whether all of these TALEs are functional [1]. Artificial TALEs required a minimal number of 6.5 repeats for gene activation *in planta* and strong activity required at least 10.5 repeats, but the minimum requirement depends on the RVD composition of the TALE [10]. A TALE with 10.5 repeats that only contained weak and intermediate RVDs (NI, NG, NH) did not induce expression of a reporter gene [33]. Either, the incorporation of strong RVDs into the 10.5 repeat-array or

extension of the array to 17.5 repeats resulted in active TALEs [33]. An alternative study showed that AvrBs3 derivatives with fewer repeats than the wild-type protein required at least 11.5 repeats for strong gene activation [30]. Both of these observations indicate that a higher number of repeats has a positive impact on DNA binding. In contrast, VP64-TALE activators with 14.5–24.5 repeats targeting different regions in the human *VEGFA* gene were active and showed no clear repeat number dependency [31]. Apparently, the number of repeats was sufficient in these cases for TALE function, and the choice of target region had a stronger impact on TALE performance than the number of repeats [31]. This suggests that TALE activity depends on the cumulative contribution of each repeat to DNA binding, but only until a threshold for maximal activity is achieved.

A high specificity of TALE binding is desired for most biotechnological applications. The fewer repeats a TALE has the more possible target boxes are statistically found in a genome, and the larger the target genome the more likely TALE target boxes occur. Less specific TALEs and TALENs might be cytotoxic, because they induce multiple genes and mutate multiple target sites, respectively [21]. In the human genome of about 3.27 giga base pairs a unique target sequence should have a minimum of $\log_4 (3.27 \times 10^9 \text{ bp}) \approx 16$ bp. A specific TALE would require at least 14.5 repeats to recognize 16 bp (including the initial thymine) without taking into account the abundance of particular genomic sequences and tolerated mismatches. Accordingly, to design orthogonal (unique) TALE transcriptional activators in human cells TALEs with 18.5 repeats (target box of 20 bp) were chosen [45]. Only target sequences with a high specificity, because of the presence of at least three mismatches to all possible 20 bp sequences in 2000 bp human genomic promoter regions, were considered. The TALEs did not contain RVDs specific for more than one base. To target G, the RVD NK was used instead of NN [45]. This bioinformatics approach demonstrates how TALE repeat arrays can be designed to be highly specific in a genomic context.

Dimeric TALENs work in pairs of tail-to-tail-oriented single-TALE-F*ok*I fusion proteins with a defined spacing of typically 12–24 bp [46–48]. The *Fok*I domains of each TALEN are required to dimerize to facilitate DNA cleavage and subsequent mutations. Because a TALEN pair is required for an active nuclease, each TALEN monomer could in principle contain less repeats than a TALE to obtain a similar site specificity for the TALEN pair. Accordingly, a systematic analysis of TALENs with 8.5–19.5 repeats showed that all TALENs were comparably active in a gene-disruption assay in human cell culture [21]. Nevertheless, TALENs with less than 11.5 repeats were significantly cytotoxic in contrast to TALENs with more repeats, possibly due to off-target activities [21].

In summary, defining a minimal repeat array length for a highly specific TALE derivative with good functional performance depends on several factors: the efficiencies of the RVDs used, the target genome size, and the use of TALE or TALEN. Notably, the possibility to choose either longer, more specific or shorter, less specific TALEs and TALENs distinguishes them from the target-size-restricted CRISPR/Cas system [49].

5 TALE Mismatch Tolerance

Natural TALEs often activate target genes by binding to boxes that are not perfect according to the RVD specificities [2, 3, 9, 10, 50–52]. This implies that TALEs naturally have a certain level of mismatch tolerance. At the same time, a few additional changes in natural target sequences can render TALEs completely incapable to function [3, 52, 53]. The cutoff where TALE binding is critically compromised likely depends on overall TALE binding efficiency (i.e., number of repeats and RVD efficiencies), the position of mismatches, and the steric implications of the individual mismatches.

Some experimental data indicate that mismatches particularly at N-terminal repeats have a higher impact than those at later positions [43, 45, 54]. The reason for this has not been determined, but it is possible that the initial TALE repeats play a particular role in the switch from general DNA binding ("scanning mode") to specific DNA binding ("target recognition") of TALEs. Possibly, TALE binding onto target sequences begins from the N-terminal domain and mismatches in the first repeat positions compromise further condensation.

TALEs with few repeats tolerate less mismatches than TALEs with many repeats [54]. In addition, for overall binding the negative impact of a mismatch is far stronger than the positive contribution of one matching repeat-DNA pair. Artificial TALEs with 17.5 repeats failed to recognize most target sequences with three mismatches or had 10 % remaining reporter activation efficiency [54]. This mismatch tolerance was further reduced if the TALE contained less repeats. An artificial TALE with 13.5 repeats and three mismatches had only about 1 % of maximum gene induction efficiency, and the 9.5 repeats TALE did not even tolerate one mismatch in its target sequence [54]. Accordingly, in another study TALEs with 14.5 repeats completely lost target recognition of boxes with three or more mismatches [10].

In contrast, the mismatch tolerance of paired TALENs at their target sites might be higher than it is the case for single-acting TALE derivatives. One TALEN bound to a target site might function as an anchor and dimerization of the nuclease domains subsequently stabilizes binding of the second TALEN even to significantly less well-matching target sites. Off-target analyses revealed that

TALENs with 14.5 repeats tolerate up to six mutations in their target sequences [55]. In another example, TALENs with 17.5 repeats tolerated four mismatches with about 30 % remaining activity [54]. In addition, TALENs with three mismatches at the 3′ end of the target sequence were active whereas TALEs were not [56]. This indicated that a correct binding of the distal repeats is more important for TALEs than for TALENs, possibly because dimerization of the *Fok*I nuclease does not require a complete binding of the repeat region to DNA. Together, these examples point to a higher mismatch tolerance of TALEN pairs in comparison to TALEs.

In summary, highly stringent target specificity can be achieved in two different ways. (1) TALEs and TALENs with few repeats will tolerate only few mismatches. (2) TALEs and TALENs with many repeats will tolerate more mismatches, but will also have a higher overall specificity in a genomic context.

6 Variations in the Repeat Backbone

Natural *Xanthomonas* TALE repeats differ with low frequency at any non-RVD amino acid except G14 and L29 [1], but the impact of these non-RVD substitutions on the TALE-DNA interaction is largely unknown. It is possible that modifications in single repeats that do not disrupt the general repeat fold are tolerated in the context of a normal repeat array.

The most frequent non-RVD variations occur at positions 4 and 24 [57] where often either charged, polar, or hydrophobic amino acids are found (e.g., position 4: D, E, A; position 24: R, A). These negatively and positively charged amino acids form electronegative and electropositive strips, respectively, along opposite sites of the superhelical TALE protein [35, 38]. These charged protein regions were suggested to control that TALE binding occurs to the DNA-leading strand only, because the negative charges repel the protein from the negative charges of the DNA-lagging strand sugar-phosphate backbone [35, 38]. Initial experiments to analyze the impact of non-RVD variations have shown that TALE specificity is determined by RVDs only and not by non-RVDs, but overall TALE performance can vary somewhat [58]. In another study, artificial TALEs with repetitive arrangements of NN RVDs were constructed that contained either an acidic amino acid (D) or a polar one (Q) at position 4 in each repeat [33]. The TALE with aspartic acid (D) at position 4 was somewhat more active in reporter gene studies *in planta* than the TALE with glutamine (Q) [33]. Recently, TALENs with either a fixed combination of amino acids at position 4 (D, E, A) and 32 (D, A) or repeat-to-repeat variations of different combinations thereof were constructed. The TALENs

with variations in non-RVDs tend to have a higher activity, but the effect was not completely consistent and the underlying rules are still unclear [59].

Another frequent non-RVD repeat variation is the number of amino acids per repeat. Most often TALE repeats are 34 amino acids in length, but variants with 33–35 amino acids length frequently occur. 33 amino acid repeats are missing the 13th amino acid (e.g., N*) and 35 amino acid repeats have an additional proline inserted after amino acid 32 [1, 60]. Both length variants function similarly to 34 amino acid repeats [10, 33]. In nature, other TALE repeat length variations (30, 39, 40, or 42 amino acids) also exist, but in very low frequency [1]. A recent functional study of these aberrant repeats revealed that they confer a surprising flexibility to TALEs and TALENs that is absent from TALEs with common 33–35 amino acid repeats [56]. A single long or short repeat within a repeat array of 34 amino acid repeats enables the TALE or TALEN to efficiently recognize two alternative target sequences: one according to the usual one-repeat-to-one-base binding mode and an alternative one with a one base pair deletion in the vicinity of the aberrant repeat [56]. The two alternative binding modes suggest that the aberrant repeat can either insert into the repeat array like a normal repeat or loop out with the following repeats shifting forward by one position. This flexibility offers the unique opportunity to build TALEs or TALENs with a dual specificity for allelic indel mutation variants.

7 TALE-Family Proteins

TALE-family proteins occur in only very few other species besides *Xanthomonas* spp [61]. TALE gene activators in plant-pathogenic *Ralstonia solanacearum* bacteria contain a highly similar repeat region with 35 amino acid repeats, but N- and C-terminal domains that are different from *Xanthomonas* TALEs. The *Ralstonia* TALE repeats differ at some non-RVD amino acid positions from *Xanthomonas* repeats as well as from repeat to repeat (Fig. 3) [33, 62, 63]. Nevertheless, the DNA-recognition specificity of *R. solanacearum* TALE RVDs is near identical to *Xanthomonas* TALE RVDs. In this case, variations in non-RVD positions apparently only have a limited impact on the repeat array structure, base preference, and RVD efficiency [62]. TALE-related proteins with 33 amino acid repeats and very short N- and C-terminal regions are present in symbiotic *Burkholderia rhizoxinica* bacteria that live inside hyphae of the Zygomycete fungus *Rhizopus microsporus* [61, 64, 65]. The repeats of these proteins are variable and share only few non-RVD amino acids with TALEs from *Xanthomonas* or *Ralstonia* (Fig. 3). Nevertheless, the repeat region forms a TALE-like superhelical structure and the repeats bind to DNA with the

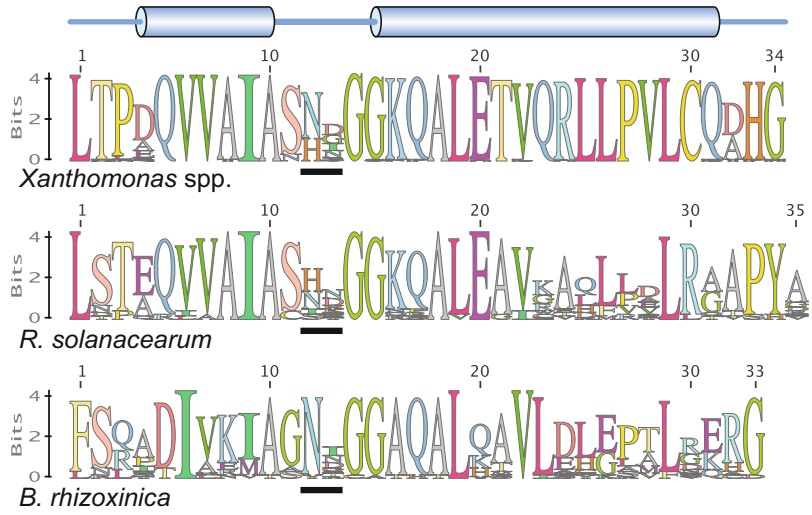


Fig. 3 Repeat conservation of TALE-family proteins. Amino acid sequence logos of repeats from *Xanthomonas* spp. TALEs (2080 repeats), *Ralstonia solanacearum* Brg11 (18 repeats), and *Burkholderia rhizoxinica* BurrH (21 repeats), respectively. The base-specifying RVDs are *underlined. Blue cylinders* indicate the two alpha helices of the TALE repeat structure [37]

same RVD code as *Xanthomonas* TALEs and can therefore be classified together with *Xanthomonas* and *Ralstonia* TALEs into a protein family with highly related DNA-binding properties.

Accordingly, they can also be exploited to generate site-specific nucleases and artificial gene activators, but only using fused functional domains, because the proteins have no apparent gene activation activity themselves [66–68]. Notably, individual repeats cannot be freely combined without compromising overall activity [66], suggesting that the variations at non-RVD positions lead to significant interdependencies of neighboring repeat architectures. One solution is the variation of RVDs with simultaneously keeping each repeat backbone at its fixed position [67, 68]. Although *Ralstonia* and *Burkholderia* TALE-family proteins are more difficult to use than *Xanthomonas* TALEs, it illustrates that non-RVDs can vary in members of this protein family, but the standard RVD code still exerts the key DNA-binding specificity.

8 Contribution of N- and C-Terminal Parts to Binding

To apply TALEs for biotechnology it is beneficial to truncate the proteins to a minimum size necessary for DNA binding. The repeat domain without N- and C-terminal protein domains could specifically bind target DNA, but with lower efficiency than the full-length protein [2].

Systematic studies were performed to optimize the length of the N- and C-terminal regions. Using heterologous activators in human cells, TALE-VP64 fusions with 63 or 121 C-terminal residues are similarly active compared to 177 C-terminal residues, whereas C-terminal domains shorter than 63 amino acids permitted only 50–60 % activity [26]. In other studies, TALE-VP16 activators displayed full activity with 17, 47, and 95, but not 22 C-terminal amino acids, respectively [47, 48]. Similarly, TALENs with 28, 47, and 63 C-terminal amino acids, respectively, performed well, considering that the optimal spacer between the TALEN pairs depends on the number of C-terminal residues, because it functions as linker to the endonuclease domain [47, 48].

In contrast to C-terminal truncations, TALEs are more sensitive to N-terminal truncations. Truncations of 158 or more amino acids of the 287-amino acid-long N-terminal part resulted in less than 60 % induction activity [26]. Similarly, TALEs and TALEN with 152 deleted N-terminal amino acids (=136 residual amino acids) work well, whereas 237 deleted (=50 residual) N-terminal amino acids permitted no activity [6, 47, 48, 69, 70] (Fig. 4a). Apparently, TALEs require a minimal N-terminal length for efficient DNA binding and deletion of the 152 N-terminal amino acids is a good compromise.

A fixed thymine (termed T_0) typically preceding target sites of natural TALEs was identified and experimentally shown to be required [9, 10]. TALE crystal structures identified four degenerated repeats in the N-terminal part directly adjacent to the canonical repeats [5]. They were termed repeat 0 to –3 and encompass 127 residues adjacent to the repeats [5, 37, 38]. The functional N-terminal truncations contain nearly all of these residues indicating that they confer a function indispensable for TALE-DNA binding. In addition to T_0 recognition, the N-terminal TALE domain was shown to mediate unspecific DNA binding [5].

Several studies addressed the possibility to change or obliterate the T_0 specificity of TALEs to avoid this base dependency for designer TALEs. Structural data suggested that the tryptophan W232 located in repeat –1 makes contact to T_0 by van der Waals interactions [37] (Fig. 4b). Targeted mutation and molecular evolution of the repeat –1 loop residues were done to achieve variants with modified specificities for T_0 [30, 71–73]. A loop variant with the tryptophan substituted for an arginine and a second substitution in the neighborhood permitted specificity for an initial guanine [72] (Fig. 4b). This observation is supported by the guanine specificity of the TALE homolog Brg11 from *Ralstonia solanacearum* which also carries an arginine instead of a tryptophan at this position [62]. Other loop mutations of repeat –1 allowed tolerance of all four nucleotides at position zero with different extent [72, 73] (Fig. 4b). The specificities and efficiencies of the

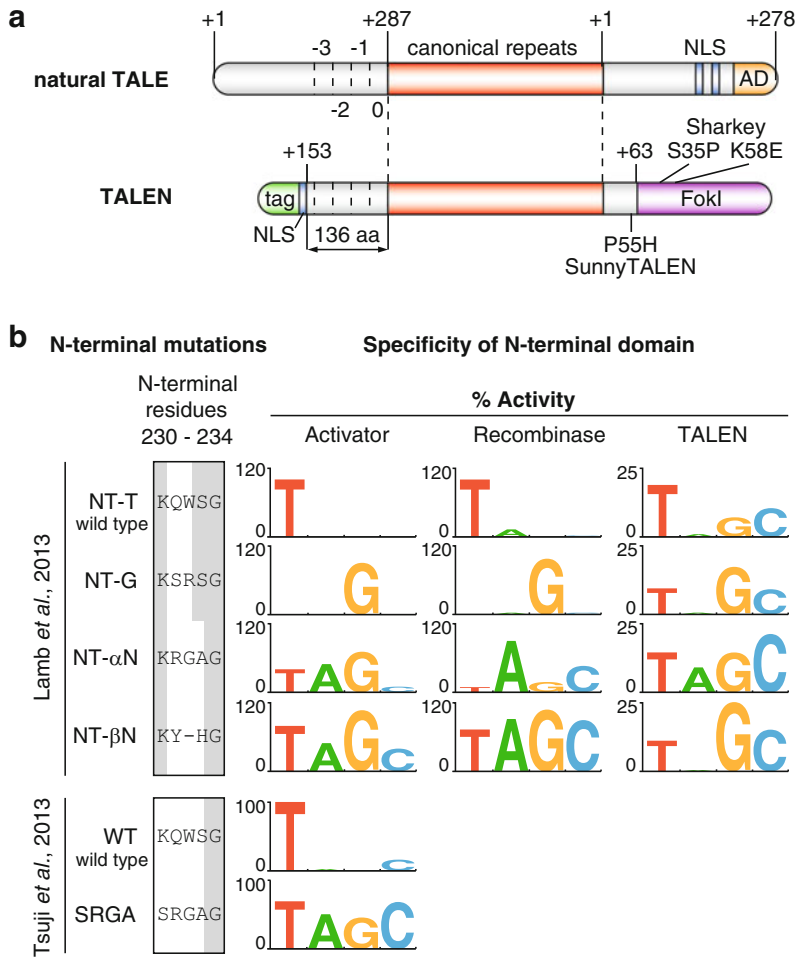


Fig. 4 Most frequently used TALEN setup and TALEs with alternative N-termini. (**a**) In a typical TALEN design, the backbone of naturally occurring TALEs is truncated to 136 residues (+153) of the N-terminal and 63 residues of the C-terminal part. A heterologous NLS to substitute for the deleted original NLS and optional tags can be attached to the N-terminal part. The *FokI* endonuclease domain is typically fused to the C-terminal part of both TALEN for dimerization in a tail-to-tail orientation. Substitutions in the C-terminal TALE domain (SunnyTALEN) and the *FokI* domain (Sharkey mutations) have been associated with higher TALEN activity. (**b**) Activity of TALE activators, recombinases, and nucleases, respectively, with alternative N-termini [72, 73] in % nucleotide-specific activity of the original N-terminus. N-terminal residues are numbered with respect to the natural TALE

novel N-termini differed slightly in in vitro binding studies and in applying them in TALE activators, TALENs and TALE recombinases, respectively [72]. No solutions for the specific recognition of an initial A or C were found. Nevertheless, TALE binding to DNA can in some cases be independent of T_0. Substitution of the first RVD of AvrBs3 from HD to other RVDs resulted in an

apparent loss of T_0 dependency, although this also depended on overall RVD composition and repeat number of the TALE [30]. This indicates that the T_0 dependency is particularly evident in the context of an HD RVD at position 1 in the canonical repeat array [30]. It also suggests a cooperation of the N-terminal domain with the first canonical repeat in TALE binding. In summary, either N-terminal variants or non-HD RVDs in the first repeat are possibilities for a broader position zero nucleotide tolerance.

A recent study aimed to improve the activity of TALENs by directed evolution of the N- and C-terminal parts of the protein [74]. A single-amino acid substitution (termed SunnyTALEN, Fig. 4) was identified that increased the TALEN activity in yeast and in human cells more than 2.5-fold in combination with the Sharkey mutations of the *Fok*I domain [74].

9 Construction Guidelines for TALEs and TALE Derivatives

1. Choose a target site of a suitable size for the given genome, approx. ($\log_4$ (bp_{genome}) = $bp_{TALEbox}$) with a minimum of 11.5 repeats and preceded by a thymine. A TALE with approx. 17.5 repeats is usually a good choice for high specificity and activity.
2. The target site should not contain elements that are found in high frequency in the target genome.
3. Choose an alternative N-terminal domain, if it is required to substitute the initial thymine to another base.
4. Design a repeat array using the standard RVDs (NI, HD, NN, NG) with at least two or better three strong RVDs (HD, NN) distributed over the array.
5. Avoid long (>4) stretches of identical RVDs.
6. Use NH to bind guanines with high specificity.
7. Use NG or N* to bind 5-methyl-cytosine.
8. Choose N- and C-terminal truncation variants depending on the desired application (e.g., nuclease, activator, repressor, recombinase).

Acknowledgements

We thank U. Bonas and T. Schreiber for discussion and comments on the manuscript. This work was supported by DFG grants BO 1496/6-1 and BO1496/7-1 to JB.

References

1. Boch J, Bonas U (2010) *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48:419–436
2. Kay S, Hahn S, Marois E, Hause G, Bonas U (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318:648–651
3. Römer P, Hahn S, Jordan T, Strauß T, Bonas U, Lahaye T (2007) Plant-pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. Science 318:645–648
4. Van den Ackerveken G, Marois E, Bonas U (1996) Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. Cell 87:1307–1316
5. Gao H, Wu X, Chai J, Han Z (2012) Crystal structure of a TALE protein reveals an extended N-terminal DNA binding region. Cell Res 22:1716–1720
6. Szurek B, Rossier O, Hause G, Bonas U (2002) Type III-dependent translocation of the *Xanthomonas* AvrBs3 protein into the plant cell. Mol Microbiol 46:13–23
7. Zhu W, Yang B, Chittoor JM, Johnson LB, White FF (1998) AvrXa10 contains an acidic transcriptional activation domain in the functionally conserved C terminus. Mol Plant Microbe Interact 11:824–832
8. Bonas U, Stall RE, Staskawicz BJ (1989) Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. Mol Gen Genet 218:127–136
9. Moscou MJ, Bogdanove AJ (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 et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512
11. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, VanDyk JK et al (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40:W117–W122
12. Grau J, Wolf A, Reschke M, Bonas U, Posch S, Boch J (2013) Computational predictions provide insights into the biology of TAL effector target sites. PLoS Comput Biol 9:e1002962
13. Pérez-Quintero AL, Rodriguez-R LM, Dereeper A, López C, Koebnik R, Szurek B et al (2013) An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. PLoS One 8:e68464
14. Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P et al (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40:e117
15. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C et al (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82
16. Geißler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens S-E et al (2011) Transcriptional activators of human genes with programmable DNA-specificity. PLoS One 6:e19509
17. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29:699–700
18. Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH et al (2011) Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res 39:6315–6325
19. Liang J, Chao R, Abil Z, Bao Z, Zhao H (2014) FairyTALE: a high-throughput TAL effector synthesis platform. ACS Synth Biol 3:67–73
20. Morbitzer R, Elsaesser J, Hausner J, Lahaye T (2011) Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res 39:5790–5799
21. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30:460–465
22. Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F (2012) A transcription activator-like effector toolbox for genome engineering. Nat Protoc 7:171–192
23. Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V (2013) A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol 31:76–81
24. Wang Z, Li J, Huang H, Wang G, Jiang M, Yin S et al (2012) An integrated chip for the high-throughput synthesis of transcription activator-like effectors. Angew Chem Int Ed Engl 51:8505–8508
25. Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S (2011) Assembly of designer TAL effectors by golden gate cloning. PLoS One 6:e19722

26. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29:149–153
27. Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. Trends Cell Biol 23:390–398
28. Mercer AC, Gaj T, Fuller RP, Barbas CF III (2012) Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic Acids Res 40:11163–11172
29. Owens JB, Mauro D, Stoytchev I, Bhakta MS, Kim MS, Segal DJ et al (2013) Transcription activator like effector (TALE)-directed *piggyBac* transposition in human cells. Nucleic Acids Res 41:9197–9207
30. Schreiber T, Bonas U (2014) Repeat 1 of TAL effectors affects target specificity for the base at position zero. Nucleic Acids Res 42:7160–7169
31. Maeder ML, Linder SJ, Reyon D, Angstman JF, Fu Y, Sander JD et al (2013) Robust, synergistic regulation of human gene expression using TALE activators. Nat Methods 10:243–246
32. Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3:968
33. Streubel J, Blücher C, Landgraf A, Boch J (2012) TAL effector RVD specificities and efficiencies. Nat Biotechnol 30:593–595
34. Yang J, Zhang Y, Yuan P, Zhou Y, Cai C, Ren Q et al (2014) Complete decoding of TAL effectors for DNA recognition. Cell Res 24:628–631
35. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK et al (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. Science 335:720–723
36. Deng D, Yan C, Wu J, Pan X, Yan N (2014) Revisiting the TALE repeat. Protein Cell 5:297–306
37. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 335:716–719
38. Stella S, Molina R, Yefimenko I, Prieto J, Silva G, Bertonati C et al (2013) Structure of the AvrBs3-DNA complex provides new insights into the initial thymine-recognition mechanism. Acta Crystallogr D Biol Crystallogr 69:1707–1716
39. Wicky BIM, Stenta M, Dal Peraro M (2013) TAL effector specificity stems from negative discrimination. PLoS One 8:e80261
40. Valton J, Dupuy A, Daboussi F, Thomas S, Marechal A, Macmaster R et al (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem 287:38427–38432
41. Deng D, Yin P, Yan C, Pan X, Gong X, Qi S et al (2012) Recognition of methylated DNA by TAL effectors. Cell Res 22:1502–1504
42. Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y et al (2012) Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS One 7:e45383
43. Meckler JF, Bhakta MS, Kim MS, Ovadia R, Habrian CH, Zykovich A et al (2013) Quantitative analysis of TALE-DNA interactions suggests polarity effects. Nucleic Acids Res 41:4118–4128
44. Wan H, Hu JP, Li KS, Tian XH, Chang S (2013) Molecular dynamics simulations of DNA-free and DNA-bound TAL effectors. PLoS One 8:e76045
45. Garg A, Lohmueller JJ, Silver PA, Armel TZ (2012) Engineering synthetic TAL effectors with orthogonal target sites. Nucleic Acids Res 40:7584–7595
46. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A et al (2010) TAL effector nucleases create targeted DNA double-strand breaks. Genetics 186:757–761
47. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF et al (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148
48. Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res 39:9283–9293
49. Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32:347–355
50. Antony G, Zhou J, Huang S, Li T, Liu B, White F et al (2010) Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene Os-*11N3*. Plant Cell 22:3864–3876
51. Römer P, Recht S, Strauß T, Elsaesser J, Schornack S, Boch J et al (2010) Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. New Phytol 187:1048–1057
52. Strauß T, van Poecke RM, Strauß A, Römer P, Minsavage GV, Singh S et al (2012) RNA-seq pinpoints a *Xanthomonas* TAL-effector activated

resistance gene in a large-crop genome. Proc Natl Acad Sci U S A 109:19480–19485

53. Römer P, Strauß T, Hahn S, Scholze H, Morbitzer R, Grau J et al (2009) Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper *Bs3* alleles. Plant Physiol 150:1697–1712
54. Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S et al (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol 31: 833–838
55. Osborn MJ, Starker CG, McElroy AN, Webber BR, Riddle MJ, Xia L et al (2013) TALEN-based gene correction for epidermolysis bullosa. Mol Ther 21:1151–1159
56. Richter A, Streubel J, Blücher C, Szurek B, Reschke M, Grau J et al (2014) A TAL effector repeat architecture for frame shift binding. Nat Commun 5:3447
57. Bonas U, Conrads-Strauch J, Balbo I (1993) Resistance in tomato to *Xanthomonas campestris* pv. *vesicatoria* is determined by alleles of the pepper-specific avirulence gene *avrBs3*. Mol Gen Genet 238:261–269
58. Morbitzer R, Römer P, Boch J, Lahaye T (2010) Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci U S A 107:21617–21622
59. Sakuma T, Ochiai H, Kaneko T, Mashimo T, Tokumasu D, Sakane Y et al (2013) Repeating pattern of non-RVD variations in DNA-binding modules enhances TALEN activity. Sci Rep 3:3379
60. Kay S, Boch J, Bonas U (2005) Characterization of AvrBs3-like effectors from a *Brassicaceae* pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. Mol Plant Microbe Interact 18: 838–848
61. Schornack S, Moscou MJ, Ward ER, Horvath DM (2013) Engineering plant disease resistance based on TAL effectors. Annu Rev Phytopathol 51:383–406
62. de Lange O, Schreiber T, Schandry N, Radeck J, Braun KH, Koszinowski J et al (2013) Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. New Phytol 199: 773–786
63. Li L, Atef A, Piatek A, Ali Z, Piatek M, Aouida M et al (2013) Characterization and DNA-binding specificities of *Ralstonia* TAL-like effectors. Mol Plant 6:1318–1330
64. Lackner G, Moebius N, Partida-Martinez L, Hertweck C (2011) Complete genome sequence of *Burkholderia rhizoxinica*, an Endosymbiont of *Rhizopus microsporus*. J Bacteriol 193:783–784
65. Partida-Martinez LP, Hertweck C (2005) Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature 437: 884–888
66. de Lange O, Wolf C, Dietze J, Elsaesser J, Morbitzer R, Lahaye T (2014) Programmable DNA-binding proteins from *Burkholderia* provide a fresh perspective on the TALE-like repeat domain. Nucleic Acids Res 42: 7436–7449
67. Juillerat A, Bertonati C, Dubois G, Guyot V, Thomas S, Valton J et al (2014) BurrH: a new modular DNA binding protein for genome engineering. Sci Rep 4:3831
68. Stella S, Molina R, Lopez-Mendez B, Juillerat A, Bertonati C, Daboussi F et al (2014) BuD, a helix-loop-helix DNA-binding domain for genome modification. Acta Crystallogr D Biol Crystallogr 70:2042–2052
69. Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M et al (2012) Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci U S A 109:17382–17387
70. Sun N, Liang J, Abil Z, Zhao H (2012) Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Biosyst 8:1255–1263
71. Doyle EL, Hummel AW, Demorest ZL, Starker CG, Voytas DF, Bradley P et al (2013) TAL effector specificity for base 0 of the DNA target is altered in a complex, effector- and assay-dependent manner by substitutions for the tryptophan in cryptic repeat −1. PLoS One 8:e82120
72. Lamb BM, Mercer AC, Barbas CF III (2013) Directed evolution of the TALE N-terminal domain for recognition of all 5′ bases. Nucleic Acids Res 41:9779–9785
73. Tsuji S, Futaki S, Imanishi M (2013) Creating a TALE protein with unbiased 5′-T binding. Biochem Biophys Res Commun 441:262–265
74. Sun N, Bao Z, Xiong X, Zhao H (2014) SunnyTALEN: a second-generation TALEN system for human genome editing. Biotechnol Bioeng 111:683–691

Chapter 3

The Development of TALE Nucleases for Biotechnology

David G. Ousterout and Charles A. Gersbach

Abstract

The development of a facile genome engineering technology based on transcription activator-like effector nucleases (TALENs) has led to significant advances in diverse areas of science and medicine. In this review, we provide a broad overview of the development of TALENs and the use of this technology in basic science, biotechnology, and biomedical applications. This includes the discovery of DNA recognition by TALEs, engineering new TALE proteins to diverse targets, general advances in nuclease-based editing strategies, and challenges that are specific to various applications of the TALEN technology. We review examples of applying TALENs for studying gene function and regulation, generating disease models, and developing gene therapies. The current status of genome editing and future directions for other uses of these technologies are also discussed.

Key words Genome engineering, Gene editing, Gene therapy, Gene targeting

1 Introduction

The emergence of transcription activator-like effector nucleases (TALENs) has made genome-editing tools widely accessible to any laboratory with basic molecular biology expertise. The development of the TALEN technology and its use in various biotechnological applications build on the considerable progress in genome editing over the previous decade with other approaches. Accordingly, the availability of the TALEN technology over the past few years has led to numerous advances in genome editing in a diverse range of cell types and organisms. This facile genome editing approach has facilitated new strategies to model disease, develop novel genetic therapies, or create desired phenotypic properties through highly specific rewriting of the genome. In this chapter, the development and use of TALEN technologies are reviewed and discussed.

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_3,

1.1 Genome-Editing Systems

Genome editing with engineered site-specific endonucleases has emerged as a technology to selectively replace or correct disrupted genes, in contrast to conventional genetic engineering methods of gene addition [1, 2]. There are numerous platforms for generating site-specific gene modifications in the genome, but to date the most successful have been based on zinc finger nucleases [3, 4], TALENs [5, 6], and, more recently, the RNA-guided CRISPR/Cas9 system [7–9]. These systems are at present the most developed publicly available platforms for robust and efficient targeted gene editing. In particular, the recent development of TALENs and CRISPR/Cas9 has dramatically advanced genome editing due to their ease of engineering and efficient genetic modification [6–20]. Other systems in development include meganucleases [21, 22], triplex-forming oligonucleotide (TFO) complexes [23], and programmable recombinases based on zinc finger protein [24–27] or TALE DNA-binding domains [28]. Historically, meganucleases have been difficult to engineer due to interdependence of the DNA-binding and cleavage domains, although recent developments in directed evolution of meganucleases [29–31] and fusion of meganucleases to TALE DNA-binding proteins [32, 33] are providing promising new opportunities with this technology. TFO complexes have thus far been limited by relatively low levels of gene modification, but oligonucleotide-mediated gene editing can be improved with the incorporation of TALENs [34]. Programmable recombinases are a promising next-generation gene editing technology, but target site requirements, overall efficiency, and unknown off-target effects are still major challenges to the widespread adoption of this technology [35].

1.2 Nuclease-Mediated Genome Editing

Engineered nucleases generate targeted genome modifications by creating a targeted double-strand break in the genome that stimulates cellular DNA repair through either homology-directed repair (HDR) or nonhomologous end joining (NHEJ) [36, 37] (Fig. 1). Briefly, HDR uses a designed synthetic donor DNA template to guide repair and can be used to create specific sequence changes to genome, including the targeted addition of whole genes. HDR has enabled integration of gene cassettes of up to 8 kb in the absence of selection at high frequency (~6 %) in human cells [38]. Generally, gene correction strategies have been based solely on HDR, the efficiency of which is dependent on the genomic target, cell type, cell-cycle state, and efficient delivery of an exogenous DNA template [39–43]. In many cases, antibiotic selection is used in tandem with genome editing for gene correction in cell types with low levels of HDR repair [40–42]. In contrast to genome modification by HDR, the template-independent religation of DNA ends by NHEJ is a stochastic, error-prone repair process that introduces random small insertions and deletions at the DNA breakpoint (Fig. 1). Gene editing by NHEJ has been used in mammalian cells

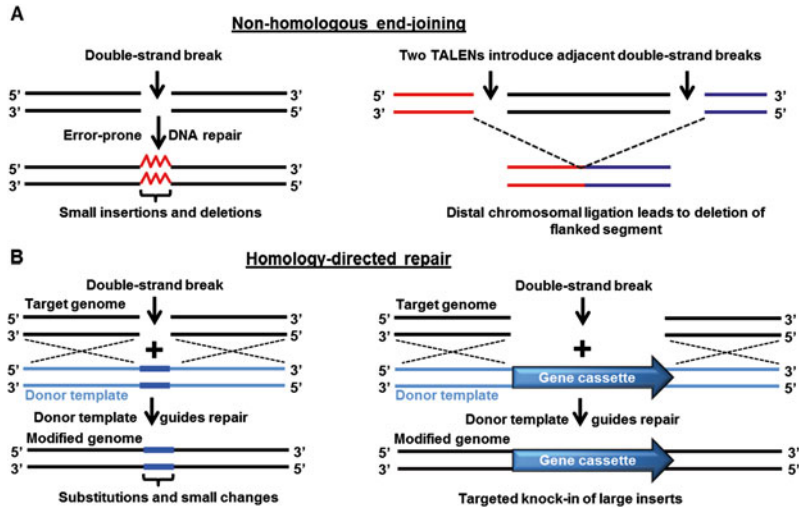


Fig. 1 Mechanisms of DNA repair following nuclease-induced double-strand breaks. (**a**) In the absence of a DNA repair template, the break is repaired by nonhomologous end joining, which is an error-prone process and can lead to small insertions or deletions. Alternatively, two adjacent nuclease-induced breaks can be used to excise the intervening chromosomal DNA from the genome. (**b**) If a DNA repair template is provided with homology to the target site surrounding the break, it will be used to guide homology-directed repair. In this way, particular small changes to the DNA sequence or the insertion of whole-gene expression cassettes can be directed to specific genome target sites

to disrupt genes [44, 45], delete chromosomal segments [46–48], or restore aberrant reading frames [49, 50]. This chapter reviews how TALENs have been used to exploit NHEJ and HDR DNA repair processes to create highly specific changes to a desired gene.

2 Development of TALENs

2.1 TALE DNA Recognition

In 2009, two landmark studies described the simple and modular TALE DNA-binding domain [14, 15]. These novel DNA-binding proteins are naturally occurring transcriptional activators from the plant pathogen *Xanthomonas*. As reported in these studies, the TALE DNA-binding domain consists of numerous tandem repeats, with each repeat specifying recognition of a single base pair of DNA. Importantly, single-base-pair recognition by each repeat is determined by alteration of only two hypervariable amino acids, termed repeat variable diresidues (RVDs), and each repeat appears to recognize DNA in a modular manner. This simple mode of DNA recognition was confirmed in structural studies of a naturally occurring TALE bound to its cognate DNA target [51, 52]. These discoveries were quickly expanded upon to create novel TALE proteins by engineering a custom RVD array to recognize a user-specified DNA target [53–55].

The only sequence requirement for TALE binding is that each target site be immediately preceded by a 5′-thymine for efficient DNA recognition, although more recently modified proteins have been developed to accept other nucleotides at this position [56, 57]. These novel DNA-binding domains were then fused to transcriptional activator domains [53, 55], nuclease catalytic domains [11, 54, 55], epigenetic modifying domains [58, 59], and recombinases [28] to generate an array of programmable enzymes for manipulating genes in complex genomes.

Although naturally occurring TALEs have a modular RVD recognition code, several studies have shown that some RVDs, specifically those targeting guanosine, display unexpected recognition of degenerate bases in the context of engineered TALE DNA-binding domains [55, 60]. However, more specific RVDs, such as NH or NK for recognition of guanosine, can result in significantly reduced activity of the reengineered TALE protein [60–62]. Recently, a publicly available web server has been developed that generates TALE targets utilizing more specific RVDs predicted to have minimal impact on activity [63]. Other publicly available web servers are available to assist in generating RVD arrays that are predicted to have high activity and specificity [64–66]. Together, these studies demonstrate the overall robustness of TALE DNA recognition and its utility in generating highly active nucleases at novel targets of interest.

2.2 Assembly of RVD Arrays to Create Customized TALE DNA-Binding Domains

Synthesizing custom TALE DNA-binding domains requires contiguous assembly of many RVD repeats, each only differing by two amino acids, into a destination TALE array. The large number of repeats, typically 15–20, makes this process difficult with conventional recombinant DNA technology. To overcome this technical challenge, several approaches have been developed that iteratively assemble new TALE arrays in a highly efficient and rapid manner. Custom TALE arrays can be rapidly created from a relatively small library of plasmids using publicly available reagents utilizing "Golden Gate" molecular cloning techniques to assemble new arrays within a few days [13, 53]. These methods are simple and only require reagents and equipment commonly found in molecular biology labs, although the overall throughput of assembly is limited. Other protocols are well suited to high-throughput generation of TALE arrays using solid-phase assembly [12, 67] or ligation-independent cloning techniques [68]. Notably, with the proper equipment, these high-throughput assembly methods are able to generate dozens to hundreds of TALEN constructs in 1 day. Alternatively, TALE arrays can also be custom ordered and pre-validated through commercial sources such as Life Technologies, Cellectis Bioresearch, Transposagen Biopharmaceuticals, and System Biosciences.

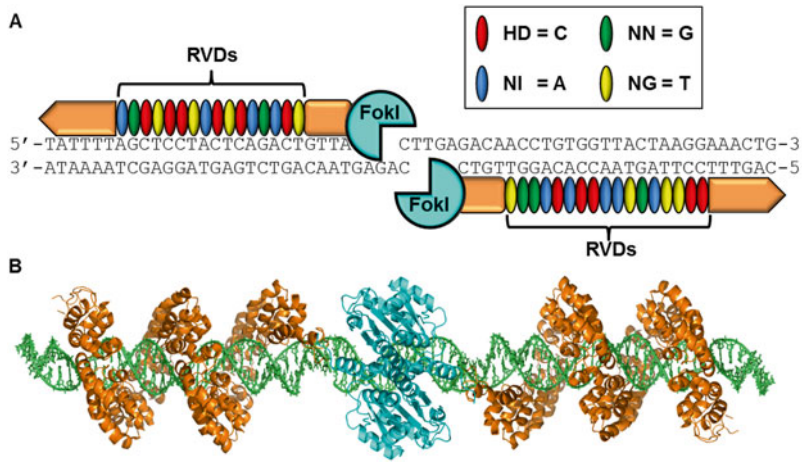


Fig. 2 TALEN architecture and structure. (**a**) The TALE DNA-binding domain consists of the array of RVDs engineered to recognize specific sequences, along with fixed N- and C-terminal domains (***orange***), fused to the catalytic domain of the ***Fok***I endonuclease (***blue***). (**b**) Schematic of the TALEN structure, with TALEs (***orange***, PDB 3UGM) fused to the ***Fok***I domain (***blue***, PDB 2FOK) on DNA (***green***)

2.3 TALE Nuclease Architectures

Conventionally, TALEN monomers are created as a fusion of the TALE DNA-binding domain to the nonspecific endonuclease catalytic domain of *Fok*I. Site-specific double-strand breaks are created when two separate nuclease monomers bind to adjacent target DNA sequences on opposite strands in a tail-to-tail fashion, thereby permitting dimerization of *Fok*I and cleavage of the target DNA (Fig. 2). Thus, since *Fok*I acts as a dimer, TALENs are designed in pairs to guide two separate *Fok*I monomers to a desired target site. Several TALEN architectures have been described that demonstrate improved nuclease activity by truncating the C-terminus of the TALE DNA-binding domain [11, 55, 69]. These studies also show that the translocation domain on the TALE N-terminus can be removed without impacting activity. Moreover, these truncations can be used to restrict the length of the sequence allowed between the TALEN monomers [55] and may be useful for restricting potential off-target mutagenesis. Directed evolution of the TALE DNA-binding domain has also yielded mutants that have higher observed gene editing activity against episomal and chromosomal targets [70]. Alternate nuclease catalytic domains are also possible; for example, fusions of TALEs to monomeric meganucleases have recently been shown to improve targeting of these enzymes [32].

2.4 Enhancement of Nuclease Activity

Several improvements have been made to enhance the specificity of the *Fok*I chimeric nucleases. A major advance was the identification of mutations that require heterodimerization of the nuclease pairs [71–73], thereby preventing potential homodimerization of nuclease monomers at unintended target loci. Furthermore, introduction

of distinct obligate heterodimer mutations can be used to create two independent TALENs by preventing unexpected interactions between monomers from either pair [48]. Introduction of inactivating mutations to the *Fok*I domain on one of the two nuclease domains in each pair can be used to generate targeted nickases. The single-strand nicks generated by these enzymes facilitate high levels of HDR but do not stimulate error-prone NHEJ repair [74, 75]. TALE nickases therefore display significantly reduced mutagenesis at off-target loci. Finally, directed evolution was utilized to find mutations that enhance the activity of *Fok*I in a target site-independent manner [76].

2.5 Relaxation of the 5′-Thymine Targeting Requirement

The range of DNA sequences that can be targeted by TALEs is constrained by a strict requirement of a thymine base at the zero base position (N0) [55]. The crystal structure of a natural TALE protein suggests that there is a cryptic repeat domain in the N-terminus of the protein that specifically recognizes thymine [51, 52]. Novel TALE architectures have been developed to overcome this requirement by engineering this region of the TALE N-terminus to recognize alternative bases at this position [56, 57] or by utilizing TALE-like domains from related plant pathogens [77, 78] that naturally recognize guanine at the N0 position. However, DNA-binding activity of these TALE architectures may be reduced, especially for targets with adenosine and cytosine bases at the N0 position. Further work in this area may yield TALE scaffolds that can readily target sequences with any base at the N0 position with high efficiency.

2.6 Targeting Methylated DNA

The methylation status of the target DNA locus is known to impact DNA binding of TALE proteins, particularly with chromosomal targets directly containing 5′-methylated cytosine (5mC) [79, 80]. As a result, DNA methylation can significantly reduce or completely eliminate TALE binding. Methylation analysis of a target loci can be used to generate TALENs targeted to open chromatin; however this further restricts the utility of TALENs for site-specific gene modification. Global demethylation of a target genome using chemical modifiers such as 5′-aza-2′-deoxycytidine can rescue TALE binding [79]; however these methods are commonly associated with undesirable toxicity. More attractive methods have been developed that substitute specific RVDs in TALE proteins to efficiently bind particular methylated and/or demethylated cytosines in the target sequence. Thus, TALE proteins can be reengineered either to be insensitive to cytosine methylation by using the N* RVD that binds to both cytosine and 5mC [81] or by utilizing RVDs that specifically recognize 5mC (NG) or cytosine (HD) [82]. By substituting these particular RVDs, TALENs can be engineered to target these sites with high efficiency. It is also noteworthy that

TALEs have been shown to target regions that are insensitive to DNase I, indicating that these proteins are able to access sites located in heterochromatin [83]. These studies were performed in dividing cells, and future work is necessary to determine the role of DNA replication in facilitating access to these target sites.

2.7 Delivery of TALENs

TALEN monomers are readily delivered by DNA expression cassettes or directly as mRNA by conventional transfection methods. However, the size of TALEN monomers and the highly repetitive array of RVD sequences present a significant challenge to viral delivery of TALEN constructs, thereby potentially limiting their utility in some gene editing applications. Adenovirus presents an attractive delivery vehicle for delivering gene constructs encoding both TALEN monomers [84], although adenovirus has limited tropism in some cell types and is highly immunogenic. Interestingly, this study also demonstrated that lentivirus was unable to deliver intact TALEN gene cassettes, due to rearrangements in the TALEN-coding region caused by the repetitive structure of RVD arrays. This limitation was overcome by the development of recoded TALEN constructs, termed re-TALEs, that can be efficiently expressed by lentiviral delivery [20], although this method may require reoptimization and synthesis of each new TALE gene. In contrast to DNA or mRNA delivery, direct protein delivery of TALENs can be achieved by utilizing cell-penetrating peptides covalently bound to purified TALEN proteins [85]. This method enables efficient genome editing in cells without the risk of spontaneous integration of the TALEN DNA expression construct into the genome that can be caused by non-viral and viral gene delivery. Furthermore, previous evidence suggests that protein delivery of gene-editing nucleases may reduce off-target activity by limiting the duration of nuclease exposure [86].

3 Applications in Basic Science and Biotechnology

Conventional genetic engineering methods involve the addition of new genes to cellular genomes by random integration of foreign genetic material into the chromosomal DNA. In contrast, genome editing using engineered nucleases enables precise manipulation at nearly any desired locus with high efficiency. Importantly, genome editing can generate a variety of genetic mutations without leaving any exogenous DNA sequences in the target genome. The development of high-throughput TALE assembly methods, in combination with high success rates of engineering highly active TALEN pairs, has resulted in the unprecedented ability to manipulate any gene of interest in a diverse array of organisms (Table 1). As one example of the breadth of TALEN assembly and applicability, libraries of TALENs have been generated to target 18,740 human protein-coding genes [80].

Table 1
Examples of biotechnology applications of TALEN-mediated gene modification

Type of modification	Organism	Genes	References
Gene disruption	Human	CCR5	[11, 55, 116, 117]
	Human	*MSTN*	[120]
	Human	Hepatitis B virus	[118, 119]
	Pig	LDL receptor	[93]
	Bovine	ACAN12, p65	[93]
	C. elegans	*ben-1*, *rex-1*, *sdc-2*	[97]
	C. elegans, *Pristionchus pacificus*, *Caenorhabditis species 9*	*ben-1*, *smo-1*, *rex-1*, *sdc-2*, *unc-119*	[96]
	Rice	OsSWEET14	[97]
	Newt	*P. waltl* tyrosinase	[98]
	Silkworm	*BmBlos2*	[99]
	Rat	*IgM*	[92]
	Mouse	*Pibf1*	[125]
	Arabidopsis thaliana	*ADH1*, *TT4*, *MAPKKK1*, *DSK2B*, and *NATA2*	[106]
	Drosophila	*ry*, *y*, *Psf2*, *Sld5*, *Pcd*, *CG12200*, *CG7224*, *CG11594*	[100]
	Barley	HvPAPhy_a	[107]
	Rice	*OsDEP1*, *OsBADH2*, *OsCKX2*, *OsSD1*	[105]
	Brachypodium	*BdABA1*, *BdCKX2*, *BdSMC6*, *BdSPL*, *BdSBP*, *BdCOI1*, *BdRHT*, *BdRHTA1*	[105]
	Mosquito (*A. gambiae*)	TEP1	[101]
	Frog	*noggin*, *ptf1a/p48*, *ets1*, *hhex*, *vpp1*, *foxd3*, *sox9*, and *grp78/bip*	[102]
Gene substitution and/or addition	Human	*OCT4*, *PITX4*, AAVS1	[10]
	Zebrafish	*th* (tyrosine hydroxylase), *fam46c*, *smad5*	[126]
	Drosophila	*ry*	[100]
	C. elegans, *Pristionchus pacificus*, *Caenorhabditis species 9*	*ben-1*, *smo-1*, *rex-1*, *sdc-2*, *unc-119*	[96]
	Mouse	*Sry*, *Ury*	[90]
Gene deletion or inversion	Human	Various microRNAs	[88]
	Arabidopsis thaliana	*GLL22a*, *GLL2b*	[106]
	Mouse	Various microRNAs	[91]
	Zebrafish	Multiple genes and noncoding regions	[47]

(continued)

Table 1 (continued)

Type of modification	Organism	Genes	References
Gene correction	Human	*DMD*	[50]
	Human	*COL7A1*	[108]
	Human	*HBB*	[109–111]
	Human	*APOB*, *SORT1*, *AKT2*, *PLIN1*	[87]
	Human	*XPC*	[112]
	Human	Mitochondrial DNA	[114]
	Human	*AAT*	[113]

A powerful application of the TALEN technology is to rapidly and efficiently generate cellular models of human disease or to interrogate disease-related mutations or genes. This approach has been exploited to create disease-associated genetic mutations in somatic and stem-cell models for a variety of human diseases [87]. Notably, in this study, few if any TALEN-associated off-target mutations were detectable in many of the modified cell populations. High-throughput TALEN assembly was also used to interrogate a large panel of genes related to epigenetic regulation or cancer, with successful modification of >85 % of targeted genes [12]. The ease of TALEN technologies has enabled researchers to rapidly generate large genomic deletions to quickly interrogate microRNA function [88, 89]. These notable examples demonstrate that TALENs are a versatile tool to interrogate and study small and large genetic elements in complex genomes.

TALENs have also enabled rapid gene modification to efficiently generate transgenic species or to knockout genes of interest. This has enabled the study of a variety of genes of interest in a diverse range of organisms, including mice [90, 91], rats [92], pigs [93], cows [93], zebrafish [47, 94, 95], *C. elegans* [96, 97], newts [98], silkworm [99], flies [100], mosquitos [101], and frogs [102]. In addition, genome engineering is an exciting method to address challenges in plant engineering [103, 104]. Many plant genes are arranged in tandem arrays, making it difficult to selectively alter single genes to study or impart new gene function. The ability of TALENs to discriminate between relatively few mismatches makes this technology particularly powerful for altering specific gene arrays. An example of this approach is the application of TALENs in rice to generate disease resistance, as well as the rapid modification of numerous other genes [105]. Other studies have demonstrated that TALENs are a powerful platform to rapidly modify plant genes, including *Arabidopsis thaliana* [106], barley [107], and *Brachypodium* [105].

4 Applications for Gene Therapies

Gene therapies using designer nucleases have shown promise to correct the genetic basis of human diseases [1, 2]. The significant advances made in the efficiency and precision of novel genome engineering technologies across the past decade have led to the development of TALENs targeted to numerous genes related to a range of human diseases (Table 1). In contrast to gene replacement therapies, genome editing can directly correct mutations associated with disease. For example, we developed TALENs to generate small insertions and deletions to restore the reading frame of the dystrophin gene as a novel method to correct the molecular basis of Duchenne muscular dystrophy [50]. TALENs have also been used to correct mutations associated with epidermolysis bullosa [108], sickle cell disease [109, 110], beta-thalassemia [111], xeroderma pigmentosum [112], and alpha-1 antitrypsin deficiency [113] by homologous recombination and to correct mitochondrial DNA disorders [114] by deletion of aberrant sequences.

Beyond correction of mutant genes, gene editing strategies have been developed to modify genes in order to modulate disease phenotypes. ZFNs targeted to the gene encoding the HIV-coreceptor CCR5 are currently in clinical trials and have laid the groundwork for genome editing as a novel treatment modality [44, 115]. Studies have demonstrated that TALENs can also introduce efficient mutations to CCR5 [11, 55, 116, 117] and present an alternative gene editing technology for this application. TALENs have also been designed to target and eliminate hepatitis B viral genomes from human cells [118, 119]. TALENs have been utilized to disrupt the myostatin gene [120], the loss of which leads to hypertrophy of skeletal muscle that could be used to treat a range of diseases, including muscular dystrophies. Collectively, these studies show that TALENs are a powerful technology to generate a variety of gene modifications to correct human diseases.

5 Discussion

Over the past 5 years, the rapid advancement of genome editing technologies has led to widespread adoption of various gene editing platforms for a diverse range of applications [1–6]. TALEN technologies have made effective gene-editing tools accessible to nearly any researcher at low cost. The robustness of this technology has enabled researchers to rapidly and efficiently interrogate a large number of genes in a range of organisms (Table 1). Importantly, TALENs have impressive observed specificity and several advances in this field have further improved the fidelity of this approach [56, 57, 60, 61, 63, 121]. The specificity and

efficiency of these approaches may be further improved as second-generation technologies are developed, such as TALE recombinases [28] and single-chain TALE-meganuclease fusions [32, 33]. The easily programmable TALE DNA-binding domain has also been a boon to creating other synthetic enzymes to regulate gene expression [53, 83, 122] and the epigenome [59, 58]. Although the recent advent of CRISPR/Cas9-based genome-engineering tools has provided an alternative facile method for gene editing [7, 123, 124], there are many differences between the two technologies and various applications could benefit from the strengths of each approach. Collectively, TALENs and other TALE-based gene-modifying tools have introduced publicly available, low-cost, efficient, and rapid gene modification that is accessible to any lab and has enabled studies for a remarkable variety of applications.

Acknowledgements

This work was supported by a US National Institutes of Health (NIH) Director's New Innovator Award (DP2OD008586), National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award (CBET-1151035), NIH R01DA036865, NIH R21AR065956, NIH UH3TR000505, NIH P30AR066527, the Duke Coulter Translational Partnership, and an American Heart Association Scientist Development Grant (10SDG3060033). D.G.O. was supported by an American Heart Association Mid-Atlantic Affiliate Predoctoral Fellowship.

References

1. Perez-Pinera P, Ousterout DG, Gersbach CA (2012) Advances in targeted genome editing. Curr Opin Chem Biol 16(3–4):268–277
2. Gaj T, Gersbach CA, Barbas CF III (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 31(7):397–405
3. Urnov FD, Rebar EJ, Holmes MC et al (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11(9):636–646
4. Gersbach CA, Gaj T, Barbas CF III (2014) Synthetic zinc finger proteins: the advent of targeted gene regulation and genome modification technologies. Acc Chem Res 47: 2309–2318
5. Mussolino C, Cathomen T (2012) TALE nucleases: tailored genome engineering made easy. Curr Opin Biotechnol 23:644–650
6. Joung JK, Sander JD (2012) TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14(1):49–55
7. Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32:347–355
8. Cong L, Ran FA, Cox D et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339(6121):819–823
9. Mali P, Yang L, Esvelt KM et al (2013) RNA-guided human genome engineering via Cas9. Science 339(6121):823–826
10. Hockemeyer D, Wang H, Kiani S et al (2011) Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29(8):731–734
11. Mussolino C, Morbitzer R, Lütge F et al (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res 39(21):9283–9293
12. Reyon D, Tsai SQ, Khayter C et al (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30:460–465

13. Cermak T, Doyle EL, Christian M et al (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39(12):e82
14. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326(5959):1501
15. Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326(5959):1509–1512
16. Cho SW, Kim S, Kim JM et al (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31(3):230–232
17. Hou Z, Zhang Y, Propson NE et al (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc Natl Acad Sci U S A 110(39):15644–15649
18. Jinek M, East A, Cheng A et al (2013) RNA-programmed genome editing in human cells. eLife 2:e00471
19. Horii T, Tamura D, Morita S et al (2013) Generation of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system. Int J Mol Sci 14(10):19774–19781
20. Yang L, Guell M, Byrne S et al (2013) Optimization of scarless human stem cell genome editing. Nucleic Acids Res 41(19):9049–9061
21. Redondo P, Prieto J, Muñoz IG et al (2008) Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. Nature 456(7218):107–111
22. Silva G, Poirot L, Galetto R et al (2011) Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. Curr Gene Ther 11(1):11–27
23. Parekh-Olmedo H, Kmiec EB (2007) Progress and prospects: targeted gene alteration (TGA). Gene Ther 14(24):1675–1680
24. Gersbach CA, Gaj T, Gordley RM et al (2011) Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. Nucleic Acids Res 39(17):7868–7878
25. Gaj T, Mercer AC, Gersbach CA et al (2011) Structure-guided reprogramming of serine recombinase DNA sequence specificity. Proc Natl Acad Sci U S A 108(2):498–503
26. Gordley RM, Gersbach CA, Barbas CF III (2009) Synthesis of programmable integrases. Proc Natl Acad Sci U S A 106(13): 5053–5058
27. Gaj T, Mercer AC, Sirk SJ et al (2013) A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. Nucleic Acids Res 41(6):3937–3946
28. Mercer AC, Gaj T, Fuller RP et al (2012) Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic Acids Res 40(21):11163–11172
29. Jarjour J, West-Foyle H, Certo MT et al (2009) High-resolution profiling of homing endonuclease binding and catalytic specificity using yeast surface display. Nucleic Acids Res 37(20):6871–6880
30. Takeuchi R, Lambert AR, Mak AN et al (2011) Tapping natural reservoirs of homing endonucleases for targeted gene modification. Proc Natl Acad Sci U S A 108(32):13077–13082
31. Takeuchi R, Choi M, Stoddard BL (2014) Redesign of extensive protein-DNA interfaces of meganucleases using iterative cycles of in vitro compartmentalization. Proc Natl Acad Sci U S A 111(11):4061–4066
32. Boissel S, Jarjour J, Astrakhan A et al (2014) megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. Nucleic Acids Res 42(4):2591–2601
33. Beurdeley M, Bietz F, Li J et al (2013) Compact designer TALENs for efficient genome engineering. Nat Commun 4:1762
34. Rivera-Torres N, Strouse B, Bialk P et al (2014) The position of DNA cleavage by TALENs and cell synchronization influences the frequency of gene editing directed by single-stranded oligonucleotides. PLoS One 9(5):e96483
35. Gaj T, Sirk SJ, Barbas CF III (2014) Expanding the scope of site-specific recombinases for genetic and metabolic engineering. Biotechnol Bioeng 111(1):1–15
36. Rouet P, Smih F, Jasin M (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol 14(12):8096–8106
37. Rouet P, Smih F, Jasin M (1994) Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A 91(13):6064–6068
38. Moehle EA, Rock JM, Lee Y-L et al (2007) Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci 104(9):3055–3060
39. Urnov F, Miller J, Lee Y et al (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435:646–651

40. Sebastiano V, Maeder ML, Angstman JF et al (2011) In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. Stem Cells 29(11):1717–1726
41. Zou J, Mali P, Huang X et al (2011) Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. Blood 118(17): 4599–4608
42. Yusa K, Rashid ST, Strick-Marchand H et al (2011) Targeted gene correction of α1-antitrypsin deficiency in induced pluripotent stem cells. Nature 478(7369):391–394
43. Li H, Haurigot V, Doyon Y et al (2011) In vivo genome editing restores haemostasis in a mouse model of haemophilia. Nature 475(7355):217–221
44. Perez EE, Wang J, Miller JC et al (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 26(7):808–816
45. Holt N, Wang J, Kim K et al (2010) Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol 28(8):839–847
46. Lee HJ, Kim E, Kim JS (2010) Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome Res 20(1): 81–89
47. Gupta A, Hall VL, Kok FO et al (2013) Targeted chromosomal deletions and inversions in zebrafish. Genome Res 23(6):1008–1017
48. Sollu C, Pars K, Cornu TI et al (2010) Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. Nucleic Acids Res 38(22):8269–8276
49. Chapdelaine P, Pichavant C, Rousseau J et al (2010) Meganucleases can restore the reading frame of a mutated dystrophin. Gene Ther 17(7):846–858
50. Ousterout DG, Perez-Pinera P, Thakore PI et al (2013) Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. Mol Ther 21(9): 1718–1726
51. Mak AN, Bradley P, Cernadas RA et al (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 335(6069):716–719
52. Deng D, Yan C, Pan X et al (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. Science 335(6069): 720–723
53. Zhang F, Cong L, Lodato S et al (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29(2):149–153
54. Christian M, Cermak T, Doyle EL et al (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186(2):757–761
55. Miller JC, Tan S, Qiao G et al (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29(2):143–148
56. Doyle EL, Hummel AW, Demorest ZL et al (2013) TAL effector specificity for base 0 of the DNA target is altered in a complex, effector- and assay-dependent manner by substitutions for the tryptophan in cryptic repeat −1. PLoS One 8(12):e82120
57. Lamb BM, Mercer AC, Barbas CF III (2013) Directed evolution of the TALE N-terminal domain for recognition of all 5′ bases. Nucleic Acids Res 41(21):9779–9785
58. Konermann S, Brigham MD, Trevino AE et al (2013) Optical control of mammalian endogenous transcription and epigenetic states. Nature 500(7463):472–476
59. Maeder ML, Angstman JF, Richardson ME et al (2013) Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat Biotechnol 31(12):1137–1142
60. Christian ML, Demorest ZL, Starker CG et al (2012) Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS One 7(9):e45383
61. Streubel J, Blucher C, Landgraf A et al (2012) TAL effector RVD specificities and efficiencies. Nat Biotechnol 30(7):593–595
62. Cong L, Zhou R, Kuo YC et al (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3:968
63. Lin Y, Fine EJ, Zheng Z et al (2014) SAPTA: a new design tool for improving TALE nuclease activity. Nucleic Acids Res 42:e47
64. Doyle EL, Booher NJ, Standage DS et al (2012) TAL effector-nucleotide targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40(Web Server issue):W117–W122
65. Sander JD, Maeder ML, Reyon D et al (2010) ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res 38(Suppl 2):W462–W468
66. Fine EJ, Cradick TJ, Zhao CL et al (2014) An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. Nucleic Acids Res 42(6):e42

67. Briggs AW, Rios X, Chari R et al (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40(15):e117
68. Schmid-Burgk JL, Schmidt T, Kaiser V et al (2013) A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol 31(1):76–81
69. Bedell VM, Wang Y, Campbell JM et al (2012) In vivo genome editing using a high-efficiency TALEN system. Nature 491(7422): 114–118
70. Sun N, Bao Z, Xiong X et al (2014) SunnyTALEN: a second-generation TALEN system for human genome editing. Biotechnol Bioeng 111:683–691
71. Miller JC, Holmes MC, Wang J et al (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25(7):778–785
72. Doyon Y, Vo TD, Mendel MC et al (2010) Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods 8(1):74–79
73. Szczepek M, Brondani V, Buchel J et al (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol 25(7):786–793
74. Ramirez CL, Certo MT, Mussolino C et al (2012) Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. Nucleic Acids Res 40(12): 5560–5568
75. Wang J, Friedman G, Doyon Y et al (2012) Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. Genome Res 22(7):1316–1326
76. Guo J, Gaj T, Barbas C III (2010) Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. J Mol Biol 400:96–107
77. de Lange O, Schreiber T, Schandry N et al (2013) Breaking the DNA-binding code of Ralstonia solanacearum TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. New Phytol 199(3):773–786
78. Li L, Atef A, Piatek A et al (2013) Characterization and DNA-binding specificities of Ralstonia TAL-like effectors. Mol Plant 6(4):1318–1330
79. Bultmann S, Morbitzer R, Schmidt CS et al (2012) Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res 40(12): 5368–5377
80. Kim Y, Kweon J, Kim A et al (2013) A library of TAL effector nucleases spanning the human genome. Nat Biotechnol 31(3):251–258
81. Valton J, Dupuy A, Daboussi F et al (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem 287(46):38427–38432
82. Deng D, Yin P, Yan C et al (2012) Recognition of methylated DNA by TAL effectors. Cell Res 22(10):1502–1504
83. Perez-Pinera P, Ousterout DG, Brunger JM et al (2013) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. Nat Methods 10(3): 239–242
84. Holkers M, Maggio I, Liu J et al (2013) Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic Acids Res 41(5):e63
85. Liu J, Gaj T, Patterson JT et al (2014) Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. PLoS One 9(1):e85755
86. Gaj T, Guo J, Kato Y et al (2012) Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods 9(8): 805–807
87. Ding Q, Lee YK, Schaefer EA et al (2013) A TALEN genome-editing system for generating human stem cell-based disease models. Cell Stem Cell 12(2):238–251
88. Kim YK, Wee G, Park J et al (2013) TALEN-based knockout library for human microRNAs. Nat Struct Mol Biol 20(12): 1458–1464
89. Zhang Z, Xiang D, Heriyanto F et al (2013) Dissecting the roles of miR-302/367 cluster in cellular reprogramming using TALE-based repressor and TALEN. Stem Cell Reports 1(3):218–225
90. Wang H, Hu YC, Markoulaki S et al (2013) TALEN-mediated editing of the mouse Y chromosome. Nat Biotechnol 31(6): 530–532
91. Takada S, Sato T, Ito Y et al (2013) Targeted gene deletion of miRNAs in mice by TALEN system. PLoS One 8(10):e76004
92. Tesson L, Usal C, Menoret S et al (2011) Knockout rats generated by embryo microinjection of TALENs. Nat Biotechnol 29(8):695–696

93. Carlson DF, Tan W, Lillico SG et al (2012) Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci U S A 109(43):17382–17387
94. Sander JD, Cade L, Khayter C et al (2011) Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29(8):697–698
95. Huang P, Xiao A, Zhou M et al (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29(8):699–700
96. Lo TW, Pickle CS, Lin S et al (2013) Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics 195(2):331–348
97. Wood AJ, Lo TW, Zeitler B et al (2011) Targeted genome editing across species using ZFNs and TALENs. Science 333(6040):307
98. Hayashi T, Sakamoto K, Sakuma T et al (2014) Transcription activator-like effector nucleases efficiently disrupt the target gene in Iberian ribbed newts (Pleurodeles waltl), an experimental model animal for regeneration. Dev Growth Differ 56(1):115–121
99. Ma S, Zhang S, Wang F et al (2012) Highly efficient and specific genome editing in silkworm using custom TALENs. PLoS One 7(9):e45035
100. Beumer KJ, Trautman JK, Christian M et al (2013) Comparing zinc finger nucleases and transcription activator-like effector nucleases for gene targeting in Drosophila. G3 (Bethesda) 3(10):1717–1725
101. Smidler AL, Terenzi O, Soichot J et al (2013) Targeted mutagenesis in the malaria mosquito using TALE nucleases. PLoS One 8(8):e74511
102. Lei Y, Guo X, Liu Y et al (2012) Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). Proc Natl Acad Sci U S A 109(43):17484–17489
103. Voytas DF (2013) Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol 64:327–350
104. Mahfouz MM, Li L, Shamimuzzaman M et al (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(6):2623–2628
105. Shan Q, Wang Y, Chen K et al (2013) Rapid and efficient gene modification in rice and Brachypodium using TALENs. Mol Plant 6(4):1365–1368
106. Christian M, Qi Y, Zhang Y et al (2013) Targeted mutagenesis of Arabidopsis thaliana using engineered TAL effector nucleases. G3 (Bethesda) 3(10):1697–1705
107. Wendt T, Holm PB, Starker CG et al (2013) TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. Plant Mol Biol 83(3):279–285
108. Osborn MJ, Starker CG, McElroy AN et al (2013) TALEN-based gene correction for epidermolysis bullosa. Mol Ther 21(6):1151–1159
109. Sun N, Liang J, Abil Z et al (2012) Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Biosyst 8(4):1255–1263
110. Voit RA, Hendel A, Pruett-Miller SM et al (2014) Nuclease-mediated gene editing by homologous recombination of the human globin locus. Nucleic Acids Res 42(2): 1365–1378
111. Ma N, Liao B, Zhang H et al (2013) Transcription activator-like effector nuclease (TALEN)-mediated gene correction in integration-free beta-thalassemia induced pluripotent stem cells. J Biol Chem 288(48): 34671–34679
112. Dupuy A, Valton J, Leduc S et al (2013) Targeted gene therapy of xeroderma pigmentosum cells using meganuclease and TALEN. PLoS One 8(11):e78678
113. Choi SM, Kim Y, Shim JS et al (2013) Efficient drug screening and gene correction for treating liver disease using patient-specific stem cells. Hepatology 57(6):2458–2468
114. Bacman SR, Williams SL, Pinto M et al (2013) Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. Nat Med 19(9):1111–1113
115. Tebas P, Stein D, Tang WW et al (2014) Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 370(10):901–910
116. Mussolino C, Alzubi J, Fine EJ et al (2014) TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. Nucleic Acids Res 42(10): 6762–6773
117. Ye L, Wang J, Beyer AI et al (2014) Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection. Proc Natl Acad Sci U S A 111:9591–9596
118. Bloom K, Ely A, Mussolino C et al (2013) Inactivation of hepatitis B virus replication in cultured cells and in vivo with engineered

transcription activator-like effector nucleases. Mol Ther 21(10):1889–1897

119. Chen J, Zhang W, Lin J et al (2014) An efficient antiviral strategy for targeting hepatitis B virus genome using transcription activator-like effector nucleases. Mol Ther 22(2): 303–311
120. Xu L, Zhao P, Mariano A et al (2013) Targeted myostatin gene editing in multiple mammalian species directed by a single pair of TALE nucleases. Mol Ther Nucleic Acids 2:e112
121. Guilinger JP, Pattanayak V, Reyon D et al (2014) Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. Nat Methods 11(4):429–435
122. Maeder ML, Linder SJ, Reyon D et al (2013) Robust, synergistic regulation of human gene expression using TALE activators. Nat Methods 10(3):243–245
123. Mali P, Esvelt KM, Church GM (2013) Cas9 as a versatile tool for engineering biology. Nat Methods 10(10):957–963
124. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. Cell 157(6):1262–1278
125. Sung YH, Baek IJ, Kim DH et al (2013) Knockout mice created by TALEN-mediated gene targeting. Nat Biotechnol 31(1):23–24
126. Zu Y, Tong X, Wang Z et al (2013) TALEN-mediated precise genome modification by homologous recombination in zebrafish. Nat Methods 10(4):329–331

Chapter 4

Online Tools for TALEN Design

Adam J. Bogdanove and Nicholas J. Booher

Abstract

Transcription activator-like effector nucleases (TALENs) can be exquisitely specific and highly effective genome editing reagents. Specificity and efficacy depend however on good design for minimal off-targeting and strong binding. Several online tools are accessible to aid in this process. Here, we tabulate those tools, noting their functions and key features.

Key words Genome-editing, Protein-DNA interaction, Off-target prediction, Web-based tools, TAL effector nucleases

1 Background

TAL (transcription activator-like) effector nucleases, or TALENs, consist of paired proteins, each a fusion of a TAL effector DNA-binding domain to the catalytic domain of the *Fok*I restriction enzyme. Binding specificity of the TAL effector domains is engineered such that the proteins bind opposing sites in DNA, bringing the *Fok*I domains together across a spacer. Since *Fok*I functions as a dimer, this results in a double-strand break in the DNA at that target, a key first step in common genome editing approaches [1].

TAL effectors are transcription factors injected into plant cells by pathogenic bacteria during infection. Their DNA-binding specificities are determined by a central region containing tandem repeats of a 33–35 amino acid sequence. Each repeat interacts with a single base in the DNA, and specificity of that interaction is determined by residues at positions 12 and 13, called the repeat variable diresidue (RVD), with residue 13 making the base-specific contacts [2–5]. Thus, the number of repeats and sequence of the RVDs determine the nucleotide sequence of the DNA target. This modularity, the TAL effector DNA-binding "code" [3] allows specificity to be engineered by assembling custom arrays of repeats

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_4,

with the appropriate polymorphisms at residue 13 (reviewed in [6]). Several moderate and high-throughput cloning methods are in place for such assembly (e.g., [7–14]).

But degeneracy in the code, i.e., qualitative and quantitative variation in the preferences of different RVDs for each of the four nucleotides [2], impacts the overall specificity of individual TAL effector DNA-binding domains. A general requirement for thymine to precede the DNA target sequence [2, 3] and the requirement of TALENs to target paired sites within a defined range of distances apart on opposing strands that depends on the particular TAL effector-*Fok*I linkage architecture used [15, 16], as well as differential contributions of different repeat-nucleotide interactions to overall affinity [17] and a polarity to mismatch tolerance [18] add further complexity.

TALENs can be exquisitely specific and highly effective genome editing reagents. Efficacy and specificity depend however on selection of targets and corresponding RVD sequences that, to the extent possible, takes into account the complexities mentioned above to achieve unique and high affinity binding in the context of the genome to be edited. Several computational tools have been developed and made available online to aid in this process. To varying degrees, these take into account degeneracy in the code, the requirement for a preceding T, and spacing constraints for different architectures. Accuracy of off-target prediction is not well characterized yet for any of the tools and will likely benefit from more precise quantification of the contributions of different RVD-nucleotide interactions to overall binding affinity and the influence of position in the array on them. Below, we tabulate the currently accessible tools, noting the functions and key features of each (Table 1).

2 Notes

1. CHOPCHOP is found at https://chopchop.rc.fas.harvard.edu. Information presented here is from Feb. 8, 2015. Output includes PCR primer sites to amplify the targeted region.
2. E-TALEN is found at http://www.e-talen.org. Information presented here is from version 2.5. E-TALEN counts off-targets as an exclusionary criterion for TALEN design. It does not output their location.
3. Mojo Hand is found at http://talendesign.org. Information presented here is from version 2. Users may specify restriction enzyme manufacturers for sites in the target amplicon.
4. PROGNOS is found at http://baolab.bme.gatech.edu/Research/BioinformaticTools/prognos.html. Information presented here is from Feb. 8, 2015. PROGNOS allows primer

Table 1
Table of online tools for TALEN design

Tool	TALEN target sequence source(s)	TALEN design considerations	Off-target prediction sequence sources	Off-target prediction method	Off-target prediction output
CHOPCHOP[1] [19]	Gene ID (RefSeq), ENSEMBL, FlyBase, WormBase, SGD, TAIR, VectorBase, Xenbase, GeneDB, FASTA	Genic context, diagnostic restriction sites, spacer length, predicted off-targets	Pre-loaded genomes	# mismatches from perfect target	Sequence, # mismatches, location
E-TALEN [20]	ENSEMBL, FASTA, plain text	Genic context, diagnostic restriction sites, experiment type, spacer length, predicted off-targets	FASTA, pre-loaded genomes	# mismatches from perfect target	n/a
Mojo Hand [21]	Gene ID (NCBI), FASTA, plain text	Genic context, spacer length, diagnostic restriction sites	n/a	n/a	n/a
PROGNOS [22]	n/a	n/a	FASTA, pre-loaded genomes	Modular position weight matrix	Sequence, score, location, annotation, genome browser link, primers for validation
SAPTA [23]	FASTA	Predicted activity based on RVD composition and site length, spacer length, diagnostic restriction sites	n/a	n/a	n/a
TALE-NT [24]	FASTA	RVD composition, spacer length, diagnostic restriction sites, predicted off-targets	FASTA, pre-loaded genomes, NCBI	Modular position weight matrix	Sequence, score, location, genome browser link
TALENdesigner	FASTA	Site length, spacer length	Human, mouse, rat genomes	Modular position weight matrix	Sequence, score, location, genome browser link
TALENoffer [25]	n/a	n/a	FASTA, pre-loaded genomes	Local mixture model	Sequence, score, location, annotation, genome browser link

design as well as specification of a minimum size difference of amplicon digest products depending on the type of analysis to be done.

5. SAPTA is found at http://baolab.bme.gatech.edu/Research/BioinformaticTools/TAL_targeter.html. Information presented here is from Jun. 7, 2014. In addition to including a design parameter to account for predicted activity, SAPTA allows a user to specify a particular location (base) for the cut.
6. TALE-NT is found at http://tale-nt.cac.cornell.edu. Information presented here is from version 2. Selection of parameters is made simple by annotated pull-down menus.
7. TALENdesigner is found at http://www.talen-design.de/index_talen.html. Information presented here is from Feb. 8, 2015. TALENdesigner also provides genome browsers for human, mouse, and rat populated with pre-designed TALEN sites.
8. TALENoffer is found at http://www.jstacs.de/index.php/TALENoffer. Information presented here is from version 1. TALENoffer predicts off-target sites for user-provided TALEN designs.

Acknowledgement

Research on TAL effectors in our laboratory is supported by grants from the National Science Foundation (IOS 1238189) and the National Institutes of Health (R01 GM098861).

References

1. Carroll D (2014) Genome engineering with targetable nucleases. Annu Rev Biochem 83:409–439. doi:10.1146/annurev-biochem-060713-035418
2. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326(5959):1501. doi:10.1126/science.1178817
3. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326(5959):1509–1512. doi:10.1126/science.1178811
4. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 335(6069):716–719. doi:10.1126/science.1216211
5. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, Shi Y, Yan N (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. Science 335(6069):720–723. doi:10.1126/science.1215670
6. Doyle EL, Stoddard BL, Voytas DF, Bogdanove AJ (2013) TAL effectors: highly adaptable phytobacterial virulence factors and readily engineered DNA-targeting proteins. Trends Cell Biol 23:390–398. doi:10.1016/j.tcb.2013.04.003
7. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39(12):e82. doi:10.1093/nar/gkr218
8. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29:149–153. doi:10.1038/nbt.1775

9. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30(5):460–465. doi:10.1038/nbt.2170
10. Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S (2011) Assembly of designer TAL effectors by Golden Gate cloning. PLoS One 6(5):e19722. doi:10.1371/journal.pone.0019722
11. Wang Z, Li J, Huang H, Wang G, Jiang M, Yin S, Sun C, Zhang H, Zhuang F, Xi JJ (2012) An integrated chip for the high-throughput synthesis of transcription activator-like effectors. Angew Chem Int Ed Engl 51(34):8505–8508. doi:10.1002/anie.201203597
12. Li L, Atef A, Piatek A, Ali Z, Piatek M, Aouida M, Sharakuu A, Mahjoub A, Wang G, Khan S, Fedoroff NV, Zhu JK, Mahfouz MM (2013) Characterization and DNA-binding specificities of Ralstonia TAL-like effectors. Mol Plant 6(4):1318–1330. doi:10.1093/mp/sst006
13. Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, Church GM (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40(15):e117. doi:10.1093/nar/gks624
14. Morbitzer R, Elsaesser J, Hausner J, Lahaye T (2011) Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res 39(13):5790–5799. doi:10.1093/nar/gkr151
15. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186:757–761. doi:10.1534/genetics.110.120717
16. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148. doi:10.1038/nbt.1755
17. Streubel J, Blucher C, Landgraf A, Boch J (2012) TAL effector RVD specificities and efficiencies. Nat Biotechnol 30(7):593–595. doi:10.1038/nbt.2304
18. Meckler JF, Bhakta MS, Kim MS, Ovadia R, Habrian CH, Zykovich A, Yu A, Lockwood SH, Morbitzer R, Elsaesser J, Lahaye T, Segal DJ, Baldwin EP (2013) Quantitative analysis of TALE-DNA interactions suggests polarity effects. Nucleic Acids Res 41(7):4118–4128. doi:10.1093/nar/gkt085
19. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E (2014) CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res 42(Web Server issue):W401–W407. doi:10.1093/nar/gku410
20. Heigwer F, Kerr G, Walther N, Glaeser K, Pelz O, Breinig M, Boutros M (2013) E-TALEN: a web tool to design TALENs for genome engineering. Nucleic Acids Res 41(20):e190. doi:10.1093/nar/gkt789
21. Neff KL, Argue DP, Ma AC, Lee HB, Clark KJ, Ekker SC (2013) Mojo Hand, a TALEN design tool for genome editing applications. BMC Bioinformatics 14:1. doi:10.1186/1471-2105-14-1
22. Fine EJ, Cradick TJ, Zhao CL, Lin Y, Bao G (2014) An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. Nucleic Acids Res 42(6):e42. doi:10.1093/nar/gkt1326
23. Lin Y, Fine EJ, Zheng Z, Antico CJ, Voit RA, Porteus MH, Cradick TJ, Bao G (2014) SAPTA: a new design tool for improving TALE nuclease activity. Nucleic Acids Res 42(6):e47. doi:10.1093/nar/gkt1363
24. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40(Web Server issue):W117–W122. doi:10.1093/nar/gks608
25. Grau J, Boch J, Posch S (2013) TALENoffer: genome-wide TALEN off-target prediction. Bioinformatics 29(22):2931–2932. doi:10.1093/bioinformatics/btt501

Chapter 5

Assembly of Customized TAL Effectors Through Advanced ULtiMATE System

Junjiao Yang, Shengjie Guo, Pengfei Yuan, and Wensheng Wei

Abstract

Transcription activator-like effectors (TALEs) have been widely applied in gene targeting. Here we describe an advanced ULtiMATE (*U*SER-based *Liga**t**i*on-*M*ediated *A*ssembly of *T*AL *E*ffector) system that utilizes USER fusion technique and archive of 512 tetramer templates to achieve highly efficient construction of TALEs, which takes only half a day to accomplish the assembly of any given TALE construct. This system is also suitable for large-scale assembly of TALENs and any other TALE-based constructions.

Key words TAL effectors, TALENs, ULtiMATE system, USER, Assembly

1 Introduction

TAL effectors (TALEs), originally identified from bacteria *Xanthomonas*, contain a central DNA-binding region of tandem repeats of 33–35 amino acids, with each repeat specifically targeting a nucleotide through repeat variable diresidues (RVDs) [1, 2]. This modular DNA binding feature has inspired custom designed TALE proteins for use in gene editing [1, 3–5]. It is important to have a robust cloning strategy in order to apply this powerful genetic tool. Here we describe an advanced ULtiMATE (*USER-based Liga**t**i*on-*M*ediated *A*ssembly of *T*AL *E*ffector) system for TALE construction, a method upgraded from our prior report [6] and different from other published protocols [4, 7–10]. ULtiMATE utilizes USER fusion technique [11] to assemble DNA fragments, which are obtained from PCR reactions using special uracil-containing primers and unique polymerases that could incorporate a deoxyadenine opposite a dU, such as PfuTurbo Cx Hotstart DNA polymerase. The USER™ enzyme mixture of glycosidase (UDG) and DNA glycosylase-lyase endo VIII was used to remove the dU residues to generate 3′-protruding sticky ends in PCR products [11].

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_5,

We designed four types of basic TALE repeat unit that differ in their DNA sequences flanking the RVD-coding region but encode the same amino acid sequence except for RVD, designated as W-, X-, Y-, and Z-type, respectively. Based on this design, total of 16 TALE repeat units were commercially synthesized in such a way that every type harbors the coding sequence of one of the four RVDs that recognize a particular DNA base (NI→A, HD→C, NN→G and NG→T). With these 16 units, we preassembled all combinations of tetramers in either **WXYZ** or **XYZW** format to make a pool of 512 templates. Time to generate 512 tetramer templates using ULtiMATE system is about 4 weeks.

With tetramer templates, construction of customized TALEs becomes much faster and easier to operate. The protocol consists of PCR reactions, DNA identification through gel electrophoresis (an optional procedure), purification of PCR products, USER fusion, and transformation. It takes only about 6 h for the construction procedure. As the PCR primers are fixed, about two dozens of TALE arrays could be assembled in one 96-well plate simultaneously, and the number of TALE arrays for assembly is only limited by the capacity of PCR reactions.

Golden Gate method could also be used for TALE assembly based on the same 512 tetramer archive. Primers for both the ULtiMATE system and the Golden Gate method are listed.

2 Materials

2.1 PCR Templates

1. Synthesizing 16 TALE monomers (*see* **Note 1**).
2. Constructing 512 TALE tetramers as PCR templates, 256 in **WXYZ** format and 256 in **XYZW** format (*see* **Note 2**).

2.2 PCR Primers

1. Normal and uracil primers for PCR reaction.
2. Primers for colony PCR and sequencing analysis.

2.3 TALE Expressing Vectors

pIRES-EGFP-TALEN (*see* **Notes 2** and **3**) (Fig. 1).

2.4 Enzymes, Chemicals, and Kits

1. PfuTurbo Cx Hotstart DNA Polymerase and buffer (Agilent Technologies).
2. dNTP mix (2.5 mM each).
3. USER™ enzyme (New England Biolabs).
4. PCR product purification kit.
5. BsmBI restriction enzyme (Thermo Fisher Scientific) (*see* **Note 4**).
6. Tango buffer (Thermo Fisher Scientific).
7. T4 DNA ligase.

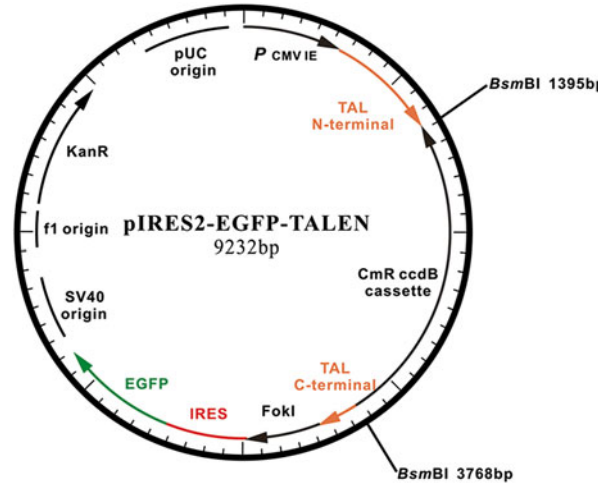


Fig. 1 pIRES2-EGFP-TALEN: Vector for TAL nuclease construction

8. 10 mM ATP.
9. 50 mM DTT.
10. Trans1-T1 competent cells.
11. 2× Easymix Taq enzyme.

3 Methods

3.1 Construction of 512 Tetramer Templates

1. Performing PCR reactions using the 16 TALE monomers (Table 1) as templates and select ULtiMATE primers (Table 2) based on Table 3. For each PCR reaction, mix the following into 0.2 mL PCR tube: TALE monomer (5 ng/μL, 3 μL), forward primer (1 μM, 3 μL), reverse primer (1 μM, 3 μL), PfuTurbo Cx DNA Polymerase (2.5 units/μL, 0.15 μL), PfuTurbo Cx Buffer (10×, 1.5 μL), dNTP mix (1.5 μL), and ddH_2O to a total volume of 15 μL. PCR steps are as follows: 95 °C, 3 min for hot start; thermal cycling (95 °C, 30 s; 60 °C, 30 s; 72 °C, 30 s; 30×); 72 °C, 10 min; and hold at 4 °C.
2. Combine four PCR reactions for one tetramer into one tube. Add 1.5 μL USERTM enzyme (1000 units/mL) into mixed PCR product, incubate at 37 °C for 15 min, and cool on ice.
3. Add 1.5 μL T4 DNA ligase (100 units) and 7 μL T4 DNA ligase buffer into USER-digested PCR products, and incubate at 16 °C for 1 h.
4. Gel-purify (2 % agarose) ligated products, extract the ~400-bp DNA band, and ligate them with a blunt vector with no kanamycin-resistant gene (*see* **Note 2**). Verify clones through Sanger sequencing.

Table 1
Archive of 16 TALE monomers

W *type*: 5′-CTGACACCAGAGCAAGTAGTGGCTATTGCAAGTNNNNNNGGTGGCAAACAAGC GCTGGAGACCGTGCAGAGGCTCCTTCCGGTGCTCTGCCAAGCACACGGT
X *type*: 5′-CTCACTCCGGAACAGGTGGTCGCAATCGCGAGCNNNNNNGGCGGCAAGCAAGC CCTTGAGACAGTCCAAAGACTTTTGCCTGTCCTTTGTCAGGCGCATGGC
Y *type*: 5′-CTTACGCCTGAGCAAGTCGTTGCGATCGCCTCCNNNNNNGGCGGAAAACAGGCT TTGGAAACCGTGCAGCGGTTGCTGCCCGTTTTGTGCCAAGCCCACGGA
Z *type*: 5′-TTGACCCCCGAACAGGTTGTAGCCATAGCTTCTNNNNNNGGAGGTAAGCAGGCA CTGGAAACCGTGCAGCGCCTGCTCCCAGTACTGTGTCAGGCTCATGGG

RVD-coding sequence NNNNNN = AACATC (NI, for A-recognition), AACGGC (NG, for T-recognition), CACGAC (HD, for C-recognition), or AACAAC (NN, for G-recognition)

Table 2
ULtiMATE primers for tetramer construction

Primer	Sequence
F-Yx	AGCAAGTCGUTGCGATCGCCTCC
F-Zy	ACGGATUGACCCCCGAACAGGTTGTAGCC
F-X-new	CTCACTCCGGAACAGGTGGTCG
F-W-new	CTGACACCAGAGCAAGTAGTGGCTATTG
R-Xy	ACGACTTGCUCAGGCGTAAGGCCATGCGCCTGACAAAGGACA
R-Yz	AATCCGUGGGCTTGGCACAAAACGGGC
R-Z-new	CCCATGAGCCTGACACAGTACTGG
R-W-new	ACCGTGTGCTTGGCAGAGC

Table 3
Selection of templates and primers for PCR reaction for TALE tetramer construction

		1st PCR	2nd PCR	3rd PCR	4th PCR
WXYZ	TALE monomer template	W	X	Y	Z
	Forward primer	F-W-new	F-Xw	F-Yx	F-Zy
	Reverse primer	R-Wx	R-Xy	R-Yz	R-Z-new
XYZW	TALE monomer template	X	Y	Z	W
	Forward primer	F-X-new	F-Yx	F-Zy	F-Wz
	Reverse primer	R-Xy	R-Yz	R-Zw	R-W-new

Table 4
ULtiMATE primers

Primer	Sequence
F-W5	tataCGTCTCaGAACCTGACACCAGAGCAAGTAGTGGCTATTG
F-Wz	ACCGUGCAGCGCCUGCUCCCAGTACTGTGTCAGGCTCATGGGCTGACA CCAGAGCAAGTAGTGG
F-Xz	AGGCTCAUGGGCTCACUCCGGAACAG
F-Xw	AACAGGTGGUCGCAAUCGC
F-Ww	AGAGCAAGUAGTGGCTAUTGCAA
F-Wy	AGCAAGUAGTGGCTATTGCAAG
F-Xy	ACTCACUCCGGAACAGGTGGTCGCAATCG
R-Zw	AGCAGGCGCUGCACGGUTTCCAGT
R-Zx	AGUGAGCCCAUGAGCCUGACACAGTA
R-Wx	ATUGCGACCACCUGTUCCGGAGTGAGACCGTGTGCTTGGCAGAG
R-Ww	ATAGCCACUACTTGCTCUGGTGTCAGACCGTGTGCTTGGCAGAGC
R-Yw	ACTTGCUCTGGTGTCAGTCCGTGGGCTTGGCACAAAACGG
R-Yx	AGTGAGUCCGTGGGCTTGGCACAAAACGG
R-W3	tataCGTCTCaTGCTCTCCAGCGCTTGTTTGCCACC
R-X3	tataCGTCTCaTGCTCTCAAGGGCTTGCTTGCCGC
R-Y3	tataCGTCTCaTGCTTTCCAAAGCCTGTTTTCCGCC
R-Z3	tataCGTCTCaTGCTTTCCAGTGCCTGCTTACCTCC

3.2 PCR Amplification for TALE Array Construction

1. Based on target sequence, select templates and primers for the PCR reaction (*see* **Notes 5–8**, Tables 4 and 5).
2. For each PCR reaction, mix the following into 0.2 mL PCR tube: tetramer template (5 ng/μL, 3 μL), forward primer (1 μM, 3 μL), reverse primer (1 μM, 3 μL), PfuTurbo Cx DNA Polymerase (2.5 units/μL, 0.15 μL), PfuTurbo Cx Buffer (10×, 1.5 μL), dNTP mix (1.5 μL), and ddH_2O to a total volume of 15 μL.
3. Perform PCR reaction as follows: 95 °C, 3 min for hot start; thermal cycling (95 °C, 30 s; 60 °C, 30 s; 72 °C, 30 s; 30×); 72 °C, 10 min; and hold at 4 °C.
4. Optional step: DNA electrophoresis with 1 μL PCR product for quality check. The size of PCR product containing N repeats is about $(102 \times N - 42)$ bp.

Table 5
Selection of templates and primers for PCR reaction to assemble TALE arrays containing 7.5–22.5 repeats

Tetramer templates		1st PCR	2nd PCR	3rd PCR	4th PCR	5th PCR	6th PCR
		WXYZ	WXYZ	XYZW	XYZW	WXYZ	WXYZ
Primers	22.5-mer					F-Ww/R-Yw	F-Wy/R-Z3
	21.5-mer						F-Wy/R-Y3
	20.5-mer						F-Wy/R-X3
	19.5-mer				F-Xw/R-Ww	F-Ww/R-Z3	–
	18.5-mer					F-Ww/R-Y3	–
	17.5-mer			F-Xz/R-Wx		F-Ww/R-X3	–
	16.5-mer					F-Ww/R-W3	–
	15.5-mer	F-W5/R-Zw	F-Wz/R-Zx		F-Xw/R-W3	–	–
	14.5-mer				F-Xw/R-Z3	–	–
	13.5-mer				F-Xw/R-Y3	–	–
	12.5-mer				F-Xw/R-X3	–	–
	11.5-mer			F-Xz/R-W3	–	–	–
	10.5-mer			F-Xz/R-Z3	–	–	–
	9.5-mer			F-Xz/R-Y3	–	–	–
	8.5-mer			F-Xz/R-X3	–	–	–
	7.5-mer		F-Wz/R-Z3	–	–	–	–

3.3 USER Digestion and Purification of PCR Products

1. Mix all the PCR products of the same TALE array. Mix all the PCR products of the same TALE array into one tube. Add 1.5 μL USER™ enzyme (1000 units/mL) directly into 60–75 μL mixed PCR product, and incubate at 37 °C for 15 min and cool on ice.
2. Add 1.5 μL T4 DNA ligase (100 units) and 7 μL T4 DNA ligase buffer into USER digested PCR products, and incubate at 16 °C for 1 h.
3. Purify USER-digested PCR products.

3.4 Cycles of BsmBI Digestion, DNA Ligation, and Transformation

1. Mix the purified PCR fragments with the following into one tube: BsmBI (Esp3I) (7.5 units), T4 DNA ligase (100 units), ATP (10 pmol), DTT (10 pmol), Tango buffer (1×), TALE expression vector (~20 ng), and ddH_2O to a total volume of 10 μL.
2. Perform thermo cycles (37 °C, 5 min; 16 °C, 5 min; 6×), and hold at 4 °C.

Table 6
Primers for colony PCR and sequencing analysis

Primer	Sequence
TAL-F	AAGAGGGGAGGCGTGACGGC
TAL-R	CAAGCCAGGGCCACCAGGT

3. Transform 5 μL products into 100 μL Trans1-T1 competent cells, and plate onto solid LB medium with 25 μg/mL of kanamycin. Other types of competent cells (except DB3.1 or any other ccdB-resistant strains) can also be used here.

3.5 Validation of TALE Clones

1. Perform colony PCR using the thermo cycles (95 °C, 30 s; 60 °C, 30 s; 72 °C, 2 min; 30×) with TAL-F and TAL-R primer pair (Table 6). If all steps are correctly performed, analyzing 5–10 colonies will be enough (*see* **Note 9**).
2. Examine the size of PCR products by electrophoresis. A DNA band from a right clone should be around the size of $(102 \times N + 107)$ bp (N indicates number of repeats in the TALE, including the last half repeat).
3. Sequence the TALE constructs using TAL-F and TAL-R primers (Table 6).

4 Notes

1. Archive of 16 TALE monomers.
2. Besides pIRES-EGFP-TALEN, other TALE vectors with non-ampicillin selection marker could be modified to perform ULtiMATE synthesis, as long as the cloning vector contains the ccdB cassette flanked by the BsmBI sites as indicated (Fig. 2). Because the tetramer templates are maintained in a vector containing ampicillin-resistant gene, ampicillin cannot be used as the selection marker for TALE vectors. All plasmids carrying tetramer templates are available to academic researchers upon request.
3. TALE expression vector pGL3-TALEN and pLentiLox3.7-TALE [11] can also be used if the purification step in Subheading 3.2 is substituted by DNA electrophoresis and gel purification.
4. Make sure that your BsmBI enzyme works at 37 °C. BsmBI from different producers may work at different temperatures.

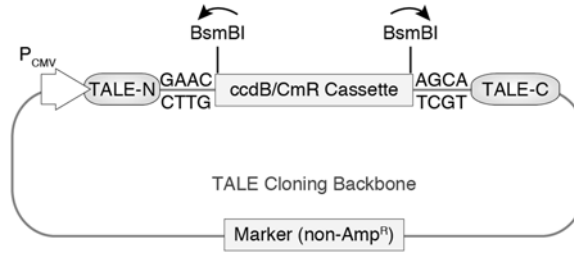


Fig. 2 Universal structure of TALE backbone vectors suitable for ULtiMATE cloning. N-terminal non-repeat region of TALE should be ended with GAAC in sense strand and a following recognition site of BsmBI, because after BsmBI digestion, a 5′-GAAC sticky end needs to be exposed to ligate the TALE repeats into backbone vector. Similarly, C-terminal non-repeat region should be beginning with AGCA. A backbone vector should contain a selection marker except ampicillin

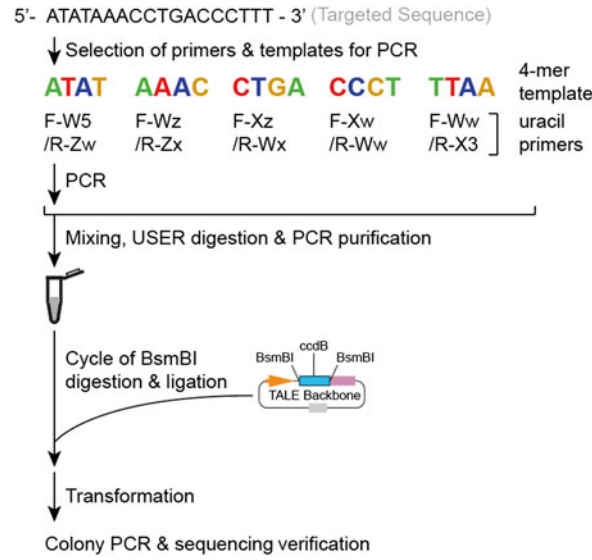


Fig. 3 Flow chart of TALE construction by ULtiMATE system with tetramer templates. Flow chart of the cloning process of a 17.5-mer TALE targeting ATATAAACCTGACCCTTT. The selection of primers are based on Table 5. The PCR amplification of tetramers can be finished within 1.5 h. Without the need of purification, the PCR products of all reactions are mixed for subsequent USER enzyme digestion. After purification, digested products are mixed with TALE cloning backbone for the cycles of BsmBI digestion and ligation, followed by the bacterial transformation (~1.5 h). The candidate clones are isolated based on colony PCR results, and verified by sequencing analysis

Table 7
Example of the assembly of a TALE array with 25.5 repeats (target sequence: TACCGTACCCGAGTTTGAACTTGTGT)

	1st PCR	2nd PCR	3rd PCR	4th PCR	5th PCR	6th PCR	7th PCR
Tetramer templates	**TACC**	**GTAC**	**CCGA**	**GTTT**	**GAAC** (or **GAAG/A/T**)	**CTTG** (or **CTTC/A/T**)	**GTGT**
Primers	F-W5/ R-Zw	F-Wz/ R-Zx	F-Xz/ R-Wx	F-Xw/ R-Ww	F-Ww/R-Yw	F-Wy/R-Yx	F-Xy/ R-W3
Expected PCR products	**TACC**	**GTAC**	**CCGA**	**GTTT**	**GAA**	**CTT**	**GTGT**

5. Select the PCR tetramer templates and primers (Table 4) for PCR reaction to construct each TALE array. TALE constructs with TALE repeat numbers from 7.5 to 22.5 are summarized in Table 5. Notably, the tetramers for the 1st, 2nd, 5th, and 6th PCR are in the WXYZ format, and the tetramers for the 3rd and 4th PCR are in the XYZW format. Figure 3 shows a schematic of a cloning process of a 17.5-mer TALE using ULtiMATE protocol.
6. For TALE constructs containing more than 22.5 repeats, the selection of tetramer templates and primers (Table 4) should be adjusted, as long as there are no identical primers used in the same ligation reaction. Table 7 shows an example of the assembly of a TALE array with 25.5 repeats.
7. With the same tetramer templates, end primers (both 5′ and 3′), and the ccdB cloning vectors, the TALE construction could also be conducted with modified Golden Gate or LIC (ligase-independent cloning) methods.
8. For the Golden Gate method, the principle of templates and primers selection, as well as the TALE assembly are the same as in Subheading 3, except that the primers are different and listed in Table 8, and the USER™ digestion step is not required. For each PCR reaction, mix the following into 0.2 mL PCR tube: tetramer template (5 ng/μL, 1 μL), forward primer (2 μM 1 μL), reverse primer (2 μM, 1 μL), TransTaq™ HiFi DNA Polymerase (5 unit/μL, 0.2 μL), TransTaq HiFi Buffer (10×, 1 μL), dNTPs (0.8 μL), and ddH_2O to a total volume of 10 μL. Perform PCR reaction as follows: 94 °C, 5 min; thermal cycling (94 °C, 30 s; 55 °C, 30 s; 72 °C, 30 s; 30×); 72 °C, 10 min; and hold at

Table 8
Primers for Golden Gate method

Primer	Sequence
F-W5	tataCGTCTCaGAACCTGACACCAGAGCAAGTAGTGGCTATTG
F-Wz′	TCTCGTCTCTCTGACACCAGAGCAAGTAGTGG
F-Xz′	GTCCGTCTCATGGGCTCACTCCGGAACA
F-Xw′	ATACGTCTCACAAGCACACGGTCTCACTCCGGAACAGGTGGT
F-Ww′	GTCCGTCTCCAGAGCAAGTAGTGGCTAT
F-Wy′	GTCCGTCTCCCCACGGACTGACACCAGAGCAAGTAGT
F-Xy′	GTGCGTCTCCCACGGACTCACTCCGGAACAGGTGG
R-Zw′	TATCGTCTCATCAGCCCATGAGCCTGACACAGTACT
R-Zx′	CCGCGTCTCGCCCATGAGCCTGACACAG
R-Wx′	AGACGTCTCGCTTGGCAGAGCACCGGAA
R-Ww′	TATCGTCTCACTCTGGTGTCAGACCGTGTGCTTGGCAGAGCA
R-Yw′	TATCGTCTCCGTGGGCTTGGCACAAAACGGGCAGC
R-Yx′	GGGCGTCTCCCGTGGGCTTGGCACAAAACGGGCAGCAACCG
R-W3	tataCGTCTCaTGCTCTCCAGCGCTTGTTTGCCACC
R-X3	tataCGTCTCaTGCTCTCAAGGGCTTGCTTGCCGC
R-Y3	tataCGTCTCaTGCTTTCCAAAGCCTGTTTTCCGCC
R-Z3	tataCGTCTCaTGCTTTCCAGTGCCTGCTTACCTCC

4 °C. Figure 4 shows a schematic of a cloning process of a 17.5-mer TALE using a Golden Gate protocol (in BsmBI digestion and ligation step, at least 16 cycles are needed).

9. When TALE construction is correctly performed, positive rate of colony PCR usually achieves 6/6 or 11/12. If positive rate is too low, please check whether there are any mistakes in the steps.

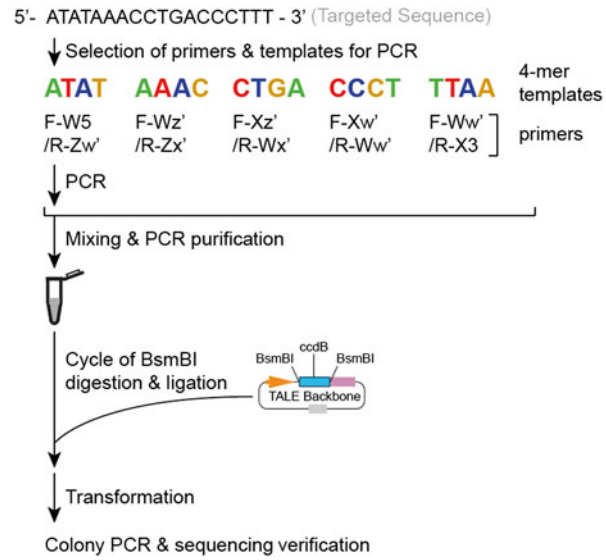


Fig. 4 Flow chart of TALE construction by Golden Gate method with tetramer templates. Flow chart of the cloning process of a 17.5-mer TALE targeting ATATAAACCTGACCCTTT. Primers are also selected based on Table 5 with a small change that Golden Gate primers are named with an inverted comma and their sequences can be found in Table 8. After PCR amplification, the PCR products of all reactions are purified and mixed with TALE cloning backbone for the cycles of BsmBI digestion and ligation

Acknowledgement

This work was supported by funds from the Major State Basic Research Development Program of China (grant number 2010CB911800), the National Science Foundation of China (grant numbers NSFC31070115, NSFC31170126), and the Peking-Tsinghua Center for Life Sciences.

References

1. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512
2. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501
3. Miller JC, Tan S, Qiao G, Barlow KA, Wang J et al (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29: 143–148
4. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y et al (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82
5. Bogdanove AJ, Voytas DF (2011) TAL effectors: customizable proteins for DNA targeting. Science 333:1843–1846
6. Yang J, Yuan P et al (2013) ULtiMATE system for rapid assembly of customized TAL effectors. PLoS One 8(9):e75649

7. Zhang F, Cong L, Lodato S, Kosuri S, Church GM et al (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29:149–153
8. Morbitzer R, Elsaesser J, Hausner J, Lahaye T (2011) Assembly of custom TALE-type DNA binding domains by modular cloning. Nucleic Acids Res 39:5790–5799
9. Geissler R, Scholze H, Hahn S, Streubel J, Bonas U et al (2011) Transcriptional activators of human genes with programmable DNA specificity. PLoS One 6:e19509
10. Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G et al (2012) A transcription activator-like effector toolbox for genome engineering. Nat Protoc 7:171–192
11. Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res 35:e55

Chapter 6

Engineering Customized TALENs Using the Platinum Gate TALEN Kit

Tetsushi Sakuma and Takashi Yamamoto

Abstract

Among various strategies for constructing customized transcription activator-like effector nucleases (TALENs), the Golden Gate assembly is the most widely used and most characterized method. The principle of Golden Gate assembly involves cycling reactions of digestion and ligation of multiple plasmids in a single tube, resulting in PCR-, fragmentation-, and purification-free concatemerization of DNA-binding repeats. Here, we describe the protocols for Golden Gate assembly-based TALEN construction using the Platinum Gate TALEN Kit, which allows generation of highly active Platinum TALENs.

Key words Platinum TALEN, Platinum Gate TALEN Kit, Golden Gate assembly, Modular assembly, 4-Module assembly

1 Introduction

Transcription activator-like effector nucleases (TALENs) recognize target DNA sequence via highly repetitive DNA-binding repeats; therefore, the construction of customized TALENs requires a sophisticated method for serial modular assembly [1]. Currently, various systems for constructing TALENs have been reported by many independent groups [2]. Among them, Golden Gate cloning is one of the most accomplished systems for laboratory-based production of TALENs. Golden Gate cloning was first reported in 2008 as a one-tube and one-step seamless cloning method [3]. In this cloning method, one only has to mix the vector plasmid and one or more insert plasmid(s) with DNA ligase and type IIs restriction enzyme in a single microtube, and perform a thermal cycling reaction, followed by additional digestion. The products can then be directly used for bacterial transformation.

Cermak et al. originally reported Golden Gate assembly-based TALEN construction in 2011 [4]. The system comprises two-step Golden Gate cloning for assembling 12–31 modules. All the plasmids for this system have been deposited in Addgene (Cambridge, MA,

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_6,

USA) as "Golden Gate TALEN and TAL Effector Kit" (cat#1000000024). We previously modified some parts of the kit and supplied the "Yamamoto Lab TALEN Accessory Pack" (cat#1000000030) through Addgene as a supplemental package of the Golden Gate kit [5]. The accessory pack contains alternative intermediate vectors for more efficient repeat assembly.

Recently, we have found that repeating patterns of non-repeat-variable di-residue (non-RVD) variations potently enhanced TALEN activity [7]. These highly active TALENs were named Platinum TALENs, and the Golden Gate-based system for construction of Platinum TALENs was established (Platinum Gate system). Materials for the Platinum Gate system have also been distributed from Addgene as the "Platinum Gate TALEN Kit" (cat#1000000043). Thus far, Platinum TALENs have been proven to work very well compared to conventional Golden Gate TALENs in various cells and organisms, such as HEK293T cells [6], HCT116 cells [7, 8], nematodes [9], sea urchins [10], ascidians [11], newts [12], frogs [6, 13], mice [14, 15], and rats [6]. In this chapter, we describe the step-by-step protocols for constructing Platinum TALENs.

2 Materials

1. Platinum Gate TALEN Kit (Addgene; *see* **Note 1**).
2. Lysogeny broth (LB) liquid media and LB-agar plate containing 100 μg/ml ampicillin or spectinomycin.
3. Plasmid Miniprep Kit (*see* **Note 2**).
4. Thermal cycler.
5. Quick Ligation Kit (New England Biolabs, Beverly, MA, USA; *see* **Note 3**).
6. T4 DNA Ligase Reaction Buffer (New England Biolabs).
7. BsaI-HF (New England Biolabs; *see* **Note 4**).
8. Esp3I (Thermo Scientific Japan, Tokyo, Japan; *see* **Note 5**).
9. MscI.
10. Dithiothreitol (DTT).
11. 5-Bromo-4-chloro-3-indolyl β-d-galactopyranoside (X-gal)/isopropyl β-d-1-thiogalactopyranoside (IPTG) solution for blue/white selection.
12. Chemically competent bacterial cells, such as XL1-Blue or XL10-Gold.
13. Standard Taq DNA polymerase.

14. Primers for colony PCR (*see* **Note 6**).
    pCR8_F1: 5′-TTGATGCCTGGCAGTTCCCT-3′
    pCR8_R1: 5′-CGAACCGAACAGGCTTATGT-3′
    TALE-Fv2: 5′-GAGCACCCCTCAACCTGACCCC-3′
    TALE-R: 5′-CTCGAAAGCTGGGCCACGATTG-3′

3 Methods

3.1 Setting Up the TALEN Construction System

1. Culture all 35 clones contained in the Platinum Gate TALEN Kit in liquid LB media with 100 μg/ml ampicillin or spectinomycin overnight at 37 °C with shaking.
2. Extract the plasmids using Plasmid Miniprep Kit.
3. Measure the DNA concentration and adjust it to 50 ng/μl for module plasmids and final destination vectors, and 25 ng/μl for intermediate array vectors.

3.2 Design of TALENs

1. Access "TALEN Targeter" [16] at Cornell University's website (https://tale-nt.cac.cornell.edu/; *see* **Note 7**).
2. Paste genome sequence in the FASTA format into the sequence box (*see* **Note 8**).
3. Select the "Provide Custom Spacer/RVD Lengths" tab and set an appropriate spacer length and repeat array (*see* **Note 9**).
4. Set "G Substitute," "Filter Options," and "Streubel et al. guidelines" as "NN," "Show all TALEN pairs (include redundant TALENs)," and "On," respectively (*see* **Note 10**).
5. Press "submit" to obtain the result (*see* **Note 11**).

3.3 First-Step Golden Gate Assembly

Assemble corresponding DNA-binding modules in the intermediate array vectors. Using the Platinum Gate TALEN Kit, TALENs with 6–21 modules can be constructed. Note that the correct combination of array vectors used in the first assembly step must be selected in accordance with the number of modules, as shown in Fig. 1. As examples, the target sequence and correspondence table for constructing *eGFP* TALEN, described previously [6, 14], are shown in Fig. 2a, b, respectively.

1. Prepare the following reaction mixture in a PCR tube. Solutions in the black box should be prepared as a premix, and 0.4 μl added to each sample (*see* **Note 12**).

First-step Golden Gate assembly

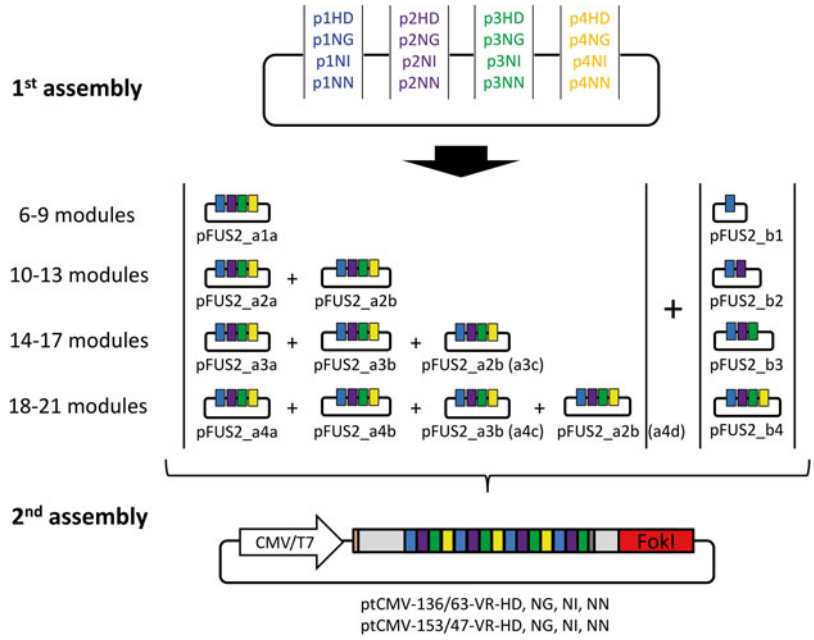


Fig. 1 Schematic overview of the Platinum Gate TALEN Kit. In the first assembly step, four or fewer modules are ligated into intermediate array vectors. In the second step, constructed arrays are joined into a mammalian expression/in vitro transcription vector. *CMV* cytomegalovirus promoter

A

Left TALEN

```
T CTTCAAGGACGACGGCAACT acaagacccgcgccg AGGTGAAGTTCGAGGGCG A
A GAAGTTCCTGCTGCCGTTGA tgttctgggcgcggc TCCACTTCAAGCTCCCGC T
```

Right TALEN

B

Left TALEN	Vector	pFUS2_a4a	pFUS2_a4b	pFUS2_a3b	pFUS2_a2b	pFUS2_b3
	Insert	p1HD p2NG p3NG p4HD (C-T-T-C)	p1NI p2NI p3NN p4NN (A-A-G-G)	p1NI p2HD p3NN p4NI (A-C-G-A)	p1HD p2NN p3NN p4HD (C-G-G-C)	p1NI p2NI p3HD (A-A-C)
Right TALEN	Vector	pFUS2_a4a	pFUS2_a4b	pFUS2_a3b	pFUS2_a2b	pFUS2_b1
	Insert	p1HD p2NN p3HD p4HD (C-G-C-C)	p1HD p2NG p3HD p4NN (C-T-C-G)	p1NI p2NI p3HD p4NG (A-A-C-T)	p1NG p2HD p3NI p4HD (T-C-A-C)	p1HD (C)

Fig. 2 Examples of TALEN target sequence (**a**) and corresponding table of vectors and inserts for the first assembly (**b**). Target sequences of left and right TALENs are *underlined*. Note that the innermost bases should be ignored in this step, because the last DNA-binding repeat is included in the destination vector for the second assembly step. *Letters in brackets* indicate recognition sequences

Number of modules	1	2	3	4
Intermediate array vector (25 ng/μl)	0.3			
Module plasmid(s) (50 ng/μl)	0.3 × 1	0.3 × 2	0.3 × 3	0.3 × 4
T4 DNA Ligase Reaction Buffer	0.2			
BsaI-HF	0.1	Premix		
Quick Ligase	0.1			
Water	1	0.7	0.4	(0.1)[a]
Total	2 (μl)			

[a]Consider this as zero

2. Place the tube in a thermal cycler and perform the following program: 37 °C for 5 min and 16 °C for 10 min for three cycles (*see* **Note 13**).
3. Remove the tube from the thermal cycler and add the following premixed solutions to each tube: 0.25 μl of NEBuffer 4, 0.25 μl of 10× BSA, and 0.1 μl of BsaI-HF (*see* **Note 14**).
4. Return the tube to the thermal cycler and perform the following program: 50 °C for 30 min, followed by 80 °C for 5 min.
5. Remove the tube and transform the product directly into competent bacterial cells.
6. Plate the transformants onto an LB/spectinomycin plate containing X-gal/IPTG.
7. Culture the plate overnight at 37 °C.

3.4 Colony PCR Screening for the First-Step Clones

1. Pick white colonies and perform colony PCR using pCR8_F1 and pCR8_R1 primers (*see* **Note 15**). Examples of band patterns amplified from correctly assembled clones are shown in Fig. 3a.

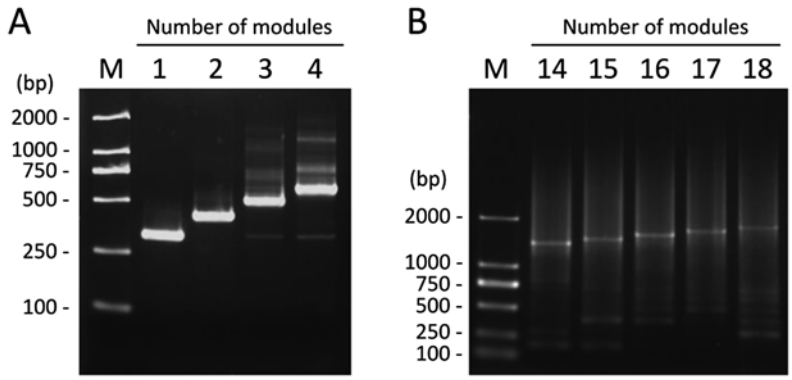


Fig. 3 Gel images of PCR screening in the first step (**a**) and the second step (**b**). In these experiments, plasmid DNAs were used as templates. Minor bands were caused by repetitive sequences in the DNA-binding repeats. *M* wide-range DNA ladder (100–2000 bp) (Takara Bio, Shiga, Japan)

2. Culture the intended clones in liquid LB media with 100 μg/ml spectinomycin overnight at 37 °C with shaking.
3. Extract the plasmids using Plasmid Miniprep Kit.
4. Measure the DNA concentration and adjust it to 50 ng/μl.

3.5 Second-Step Golden Gate Assembly

Assemble 1- to 4-module-assembled plasmids in the destination vectors. Destination vectors contain the last half repeat; therefore, you should choose the appropriate vector in accord with each target sequence. Although the basic procedures are similar to the first assembly step, enzymes and reaction conditions differ slightly.

1. Prepare the following reaction mixture in a PCR tube. Solutions in the black box should be prepared as a premix, and 0.8 μl added to each sample (*see* **Note 16**).
   Second-step Golden Gate assembly

Number of modules	6–9	10–13	14–17	18–21
pFUS2_a plasmid(s) (50 ng/μl)	0.6 × 1	0.6 × 2	0.6 × 3	0.6 × 4
pFUS2_b plasmid (50 ng/μl)	0.6			
Destination vector (50 ng/μl)	0.3			
T4 DNA Ligase Reaction Buffer	0.4			
Esp3I	0.2	Premix		
Quick Ligase	0.2			
Water	1.7	1.1	0.5	(−0.1)[a]
Total	4 (μl)			

[a]Consider this as zero

2. Place the tube in the thermal cycler and perform the following program: 37 °C for 5 min and 16 °C for 10 min for six cycles (*see* **Note 17**).
3. Remove the tube from the thermal cycler and add the following premixed solutions to each tube: 0.5 μl of 10× Tango buffer, 0.5 μl of 10 mM DTT, and 0.2 μl of Esp3I (*see* **Note 18**).
4. Return the tube to the thermal cycler and perform the following program: 37 °C for 1 h, followed by 80 °C for 5 min.
5. Remove the tube and transform the product directly to competent bacterial cells.
6. Plate the transformants onto an LB/ampicillin plate containing X-gal/IPTG.
7. Culture the plate overnight at 37 °C.

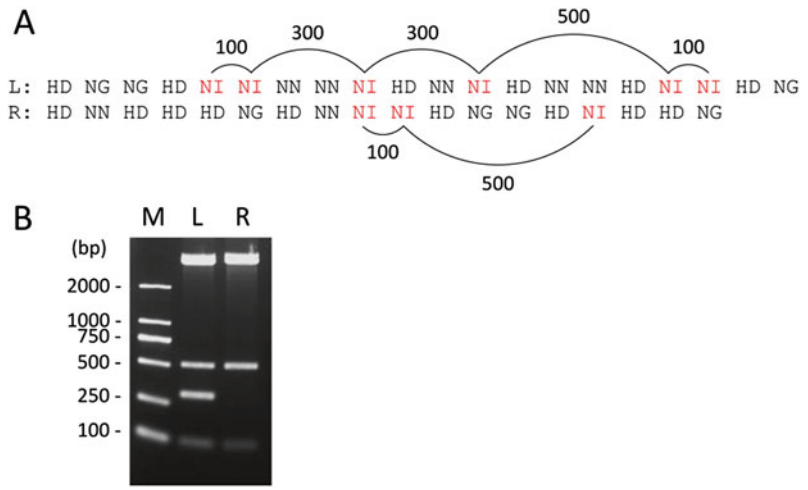


Fig. 4 Schematic illustration (**a**) and gel image (**b**) of MscI digestion. *Red letters* indicate positions of DNA-binding modules harboring NI as the repeat-variable di-residue (RVD), which are cut by MscI. Note that NI modules at the last repeat cannot be cut. Approximate sizes of fragments generated by MscI digestion are illustrated in panel **a**. *L* left TALEN, *R* right TALEN, *M* wide-range DNA ladder (100–2000 bp)

3.6 Colony PCR Screening for the Second-Step Clones

1. Pick white colonies and perform colony PCR using TALE-Fv2 and TALE-R primers (*see* **Note 19**). Examples of band patterns amplified from correctly assembled clones are shown in Fig. 3b.
2. Culture the intended clones in liquid LB media with 100 μg/ml ampicillin overnight at 37 °C with shaking.
3. Extract the plasmids using Plasmid Miniprep Kit.
4. Measure the DNA concentration and adjust it appropriately.

3.7 Checking the TALEN Plasmid by MscI Digestion (Optional)

The Golden Gate cloning enables highly accurate modular assembly; therefore, colony PCR screening is generally sufficient for checking the clones. Nonetheless, restriction digestion may be used as another simple verification method. As shown in Fig. 4a, b, MscI enzyme cuts modules harboring NI as RVD only, and we can easily verify the accuracy of assembly.

4 Notes

1. The Platinum Gate TALEN Kit contains 16 module plasmids (p[1–4]HD, p[1–4]NG, p[1–4]NI, p[1–4]NN), 11 intermediate array vectors (pFUS2_[a1a, a2a, a2b, a3a, a3b, a4a, a4b, b1, b2, b3, b4]), and 8 destination vectors (ptCMV-136/63-VR-[HD, NG, NI, NN], ptCMV-153/47-VR-[HD, NG, NI, NN]) (Fig. 1). Module plasmids and destination vectors

contain an ampicillin resistance gene, whereas intermediate array vectors contain a spectinomycin resistance gene.

2. The success rate of Golden Gate assembly is greatly affected by the quality of the plasmid DNA. We have confirmed that the GenElute™ HP Plasmid Miniprep Kit (Sigma-Aldrich, St. Louis, MO, USA) and the ChargeSwitch® Pro Plasmid Miniprep Kit (Life Technologies, Carlsbad, CA, USA) could be used.
3. The Quick Ligation Kit contains Quick Ligase and 2× reaction buffer. In our protocol, only the Quick Ligase is used for the first and second assembly, and the separately sold 10× T4 DNA Ligase Reaction Buffer is used in these assembly steps instead of 2× buffer.
4. New England Biolabs sells conventional BsaI and high-fidelity BsaI-HF. Both enzymes can be used for the first assembly, but we recommend the use of BsaI-HF (#R3535S or #R3535L), because conventional BsaI is likely to become less active during long storage in our experience.
5. Thermo Scientific sells conventional Esp3I and FastDigest Esp3I. FastDigest enzymes are not suitable for Golden Gate cloning; therefore, you should choose the conventional enzyme (#ER0451 or #ER0452).
6. The optimal annealing temperature is 55 °C for pCR8_F1 and pCR8_R1, and 66 °C for TALE-Fv2 and TALE-R. Note that TALE-Fv2 and TALE-R primers are designed to amplify constructed TALEN plasmids and cannot be used for empty destination vectors before assembling DNA-binding repeats.
7. TALEN Targeter is one of the tools on the TAL Effector Nucleotide Targeter (TALE-NT) web resource. Two versions of the TALEN Targeter program, with or without design guidelines, are currently available on TALE-NT. We have previously used the old version with design guidelines; however, because of the high efficiency of Platinum TALENs, we use the newer version of TALEN Targeter without design guidelines currently (https://tale-nt.cac.cornell.edu/node/add/talen).
8. The required length of the genomic region depends on the base sequence and conditions set, such as spacer region and repeat number. Typically, several hundred bases are enough to design multiple TALEN sets.
9. Regarding spacer length, we recommend 15–19 (around 17) for ptCMV-136/63 vectors and 12–16 (around 15) for ptCMV-153/47 vectors. Optimal numbers of repeat arrays are 15–20 (around 17) for both scaffolds.
10. The Platinum Gate TALEN Kit contains only NN modules for guanine recognition. Turn the Streubel et al. guidelines off when no results matching this criterion are found.

11. Other parameters such as "Upstream Base" and "Expires" need not be changed if there is no reason to do so. After submission, you will find the table of search results on the web browser. However, the full information is not displayed on the table; therefore, we strongly recommend the text file labeled as "Result File (Tab-Delimited)" is saved and opened with spreadsheet software such as Microsoft Excel.
12. If the assembly does not work well with these conditions, increase the total reaction volume up to 10 μl.
13. Increase the number of cycles if the efficacy of the assembly is low.
14. This additional digestion with fresh enzyme and buffer dramatically reduces the number of blue colonies.
15. Any standard DNA polymerase can be used. Two colonies for each sample are sufficient to screen the correct clones.
16. If the assembly does not work well with these conditions, increase the total reaction volume up to 10 μl.
17. Increase the number of cycles if the efficacy of the assembly is low.
18. This additional digestion with fresh enzyme and buffer dramatically reduces the number of blue colonies.
19. Any standard DNA polymerase can be used. Two colonies for each sample are sufficient to screen the correct clones.

References

1. Sakuma T, Woltjen K (2014) Nuclease-mediated genome editing: at the front-line of functional genomics technology. Dev Growth Differ 56:2–13
2. Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14:49–55
3. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. PLoS One 3:e3647
4. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82
5. Sakuma T, Hosoi S, Woltjen K, Suzuki K, Kashiwagi K, Wada H, Ochiai H, Miyamoto T, Kawai N, Sasakura Y, Matsuura S, Okada Y, Kawahara A, Hayashi S, Yamamoto T (2013) Efficient TALEN construction and evaluation methods for human cell and animal applications. Genes Cells 18:315–326
6. Sakuma T, Ochiai H, Kaneko T, Mashimo T, Tokumasu D, Sakane Y, Suzuki K, Miyamoto T, Sakamoto N, Matsuura S, Yamamoto T (2013) Repeating pattern of non-RVD variations in DNA-binding modules enhances TALEN activity. Sci Rep 3:3379
7. Ochiai H, Miyamoto T, Kanai A, Hosoba K, Sakuma T, Kudo Y, Asami K, Ogawa A, Watanabe A, Kajii T, Yamamoto T, Matsuura S (2014) TALEN-mediated single-base-pair editing identification of an intergenic mutation upstream of BUB1B as causative of PCS (MVA) syndrome. Proc Natl Acad Sci U S A 111:1461–1466
8. Ninagawa S, Okada T, Sumitomo Y, Kamiya Y, Kato K, Horimoto S, Ishikawa T, Takeda S, Sakuma T, Yamamoto T, Mori K (2014) EDEM2 initiates mammalian glycoprotein ERAD by catalyzing the first mannose trimming step. J Cell Biol 206:347–356
9. Sugi T, Sakuma T, Ohtani T, Yamamoto T (2014) Versatile strategy for isolating TALEN-mediated knockout mutants in *Caenorhabditis elegans*. Dev Growth Differ 56:78–85

10. Hosoi S, Sakuma T, Sakamoto N, Yamamoto T (2014) Targeted mutagenesis in sea urchin embryos using TALENs. Dev Growth Differ 56:92–97
11. Treen N, Yoshida K, Sakuma T, Sasaki H, Kawai N, Yamamoto T, Sasakura Y (2014) Tissue-specific and ubiquitous gene knockouts by TALEN electroporation provide new approaches to investigating gene function in *Ciona*. Development 141:481–487
12. Hayashi T, Sakamoto K, Sakuma T, Yokotani N, Inoue T, Kawaguchi E, Agata K, Yamamoto T, Takeuchi T (2014) TALENs efficiently disrupt the target gene in Iberian ribbed newts (*Pleurodeles waltl*), an experimental model animal for regeneration. Dev Growth Differ 56:115–121
13. Sakane Y, Sakuma T, Kashiwagi K, Kashiwagi A, Yamamoto T, Suzuki K (2014) Targeted mutagenesis of multiple and paralogous genes in *Xenopus laevis* using two pairs of transcription activator-like effector nucleases. Dev Growth Differ 56:108–114
14. Nakagawa Y, Yamamoto T, Suzuki K, Araki K, Takeda N, Ohmuraya M, Sakuma T (2014) Screening methods to identify TALEN-mediated knockout mice. Exp Anim 63:79–84
15. Nakagawa Y, Sakuma T, Nakagata N, Yamasaki S, Takeda N, Ohmuraya M, Yamamoto T (2014) Application of oocyte cryopreservation technology in TALEN-mediated mouse genome editing. Exp Anim 63:349–355
16. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40: W117–W122

Chapter 7

Design, Assembly, and Characterization of TALE-Based Transcriptional Activators and Repressors

Pratiksha I. Thakore and Charles A. Gersbach

Abstract

Transcription activator-like effectors (TALEs) are modular DNA-binding proteins that can be fused to a variety of effector domains to regulate the epigenome. Nucleotide recognition by TALE monomers follows a simple cipher, making this a powerful and versatile method to activate or repress gene expression. Described here are methods to design, assemble, and test TALE transcription factors (TALE-TFs) for control of endogenous gene expression. In this protocol, TALE arrays are constructed by Golden Gate cloning and tested for activity by transfection and quantitative RT-PCR. These methods for engineering TALE-TFs are useful for studies in reverse genetics and genomics, synthetic biology, and gene therapy.

Key words Protein engineering, Transcription factors, Synthetic biology, Genome engineering, Gene regulation, Gene activation, Gene repression

1 Introduction

Programmable synthetic transcription factors are a powerful tool for targeted control of gene expression. Synthetic transcription factors are typically proteins that consist of a DNA-binding domain linked to an effector molecule that modifies transcription at the targeted locus. Zinc finger proteins [1] and transcription activator-like effector (TALE) proteins [2, 3] are two commonly used programmable DNA-binding protein domains. TALEs are naturally occurring proteins from the pathogenic bacteria genus *Xanthomonas* that target and regulate genes in host plants. The DNA-binding domains of these proteins contain 33–35 amino acid repeats in which the 12th and 13th positions within each repeat, termed the repeat variable di-residue (RVD), define single-nucleotide specificity (Fig. 1) [4, 5]. The RVD-base pair recognition follows a simple code, and individual TALE repeats can be genetically linked together to recognize a series of contiguous DNA nucleotides as a single-TALE protein. Thus this technology is a platform for engineering customizable DNA-binding proteins [6–8]. The original

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_7,

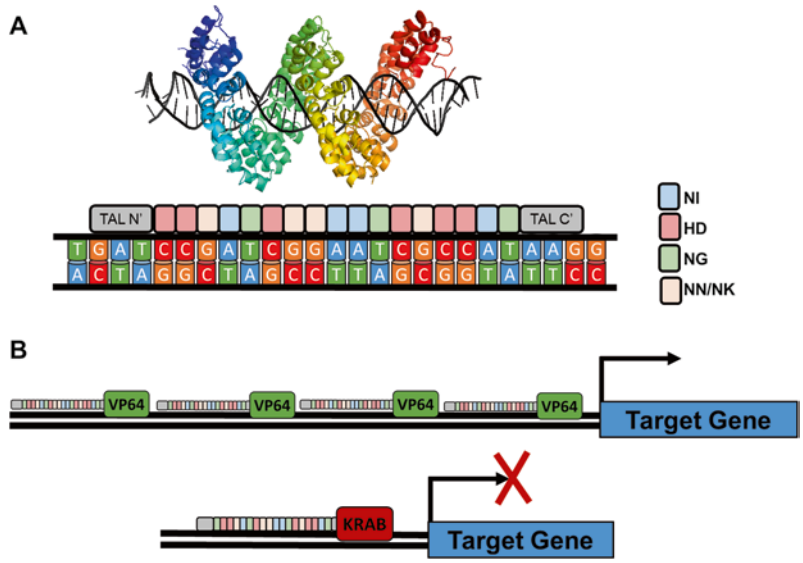


Fig. 1 (**a**) TALE arrays consist of RVD modules that bind single nucleotides according to a predefined code (PDB files 3UGM for TALE protein). (**b**) TALE arrays can be fused to effector domains to create synthetic transcription factors. Multiplexing TALE-VP64 fusions can result in activation of a target gene. Repressor domains, such as KRAB, can also be genetically linked to TALEs to repress target gene expression

Table 1
DNA recognition cipher for TALE RVDs

Target nucleotide	RVD
A	NI
C	HD
G	NN, NK
T	NG

amino acid code for recognition was described as NI for adenine, HD for cytosine, NG for thymine, and NN for guanine or adenine (Table 1), but alternatives such as NH or NK have since been developed for more specific guanine recognition [9–11].

TALE transcription factors (TALE-TFs) are highly versatile and can be designed to target almost any sequence for control of gene expression with a variety of different regulatory domains (Fig. 1b). TALE transcription factors (TALE-TFs) have been used to activate or repress gene expression in a variety of organisms, including plants, drosophila, mammalian cells, and mice following delivery in vivo [7, 12–17]. Libraries of TALE-TFs orthogonal to the human genome have also been developed to facilitate the use of these artificial transcription factors in synthetic gene circuits [18, 19].

For gene activation with TALE-TFs, a commonly used effector module is VP64, a tandem repeat of four copies of the minimal acidic activation domain of the VP16 transactivator from herpes simplex virus [7, 20, 21]. Other domains such as the activation domain of the human p65 subunit of NFκB have also been used to create TALE activators [20, 21]. For gene inhibition, Kruppel-associated box (KRAB) and Sid4 repressors have been linked to TALEs to achieve sequence-specific repression in mammalian cells [9, 13]. More recently, epigenetic modifiers, including enzymes that manipulate histone posttranslational modifications and DNA methylation, have been fused to TALEs to control gene expression [14, 22, 23].

TALEs are versatile DNA-binding proteins with the potential to target any nucleotide sequence of interest according to a well-described RVD code [2, 3]. Optimal target site selection, however, is still an active area of research for synthetic transcription factors in general. Although TALE RVD subunits bind single nucleotides with a relatively simple code, different TALE transcription factors designed according to the same code result in different levels of activity, perhaps because of differences in binding efficiency or differences in chromatin state or conformation at the target site. For the methods described here, the TALE target site must begin with a T for optimal activity, although TALE architectures that can bind any 5′ nucleotide with similar efficiencies have been developed through directed evolution [24] or expanding the technology to TALE domains from other bacterial species [25]. For TALE activators, it has been generally observed that having multiple transcription factors acting at a single promoter is necessary for robust gene activation [20, 21]. For example, multiplexing up to six TALE transcription factors at a single endogenous promoter has resulted in synergistic activation for multiple different gene targets [20, 21]. Target sites selected up to 700 base pairs upstream of the transcription start site of the gene of interest have resulted in activation of downstream gene expression [7, 20, 21]. TALEs are often designed to bind within DNase I-hypersensitive regions to increase targeting efficiency [21], but studies have shown that TALE-TFs can also act in regions characterized by closed chromatin [20]. TALE targeting is also not limited to gene promoters: TALE activators and inhibitors directed to enhancers have been shown to regulate gene expression in mammalian cells [16] and Drosophila [13].

Rapid assembly of custom, sequence-specific TALEs is possible with Golden Gate assembly [8], solid-phase synthesis [26, 27], or ligation-independent assembly [28]. In Golden Gate assembly, type IIS restriction enzymes create distinct, position-dependent overhangs to concatemerize TALE monomers with distinct RVDs in predefined orders. Solid-phase synthesis and ligation-independent strategies have been configured for high-throughput, automated

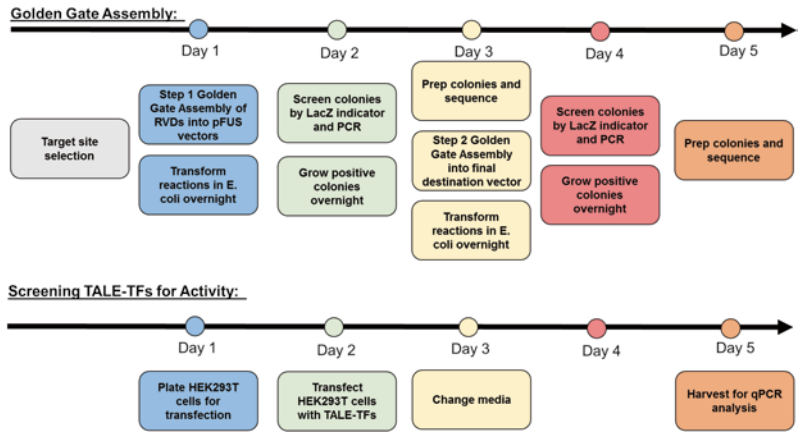


Fig. 2 Timeline for TALE-TF assembly and screening

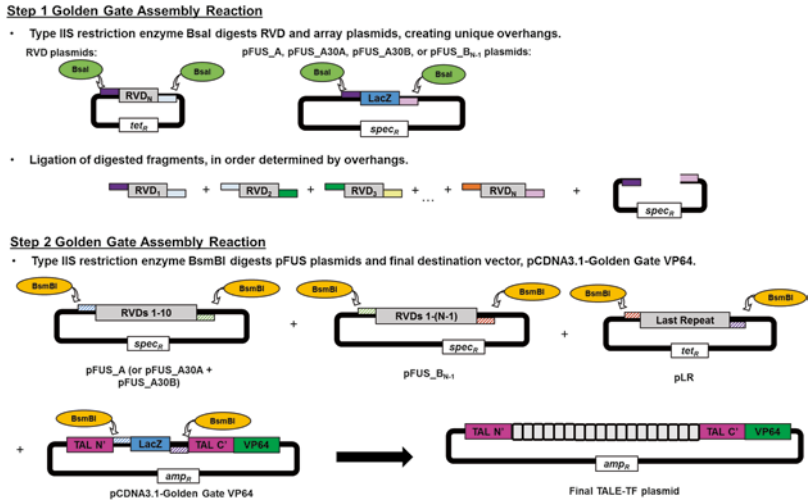


Fig. 3 Schematic of Golden Gate TALE-TF assembly strategy

TALE production using specialized protocols and equipment. The method described here is the Golden Gate strategy for cloning of TALE-TFs into mammalian expression plasmids with medium throughput, as pioneered by the Bogdanove and Voytas groups [8]. The Golden Gate method can be used to create custom TALE arrays of 13–33 RVDs in length within 5 days after target site selection (Figs. 2 and 3). Throughput is limited by the number of cloning reactions handled by a single researcher, but the concurrent assembly of 12 TALE arrays is possible. We also describe a protocol to deliver and assay assembled TALE-TFs for activity by transfection and quantitative RT-PCR (qRT-PCR).

2 Materials

2.1 Golden Gate Cloning of TALE-TFs

1. Oligonucleotide primers for polymerase chain reaction and DNA sequencing:
   (a) Step 1 Fwd: 5′ ttgatgcctggcagttccct.
   (b) Step 1 Rev: 5′ cgaaccgaacaggcttatgt.
   (c) Step 2 Fwd: 5′ ttggcgtcggcaaacagtgg.
   (d) Step 2 Rev: 5′ ggcgacgaggtggtcgttgg.
   (e) Step 2 Seq: 5′ catcgcgcaatgcactgac.
2. Golden Gate TALEN and TAL effector kit 2.0 (Addgene TALEN Kit 1000000024), including:
   (a) RVD module plasmids: pNN(1–10), pHD(1–10), pNG(1–10), pNI(1–10), pNK(1–10).
   (b) pFUS shuttle vectors: pFUS_A, pFUS_A30A, pFUS_A30B, pFUS_B.
   (c) RVD last half-repeat plasmids: pLR-NN, pLR-HD, pLR-NG, pLR-NI.
3. TALE expression plasmid pcDNA3.1-GoldenGate (Addgene 47388).
4. TALE-VP64 expression plasmid pcDNA3.1-GoldenGate-VP64 (Addgene 47389).
5. Restriction enzymes and appropriate buffers:
   (a) BsmBI Fast Digest (Thermo Scientific).
   (b) BsaI (New England Biolabs).
6. PlasmidSafe nuclease and appropriate buffers (Epicentre).
7. Standard equipment and reagents for gel electrophoresis.
8. T4 DNA Ligase and appropriate buffer (New England Biolabs).
9. Chemically competent DH5α *E. coli* (Invitrogen).
10. SOC Medium (Invitrogen).
11. LB-agar plates containing 100 μg/mL ampicillin.
12. LB media containing 100 μg/mL ampicillin.
13. LB-agar plates containing 50 μg/mL spectinomycin.
14. LB media containing 100 μg/mL spectinomycin.
15. 0.1 M Isopropyl-beta-D-thiogalactoside (IPTG) in filtered H_2O.
16. 20 mg/mL 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside (X-gal) in dimethyl sulfoxide (DMSO).
17. Taq DNA Polymerase and appropriate buffers (New England Biolabs).
18. Standard equipment and reagents for purification of bacterial plasmid DNA, such as the Qiagen Spin Miniprep Kit.

2.2 Testing TALE-TFs for Activity

1. Mammalian cell line that can be readily transfected, such as HEK293T cells.
2. Lipofectamine 2000 Reagent (Invitrogen), or other suitable transfection reagent.
3. OptiMEM Reduced Serum Media (Life Technologies).
4. DMEM supplemented with 10 % fetal bovine serum (FBS).
5. DMEM supplemented with 10 % FBS and 1 % penicillin/streptomycin (P/S).
6. Dulbecco's phosphate-buffered saline (DPBS).
7. 0.25 % Trypsin.
8. 24-Well tissue culture-treated plates.
9. Transfection plasmids.
   (a) "Empty" expression plasmid (i.e., plasmid containing a promoter but no expressed gene).
   (b) pMAX-GFP to test transfection efficiency (optional).
   (c) Negative control TALE-TF plasmid (*see* **Note 1**).
10. RNase/DNase-free H_2O.
11. Quantitative PCR primers for target gene and control gene.
12. Standard equipment and reagents for purification of total RNA from mammalian cells, such as Qiagen RNeasy Plus Mini Kit.
13. Standard equipment and reagents for cDNA synthesis by reverse transcription, such as SuperScript Vilo cDNA Synthesis Kit (Invitrogen).
14. Standard equipment and reagents for quantitative PCR, such as Perfecta SYBR Green Supermix (Quanta Biosciences).

3 Methods

3.1 Target Site Selection

1. Determine transcription start site of target gene by searching gene annotation on the National Center for Biotechnology Information (NCBI) online database (http://www.ncbi.nlm.nih.gov/) or the UCSC Genome Browser (https://genome.ucsc.edu/).
2. Select region within 500 base pairs of transcription start site. Enter selected region into TAL Effector Targeter (https://tale-nt.cac.cornell.edu/node/add/single-tale) [29] (*see* **Note 2**).
3. Select the following default design guidelines (*see* **Note 3**):
   (a) The target sequence must begin with a T.
   (b) GC composition.
   (c) Use NH as the G substitute.
   (d) Do not allow sites to end in G.

Table 2
Example genome sizes and recommended minimum target sizes for common model organisms

Species	Haploid genome size (bp)	Recommended minimum target size (bp)
Saccharomyces cerevisiae	1.2×10^7	12
Drosophila melanogaster	1.3×10^8	14
Mus musculus	2.7×10^9	16
Homo sapiens	3.2×10^9	16

(e) Require A, C, or G at position 1.

(f) Require a T at position N.

4. Determine desired sequence length. Statistically, n is the minimum target size to achieve specificity in a genome of random DNA sequence of size equal to the target species genome, where $4n$ is greater than the haploid genome size (Table 2). Target sequences for this Golden Gate cloning kit must be at least 13 base pairs, including the first T base pair.
5. Examine RVD sequences for specificity in target genome. Off-target scoring can be performed using the Target Finder application on the TAL Effector Nucleotide Targeter 2.0. Three or more mismatches to potential off-target sites are recommended for designing specific TALE-TFs [30] (*see* **Note 4**).
6. Select target sequences evenly distributed throughout the region of interest. It is recommended that at least four target sites are selected per target gene based on studies of synergistic effects between multiple TALE-TFs [20, 21].

3.2 Golden Gate Assembly of TALE-TFs

1. Collect RVD plasmids corresponding to target sequence (*see* **Note 5**). For the cloning method described here, the 5′ T is recognized by a UPA box element in the N′ terminus of the TALE protein [4, 5, 8] and is thus not considered in the assembly process. For each RVD in the target site, select the module plasmid that matches the RVD and its corresponding position. Omit the last half-repeat RVD in the target sequence, as that will be incorporated in the final Golden Gate reaction.
2. Prepare Step 1 Golden Gate digestion-ligation reaction. For TALE arrays of N RVDs, where N is 12–21, RVDs 1–10 will be shuttled into pFUS_A and RVDs 11 to $N-1$ will be shuttled into pFUS_B($N-11$). If N is 22–31, then RVDs 1–10 are cloned into pFUS_A30A, RVDs 11–20 are cloned into pFUS_A30B, and RVDs 21 up to $N-1$ are cloned into the pFUS_B($N-21$) plasmid.

Golden Gate Reaction Step 1:

1 μL BsaI.

1 μL T4 DNA Ligase.

2 μL 10× T4 DNA Ligase buffer.

150 ng of appropriate pFUS vector.

150 ng of EACH RVD module vector.

H_2O up to 20 μL final reaction volume.

	Shuttle vector	**RVD vectors**
Reaction A	pFUS_A	RVD module vectors 1–10
Reaction B	pFUS_B($N-11$)	RVD module vectors 11 – ($N-1$)

	Shuttle vector	**RVD vectors**
Reaction A	pFUS_A30A	RVD module vectors 1–10
Reaction B	pFUS_A30B	RVD module vectors 11–20
Reaction C	pFUS_B($N-21$)	RVD module vectors 21 – ($N-1$)

3. Run Step 1 Golden Gate reactions under the following thermocycler conditions for plasmid digestion and ligation:

Temperature (°C)	**Time (min)**	**Repeat**
37 16	5 10	Steps 1 and 2, 10×
50	5	1×
80	5	1×

4. Treat each Step 1 Golden Gate reaction with Plasmid-Safe nuclease in order to process unligated linear dsDNA fragments. This will destroy incomplete ligation products that do not contain all of the RVD repeats and lower the background of undesired transformants.

   To each Step 1 Golden Gate reaction, add:

   1 μL 10 mM ATP.

   1 μL Plasmid-Safe nuclease.

   Incubate at 37 °C for 1 h.

5. Transform 5 μL of each Step 1 Golden Gate reaction into 50 μL of DH5α chemically competent bacteria (*see* **Note 6**). Plate transformations on spectinomycin-agar plates, with 40 μL of X-gal and 40 μL of IPTG. Spectinomycin resistance ensures that only colonies transformed by pFUS-based plasmids will grow. Incubate for 12–14 h at 37 °C.

6. Screen colonies for correct RVD array insertions using colony PCR. Plates should have a mix of blue and white colonies, with the ratio of blue to white depending on the efficiency of the Step 1 Golden Gate reactions. White colonies indicate that the LacZ expression cassette contained in the pFUS backbone has been exchanged for the RVD modules. Pick one to three white colonies to check by colony PCR, using Step 1 Fwd and Step 2 Rev primers. Add colony sequentially to the PCR reaction and then to the starter LB culture. Include negative control PCR reactions with empty pFUS_A and pFUS_B plasmids as the DNA template.

   PCR reaction:

   2.5 μL ThermoPol Buffer.

   0.1 μL Taq DNA Polymerase.

   1 μL 10 mM dNTP.

   1 μL primer 1.

   1 μL primer 2.

   19.4 μL H_2O.

7. Run the PCR with the following thermocycler conditions:

Temperature (°C)	Time (s)	Repeat
95	30	1×
95	30	Steps 2–4, 30×
55	30	
68	105	
68	300	1×

8. Screen for positive colonies by gel electrophoresis of PCR products. A pFUS plasmid containing ten RVD repeats will yield a 1.2 kb product with primers Step 1 Fwd and Step 1 Rev (Fig. 4) (*see* **Note 7**).
9. Inoculate the corresponding starter cultures for positive colonies for overnight culture in 5 mL LB containing spectinomycin.
10. Purify plasmid DNA from overnight culture using standard DNA miniprep kit, according to the manufacturer's protocol. Submit for DNA sequencing using primers Step 1 Fwd and Step 1 Rev. For TALE-TFs of N RVDs, after this step, there should be the following vectors:

    Vector 1A: pFUSA with RVDs 1–10.

    Vector 1B: pFUSB($N-1$) plasmid with RVDs 11 through $N-1$.

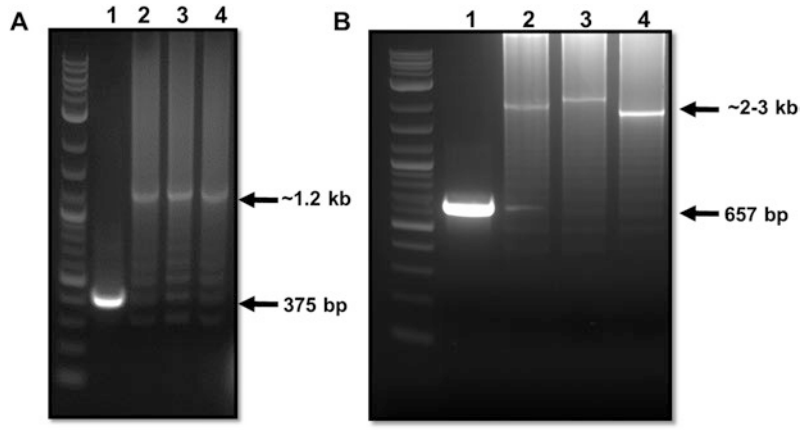


Fig. 4 Example of colony PCR for (**a**) Step 1 and (**b**) Step 2 Golden Gate reactions. (**a**) *Lane 1* demonstrates a negative colony PCR result for a pFUS plasmid without any inserted RVDs. *Lanes 2–4* show positive colonies for a pFUS plasmid containing ten RVDs. The products have a band at ~1.2 kb, accompanied by a smear and laddering effect due to the repetitive nature of the TALEs. (**b**) *Lane 1* is an example of a negative clone in the Step 2 Golden Gate reaction. *Lanes 2–4* demonstrate positive clones for a pCDNA3.1-Golden Gate destination vector with 12–14 RVDs inserted. In a positive clone, a faint band appears between 2 and 3 kb, depending on the length of the TALE, accompanied by a smear and laddering effect

OR

Vector 1A: pFUSA30A with RVDs 1–10.

Vector 1B: pFUSA30B with RVDs 11–20.

Vector 1C: pFUSB($N-1$) plasmid with RVDs 21 through $N-1$.

11. Prepare Step 2 Golden Gate digestion-ligation reaction. In this reaction, you will combine the intermediate TALE arrays made in the Step 1 Golden Gate reaction and the last half-repeat of the target RVD sequence with the destination vector, which contains the TALE N-terminus and the C-terminus fused to an effector domain. For activation, we use a TALE-VP64 fusion: pcDNA3.1-GoldenGate-VP64 (*see* **Note 8**). There are five last repeat vectors: pLR-HD, pLR-NG, pLR-NI, pLR-NN, and pLR-NH. These correspond to the last, or Nth, nucleotide of your target sequence.

Step 2 Golden Gate reaction:

150 ng each of vectors 1, 2, and 3 from previous step.

150 ng of appropriate pLR vector.

75 ng of destination vector.

1 μL BsmBI restriction enzyme.

1 μL T4 DNA Ligase.

2 μL 10× T4 DNA Ligase Buffer.

H_2O up to a 20 μL final reaction volume.

12. Run Step 2 Golden Gate digestion reaction under the following thermocycler conditions:

Temperature (°C)	Duration (min)	Repeats
37 16	5 10	Steps 1 and 2, 10×
37	15	1×
80	5	1×

13. Transform 5 μL of the Step 2 reaction in 50 μL of DH5α chemically competent bacteria. Plate transformation on ampicillin-agar plates with 40 μL of X-gal and 40 μL of IPTG (*see* **Note 9**).
14. Incubate plates for 14 h overnight at 37 °C.
15. The next morning, perform colony PCR on one to three white colonies with primers Step 2 Fwd and Step 2 Rev:

    2.5 μL ThermoPol Buffer.

    0.1 μL Taq DNA Polymerase.

    1 μL 10 mM dNTP.

    1 μL primer 3.

    1 μL primer 4.

    19.4 μL H_2O.

16. Run PCR with the following thermocycler conditions:

Temperature (°C)	Time (s)	Repeat
95	30	1×
95 55 68	30 30 180	Steps 2–4, 30×
68	300	1×

17. Verify clones by gel electrophoresis using a 1.2 % agarose gel. A correct product displays a faint band between 2 and 3 kb, for TALEs of 12–22 RVDs, accompanied by a smear and ladder effect due to the repetitive nature of the RVD array (Fig. 4). Longer TALEs may not fully amplify, in which case only a ladder and smear will be seen.
18. Inoculate positive colonies in LB with ampicillin overnight at 37 °C for 14–16 h, shaking at 225 rpm.
19. Purify plasmid DNA from overnight culture using standard DNA miniprep kit, according to the manufacturer's protocol. Verify the sequence of the assembled TALE-TF with primers Step 2 Seq and Step 2 Rev (*see* **Note 10**).

3.3 Evaluating Activity of TALE-TFs at Endogenous Targets by Quantitative RT-PCR (See Note 11)

1. Design PCR primers to specifically amplify mRNA of the gene of interest and of a housekeeping gene. When analyzing gene expression by qRT-PCR using the $\Delta\Delta C_t$ method [31], a stably expressed housekeeping gene must also be tested as a control. GAPDH is a suitable housekeeping gene for many cell types. Primers can be designed using online software such as Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) (*see* **Note 12**).
2. Select cell types for testing TALE-TFs (*see* **Note 13**). To screen TALE-TFs for gene activation of human targets, HEK293T cells are a recommended and efficiently transfected cell line that have been routinely used for this purpose [7, 18–21], although activation will be most readily observable in a cell type that does not highly express the target gene [20]. When testing TALE-TFs for gene inhibition, select a cell line that expresses the target gene of interest.
3. The day before transfection (day 0), seed cells on 24-well plates such that confluency will be ~80 % in 24 h. If using HEK293T cells, seed 400,000 cells/well. If using Lipofectamine reagents, seed and incubate cells in medium that does not contain antibiotic; presence of antibiotic during transfection will decrease transfection efficiency and cell viability.
4. The next day (day 1), transfect cells according to the transfection reagent manufacturer's instructions. For Lipofectamine 2000, transfect each well of HEK293T cells using 2 μL of reagent and 800 ng total plasmid DNA in 100 μL OptiMEM Reduced Serum Medium. If testing combinations of multiple TALE-TFs, the total plasmid transfected should still be 800 ng and typically equal amounts of each plasmid are combined. The following experimental controls are recommended:
   (a) Control 1: Mock-transfected cells (cells that receive no plasmid or an "empty" expression plasmid).
   (b) Control 2: Negative control TALE-TF plasmid (*see* **Note 1**).
   (c) Control 3: GFP expression plasmid (optional to determine transfection efficiency).
5. Exchange the media containing transfection reagent with fresh media containing antibiotic 4–6 h post-transfection.
6. Harvest samples at desired time point. For TALE-based activators, we typically harvest at 72 h post-transfection. Wash cells with DPBS and trypsinize using 0.25 % trypsin-EDTA.
7. After no more than 5 min, neutralize trypsin with equal volume of media and spin samples down in microcentrifuge tube at 1000 × *g* for 5 min.
8. Aspirate and discard supernatant.

9. Isolate total RNA from pellet using Qiagen RNeasy Plus Kit or equivalent product, according to the manufacturer's protocol (*see* **Notes 14** and **15**).
10. Synthesize cDNA by reverse transcription. Load the same mass amount of total RNA into each reaction. If using the VILO cDNA Synthesis Kit, a recipe for each reaction is as follows:

    2 μL 5× VILO Buffer.

    1 μL SuperScript II reverse transcription enzyme.

    500 ng of total RNA.

    H_2O up to 10 μL total reaction volume.
11. Run the reverse transcription reaction according to the following thermocycler protocol:

Temperature (°C)	Duration (min)	Repeat
25	10	1×
42	60	1×
80	5	1×

    Store at 4 °C overnight, –20 °C for long term, or use immediately.
12. Dilute cDNA 1:4 by adding 30 μL of H_2O. This reduces pipetting error for setting up qPCR reactions.
13. Set up qPCR reactions according to the manufacturer's protocol. Make a master reaction mix for each gene assayed. The per-reaction recipe for the Quanta Perfecta SYBR Green reagent is as follows:

    1 μL 10 μM forward primer.

    1 μL 10 μM reverse primer.

    10 μL Perfecta SYBR Green reagent.

    4 μL H_2O.
14. Aliquot 16 μL master mix into qPCR-compatible tubes. Add 4 μL of cDNA from **step 12** to each corresponding reaction.
15. Run qPCR reaction according to the manufacturer's recommended thermocycling conditions. The following two-step thermocycler protocol is used for Quanta Perfecta SYBR green:

Temperature (°C)	Time (s)	Repeat
95	30	1×
95 55–65	3–5 20–30	Steps 2–3, 45×

16. Analyze changes in gene expression by the $\Delta\Delta C_t$ method [31]. For each sample, normalize expression to the housekeeping gene by the following calculation:
$\Delta C_t = 2\text{^}-\left(C_{t,\text{TARGETGENE}} - C_{t,\text{GAPDH}}\right)$, where C_t is the number of cycles to reach a threshold level of fluorescence above background. The SYBR Green reagent contains a fluorescent dye that binds and detects double-stranded DNA during PCR amplification. Samples with a higher amount of cDNA encoding the target gene will therefore demonstrate a lower C_t value compared to samples with lower expression of the target gene.
17. Normalize expression of samples treated with TALE-TF to a mock-transfected control. If experiment was done with replicates, normalize expression of each sample treated to the average of the mock-transfected controls:
$\Delta\Delta C_{t\ \text{TALE-TF1}} = \Delta C_{t\ \text{TALE-TF1}} / \left(\Delta C_{t\,\text{Replicate Average, Mock}}\right)$
A $\Delta\Delta C_t$ significantly greater than 1 by statistical analysis indicates target gene activation compared to a mock-transfected control. Similarly, a $\Delta\Delta C_t$ significantly lower than 1 indicates target gene repression. These experiments can be used to identify TALE-TFs or combinations of TALE-TFs that yield the greatest desired changes in gene expression.

4 Notes

1. This optional, but recommended, control may be a TALE-TF targeted to a sequence not present in the target promoter and that does not indirectly affect target gene expression. This controls for any nonspecific effects of TALE-TF overexpression.
2. TALE-TFs can also be targeted to DNase I hypersensitivity sites located within proximity to the transcription start site using data available on the UCSC Genome Browser, such as that from the ENCODE project [32, 33]. Open chromatin, as measured by DNase I hypersensitivity, has been shown to affect the activity of some DNA-binding proteins, but TALE-TFs can activate gene expression when targeted to chromatin that is inaccessible to DNase digestion [20].
3. TALE target design is an ongoing area of research. The TALE architecture used in the Golden Gate Cloning Kit requires a T nucleotide to start the sequence for efficient targeting, although other TALEs have been developed through directed evolution that can target other sequences [24]. Guidelines b through e are based on the target sites of TALE proteins found in nature [8]. If more target site flexibility is desired, guidelines b through e can be relaxed, as TALE nuclease-based studies have shown that these guidelines are not necessarily required for

TALE functionality [26]. Other work in predicting optimal TALE target sites is ongoing, although much of these efforts to date have focused on TALE nucleases, rather than TALE-TFs [34, 35].

4. These guidelines are based on research done with TALE nucleases. Study of synthetic transcription factor specificity is ongoing. Studies evaluating the specificity of TALE-based activators, however, have shown that TALE-VP64 fusions can be tolerant of up to one mismatch in a reporter-based assay and are more sensitive to mismatches on the 5′ end of the TALE-binding site [18]. Off-target effects of specific TALE-TFs in cells can also be assayed using RNA-seq [36].
5. This protocol can be used to supplement other protocols published by Voytas and Bogdanove, the originators of this kit [8].
6. Use an *E. coli* cell line that supports LacZ-based, blue-white screening, such as DH5α or XL1-Blue. Other commonly used bacterial strains, such as Stbl3, express the LacZ gene and will result in a plate of blue colonies, regardless of whether the reaction was successful.
7. Due to the repetitive nature of RVD arrays, a correct clone will present as a ladder of bands on the gel, starting at approximately 200 bp up to 500 bp. Empty pFUS vectors exhibit a band between 300 and 400 bp.
8. TALE repressors have been developed using KRAB and SID effector domains [9]. KRAB and SID effector domains can be inserted in the empty pCDNA3.1-Golden Gate empty plasmid using BamHI and EcoRI sites. Additionally, TALE fusions to the LSD1 histone demethylase [23] are also available on Addgene (plasmids 49042-4).
9. Because the Step 2 Golden Gate reaction is more efficient than Step 1, it is recommended that the transformation be diluted before plating and only half or less of the transformation be plated in order to reduce the total number of colonies.
10. Standard DNA sequencing reads may not fully cover the length of the TALE region, and the repetitive nature of RVD modules presents a challenge when designing unique primers within the TALE. Sequencing can verify that the TALE proteins inserted correctly at BsaI sites, and if Step 1 plasmids were sequence-confirmed, additional verification is generally not necessary.
11. There are many possible assays for evaluating TALE-TF activity, including the generation of reporter plasmids in which TALE target sites are inserted upstream of GFP or luciferase [7, 18–21]. However one of the most powerful and broadly used applications of TALE-TFs is regulating endogenous genes in chromosomal DNA. Many reports have observed that

TALE-TF activity on reporter plasmids does not perfectly correlate to activity at the chromosomal target locus, likely due to effects of chromatin, local genome structure, and copy number of available targets. Thus confirming regulation of the endogenous gene is critical, and we focus on qRT-PCR here for that purpose.

12. Primers and qPCR conditions should be optimized to amplify target cDNA twofold for each qPCR cycle. Primer efficiency can be tested by creating a standard curve of template target dilutions [31].
13. The optimal TALE-TF delivery method may depend on the cell type selected for screening TALE-TFs. If effective transfection efficiencies are not obtained in the cell line selected with Lipofectamine or other chemical transfection reagents, electroporation may be optimized for delivery of TALE expression plasmids [37]. Adenoviral vectors can also be used for TALE delivery [38].
14. Pellet can also be snap-frozen in liquid N_2 and stored at –80 °C to perform RNA isolation at a later time.
15. Performing the RNA isolation in an RNase/DNase-free environment is critical to protect RNA samples from degradation. For the RNA isolation and setting up the reverse transcription reaction, we recommend treating equipment and bench space with an RNase inhibitor, such as RNase Zap (Life Technologies).

Acknowledgements

This work was supported by a US National Institutes of Health (NIH) Director's New Innovator Award (DP2OD008586), National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award (CBET-1151035), NIH R01DA036865, NIH R21AR065956, NIH UH3TR000505, NIH P30AR066527, the Duke Coulter Translational Partnership, and an American Heart Association Scientist Development Grant (10SDG3060033). P.I.T. was supported by a National Science Foundation Graduate Research Fellowship and an American Heart Association Mid-Atlantic Affiliate Predoctoral Fellowship.

References

1. Gersbach CA, Gaj T, Barbas CF III (2014) Synthetic zinc finger proteins: the advent of targeted gene regulation and genome modification technologies. Acc Chem Res 47(8):2309–2318
2. Kabadi AM, Gersbach CA (2014) Engineering synthetic TALE and CRISPR/Cas9 transcription factors for regulating gene expression. Methods 69(2):188–197
3. Joung JK, Sander JD (2013) TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Biol 14(1): 49–55
4. Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326(5959): 1509–1512

5. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326(5959):1501
6. Miller JC, Tan S, Qiao G et al (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29(2):143–148
7. Zhang F, Cong L, Lodato S et al (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29(2):149–153
8. Cermak T, Doyle EL, Christian M et al (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39(12):e82
9. Cong L, Zhou R, Kuo YC et al (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3:968
10. Morbitzer R, Romer P, Boch J et al (2010) Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci U S A 107(50):21617–21622
11. Streubel J, Blucher C, Landgraf A et al (2012) TAL effector RVD specificities and efficiencies. Nat Biotechnol 30(7):593–595
12. Mahfouz MM, Li L, Piatek M et al (2012) Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. Plant Mol Biol 78(3):311–321
13. Crocker J, Stern DL (2013) TALE-mediated modulation of transcriptional enhancers in vivo. Nat Methods 10(8):762–767
14. Konermann S, Brigham MD, Trevino AE et al (2013) Optical control of mammalian endogenous transcription and epigenetic states. Nature 500(7463):472–476
15. Bultmann S, Morbitzer R, Schmidt CS et al (2012) Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res 40(12): 5368–5377
16. Gao X, Yang J, Tsang JC et al (2013) Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. Stem Cell Reports 1(2):183–197
17. Hu J, Lei Y, Wong WK et al (2014) Direct activation of human and mouse Oct4 genes using engineered TALE and Cas9 transcription factors. Nucleic Acids Res 42(7):4375–4390
18. Garg A, Lohmueller JJ, Silver PA et al (2012) Engineering synthetic TAL effectors with orthogonal target sites. Nucleic Acids Res 40(15):7584–7595
19. Li Y, Moore R, Guinn M et al (2012) Transcription activator-like effector hybrids for conditional control and rewiring of chromosomal transgene expression. Sci Rep 2:897
20. Perez-Pinera P, Ousterout DG, Brunger JM et al (2013) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. Nat Methods 10(3): 239–242
21. Maeder ML, Linder SJ, Reyon D et al (2013) Robust, synergistic regulation of human gene expression using TALE activators. Nat Methods 10(3):243–245
22. Maeder ML, Angstman JF, Richardson ME et al (2013) Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat Biotechnol 31(12):1137–1142
23. Mendenhall EM, Williamson KE, Reyon D et al (2013) Locus-specific editing of histone modifications at endogenous enhancers. Nat Biotechnol 31(12):1133–1136
24. Lamb BM, Mercer AC, Barbas CF III (2013) Directed evolution of the TALE N-terminal domain for recognition of all 5′ bases. Nucleic Acids Res 41(21):9779–9785
25. de Lange O, Schreiber T, Schandry N et al (2013) Breaking the DNA-binding code of Ralstonia solanacearum TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. New Phytol 199(3):773–786
26. Reyon D, Tsai SQ, Khayter C et al (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30(5):460–465
27. Reyon D, Maeder ML, Khayter C et al (2013) Engineering customized TALE nucleases (TALENs) and TALE transcription factors by fast ligation-based automatable solid-phase high-throughput (FLASH) assembly. In: Ausubel FM et al (ed) Current protocols in molecular biology, Chapter 12:Unit 12 16
28. Schmid-Burgk JL, Schmidt T, Kaiser V et al (2013) A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol 31(1):76–81
29. Doyle EL, Booher NJ, Standage DS et al (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40(Web Server issue):W117–W122
30. Meckler JF, Bhakta MS, Kim MS et al (2013) Quantitative analysis of TALE-DNA interactions suggests polarity effects. Nucleic Acids Res 41(7):4118–4128
31. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4):402–408
32. Thurman RE, Rynes E, Humbert R et al (2012) The accessible chromatin landscape of the human genome. Nature 489(7414):75–82

33. Boyle AP, Davis S, Shulha HP et al (2008) High-resolution mapping and characterization of open chromatin across the genome. Cell 132(2):311–322
34. Lin Y, Fine EJ, Zheng Z et al (2014) SAPTA: a new design tool for improving TALE nuclease activity. Nucleic Acids Res 42(6):e47
35. Guilinger JP, Pattanayak V, Reyon D et al (2014) Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. Nat Methods 11(4):429–435
36. Perez-Pinera P, Kocak DD, Vockley CM et al (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat Methods 10(10):973–976
37. Ousterout DG, Perez-Pinera P, Thakore PI et al (2013) Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. Mol Ther 21(9): 1718–1726
38. Holkers M, Maggio I, Liu J et al (2013) Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. Nucleic Acids Res 41(5):e63

Chapter 8

A New Approach to Dissect Nuclear Organization: TALE-Mediated Genome Visualization (TGV)

Yusuke Miyanari

Abstract

Spatiotemporal organization of chromatin within the nucleus has so far remained elusive. Live visualization of nuclear remodeling could be a promising approach to understand its functional relevance in genome functions and mechanisms regulating genome architecture. Recent technological advances in live imaging of chromosomes begun to explore the biological roles of the movement of the chromatin within the nucleus. Here I describe a new technique, called TALE-mediated genome visualization (TGV), which allows us to visualize endogenous repetitive sequence including centromeric, pericentromeric, and telomeric repeats in living cells.

Key words Nuclear architecture, Live imaging, TALE, TGV

1 Introduction

Chromatin is organized in a nonrandom fashion within three-dimensional nuclear space. During developmental process, nuclear architecture is dramatically reconstructed, resulting in establishment of cell-type specific nuclear organization. Defects in structural components of the nucleus are responsible for developmental aberrations and several human diseases [1]. Remodeling of nuclear architecture leads to spatial arrangement of genes, which could affect genome functions including gene expression. However regulatory mechanism underlying nuclear reorganization during cell-fate decision remains largely unknown, possible due to technical limitations to study nuclear dynamics in living cells. DNA fluorescent in situ hybridization (DNA-FISH) is the method of choice to study subnuclear positioning of endogenous target sequence, but it is performed on fixed samples. Live imaging of nuclear remodeling would be a breakthrough toward uncovering the functional relevance and mechanisms regulating

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_8,

genome architecture. I describe here a new approach to visualize genomic sequences in living cells using transcription activator-like effector (TALE) technology [2].

TALE, a protein discovered in the plant pathogenic bacteria Xanthomonas, has emerged as a powerful platform for engineering DNA-binding proteins [3, 4]. Central region of TALE proteins consists of tandem repeat of ~34 amino (TALE repeats), which contains two variable amino acids referred as repeat-variable diresidue (RVDs) (Fig. 1a, b). Each repeat displays a single nucleotide binding specificity determined by the RVDs (NI to A, NG to T, HD to C, and NH to G). The simple and flexible feature of designing TALEs allows us to engineer artificial DNA binding proteins with specific functionalities upon fusion of TALEs with operative proteins such as nucleases [5], transcriptional modulators [6–8], and recombinases [9]. We integrated the TALE technology to visualize nuclear dynamics of endogenous repetitive sequences by targeting TALEs fused with fluorescent proteins, and developed TALE-mediated genome visualization (TGV) [2].

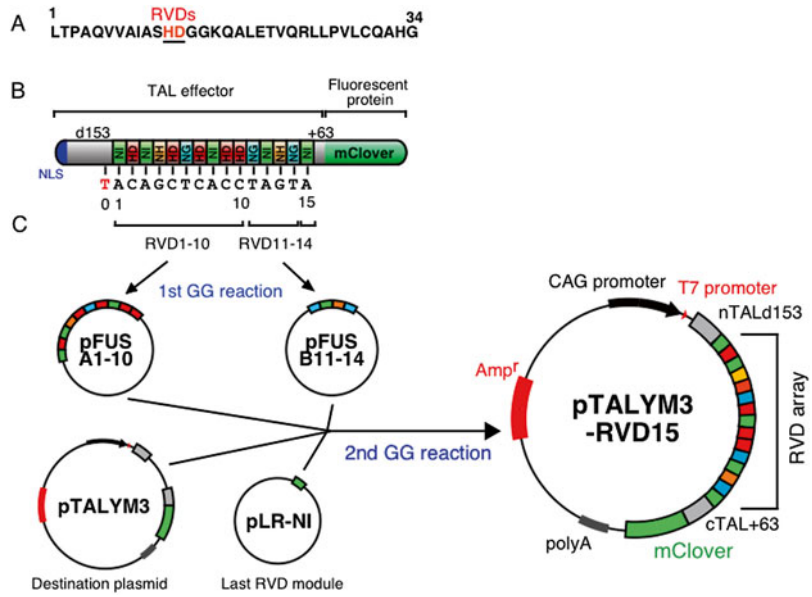


Fig. 1 (**a**) Amino acid sequence of TALE-repeat harboring repeat-variable diresidue (RVDs, *underlined*). (**b**) Schematic diagram of TALE-mClover and target sequence. Truncated TALE (N153AA and C63AA) is fused with nuclear localization signal (NLS) at N-terminus and monomeric green fluorescent protein, mClover, at C-terminus. N-terminus domain of TALE binds to 5′ T. Following nucleotides are recognized by corresponding TALE repeats. (**c**) Golden gate assembly of custom fluorescent TALE constructs. RVD arrays targeting initial ten nucleotides and following four nucleotides are assembled by first golden gate reaction, resulting in generation of pFUSA1-10 and pFUSB11-14, respectively. These plasmids are used for second GG reaction with pTALYM3 and pLR-NI to generate final construct, pTALYM3-RVD15

2 Materials

2.1 Plasmids for Assembly of Fluorescent TALEs

1. TAL Effector Kit 2.0 (Addgene, cat#1000000024) [10].
2. Destination plasmids: pTALYM3 and pTALYM4 (Addgene, Plasmid# 47874 and 47875) encoding TALE-mClover and TALE-mRuby2, respectively [2]. Final construct is designed to contain chicken beta-actin promoter, T7 promoter, truncated TALE (N153AA and C63AA) fused with nuclear localization signal (NLS) at N-terminus, and a monomeric fluorescent protein such as mClover and mRuyb2 at C-terminus (Fig. 1b, c) (*see* **Note 1**).

2.2 TALE Assembly

1. T4 DNA ligase (New England Biolabs, M0202M).
2. 10× T4 DNA ligase buffer: 50 mM Tris–HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM DTT, and 1 mM ATP.
3. Restriction enzymes: BsaI, BsmBI, XbaI, AflII, BspEI, and StuI.
4. 10× BSA, 10 mg/ml of bovine serum albumin.
5. Plasmid-Safe DNase (Epicenter, E3105K).
6. Luria Broth (LB) agar plate. To prepare 500 ml of LB agar, dissolve 5 g of tryptone, 2.5 g yeast extract, 5 g of NaCl, and 7.5 g of agar into 500 ml of water. After autoclaving the mixture, pour a thin layer (5 mm) of LB Agar (~20 ml) into a 10 cm plastic dish.
7. Antibiotics: Tetracycline, spectinomycin, or ampicillin.
8. X-gal: 5-Bromo-4-chloro-3-indolyl beta-d-galactopyranoside.
9. IPTG: Isopropyl β-d-1-thiogalactopyranoside.
10. LB medium: To prepare 500 ml of LB medium, dissolve 5 g of tryptone, 2.5 g yeast extract, 5 g of NaCl, and 7.5 g of agar into 500 ml of water and autoclave it for 20 min at 120 °C.
11. *Escherichia coli* SURE2 (200152, Agilent).

2.3 ES Cell Culture

1. Culture medium: DMEM with GlutaMAX (Life Technologies, 10566-024), 15 % fetal calf serum, 1000 U/ml of leukemia inhibitory factor, 1 mM sodium pyruvate, penicillin/streptomycin, 0.1 mM 2-mercaptoethanol, 3 mM CHIR99021 (a GSK3b inhibitor), and 1 mM PD0325901 (an MEK inhibitor).
2. PBS: Phosphate-buffered saline. To prepare 1 l PBS, dissolve 8 g NaCl, 0.2 g KCl, 1.15 g Na_2HPO_4 (anhydrous), and 0.2 g KH_2PO_4 (anhydrous) into 1 l water.
3. Trypsin solution: 0.25 % trypsin dissolved in PBS.
4. Lipofectamine 2000 (Life Technologies, 11668027).
5. Laminin-511 (BioLamina, LN-511).
6. Glass-bottom dishes.

3 Methods

3.1 Design of TALEs

1. Identify candidates of TALE target site from the sequence of interest using the TAL effector Nucleotide Targeter (TALE-NT) 2.0 website (https://tale-nt.cac.cornell.edu/about) [10, 11]. Sequences of 15 nt or longer without 5′ T are appropriate for specific labeling of target sequences, since TALE against shorter sequence displayed higher background signal in the nucleoplasm possibly due to lower binding affinity or lower specificity to the target sequences [2]. Use NH RVD module for recognizing guanine, because it displays higher specificity than NN RVD module, which has affinity for both guanine and adenine [12, 13].
2. Select target sequences without CpG sites, since DNA methylation on the site potentially reduce binding affinity of TALEs [14] (*see* **Notes 2** and **3**).

3.2 Assembly of Fluorescent TALEs

RVD modules of fluorescent TALEs are assembled by golden gate cloning approach as described before [10], with slight modifications. Here I depict a protocol for the assembly of TALE-mClover by taking a TALE targeting the sequence 5′-TACAGCTCACCTAGTA for example (Fig. 1c). Note that 5′ T (here designated as T0) should not be considered for the assembly since N-terminal region of TALEs (d153 in Fig. 1c) but not RVD array recognize it.

1. Split the target sequence into initial 10 nucleotides (ACAGCTCACC) and following 4 nucleotide (TAGT), and consider corresponding RVDs for each sequence: RVD1-10 (NI HD NI NH HD NG HD NI HD HD), RVD11-14 (NG NI NH NG). These two RVD modules are assembled separately into pFUSA and pFUSB plasmid, respectively, in the following steps (*see* **Note 4**).
2. Mix all plasmids required for assembly of RVD1-10 array containing 25 μg of each RVD module (pNI1, pHD2, pNI3, pNH4, pHD5, pNG6, pHD7, pNI8, pHD9, and pHD10) and 25 μg of pFUSA in a 200 μl PCR tube, and then concentrate it by ethanol precipitation to a final volume of 3 μl of H_2O. For assembling RVD11-14, mix RVD modules (pNG1, pNI2, pNH3, and pNG4) and pFUSB4 in a tube and concentrate it to 3 μl of H_2O as well (*see* **Note 5**).
3. First golden gate reaction: Add the following reagents to each tube: 0.6 μl of 10× T4 DNA ligase buffer, 0.6 μl of 10× BSA, 0.3 μl of 10 U/μl BsaI, 0.3 μl of T4 DNA ligase, and 1.2 μl of H_2O. Incubate the reactions in a thermocycler for 40 cycles of (5 min at 37 °C and 10 min at 16 °C), at 50 °C for 5 min, and 80 °C for 5 min (*see* **Notes 6** and **7**).

4. Add 0.3 μl of 25 mM ATP and 0.3 μl of Plasmid Safe DNase to the reactions and incubate them at 37 °C for 30 min (*see* **Note 8**).
5. Transform each of the reaction into *Escherichia coli* SURE2 cells as per the manufacturer's instruction. Plate each transformed cells onto LB agar plates containing 50 mg/ml of spectinomycin and Xgal/IPTG, and culture them at 37 °C overnight.
6. Pick up a few white colonies from each transformation and start overnight culture in 2 ml of LB medium containing 50 mg/ml of spectinomycin (*see* **Note 9**).
7. Purify plasmid DNA and identify resulting clones pFUSA1-10 and pFUSB11-14 by double digestion with AflII and XbaI. 1048 bp and 430 bp are released from pFUSA1-10 and pFUSB11-14 by the digestion, respectively.
8. Second golden gate reaction: Mix the following reagents in a total volume of 5 μl: 30 μg of pFUSA1-10, 30 μg of pFUSB11-14, 30 μg of pLR-NI, 30 μg of pTALYM3, 0.25 μl of BsmBI, 0.25 μl of T4 DNA ligase, 0.5 μl of 10× T4 DNA ligase buffer, 0.5 μl of 10× BSA, and H_2O. Incubate the reactions in a thermocycler for 4 cycles of (5 min at 37 °C and 10 min at 16 °C), at 50 °C for 5 min, and 80 °C for 5 min (*see* **Note 10**).
9. Transform the reaction into SURE2 cells, and plate the cells onto a LB-agar plate containing 50 mg/ml of ampicillin and Xgal/IPTG. Incubate the plate at 37 °C overnight (*see* **Note 11**).
10. Pick up a few white colonies and start overnight culture in 2 ml of LB medium containing 50 mg/ml of ampicillin.
11. Isolate plasmid DNA and identify clones containing the final, full-length repeat array by double digestion with StuI and AatII, which cut just outside of the repeat resulting in release of 1724 bp fragment in this case. To check alignment of RVD array, digest clones with BspEI, which cuts only in HD modules 2–10. The array can also be characterized by DNA sequencing.

3.3 Imaging of Repetitive Sequence by TGV

While flexibility of TALE design can potentially allow detection of any sequences in the genome, this application is limited to repetitive sequences so far (*see* **Note 12**) [2]. Here I describe a method to visualize mouse repetitive sequence including pericentromeric major satellite, centromeric minor satellite, and telomeric repeats (Fig. 2).

1. TALEs targeting following repetitive sequences were assembled as described above [2]: major satellite sequence (TTGCCATATTCCACGT), minor satellite sequence (TCAT TGATATACACTG), and telomere sequence (TAGGGTTAG GGTTAGG).

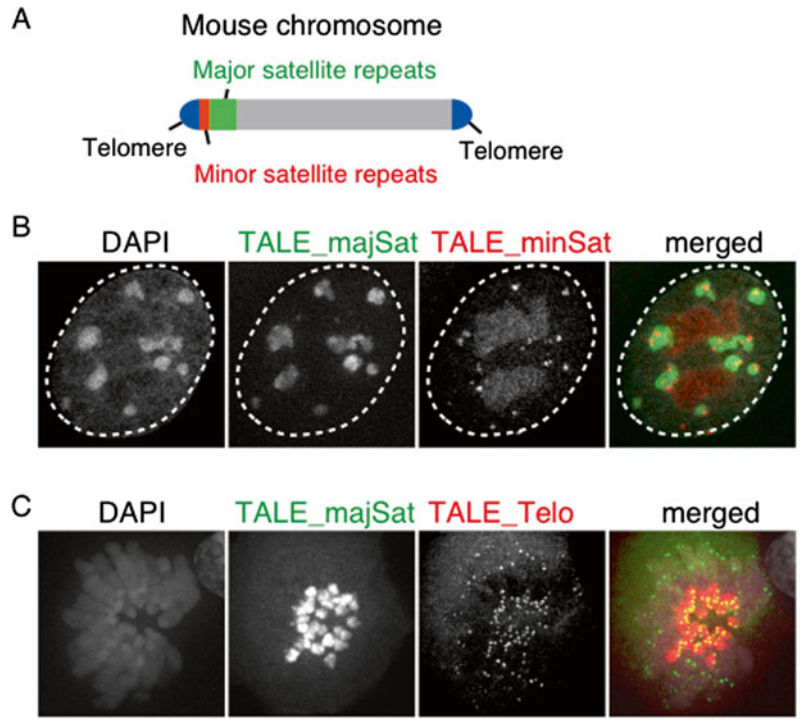


Fig. 2 (**a**) Schematic diagram of mouse chromosome. The location of telomeres (*blue*), pericentromeric major satellites (*green*), centromeric minor satellites (*red*), and the long arm of the chromosome (*gray*) are indicated. (**b**) Fluorescent images of ES cells expressing TALE-mClover targeting major satellite repeat (TALE_majSat, *green*), TALE-mRuby2 against minor satellite repeat (TALE_minSat, *red*). DNA was stained with DAPI (*gray*). Major satellite regions, which are highly colocalized with DAPI dense heterochromatic domain, were specifically labeled by TALE_majSat. Small fluorescent spots for TALE_minSat were detected at centromeric regions adjacent with major satellite regions. (**c**) Mitotic chromosomes of ES cells expressing TALE-mClover_majSat (TALE_majSat, *green*) and TALE-mRuby2 targeting telomere sequence (TALE_Telo, *red*). Mitotic chromosomes were stained with DAPI (*gray*). Both chromosome ends were efficiently labeled with TALE_Telo

2. Transfect plasmid DNA encoding fluorescent TALE against the repetitive sequences to mouse ES cells using Lipofectamine 2000 as per the manufacturer's instruction, and culture the cells at 37 °C under a 5 % CO_2 atm overnight (*see* **Notes 13** and **14**).
3. Next day, seed the cells onto 3 cm glass-bottomed dishes coated with Laminin-511. 4 h after the seeding, place the dish in an incubation chamber on the fluorescent microscope stage at 37 °C under a 5 % CO_2 atm (*see* **Note 15**).
4. Fluorescent signals can be acquired with a conventional epi-fluorescent microscopy or confocal microscopy (*see* **Note 16**). mClover, mRuby2, and 4,6-diamidino-2-phenylindole (DAPI)

are excited with 488 nm, 560 nm, and 405 nm lasers, respectively (*see* **Notes 17** and **18**).

4 Notes

1. Some fluorescent proteins such as EGFP indeed display weak dimerization activity [15], which could potentially have abnormal effects to its fusion proteins. To avoid any effect from fluorescent proteins, monomeric fluorescent proteins such as mClover and mRuby2 are applicable for TGV.
2. Alternatively, N* RVD module could be used for recognizing methylated cytosine [14].
3. It is recommended to pick up a few sequences for single-target locus because not every TALE works as expected.
4. The last RVD (NI for A in this case) is added later in second golden gate reaction (Subheading 3.2, **step 8**).
5. If you use less purified plasmids for RVD modules which were prepared by miniprep kit, ethanol precipitation of these plasmids will enhance efficiency of subsequent golden gate reaction.
6. During the reaction, RVD region on each plasmid flanked by BsaI site is cleaved out at 37 °C and then ligated with neighboring RVD modules in a proper alignment at 16 °C.
7. Increasing the number of cycles in the golden gate reaction will facilitate the successful cloning compared to the original protocol which mentions only 10 cycles [10].
8. Linear DNA containing partial RVD arrays is specifically digested in this step, resulting in purification of desired circular plasmids.
9. Two colonies are practically enough to identify positive clones.
10. During this step, RVD1-10, RVD11-14, and the last RVD NI are concatenated together and cloned into a destination plasmid, pTALYM3.
11. It is recommended to use a strain such as SURE2 which is compatible for expansion of repetitive sequence since I have experienced intra-plasmid recombination in RVD tandem array using DH5alpha which is used in the original paper [10].
12. Chen et al. applied catalytic inactive Cas9, an RNA-guided DNA-binding protein, to label specific DNA sequence in vivo as we did with fluorescent TALEs [16]. They succeeded to label non-repetitive genomic locus by targeting around 30 molecules of fluorescent Cas9, indicating that at least 30 copies of repetitive binding sites are required to detect fluorescent signals on target genomic loci by TGV.

13. We have succeeded to detect TGV signals with several different cell lines including mouse ES cells, NIH3T3 cells, and Hela cells, suggesting that any types of cells are potentially labeled by this technique.
14. An another fluorescent marker such as Histone H2B-mRFP is useful to label nuclei.
15. ES cells can grow on Laminin511-coated glass-bottom dish in a monolayer, which is advantageous for live imaging of nuclear architecture because morphologies of nuclei are more uniform between cells than that of cells in colonies grown onto gelatin-coated dish.
16. We practically use an inverted confocal microscope attached to a Nipkow disk with an EMCCD camera, which is more sensitive than conventional epifluorescent microscopies.
17. Over-expression of fluorescent TALE reads to high background signals in nucleoli (*see* Fig. 2a, TALE_minSat). It is recommended to keep the expression level of fluorescent TALE at minimum enough for imaging.
18. It is important to check the specificity of fluorescent TALEs. Validation of binding specificity to endogenous target sequences is performed by comparison between fluorescent signals from TGV and DNA-FISH [2]. ChIp-seq against TALE is an alternative approach to confirm the specificity.

References

1. Misteli T (2010) Higher-order genome organization in human disease. Cold Spring Harb Perspect Biol 2:a000794
2. Miyanari Y, Ziegler-Birling C, Torres-Padilla ME (2013) Live visualization of chromatin dynamics with fluorescent TALEs. Nat Struct Mol Biol 20:1321–1324
3. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512
4. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501
5. Sun N, Zhao H (2013) Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. Biotechnol Bioeng 110:1811–1821
6. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29:149–153
7. Perez-Pinera P, Ousterout DG, Brunger JM, Farin AM, Glass KA, Guilak F, Crawford GE, Hartemink AJ, Gersbach CA (2013) Synergistic and tunable human gene activation by combinations of synthetic transcription factors. Nat Methods 10:239–242
8. Maeder ML, Linder SJ, Reyon D, Angstman JF, Fu Y, Sander JD, Joung JK (2013) Robust, synergistic regulation of human gene expression using TALE activators. Nat Methods 10:243–245
9. Mercer AC, Gaj T, Fuller RP, Barbas CF 3rd (2012) Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic Acids Res 40:11163–11172
10. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82
11. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ (2012) TAL Effector-Nucleotide Targeter

(TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40(Web Server issue):W117–W122

12. Streubel J, Blucher C, Landgraf A, Boch J (2012) TAL effector RVD specificities and efficiencies. Nat Biotechnol 30:593–595
13. Cong L, Zhou R, Kuo YC, Cunniff M, Zhang F (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3:968
14. Valton J, Dupuy A, Daboussi F, Thomas S, Marechal A, Macmaster R, Melliand K, Juillerat A, Duchateau P (2012) Overcoming transcription activator-like effector (TALE) DNA binding domain sensitivity to cytosine methylation. J Biol Chem 287:38427–38432
15. Zacharias DA, Violin JD, Newton AC, Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science 296:913–916
16. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell 155:1479–1491

Chapter 9

TALEN-Induced Translocations in Human Cells

Marion Piganeau, Benjamin Renouf, Hind Ghezraoui, and Erika Brunet

Abstract

Induction of chromosomal translocations in human cells is of a great interest to study tumorigenesis and genome instability. Here, we explain in detail a method to induce translocations using the transcription activator-like effector nucleases (TALENs). We describe how to detect translocation formation by PCR, calculate translocation frequency by 96-well PCR screen, and analyze breakpoint junctions. When inducing cancer translocations, it is also possible to detect the fusion gene by FISH analysis or western blot.

Key words TALEN induced translocation, 96-Well plate PCR screen, T7-endonuclease assay, Breakpoint junction analysis, Fusion gene

1 Introduction

Chromosomal translocations are genome modifications found in various cancers and diseases. To date molecular mechanisms leading to translocation formation in human cells are poorly understood. Likewise cellular consequences and direct tumorigenesis induced by the presence of translocated chromosomes are not always described. One explanation is the lack of study tools. In human cells, studies are mostly based on ectopically expressing or silencing the fusion gene induced by the translocation. Overexpression requires expressing the fusion protein from a chosen promoter. However this promoter is not always suitable when the expression level of potential oncogene is crucial [1, 2]. As well, residual expression of the fusion gene in silencing strategies can have residual effects.

DNA double-strand breaks (DSBs) are considered to be causative lesions for many genomic rearrangements, including chromosomal translocations [3]. We used engineered nucleases to generate translocations in human cells at specific loci. In former studies, we used I-SceI endonuclease and zinc finger nuclease (ZFN) technology to induce two DSBs at chosen chromosomes [4]. However the design of these nucleases remains complex. The more recent

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_9,

generation of transcription activator-like effector nucleases (TALENs) is easier to design and allows higher sequence flexibility, making them a better tool for translocation induction. With two pairs of TALENs, we can induce cancer-related translocations in human cells with good efficiency and succeed in expressing the fusion gene in situ, in its genomic context [5].

In this protocol chapter, we describe in detail the method to induce chromosomal translocations in human cells using TALENs. Briefly two TALENs are designed to target sequences at the chosen translocation breakpoints. The TALEN cleavage efficiency is estimated by a T7 Endonuclease I assay, and the breakpoint junctions of the two derivative chromosomes are PCR amplified and sequenced. We also developed a 96-well screen assay to estimate translocation frequency. Furthermore, TALEN-induced cancer translocations can be detected by FISH experiments and fusion protein expression detected by simple western blot analysis.

2 Materials

2.1 Expression Plasmid Preparation

1. TALEN expression plasmids (homemade from Addgene plasmids for example (https://www.addgene.org/TALEN/), or ordered from companies).
2. Competent bacteria.
3. DNA Maxiprep kit.
4. UV spectrophotometer.

2.2 Cell Culture and Transfection

1. RPE1 (hTert-RPE1) cells or any cells of interest.
2. Dulbecco's modified Eagle medium: Nutrient Mixture F-12 (DMEM/F-12), supplemented with 10 % fetal bovine serum (FBS).
3. TALEN expression plasmids, at concentration equal to or greater than 2 μg/μL.
4. 0.05 % Trypsin-EDTA solution.
5. Phosphate-buffered saline (PBS).
6. Nucleofector II Device (Lonza).
7. Cell Line Nucleofector Kit V (Lonza).

2.3 PCR Analysis

1. Genomic isolation kit.
2. Primers for nested PCR (two sets of primers, *see* **Note 1**).
3. Taq polymerase.
4. Agarose gel with ethidium bromide (EtBr).
5. 0.5× TBE running buffer (Tris base, borate, EDTA).
6. UV station.

2.4 T7 Endonuclease I Assay

1. Taq polymerase.
2. Primers (*see* **Note 2**).
3. T7 Endonuclease I (New England Biolabs).
4. NEBuffer 2.1 (New England Biolabs).
5. 2× T7 loading buffer: 50 % sucrose, bromophenol blue, 250 μg/mL proteinase K.
6. Agarose gel with EtBr.
7. 0.5× TBE running buffer (Tris, borate, EDTA).
8. UV station.

2.5 Translocation Quantification

1. 1× Lysis buffer: 10 mM Tris–HCl pH 8, 0.45 % NP40, 0.45 % Tween20.
2. Proteinase K solution.
3. 2× Master Mix 1: 1× GC-RICH solution (Roche FastStart Taq), 2× PCR Buffer with 20 mM $MgCl_2$ (Roche FastStart Taq), 400 μM dNTP mix (Roche FastStart Taq), 4 % DMSO, 0.01 % Tween20, 0.01 % NP40. Store at 4 °C, up to 2 weeks.
4. 2× Master Mix 2: 1× GC-RICH solution (Roche FastStart Taq), 2× PCR Buffer with 20 mM $MgCl_2$ (Roche FastStart Taq), 400 μM dNTP mix (Roche FastStart Taq), 4 % DMSO, 0.01 % Tween20, 0.01 % NP40, SYBR Green and Rox (concentrations depend on the real time PCR machine, e.g., for MX005P (Agilent) 30 nM is used). Store at 4 °C, up to 2 weeks.
5. Primers for nested PCR (two sets of primers, *see* **Note 1**).
6. FastStart Taq DNA polymerase (Roche).
7. Real-time quantitative PCR machine.

2.6 FISH Analysis

1. Hypotonic solution: 75 mM KCl.
2. Fixative solution: Ethanol:acetic acid (3:1) (has to be freshly prepared).
3. Colcemid.
4. Microscope slides.
5. Cover slips.
6. Moist towels.
7. Optical microscope.
8. Fluorescent probe.
9. 2× SSC (pH 6.3): 300 mM NaCl, 30 mM citric acid trisodium salt dihydrate.
10. DAPI solution.
11. PBS.
12. Mounting medium.
13. Fluorescence microscope.

2.7 Western Blot

1. Lysis buffer: 50 mM Tris–HCl pH 7, 1 % Triton X-100, 0.1 % SDS, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, Protease inhibitor cocktail.
2. Protein assay kit.
3. 4× Laemmli sample buffer: 8 % SDS, 240 mM Tris–HCl pH 6.8, 40 % glycerol.
4. DTT.
5. SDS-PAGE gel.
6. Running buffer: 25 mM Tris base, 192 mM glycine, 0.1 % SDS, pH 8.5.
7. Nitrocellulose membrane.
8. Blocking solution: TBS-T 5 % bovine serum albumin (BSA).
9. Tris-buffered saline, Tween 20 (TBS-T): 50 mM Tris base, 150 mM Nacl, 0.05 % Tween 20, pH 7.4.
10. Primary antibody directed against the protein of interest.
11. HRP-conjugated secondary antibody.
12. Chemiluminescent substrate solution.
13. X-ray films or chemiluminescence detection systems.

3 Methods

3.1 TALEN Design and Preparation of the Expression Plasmids

1. After choosing one translocation of interest, choose the two genomic regions to find a target site. When the translocation is cancer related, specific location of reported breakpoint junctions found in patient cells can be used as target sites.
2. Design pairs of TALENs (TAL^A and TAL^B), targeting sequences A and B respectively at the two chosen breakpoints (Fig. 1) (*see* **Note 3**) by using available TALEN design tools (for example https://tale-nt.cac.cornell.edu/node/add/talen or http://www.e-talen.org/E-TALEN/designtalens.html), or by finding target sequences following the simple rules described in [6–9]: we used T-N_{17}-$Linker_{15\text{-}20}$-N_{17}-A. You can check for off-target sites with specific designed tools such as "TALENoffer" http://galaxy2.informatik.uni-halle.de:8976/ [10] (*see* **Note 4**). TALEN expression plasmids can be constructed following the method described in [11] (*see* **Note 5**).
3. Transform competent bacteria with each TALEN expression plasmid. Spread on a LB agar plate and incubate overnight at 37 °C.
4. Start a 200 mL LB culture with one clone of transformed bacteria. Incubate overnight at 37 °C.

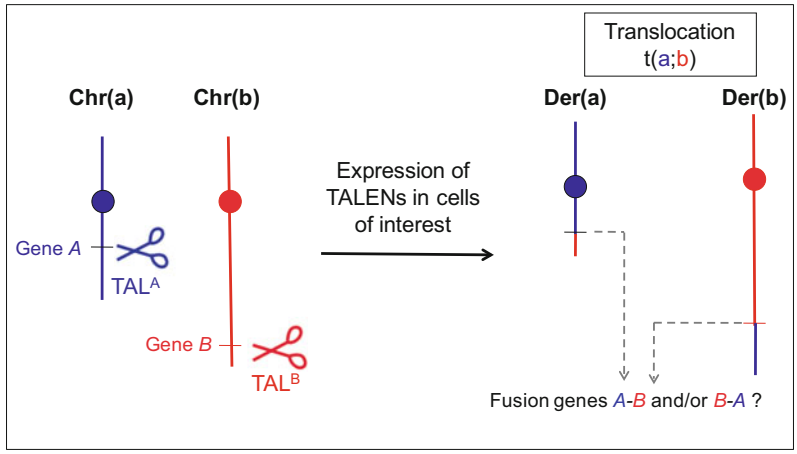


Fig. 1 Induction of translocation t(a;b) in human cells using the TALEN strategy. TAL^A and TAL^B (*scissors*) targeting genes *A* and *B* on chromosomes Chr(a) and Chr(b), respectively, are expressed in cells of interest. DSBs induced by TALENs can lead to the formation of two derivative chromosomes Der(a) and Der(b). Translocation formation can induce expression of a fusion gene (*A-B* and/or *B-A*)

5. Proceed to plasmid isolation with DNA Maxiprep kit following the manufacturer's instructions. Resuspend DNA in a small final volume (typically 150 µL).
6. Measure DNA concentration using a UV spectrophotometer. The final DNA plasmid concentration should be greater than 2 µg/µL to be optimal for transfection. The total volume of transfected DNA should not exceed 10 µL.
7. Make aliquots of the TALEN expression plasmids (typically 10 µL per tube) to avoid degradation by repeat freezing and thawing.

3.2 TALEN Transfection

The protocol described here has been optimized for RPE1 cells; however this translocation induction system can be used for a number of cell lines (*see* **Note 6**).

1. RPE1 cells are cultured in 20 mL of DMEM/F-12 medium supplemented with 10 % FBS at 37 °C and 5 % CO_2. Passage them every 2–3 days at a split ratio of 1:10, never allowing cells to reach more than 80 % confluency.
2. The day before transfection split the cells to obtain about 70 % confluency the next day (*see* **Note 7**).
3. The next day, prepare a tube containing 5 µg of each of the four TALEN expression plasmids. If one TALEN induces high toxicity, you can decrease the amount to 2 µg. We advise you to transfect each TALEN pair separately as control, making three transfection samples: TAL^A, TAL^B, and $TAL^A + TAL^B$. The total volume of transfected DNA should not exceed 10 µL.

4. Pre-fill 3 wells of a 6-well plate with 2 mL of supplemented DMEM/F-12 medium for each transfection (TAL^A, TAL^B, and $TAL^A + TAL^B$). Pre-warm at 37 °C (*see* **Note 8**).
5. To determine translocation frequency of the $TAL^A + TAL^B$ transfection, pre-fill additionally 3 × 96-well plates with 50 µL of medium per well. Pre-warm at 37 °C.
6. Trypsinize and resuspend the RPE1 cells in supplemented DMEM/F-12 medium (typically 10 mL).
7. Count and aliquot 7.5×10^5 cells for each transfection.
8. Centrifuge for 10 min at 90 × *g* at room temperature (*see* **Note 8**).
9. Carefully remove medium without aspirating the cells (*see* **Note 9**).
10. Resuspend cells in 100 µL of Cell Line Nucleofector Kit V solution (*see* **Note 10**).
11. Transfer the cell solution into the tube containing DNA, then to the provided Amaxa DNA cuvettes.
12. Proceed to electroporation using Nucleofector II system with program X-001 (*see* **Note 11**).
13. After TAL^A (or TAL^B) transfection, transfer directly cells in the 3 wells of a 6-well plate prepared in **step 4**, for DNA, protein, and FISH analysis (Fig. 2) (*see* **Note 12**).
14. After $TAL^A + TAL^B$ transfection, transfer cells in 5 mL of pre-warmed medium. To calculate translocation frequency, dilute further 1 mL, 500 µL and 250 µL of cell suspension in a new tube with 5 mL of medium (making 1/5, 1/10 and 1/20 dilutions). Plate 50 µL per well of each dilution in a 96-well plate from **step 5**. Split the rest of the cells (i.e., 3.750 mL) in three separate wells of the 6-well plate prepared in **step 4** for DNA, protein, and FISH analysis (Fig. 2) (*see* **Note 12**).
15. Incubate at 37 °C and 5 % CO_2.

3.3 PCR-Based Translocation Detection

1. Any time after 48 h after the transfection, trypsinize the cells from a well of the 6-well plates prepared in **steps 13** and **14** of Subheading 3.2.
2. Centrifuge at 200 × *g*, at 4 °C for 5 min.
3. Remove supernatant and wash the pellet with 1 mL of PBS.
4. Centrifuge the same way.
5. Remove supernatant. At this step, cell pellets can be stored at −80 °C.
6. Proceed to DNA extraction with any genomic DNA extraction kits following the manufacturer's instructions. Genomic DNA can be stored at −20 °C.

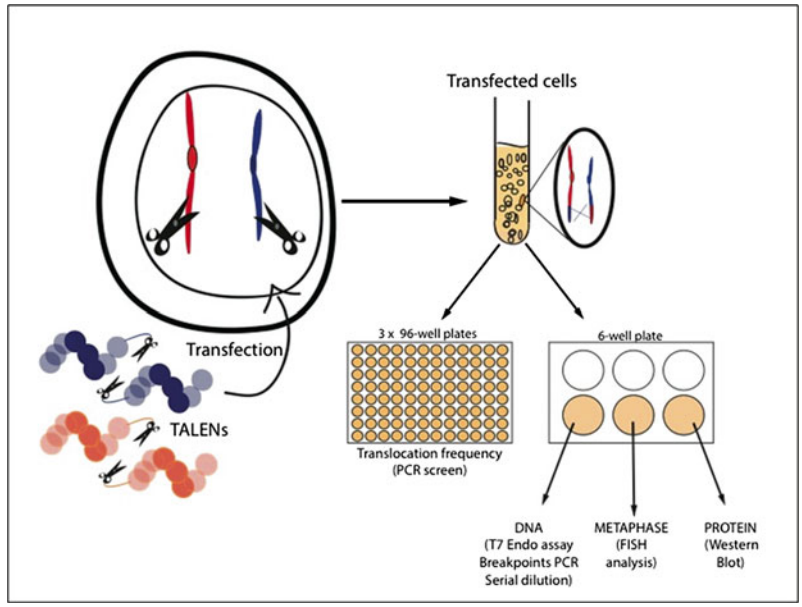


Fig. 2 Experimental setting of transfection. TALEN expression plasmids are transfected into chosen cells to induce translocation formation. After transfection, cells are resuspended in 5 mL medium and split to (1) 3 × 96-well plates for PCR screen (translocation frequency quantification (Subheading 3.5)); (2) 1 well of 6-well plate for DNA analysis (PCR translocation amplification (Subheading 3.3), T7 Endonuclease I assay (Subheading 3.4), serial dilution PCR to estimate translocation frequency (Subheading 3.6)); (3) 1 well of 6-well plate for metaphase spreading (FISH translocation detection (Subheading 3.7)); (4) 1 well of 6-well plate for protein extraction (western blotting (Subheading 3.8))

7. Because of low translocation frequency, you may need to proceed to a nested PCR (*see* **Note 13**). Design two sets of primers on each side of the translocation junction, the second set is located within the first product (Fig. 4b) (*see* **Note 1**). Primers should amplify a 600–1000 bp product around the breakpoint to recover potential deletions.
8. Perform a first 23-cycle PCR on 100–150 ng of genomic DNA with the external set of primers. We classically use FastStart Taq DNA polymerase (Roche).
9. Perform a 40-cycle PCR on 0.5–1 μL of the first PCR product, with the internal set of primers (*see* **Note 14**).
10. Load the PCR product on an agarose gel with EtBr, in 0.5× TBE buffer.
11. Capture a picture of the gel with a UV imaging station. A product should be amplified only for the $TAL^{A}+TAL^{B}$ transfection (Fig. 3a) (*see* **Note 15**).

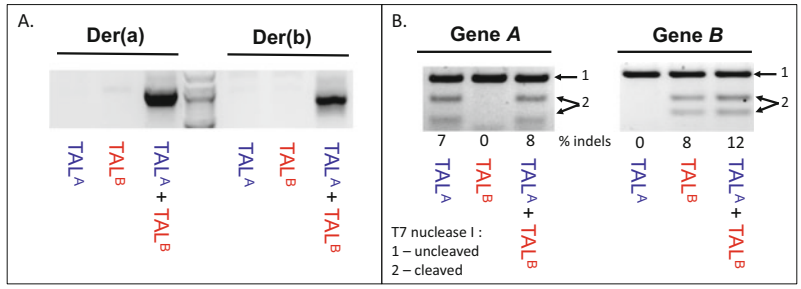


Fig. 3 PCR translocation detection and T7 Endonuclease I assay. (**a**) Nested PCR to detect derivative chromosomes Der(a) and Der(b) from a pool of transfected cells. A product should be amplified only in cells transfected with both TAL^A and TAL^B. (**b**) T7 Endonuclease I assay on gene *A* and gene *B*. T7 Endonuclease I cleavage efficiency reflects TALEN-induced DSB repair by NHEJ (% of indels). The PCR product corresponding to target site (*1*) must be cleaved in two products (*2*) when the corresponding TALEN is expressed and cuts

3.4 Estimation of TALEN Efficiency by a T7 Endocuclease I Assay (Adapted from [12])

This assay quantifies indels (insertions and deletions) resulting from DSB repair via nonhomologous end joining. This assay indirectly estimates the TALEN cleavage efficiency.

1. Design a set of primers amplifying a 300–500 bp region encompassing the TALEN target sites on non translocated chromosomes (*see* **Note 2**). TALEN target sites should not be located at the center of the amplicon, so that the T7 Endonuclease I cleavage results in two distinct digested bands during PAGE.
2. Set up a 35-cycle PCR on 50 ng of genomic DNA (extracted in **step 6** of Subheading 3.3) in a total of 25 μL to amplify region encompassing the TALEN target site. We classically use Phusion High-Fidelity DNA Polymerase (Thermo Scientific).
3. Verify amplification of only one PCR product by running 5 μL of PCR reaction on an agarose gel with EtBr in 0.5× TBE buffer.
4. Mix 10 μL of PCR reaction and 10 μL of 2× NEB Buffer 2.1 in duplicate.
5. Melt and randomly reanneal amplicons using the following denaturation program: 5 min at 95 °C, 95–25 °C at −0.5 °C/s, and 15 min at 4 °C. This step converts TALEN-induced mutations into mismatched duplex DNA.
6. Add 1.5 U of T7 Endonuclease I in one of the two duplicate tubes. Add the same volume of buffer in the second one as control.
7. Incubate at 37 °C for 20 min. This step allows the T7 Endonuclease I to cleave mismatched DNA duplexes.
8. Add 10 μL of 2× T7 loading buffer containing Proteinase K to 10 μL of DNA.

9. Incubate at room temperature for 5 min for T7 Endonuclease I degradation by Proteinase K.
10. Load on an agarose gel with EtBr in TBE buffer.
11. Capture the gel image with a UV imaging station, without saturating the picture (Fig. 3b). Open the file with an image quantification software (we typically use ImageJ open-source program) to quantify the ratio between the undigested and digested band intensity. The rate of indels induced by TALEN is calculated following the equation: % of indels = 100 × (1 – (1 – digested fraction)$^{1/2}$) [12] (*see* **Note 16**).

3.5 Quantification of Translocation Frequency by 96-Well Plate PCR Screen [4]

1. From 6 to 24 h after transfection (*see* **Note 17**), trypsinize 3 wells of each 96-well plate prepared in **step 14** of Subheading 3.2 in a small volume of trypsin (typically 30 µL).
2. Count the trypsinized cells and calculate the total number of plated cells, remaining alive after transfection.
3. After 48 h, remove the medium from the 96-well plates. Store them at –80 °C (*see* **Note 18**).
4. Prepare 2.5 mL of 1× lysis buffer. Add 100 µg/mL of Proteinase K. Fill the plate with 25 µL of buffer per well (*see* **Note 19**).
5. Incubate at 55 °C for at least 120 min in humid chamber (*see* **Note 20**).
6. Transfer the lysate into 96 PCR tubes.
7. Incubate at 95 °C for 10 min to inactivate Proteinase K.
8. Prepare 5 mL of 1× PCR Master Mix 1, with 200 nM of the external primers (same primers as in **step 8** of Subheading 3.3) and 25 µL of FastStart Taq Polymerase. Perform the first round 23-cycle PCR with 4–7 µL of cell lysate from each well in a total of 50 µL.
9. Prepare 1 mL of 1× PCR Master Mix 2, with 200 nM of the internal primers (same primers as in **step 9** of Subheading 3.3) and 10 µL of FastStart Taq Polymerase. Perform the second 40-cycle PCR with SYBR Green in real-time PCR conditions, with 0.5–1 µL of the first PCR in a total of 10 µL (*see* **Note 14**). The PCR has to contain a denaturation curve cycle.
10. Nested PCR fragments corresponding to translocation junctions are typically 600–900 bp, having melting temperatures (Tm) superior to 80 °C. All wells with a Tm >80 °C are considered to be positive, corresponding to fragments >100–150 bp. These PCR amplified fragments can include deletions and/or insertions and can be further sent for sequencing (*see* **step 13** below) (Fig. 4).

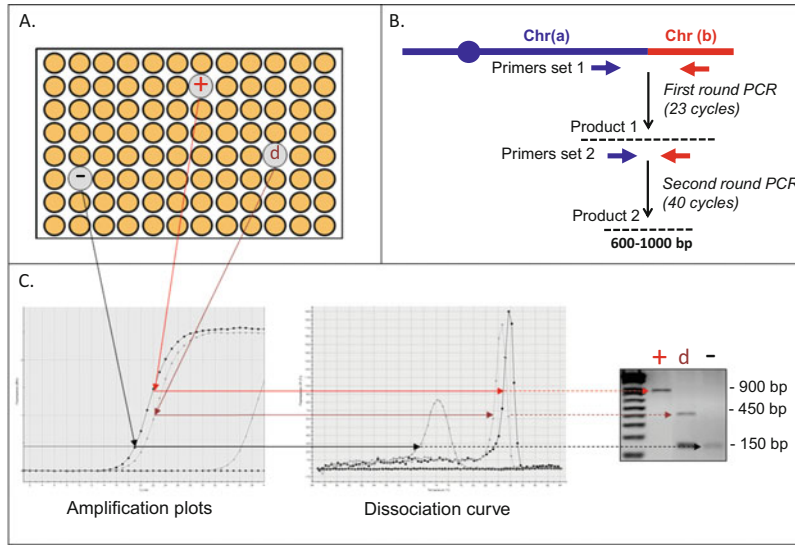


Fig. 4 96-Well plate PCR screen for translocation frequency calculation (modified from [4]). (**a**) After TALEN transfection, cells are plated in a 96-well plate. (**b**) Translocation formation is detected by nested PCR. Two rounds of PCR are needed for amplification of rare events like translocation junctions. The second PCR round uses primers (*arrows*) within the amplicon obtained in the first PCR round. A final 600–1000 bp product is amplified corresponding to the breakpoint junction. (**c**) A well is considered as positive (+) when the amplification plot shows an early amplification signal with a dissociation curve superior to 80 °C (corresponding to the melting temperature (Tm) of the PCR product). In case of late amplification with a low Tm, the well is considered as negative (–), often due to primer dimers. A Tm shift may correspond to a breakpoint junction with deletion (*d*) or insertion as shown in the gel. PCR products from the 96-well plate can be sent for sequencing to analyze the junctions (*see* Fig. 5)

11. If the number of positive wells is less than or equal to 12 positive wells per plate, we can surmise that each well contains no more than one translocation. In that case, the translocation frequency is the number p of positive wells divided by the number of plated cells calculated in **step 2**.
12. If the number of positive wells exceeds 12–14 translocations, frequency has to be corrected following a beta cumulative distribution function $k(x,a,b)$ (k = number of translocations per well, p = number of positive wells per number of plated cells, n = number of cells per well and $x = 1 - p$, $a = n - k$, $b = k + 1$). For example, if 35 positive wells are PCR amplified for 1×10^5 transfected cells per plate, the corrected number of translocations is 44 and the frequency is 2.3×10^{-5} (in terms of probability 6.4 wells have two translocations and 1 well has three translocations).
13. Positive wells can be sent to sequencing for further analysis of breakpoint junctions to identify the presence of deletions, insertions, or microhomologies (Fig. 5).

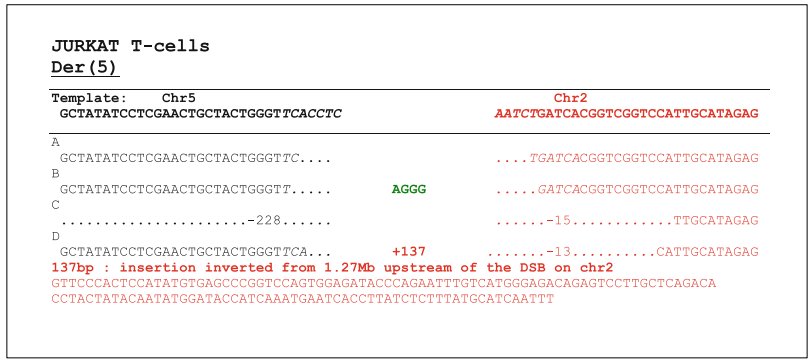


Fig. 5 Examples of translocation breakpoint sequences after expression of TALENs targeting *NPM* (Chr5) and *ALK* (Chr2) genes in JURKAT T-cell line (modified from [5]). Junctions were obtained from individual wells of a 96-well plate PCR screen. The chromosome 5 sequence is in *black* and the chromosome 2 sequence is in *red*, both shown with fill-in of the 5′ overhang (*italics*) and with the TALEN-binding sequence underlined. (**a**) Translocation deletion restricted to the 5′ overhang, no insertion and a 2 bp microhomology (*underline*). (**b**) Translocation with a 4 bp insertion (in *green*). (**c**) Translocation with a 243 bp deletion (*dots*) and a 4 bp microhomology (*underline*). (**d**) Translocation with a 137 bp insertion from chromosome 2 (sequence of the insertion in *red*)

14. Adjust cell dilutions after transfection in **step 14** of Subheading 3.2, to obtain plates with less than 30 positive wells to keep accuracy and allow correct sequencing (*see* **Note 21**).

3.6 Calculation of Frequency by Serial PCR Dilutions

We suggest an alternative method to estimate translocation frequency (when frequency is higher than 10^{-4}) specifically for cells grown in suspension when medium aspiration leads to cell loss.

1. Perform serial dilutions (1–2 for example) on genomic DNA extracted in **step 6** of Subheading 3.3, typically starting from 100 ng of DNA.
2. Perform a 37-cycles PCR on each dilution in triplicates as described in **step 8** of Subheading 3.3.
3. Load on an agarose gel with EtBr, in 0.5× TBE buffer.
4. Capture the gel image with a UV imaging station. Identify the smallest dilution of DNA where a PCR product can be amplified, i.e., including at least one translocated cell. Considering that one human diploid cell contains 6 pg of DNA, 6.25 ng correspond to about 10^3 cells. At the limit dilution, PCR product may be amplified in only one or two samples of the triplicate. Consider all the triplicates to determine the translocation frequency (Fig. 6).

3.7 Translocation Detection by FISH Analysis

In the case of cancer-related translocations, fluorescent in situ hybridization (FISH) probes are often available as diagnostic tools (e.g., break-apart probes) (Fig. 7a). When translocation frequency is $>2\times10^{-3}$ (calculated in Subheadings 3.5 or 3.6), translocations can be directly visualized by using FISH probes (*see* **Note 22**).

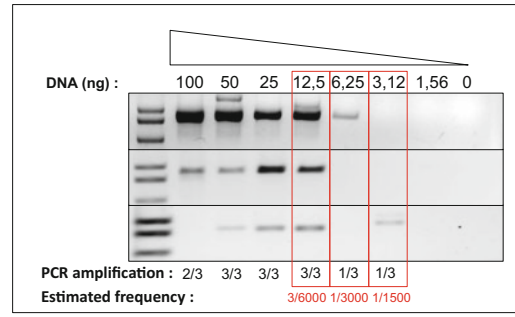


Fig. 6 Serial dilutions for translocation frequency calculation. Translocation junctions are PCR amplified from dilution of DNA in triplicates. The highest dilution from which a product is amplified corresponds to the smallest number of cells containing at least one translocated cell. Considering that one human diploid cell contains 6 pg of DNA, here we amplified 3 products out of 6000 cells, 1 product out of 3000 cells, and 1 product out of 1500 cells, corresponding to an average frequency of $5 \times 10^{-4} \pm 1.7 \times 10^{-5}$

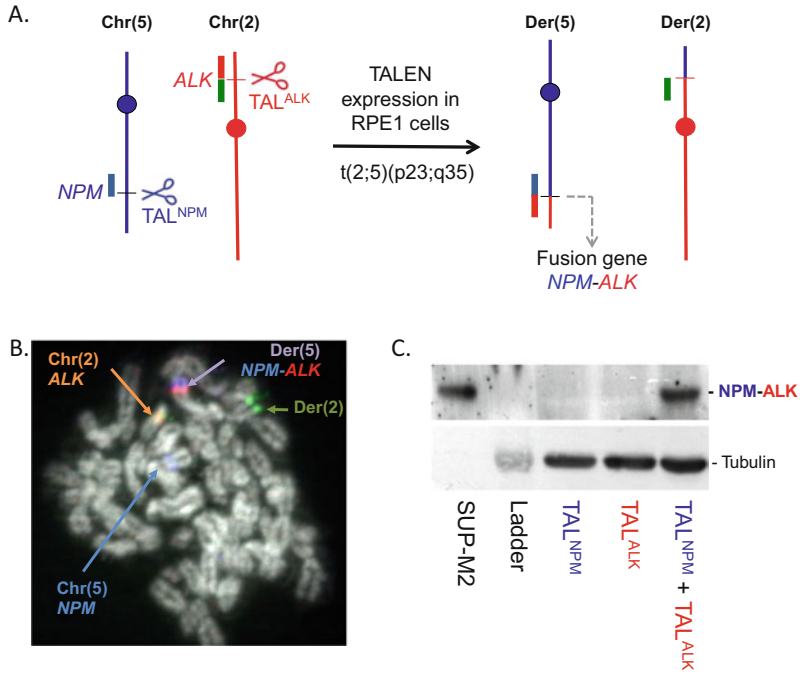


Fig. 7 FISH detection of translocation and western blot detection of induced fusion protein (modified from [5]). (**a**) We used TALENs targeting *NPM* and *ALK* genes to induce the anaplastic large cell lymphoma (ALCL)-related translocation, t(2;5)(p23;q35). The FISH probes are indicated by color lines. (**b**) We used a *red-green* commercial FISH probe targeting *ALK* (ALK Break Apart Probe, Cytocell) and a *blue* homemade probe (from BAC RP11-546B8) targeting *NPM* to visualize the translocation. Out of 70 screened metaphases, two exhibited translocations. (**c**) We confirmed the fusion gene expression by detecting the fusion protein NPM-ALK with a C-term ALK antibody (Cell Signaling Technology—ALK(31F12)) by western blot. After TAL^{NPM} and TAL^{ALK} expression, we could detect a protein with a size similar to the NPM-ALK protein from the ALCL cell line SUP-M2, used as control

1. Pre-warm the hypotonic solution at 37 °C.
2. Prepare fresh fixative solution. Precool at 4 °C.
3. 48 h after transfection, enrich cell population in metaphase by adding 50 ng/mL colcemid in one well of the 6-well plates prepared in **steps 13** and **14** of Subheading 3.2 (*see* **Note 23**).
4. Incubate at 37 °C for 2 h for RPEI cells (*see* **Note 24**).
5. Trypsinize cells. Be careful when aspirating medium and washing with PBS, dividing cells detach easily.
6. Centrifuge at 150 × *g* for 5 min at room temperature. Remove supernatant without aspirating the cells.
7. Proceed to hypotonic shock by adding first two drops of the hypotonic solution. Mix gently, and add drop by drop the hypotonic solution up to 2 mL.
8. Incubate at 37 °C for 10 min.
9. Centrifuge at 150 × *g* for 5 min at room temperature. Remove supernatant without aspirating the cells. Leave about 200 μL of hypotonic solution.
10. Proceed to pre-fixation by adding two drops of fixative solution. Mix gently. This step is important to avoid aggregates.
11. Centrifuge at 150 × *g* for 5 min at room temperature. Remove supernatant without aspirating the cells.
12. Gently resuspend cells in 2 mL of fixative solution.
13. Incubate for at least 24 h at 4 °C. Fixed cells can be stored at −20 °C for a few months.
14. The day of metaphase spreading, centrifuge fixed cells at 150 × *g* for 5 min at room temperature. Resuspend the pellet in a freshly prepared fixative solution.
15. Wash with 100 % ethanol and air-dry microscope slides. Dispose them on moist towels on a flat surface (e.g., a large glass or wood plate).
16. Release one drop of fixed cells from **step 14** on a slide. Allow to air-dry for 5 min. Verify metaphase concentration, spread quality and chromosome compaction with an optical microscope. More than hundreds of metaphases are needed per slide.
17. We advise you to mark drop edges with a diamond tip pen.
18. Allow slides to air-dry for 24 h at room temperature, protected from dust. Slides can then be stored at −20 °C for months.
19. From this step, work in the dark. You may use two probes, one labeling each chromosome implicated in the translocation. Mix the two probes and spot 10 μL of probe mixture on the sample. Using BAC resource to make your own FISH probe is also possible when one probe is not directly available as a diagnostic tool (see http://www.ncbi.nlm.nih.gov/genome/cyto/hbrc.shtml).

20. Apply a cover slip without making bubbles and seal with a rubber solution glue.
21. Denature the sample and probe by heating the slide at 75 °C for 3 min.
22. Perform hybridization following the manufacturer's instructions. We typically incubate overnight at 37 °C in a humid chamber (*see* **Note 20**).
23. Remove the glue and the cover slip carefully.
24. Wash the sample in 2× SSC (pH 6.3) at 72 °C for 2 min, without agitation.
25. Wash the sample in 2× SSC (pH 6.3) at room temperature for 30 s, without agitation.
26. Drain the slide. Apply a 100 ng/μL DAPI solution for 5 min.
27. Wash carefully with PBS and drain the slide.
28. Spot 10 μL of mounting medium and carefully apply a cover slip without making bubbles. Seal the slide with varnish.
29. Screen hundreds of metaphases with a fluorescence microscope to find the translocation induced by TALEN (Fig. 7b). Slide scanning platform systems can be used to facilitate the screening.
30. If the translocation frequency is too low and recovery of metaphases with translocated cells too difficult, you can enrich your population of translocated cells by using sib-selection over several cell generations using PCR screening for breakpoint junctions (Subheading 3.5) (as described in [4]). Briefly make a replica of the 96-well plate before running a first PCR screen and identify a positive well for translocation. Re-plate the cells from this well in a new 96-well plate at a lower density than in the first plate, in order to obtain an enrichment of at least tenfold. Grow to confluency, and then cells are replica plated one more time. One replica plate is used for PCR analysis to detect a positive well. This well (1/96), containing the translocation, is then grown for FISH analysis to identify translocation-positive metaphase spreads.

3.8 Western Blot Detection of the Fusion Protein

Chromosome translocation formation often induces either the expression of a new chimer protein, or the deregulation of preexistent gene. The fusion gene expression or the gene deregulated can be detected by western blot. Expression of a new chimer protein can be observed even for translocation frequency around 10^{-4} (calculated in Subheadings 3.5 or 3.6).

1. 48 h after transfection trypsinize one well of the 6-well plates prepared in **steps 13** and **14** of Subheading 3.2, and pellet cells as described in **steps 2–5** of Subheading 3.3. During all the cell lysis and when manipulating proteins, work on ice.
2. Resuspend cells in a small volume of lysis buffer (typically 50 μL), to obtain a protein concentration higher than 2 μg/μL.
3. Incubate on ice for 20 min.
4. Centrifuge at maximal speed, at 4 °C, for 15 min.
5. Transfer the supernatant in a new tube. The lysate can be stored at −80 °C.
6. Determine the protein concentration using a protein assay kit, following the manufacturer's instructions.
7. Prepare 30 μg of proteins in 1× Laemmli sample buffer with 0.1 M DTT, in a final volume of 20 μL.
8. Incubate at 95 °C for 5 min, and cool on ice.
9. Run the samples on a SDS-PAGE gel in running buffer. Gel percentage will depend on the size of the fusion protein.
10. Transfer on a nitrocellulose membrane.
11. Incubate the membrane for 1 h at room temperature in the blocking solution.
12. Incubate with the primary antibody following the manufacturer's instructions. Several antibodies may be available for one or the other protein implicated in the translocation. Try several if necessary.
13. Wash the membrane for 10 min in TBS-T at room temperature. Repeat this step twice.
14. Incubate with the HRP-conjugated secondary antibody following the manufacturer's instructions.
15. Wash the membrane for 10 min in TBS-T at room temperature. Repeat this step twice.
16. Develop the blot with chemiluminescent substrate solution (Fig. 7c) (*see* **Note 25**).

3.9 Translocation Induction with Other Different Types of Engineered Nucleases

All this procedure can be done with different types of engineered nucleases. We have tested this strategy with ZFNs, with the CRISPR/Cas9 system (DSB nuclease (*see* review [13]) and with modified Cas9-D10A nickase [14]), and by combining two different types of nucleases like a TALEN and a ZFN. Translocations can be induced with all these different types of nucleases (Fig. 8). No derivative chromosome must be PCR amplified when cells are transfected with only one TALEN or one ZFN or the Cas9 nuclease/nickase without all the RNA guides needed.

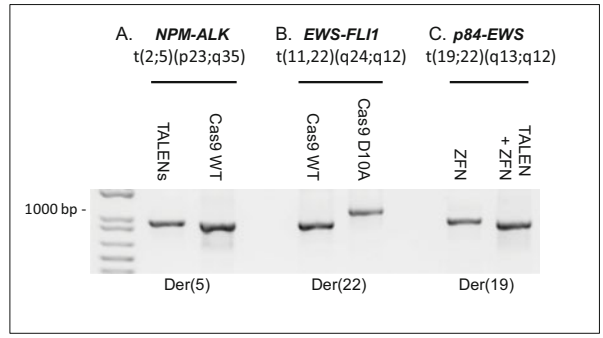


Fig. 8 Induction of translocations with other engineering nucleases. To generate DSBs and induce chromosomal translocations we used: ZFNs, TALENs, and the latest generation of the CRISPR/Cas9 system. For the wild type Cas9 (Cas9 WT), we used two RNA guides; for the double-nicking mutated version (Cas9 D10A), we used 2×2 RNA guides (two on each chromosome), as described in [17]. For all nucleases, we could recover derivative chromosomes for (**a**) *NPM-ALK* in primary T-cells using TALENs or Cas9 WT; (**b**) *EWS-FLI* in mesenchymal stem cells (hMSCs) using Cas9 WT or Cas9 D10A; (**c**) *p84-EWS* in HCT116 cell line using ZFNs, or TALEN targeting *p84* + ZFN targeting *EWS*

4 Notes

1. Primers may be designed using a variety of web design programs such as Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/). Settings are chosen to yield 20 bp long primers with melting temperatures around 60 °C.
2. Primers may be designed using a variety of web design programs such as Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/). Settings are chosen to yield 22 bp long primers with melting temperatures around 62 °C.
3. If the translocation generates a fusion gene, we advice you to choose target sites within intronic regions of the genes involved in the translocation.
4. Web design tools offer to exclude off-target sites as well.
5. Two pairs of TALENs are used to cut at two different loci. To prevent undesired dimerization and increase TALEN specificity, you can use mutant Fok1 domains, forming obligate heterodimers [15].
6. Transfection can induce high cell mortality. This mortality can be explained by (1) inappropriate cell culture conditions; (2) inappropriate transfection conditions (follow manufacturer's conditions); (3) transfected DNA volume over 10 μL; (4) high

number of TALEN off-target sites; and (5) induction of a translocation inducing cell death. Check all these parameters to optimize translocation induction. Cell mortality can also be decreased by transfecting TALEN mRNAs after in vitro transcription instead of expression plasmids [16].

7. The use of antibiotics increases cell mortality and can interfere with transfection agents. If you use antibiotics, remove them at this step.
8. Working with pre-warmed medium and solutions at room temperature during the transfection promotes cells recovery. Do not use any 4 °C solutions.
9. It is important to remove medium before transfection. FBS supplementing the medium contains several transfection inhibitors. You can also proceed to a supplemental wash with PBS to increase the transfection rate.
10. To prevent cell mortality, avoid at maximum pipette aspiration movements.
11. Transfection conditions such as the program or the cell number may vary depending on the cell line. Conditions described here are optimized for the RPE1 cell line. For other cell types, we advise you to refer to Lonza Nucleofector Protocols, or to test several transfection programs using eGFP plasmid expression. After 24 h, analyze the rate of cells expressing eGFP with flow cytometry. A low transfection rate may explain low translocation efficiency.
12. You can perform several transfections to increase cell quantity needed for each analysis. We advise you to seed more cells for protein analysis, and less for FISH analysis, so that they are 80 % confluency after 48 h.
13. With frequency higher than 10^{-4}, it is possible to detect translocations by a classical one-run PCR, as shown in Subheading 3.6.
14. Nested PCR is a highly sensitive method, presenting risks of contamination by other PCR products. Work carefully in a dedicated place, wear gloves, take care not to contaminate a tube with already amplified PCR products, and clean up your bench and material after each PCR.
15. PCR amplification of bigger or smaller products than expected may appear, due to the presence of insertions and deletions generated during translocations formation. This can be confirmed by sequencing.
16. We consider that a TALEN is efficient as soon as indels are observed (i.e., >3 % of indels). However, we have already detected translocations despite a negative T7 Endonuclease I assay, due to the high sensitivity of nested PCR method.

In case of negative test, you can try to increase the quantity of transfected TALEN expression plasmid in **step 3** of Subheading 3.2.

17. Do not exceed this time lapse to count cells, before they divide, to ensure an accurate translocation frequency calculation.
18. If cells grow slow, it is possible to wait more than 48 h to freeze the 96-well plates, to let them time to divide. This could improve the 96-well plate PCR screen sensitivity.
19. Prepare a 10× lysis buffer that you can store at 4 °C.
20. The humid chamber consists in a closed plastic container with moist towels at the bottom.
21. In RPE1 cells, we typically obtain a translocation frequency of 10^{-3}. Translocation frequency mainly depends on each TALEN cleavage efficiency, cell type, and transfection conditions.
22. We advise you to first adjust the protocol using non transfected cells to check the background of the probes.
23. Cell confluency should not exceed 80 % at the time of colcemid treatment.
24. Incubation time depends on cell division rate (can be up to 6 h). The cells round up upon colcemid treatment.
25. When studying a cancer translocation, proteins from the corresponding cancer cells can be run on the same gel and be used as a control (Fig. 7c).

Acknowledgements

We thank Maria Jasin (MSKCC, NY), Anne De Cian, Carine Giovannangeli, and Jean-Paul Concordet (MNHN) for helpful discussions. This work is supported by the "ANR" grant (ANR-12-JSV6-0005), M.P. is supported by the "Canceropole Ile-de-France," and H.G. is supported by "La ligue contre le cancer."

References

1. Collado M, Serrano M (2010) Senescence in tumours: evidence from mice and humans. Nat Rev Cancer 10(1):51–57. doi:10.1038/nrc2772, nrc2772 [pii]
2. Evan GI, d'Adda di Fagagna F (2009) Cellular senescence: hot or what? Curr Opin Genet Dev 19(1):25–31. doi:10.1016/j.gde.2008.11.009, S0959-437X(08)00191-3 [pii]
3. Richardson C, Jasin M (2000) Frequent chromosomal translocations induced by DNA double-strand breaks. Nature 405(6787):697–700. doi:10.1038/35015097
4. Brunet E, Simsek D, Tomishima M, DeKelver R, Choi VM, Gregory P, Urnov F, Weinstock DM, Jasin M (2009) Chromosomal translocations induced at specified loci in human stem cells. Proc Natl Acad Sci U S A 106(26):10620–10625. doi:10.1073/pnas.0902076106, 0902076106 [pii]
5. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, Rene O, Katibah G, Zhang L, Holmes M, Doyon Y, Concordet JP, Giovannangeli C, Jasin M, Brunet E (2013) Cancer translocations in

human cells induced by zinc finger and TALE nucleases. Genome Res 23:1182. doi:10.1101/gr.147314.112

6. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39(12):e82. doi:10.1093/nar/gkr218
7. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, Yang B (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. Nucleic Acids Res 39(1):359–372. doi:10.1093/nar/gkq704
8. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29(2):143–148. doi:10.1038/nbt.1755, nbt.1755 [pii]
9. Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res 39(21):9283–9293. doi:10.1093/nar/gkr597
10. Grau J, Boch J, Posch S (2013) TALENoffer: genome-wide TALEN off-target prediction. Bioinformatics 29(22):2931–2932. doi:10.1093/bioinformatics/btt501
11. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29(8):699–700. doi:10.1038/nbt.1939
12. Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ (2010) A rapid and general assay for monitoring endogenous gene modification. Methods Mol Biol 649:247–256. doi:10.1007/978-1-60761-753-2_15
13. Mali P, Esvelt KM, Church GM (2013) Cas9 as a versatile tool for engineering biology. Nat Methods 10(10):957–963. doi:10.1038/nmeth.2649
14. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell 154(6):1380–1389. doi:10.1016/j.cell.2013.08.021
15. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods 8(1):74–79. doi:10.1038/nmeth.1539, nmeth.1539 [pii]
16. Dahlem TJ, Hoshijima K, Jurynec MJ, Gunther D, Starker CG, Locke AS, Weis AM, Voytas DF, Grunwald DJ (2012) Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. PLoS Genet 8(8):e1002861. doi:10.1371/journal.pgen.1002861
17. Ghezraoui H, Piganeau M, Renouf B, Renaud JB, Sallmyr A, Ruis B, Oh S, Tomkinson AE, Hendrickson EA, Giovannangeli C, Jasin M, Brunet E (2014) Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol Cell 55(6):829–842. doi:10.1016/j.molcel.2014.08.002

Chapter 10

Mutagenesis in Newts: Protocol for Iberian Ribbed Newts

Toshinori Hayashi and Takashi Takeuchi

Abstract

Newts have the remarkable capability of organ/tissue regeneration, and have been used as a unique experimental model for regenerative biology. The Iberian ribbed newt (*Pleurodeles waltl*) is suitable as a model animal. We have established methods for artificial insemination and efficient transgenesis using *P. waltl* newts. In addition to the transgenic technique, development of TALENs enables targeting mutagenesis in the newts. We have reported that TALENs efficiently disrupted targeted genes in newt embryos. In this chapter, we introduce a protocol for TALEN-mediated gene targeting in Iberian ribbed newts.

Key words Newts, Animal model, Molecular genetics, Regeneration, Transgenesis, Genome editing

1 Introduction

Amphibians have a long history as model animals in basic biology such as developmental and cell biology. Newts have the remarkable capability to regenerate organs and/or tissues, and have been used as a unique experimental model for regenerative biology. Newts can regenerate various body parts such as limbs, optical tissues, brain, spinal cord, intestine, and even the heart throughout their lives [1, 2]. However, the mechanisms regulating regeneration are not well understood. In order to investigate the mechanisms of regeneration from the perspective of gene function, an experimental model system for genetic manipulation in newts was needed. For that purpose, we introduced the Iberian ribbed newt (*Pleurodeles waltl*) as a model animal. We observed that *P. waltl* spawned fertilized eggs all year around in the laboratory. We were able to shorten the period for sexual maturation from 18 to 6 months. In addition, each female laid more than 150–600 eggs per spawning and they spawned every 2–4 weeks [3]. Finally, we established methods for artificial insemination and efficient transgenesis. Although gene knockout newts, which are very useful for understanding gene function, were still not available in newts until a few years ago, genome-editing technology

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_10,

changed this situation completely. We have reported that TALENs efficiently disrupted targeted genes in newt embryos [4]. In this chapter, we introduce a protocol for TALEN-mediated mutagenesis in Iberian ribbed newts.

2 Materials

2.1 TALEN Vectors

Plasmids cording TALENs, which recognize the target gene (*see* **Notes 1** and **2**).

2.2 Equipment (Fig. 1)

1. Injection dish (*see* **Note 3**).
2. Microinjector and micromanipulator (*see* **Note 4**).
3. Stereomicroscope with flexible lighting (*see* **Note 5**).
4. Low-temperature incubator (temperature range; 8–25 °C).
5. Glass capillaries and puller (*see* **Note 6**).

2.3 Reagents and Solutions (See Note 7)

1. 10× Holtfreter's solution (HF, *see* **Note 8**): Dissolve 35 g NaCl (600 mM), 0.5 g KCl (7 mM) and 2 g $NaHCO_3$ (24 mM) in 1 L of distilled water (*see* **Note 9**).
2. 1× HF solution: Add 100 mL 10×HF solution to 900 mL distilled water. Sterilize by autoclaving and allow it to cool at room temperature. Add 5 mL of 1 M HEPES (5 mM), 1.7 mL of 1 M $CaCl_2$ (1.7 mM) and 1.7 mL 1 M $MgSO_4$ solution (1.7 mM) to this solution. Store at room temperature (*see* **Note 10**).
3. 0.25×HF solution: Add 25 mL 10×HF solution to 975 mL distilled water and sterilize by autoclaving. Allow the solution to cool at room temperature and add 5 mL of 1 mL HEPES, 1.7 mL of 1 M $CaCl_2$, 1.7 mL 1 M $MgSO_4$, and 5 mL of

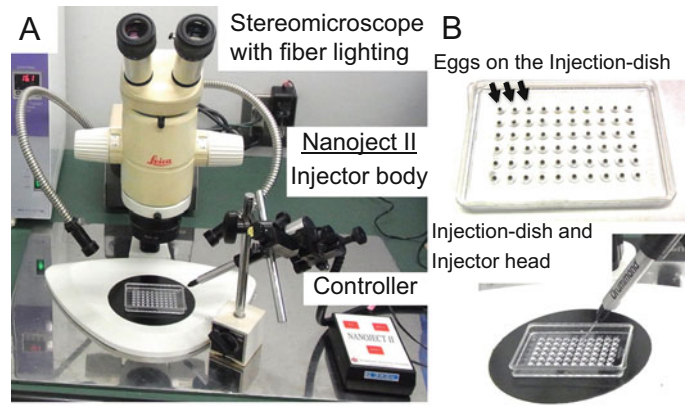


Fig. 1 Setting up the equipment for the microinjection. (**a**) Microinjector, micromanipulator, and stereomicroscope. (**b**) Fertilized eggs in the injection dish

(50 mg/mL) gentamicin sulfate. Store at 4 °C is sufficient. (*see* **Note 10**).

4. 0.1 % MS222 solution for anesthesia: Add 0.4 g to 400 mL 0.25 × HF solution (*see* **Note 11**).
5. Ficoll solution: Dissolve 4 g of Ficoll in 100 mL 0.25 × HF solution containing gentamicin. Sterilize by filtration using membrane filter and clean bottle. Store at 4 °C.
6. De-jelly solution: Dissolve 1 g of sodium thioglycolate in 50 mL 25 × HF solution, adjust the to pH 7.9 with 5 N NaOH (*see* **Note 11**).
7. Human chorionic gonadotropin (hCG): Dissolve 1000 units in 1 mL 1 × HF solution (*see* **Note 12**).

3 Methods

3.1 mRNA Synthesis for TALENs

1. Linearize 2–10 μg of plasmid vectors with appropriate restriction enzymes.
2. Extract the plasmids by phenol/chloroform, and ethanol precipitation, and dissolve the plasmid in RNase-free water (500 ng/μL).
3. Synthesize TALEN mRNAs by in vitro transcription using RNA polymerase (*see* **Note 13**).
4. After ethanol precipitation, dissolve in nuclease-free water. Store at −80 °C.

3.2 Fertilized Egg Preparation

1. Fourteen to 18 h prior to egg collection, subcutaneously inject 100–200 units of hCG into the lower jaw of mature females (older than 9 months).
2. On the next day, anesthetize sexually mature male animals (older than 6 months) with 0.1 % MS222 solution.
3. Squeeze out sperm by pressing the abdomen of the male and dilute with 1 × HF to 2000–5000 sperm/μL (*see* **Note 14**).
4. When the female started laying eggs, anesthetize the female animal with 0.1 % MS222 solution.
5. Push an abdomen gently and collect the eggs in a dry Petri dish (100 mm diameter). Start insemination immediately (*see* **Note 15**).
6. Inseminate the eggs with diluted sperm suspension (5–20 μL per 20 eggs) and mix gently using pipette tips. Incubate the inseminated eggs at room temperature for 15 min (*see* **Note 16**).
7. Pour 0.25 × HF into the dish, and then incubate the eggs for a further 15 min to allow water absorption by the jelly.

8. Soak the eggs in de-jelly solution for 5 min with gentle shaking. The egg jelly may dissolve (*see* **Note 17**).
9. Rinse de-jellied eggs 3–5 times in 0.25 × HF.
10. Transfer the eggs to Injection-dish (Fig. 1b), using a plastic pipet (*see* **Note 18**).
11. Store the eggs at 8–10 °C until use (*see* **Note 19**).

3.3 Microinjection and Rearing the Embryos

1. Prepare micro needles (diameter of outer tip: 10–20 μm) (*see* **Note 20**).
2. Dilute RNAs with RNase-free water to 10–100 ng/μL (*see* **Note 21**).
3. Set a micro needle on the injector head, and fill the micro needle with the diluted RNA.
4. Stab an egg with a glass needle, and push "inject" switch (*see* **Note 22**).
5. After microinjection, incubate the eggs overnight at 25 °C.
6. Sort normally developing embryos (*see* **Note 23**) and transfer them into a new dish filled with 0.25 × HF.
7. Wash out the Ficoll in 0.25 × HF for 30 min twice (*see* **Note 24**).
8. Incubate the embryos in 0.25 × HF until hatching stage, and then transfer the larvae into tap water.
9. Start feeding after 10–12 days of post fertilization (*see* **Note 25**).

3.4 Detection of the Mutation

1. For genotyping, collect the tail tips or limb tips under anesthesia (*see* **Note 26**).
2. Prepare lysates from the newt tissue samples, and amplify the target site by PCR.
3. Screen the mutation by DNA sequencing, RFLP (restriction fragment length polymorphism) assay, or Cel-I assay.

4 Practical Examples: Phenotypes, Detection of the Mutations

4.1 Targeted Disruption of the Tyrosinase Gene (Fig. 2)

Tyrosinase is involved in melanin synthesis. The disruption of this gene causes an albino phenotype. Normal melanophores were not observed in the animals injected with equal amounts of right TALEN (R-TALEN) and left TALEN (L-TALEN) mRNAs. In contrast, injection with only R-TALEN as negative control showed wild-type phenotype.

4.2 Detection of the Targeted Mutation by RFLP Assay and DNA Sequencing (Fig. 3)

In order to confirm disruption of tyrosinase gene, the target site which containing a BspMI site (Fig. 3b) was amplified from albino animals by PCR. And then, the PCR products were digested with BspMI (*see* **Note 27**). When the TALENs destroy tyrosinase gene,

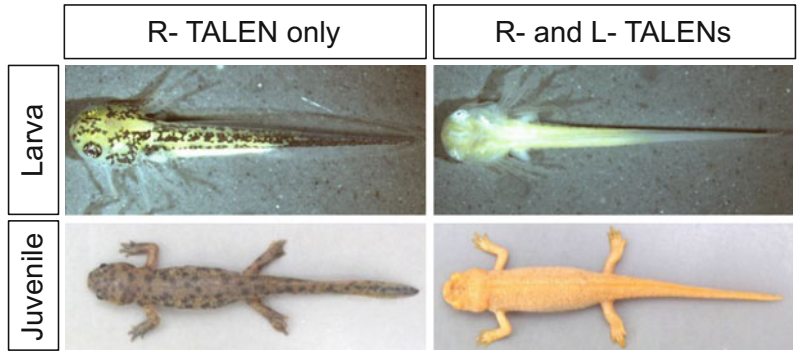
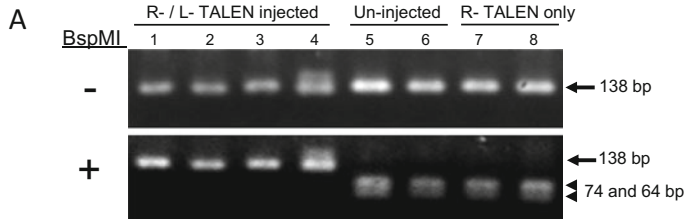


Fig. 2 Phenotypes of tyrosinase gene knockout newts. Two-week post-fertilization larvae (*upper*) and 3-month-old juveniles (*lower*). Animals were injected with 200 pg of R-TALEN only (*left column*) or total of 400 pg of R- and L-TALENs (*right column*)



B Genomic sequences of *tyrosinase* in TALEN-injected albino newt

	L-TALEN binding site		BspMI	R-TALEN binding site	
Wild-type	CGCCCGCCT	TCCTGCCGTGGCACCGGA	TCTACCTGCTCCTCT	GGGAACGCGAGCTCCAGA	AGGTGACCGGA
Albino animals	CGCCCGCCT	TCCTGCCGTGGCACCGGA	TCT------TCCTCT	GGGAACGCGAGCTCCAGA	AGGTGACCGGA
	CGCCCGCCT	TCCTGCCGTGGCACCGGA	TCT----------CT	GGGAACGCGAGCTCCAGG	AGGTGACCGGA
	CGCCCGCCT	TCCTGCCGTGGCACCGGA	TCTACC------TCT	GGGAACGCGAGCTCCAGA	AGGTGACCGGA
	CGCCCGCCT	TCCTGCCGTGGCACCGGA	TCTACC-----CC--T	GGGAACGCGAGCTCCAGG	AGGTGACCGGA

"-" indicates deletion of nucleotide.

Fig. 3 (**a**) A gel electrophoresis analysis of genomic PCR product of R- and L-TALEN-injected (*lanes 1–4*), un-injected (*lanes 5* and *6*), and R-TALEN-only-injected (*lanes 7* and *8*) larvae. *Upper image* shows undigested products (*arrow*, 138 bp), *lower image* shows BspMI-digested products. *Arrowheads* indicate the fragments cleaved by BspMI (64 and 74 bp). (**b**) Genomic sequences observed in tyrosinase knockout larvae. The wild-type sequence is shown at the *top*. TALEN-binding sequences are *boxed*

the PCR products became BspMI tolerant (Fig. 3a lanes 1–4). In contrast, samples amplified from un-injected and R-TALEN only injected animals were digested into two bands. The disruption was confirmed by sequencing of the PCR products. The results indicated mutations in all clones examined (Fig. 3b).

5 Notes

1. It is suggested that an important domain such as the catalytic core of an enzyme or first ATG should be chosen as the target site of the TALENs in order to destroy the function of the protein by the deletion of a small number of nucleotides or generating frameshifts.
2. Obtain or construct TALEN plasmids for your targeting genes. Many laboratories and companies provide various types of TALEN scaffolds. We recommend Platinum Gate TALENs [5], as high activity had been confirmed in the newt embryos [4]. The platform is available from Addgene, and the protocol is described by Sakuma et al. [5], also Chapter 6 in this issue.
3. We prefer to use Nunc minitray (60 wells). Since the dechorionized eggs settle in the well at the bottom, it helps to hold the egg during the injection (Fig. 1b). Fill the dish with Ficoll solution.
4. We prefer to use a NANOJECT II microinjector (Drummond, Broomall, PA). As this injector is driven not by compressed air but by an electric motor, it is compact and easy to operate. You can use other types of microinjectors. Refer to the operating instructions for detailed directions before using. The injector head should be mounted in a suitable micromanipulator.
5. Although it is in common with frog and newt, the working distance for the injector must be established.
6. A puller is a device for making the micro needles from the glass capillaries. Select appropriate type of capillary for your microinjector according to the company's instruction.
7. To examine the disruption of target genes, PCR primer sets amplifying the target sites (optional; restriction enzymes to digest the PCR products) are required. These reagents should be prepared before a gene-targeting experiment.
8. Holtfreter's solution is PBS modified for the newts.
9. Sterilize by autoclaving and store at room temperature.
10. After autoclaving, add HEPES, $CaCl_2$, $MgSO_4$, and gentamicin sulfate.
11. Prepare this solution just before use.
12. Store at –20 °C. It should be stable after several times of freeze-thawing.
13. The 5-prime ends of RNAs must be caped. It is easy to use a kit for caped-mRNA synthesize (e.g., mMESSAGE mMACHINE T7 Ultra Kit, Life Technologies).
14. The sperm suspension can be stored up to 8 h at room temperature. You may collect sperm from the same animals every 2–4 weeks.

15. You may collect the eggs from the same animals 2–3 times with 2 h intervals.
16. The lid of the Petri dish is shut to keep humidity.
17. Since the de-jelly solution is harmful, Keep for 5 min or less. And then, rinse the eggs well at the next step.
18. When the eggshells are remained, remove them using fine forceps.
19. Although you can delay the timing of the first cleavage at 8–10 °C, a good result will be obtained if injection is performed as quickly as possible after fertilization.
20. The micro needle tips should not be ground. Prepare spear needles and stocked for future use. You can refer to Casco-Robles et al. [6].
21. Appropriate amounts of TALENs depend on their activity. Optimize total amounts of TALEN mRNAs between 10 and 400 pg per egg. The injection volume must be set between 9 and 20 nL per egg.
22. Support the egg using the forceps.
23. The embryos should be between late morula to blastula stages. The criteria for the developing stages were defined by Shi and Boucaut [7].
24. Ficoll prevents gastrulation: remove completely.
25. Animals should be fed at least five times per week with hatched blain shrimp for larvae and compound feeds for juvenile/adult newts. *P. waltl* newts eat compounds feeds for Xenopus, catfish, and trout.
26. Do not worry, they regenerate the tail or limbs. 0.01 % MS222 is suitable for the larvae.
27. Some minor restriction enzymes such as BspMI have unstable activity. It is recommended to use a major restriction enzyme as possible.

6 Discussion

The amphibians have a long history of research in early developmental biology because their embryos are easy to manipulate. It is also easy to obtain numerous single cell-stage fertilized eggs from Iberian ribbed newts. In addition, we have already revealed that TALENs worked effectively in these newts, and that knockout animals were generated at F0 generation with high efficiency. Moreover, we have previously demonstrated precise TALEN-mediated insertion of the loxP site into the newt genome [4]. Establishment of genome editing technology enables more advanced manipulation of gene expression such as conditional knockout, or exchange of promoter sequence.

Adding to the TALENs, Flowers et al. reported that the CRISPR/Cas9 system also works efficiently in another urodele, axolotl [8]. We prefer to use TALENs for the newts, because it is believed that TALENs have low off-targeting frequency compared to the CRISPR/Cas9 system. As urodeles have a huge genome, the low off-targeting is advantageous as a tool for genome editing. However, it is relatively easy to generate target-specific sgRNAs in CRISPR/Cas9. Hereafter, TALENs or CRISPR/Cas9 would be chosen according to the experimental design. In addition to regeneration, urodeles would also represent valuable experimental models for studies involving development, stem cells, and reprogramming. These nuclease-mediated genome alterations make it possible to perform genetic studies that were previously difficult or impossible in newts.

Acknowledgements

The TALENs were designed and constructed by Drs. Takashi Yamamoto and Tetsushi Sakuma (Hiroshima University, Hiroshima, Japan). Dr. Ken-Ichi Suzuki (Hirosima University) provided useful suggestions during the preparation of this chapter. This work was supported by JSPS KAKENHI and MEXT. We would like to thank Kyorin Corporation (Hyogo, Japan) for providing the feed for the newts.

References

1. Agata K, Inoue T (2012) Survey of the differences between regenerative and non-regenerative animals. Dev Growth Differ 54:143–152
2. Brockes JP, Kumar A (2002) Plasticity and reprogramming of differentiated cells in amphibian regeneration. Nat Rev Mol Cell Biol 3:566–574
3. Hayashi T, Yokotani N, Tane S, Matsumoto A, Myouga A, Okamoto M, Takeuchi T (2013) Molecular genetic system for regenerative studies using newts. Dev Growth Differ 55:229–236
4. Hayashi T, Sakamoto K, Sakuma T, Inoue T, Kawaguchi E, Agata K, Yamamoto T, Takeuchi T (2014) Molecular genetic system for regenerative studies using newts. Dev Growth Differ 56:115–121
5. Sakuma T, Ochiai O, Kaneko T, Mashimo T, Tokumasu D, Sakane Y, Suzuki K, Miyamoto T, Sakamoto N, Matsuura S, Yamamoto T (2014) Repeating pattern of non-RVD variations in DNA-binding modules enhances TALEN activity. Sci Rep 3:3379
6. Casco-Robles MM, Yamada S, Miura T, Nakamura K, Haynes T, Maki N, Del Rio-Tsonis K, Tsonis PA, Chiba C (2011) Expressing exogenous genes in newts by transgenesis. Nat Protoc 6:600–608
7. Shi DL, Boucaut JC (1995) The chronological development of the urodele amphibian Pleurodeles waltl (Michah). Int J Dev Biol 39: 427–441
8. Flowers GP, Timberlake AT, McLean KC, Monaghan JR, Crews CM (2014) Highly efficient targeted mutagenesis in axolotl using Cas9 RNA-guided nuclease. Development 141: 2165–2171

Chapter 11

Targeted Mutagenesis in *Bombyx mori* Using TALENs

Yoko Takasu, Toshiki Tamura, Marian Goldsmith, and Michal Zurovec

Abstract

Bombyx mori is a valuable model organism of high economic importance. Its genome sequence is available, as well as basic genetic and molecular genetic tools and markers. The introduction of genome editing methods based on engineered nucleases enables precise manipulations with genomic DNA, including targeted DNA deletions, insertions, or replacements in the genome allowing gene analysis and various applications. We describe here the use of TALENs which have a simple modular design of their DNA-binding domains, are easy to prepare and proved to be efficient in targeting of a wide range of cleavage sites. Our procedure often allows the production of individuals carrying homozygous mutations as early as in the G_1 generation.

Key words Silkworm, pBlue-TAL, Engineered nucleases, Golden Gate assembly, Nonhomologous end joining, NHEJ

1 Introduction

Bombyx mori has a long tradition as a valuable model organism for biochemical, physiological, and genetic studies. Early biochemical research included protein structure, leading to the first discovery of the β-sheet in the protein fibroin, the major constituent of silk [1]. In addition, the availability of large amounts of silkworm pupae allowed the isolation of 25 mg of the pure crystalline ecdysone, the first insect hormone identified—from 500 kg of silkworm material [2]. The abundance, large size and high GC content of fibroin mRNA enabled its isolation as the first chemically identified messenger RNA in eukaryotes [3]. Silk glands served as a crucial model for tissue specific gene expression [4] and for studies of the genetic and physiological basis of massive protein synthesis during silk production [5]. Discovery of the high stability of the mRNA for cocoonase, the "hatching" enzyme which allows the moth to escape from its cocoon, and clarification of the cluster organization of the chorion genes, one of the earliest studied multigene families, were also important landmarks in molecular biological research [6, 7].

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_11,

Sequencing of the silkworm genome (~432 Mb) in 2008 [8] and rapid progress in the development of techniques for the production of transgenic animals in the last several years confirmed its position as an important insect model organism (for reviews *see* [9, 10]).

B. mori provides a number of advantages as a model system. It is a fully domesticated organism which cannot fly and needs human help for its reproduction. It is well adapted to artificial diet, is easy to breed, and produces large numbers of eggs. It has a generation time of 40–55 days at ~24 °C, including 11–14 days for embryonic development, 18–25 days for larval instars (3–4 days for first larval instar, 2–3 days for second, 3–4 days for third, 4–6 days for fourth and 6–8 days for fifth larval instar) and 10–15 days for the pupal stage. The adult silkmoth also needs at least 1 day for mating and oviposition.

There are more than 600 mutant strains of *B. mori* available in Japanese (http://www.shigen.nig.ac.jp/silkwormbase/, [11]) and Chinese (Southwest University) stock centers with more than 350 mutant loci (http://www.fao.org/docrep/005/ad108e/ad108e08.htm#TopOfPage). The silkworm has 28 chromosomes with well covered genetic linkage maps of molecular markers including RAPD [12], SSR [13], and SNPs [14, 15], large EST [16], and full-length cDNA [17] collections, and extensive BAC libraries [18]. These excellent resources have resulted in efficient positional cloning of more than 45 morphological mutations (*see* for example, [19–23]) Introduction of advanced molecular biology methods including transgenesis (*piggyBac* and *Minos*) [24, 25], transposon mutagenesis [26], and functional binary expression systems [27] make *B. mori* an attractive model for functional genomics.

The establishment of genome editing methods in the last 5 years represents another important step in silkworm research. Genome editing with engineered nucleases allows precise manipulations with genomic DNA, including targeted DNA deletions, insertions, or replacements using custom-prepared engineered nucleases. These enzymes induce double-stranded breaks at desired genomic locations, and subsequent interference with the repair process allows the insertion of donor DNA fragments into the mutated region. Three types of engineered nucleases have been tested in the *B. mori* system, including Zinc finger nucleases (ZFNs) [28], Transcription Activator-like nucleases (TALENs) [28–32], and nucleases from the adaptive prokaryotic immune system based on Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) [33]. We describe here the use of TALENs which have a simple modular design for their DNA-binding domains, are easy to prepare, and have a high success rate and efficiency in targeting a wide range of candidate cleavage sites in addition to low off-target effects.

2 Materials

2.1 Silkworms

The research community employing *B. mori* is relatively small and most resources are available from Kyushu University (http://www.shigen.nig.ac.jp/silkwormbase/index.jsp; [11]). Various *B. mori* strains are also kept at the National Institute of Agrobiological Sciences (NIAS; Tsukuba, Japan), and at Tokyo University (http://silkbase.ab.a.u-tokyo.ac.jp/cgi-bin/index.cgi). In addition, a number of *B. mori* strains including more than 700 silkworm germplasms for mutations and geographic races are kept at the Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhengjiang, China (http://www.sricaas.com/eng/enindex.htm), more than 300 strains are kept at The National Academy of Agricultural Science in Korea (http://www.naas.go.kr/english/; [34]), and about 200 *B. mori* strains are kept in Italy (http://api.entecra.it/index.php?c=29&ln=eng). Information on *B. mori* genes can be found at two databases whose resources are complementary and connected, Kaikobase in Japan (http://sgp.dna.affrc.go.jp/KAIKObase/) and SilkDB in China (http://www.silkdb.org/silkdb/).

Several silkworm strains are widely used in biological experiments. Daizo (p50) is a diapausing Chinese silkworm strain with a relatively low level of polymorphism which was used for the genome sequencing project. Daizo is, however, not an ideal strain for transgenesis since it is not polyvoltine. We chose a different strain, w1-pnd, which is nondiapausing. This strain seems to be well adapted to the injection of DNA and rearing by artificial diet. There are significant sequence differences between Daizo and the w1-pnd strain, which also seems to be polymorphic within substrains. Therefore, it is practical to check the DNA sequence of the target area as well as the primer functions in the w1-pnd strain before conducting experiments with it. Another polyvoltine *B. mori* strain successfully used for TALEN mutagenesis is the Indian strain, Nistari [29].

2.2 Equipment

1. Binocular microscope.
2. Micromanipulator and microinjection system.
3. Tungsten needle.
4. Glass capillary.

2.3 Reagents

1. Plasmids from Golden Gate TALEN and TAL Effector Kit 2.0 (Addgene).
2. pBlue-TAL vector (plasmid 49401, Addgene).
3. Competent *E. coli* cells.
4. Proteinase K (0.2 mg/ml) in a buffer containing 0.5 % SDS (w/v).

5. Restriction enzymes *Bsa*I, *Esp*3I, and *Xba* I.
6. T4 DNA Ligase (400 units/μl).
7. Ligation/restriction buffer 10×(300 mM Tris–HCl, pH 7.8, 100 mM $MgCl_2$, 100 mM dithiothreitol, 10 mM ATP).
8. Plasmid-safe nuclease.
9. PCR primers for array amplifications from pFus and pBlue-TAL, respectively: pCR8_F1: ttgatgcctggcagttccct; pCR8_R1: cgaaccgaacaggcttatgt; sTALseqF: gtgcatgcttggcgtaac; TAL_R: agccacgaggtggtcgttgg
10. PCR enzymes: Taq and KOD DNA Polymerase.
11. Primers for amplification of the target region from genomic DNA.
12. Agarose mix for the resolution of small DNA fragments (*see* **Note 1**).
13. Transfection-grade plasmid DNA purification kit; plasmid miniprep kit.
14. DNAzol or other genome extraction reagents.
15. T7 In Vitro Transcription Kit (*see* **Note 2**).
16. Injection buffer: 0.5 mM potassium phosphate buffer (pH 7.0), 5 mM KCl.
17. Sequencing reagents.

3 Methods

Gene targeting using TALENs involves the selection of a target DNA sequence in the genome, the assembly of TALEN DNA binding modules using the Golden Gate or Serial ligation method, in vitro synthesis of TALEN mRNA, microinjection of *B. mori* embryos (we usually test the TALEN efficiency in a pilot test first and follow up with germline mutagenesis), the detection of mutations, and construction of the mutant lines. Since the strategy for selecting candidate targeting positions and TALEN construct assembly has been discussed in earlier chapters, they are covered here only briefly. *See* **Notes 3–5** for different types of experiments.

3.1 TALEN Construction in pBlue-TAL

TALEN design has been discussed in Chapters 2–5 of this volume. The selection of a target site in the *B. mori Bm-re* gene is shown in Fig. 1.

1. Select a TALEN target surrounding a suitable natural restriction site in the DNA of the desired gene (preferably near the 5′ end within an ORF) (Fig. 1) (*see* **Notes 6** and 7).
2. TALEN construction involves custom assembly of a specific DNA binding domain consisting of an array of TAL repeats.

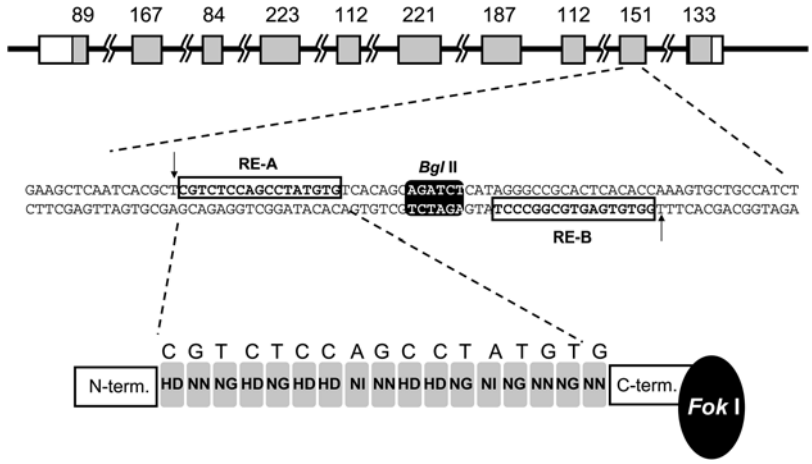


Fig. 1 Selection of target sequence in the *Bm-re* gene and in silico TALEN design. *Top line* with *open rectangles* and *numbers* above represents *Bm-re* exons and ORF lengths in bp. The TALEN target site was selected in the ninth exon in the *Bm-re* gene around a unique *Bgl* II restriction site (*highlighted* in *black*). The spacer region between the half sites is 16 bp long. The 5′ T residue flanking the TALEN target is denoted by an *arrow*. TALEN RVDs recognizing the RE-A target half site are shown below together with the nucleotide sequence they recognize (*bottom*)

Construct the primary TALEN arrays by combining the proper plasmids carrying RVD modules from the Golden Gate TALEN kit [35] in the first step restriction/cloning reaction using *Bsa*I and T4 DNA ligase as described in Chapter 6 (Fig. 2). Transform the reaction mix into competent *E. coli*, and select the colonies carrying assembled RVDs using the primers pCR8-F1 and pCR8-R1 (Fig. 3). Isolate the plasmids using a standard plasmid miniprep protocol (*see* **Note 8**).

3. Release the inserts from step 1 plasmids with the *Esp* 3I restriction enzyme, ligate the final assembly into the backbone vector pBlue-TAL (*see* **Note 9**), and transform competent *E. coli* (Fig. 2).
4. Select the colonies carrying assembled RVDs by PCR using the primers sTALseqF and TAL_R, and test the resulting plasmids for the presence of the assembled TALEN construct by restriction digestion and/or sequencing.

3.2 RNA Synthesis

1. Isolate plasmids carrying TALEN constructs and linearize 15 μg of plasmid DNA encoding a TALEN construct using fourfold excess of restriction endonuclease *Xba* I in 1 × buffer supplied with the enzyme.
2. Treat the reaction with proteinase K for 30 min at 50 °C.
3. Extract the digested DNA by adding an equal volume of phenol/chloroform, separate the phases, and precipitate the DNA with ethanol.

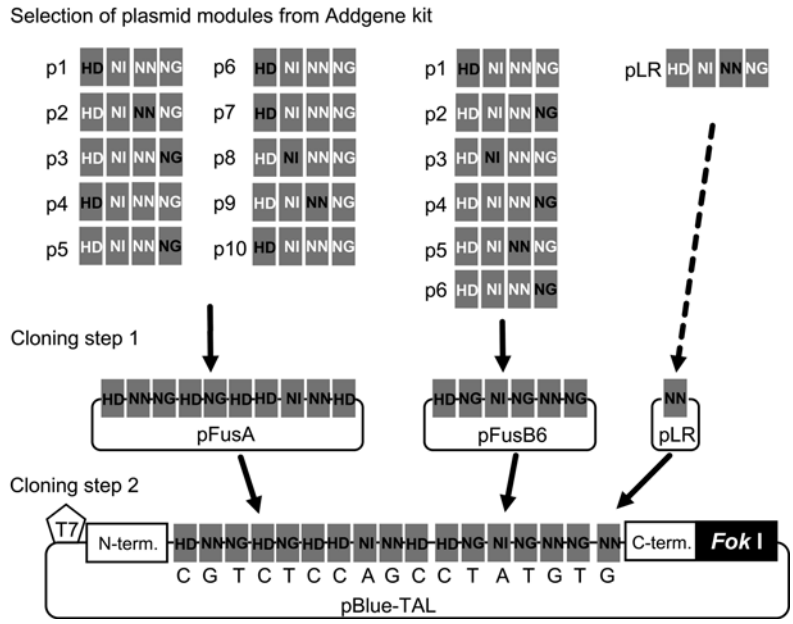


Fig. 2 Two step TALEN assembly using the Addgene Golden Gate cloning kit. The kit includes a pool of 40 repeat modules encoding RVDs and four pLR plasmid modules encoding the terminal half-repeat (the kit modules used for the cloning step 1 are *highlighted*). Plasmids pFusA and pFusB1-10 are used for array cloning in the two step cloning protocol. In the first cloning step the first ten modules will be released by *Bsa* I assembled into the pFusA plasmid, whereas the following four modules will be cloned into pFusB4. The cloning modules are then released by the *Esp* 3I restriction enzyme, assembled and ligated in the second step together with the terminal module into the backbone vector pBlue-TAL

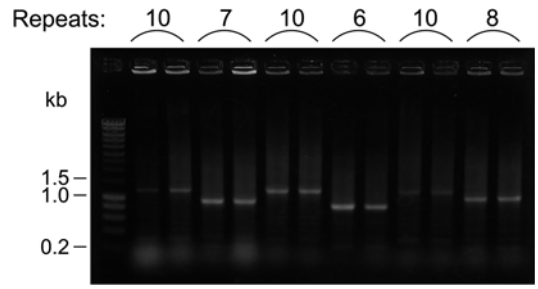


Fig. 3 Detection of assembled TALEN arrays. The TAL repeat assemblies in pFusA and pFusB plasmids are verified by PCR colony screening. The length of the PCR product containing ten TAL repeats is around 1.2 kb long: seven TAL repeats of 0.9 kb, six repeats of 0.8 kb and eight repeats of 1 kb

4. Spin down the DNA, wash the pellet three times with 70 % ethanol, air-dry, and resuspend in 11 μl of nuclease-free water supplied with the kit. Measure the absorbance at 260 nm using 0.5–1 μl of the DNA solution to estimate the DNA concentration.

5. Set up a 20 μl RNA synthesis reaction using T7 In Vitro Transcription Kit (*see* **Note 2**) by mixing:

   1 μg of plasmid DNA.

   Nuclease-free water to 6 μl.

   10 μl of 2× NTP/CAP.

   2 μl of 10× reaction buffer.

   2 μl enzyme mix.

6. Incubate the reaction at 37 °C for 2 h.
7. Increase the volume to 50 μl by nuclease-free water and precipitate the RNA by adding 30 μl 7.5 M LiCl and incubating for at least 30 min at −20 °C. The RNA is pelleted by centrifugation at 16,000 × *g* at 4 °C for 20 min, washed three times with ice cold 70 % ethanol, and then dried.

3.3 RNA Microinjection

1. Combine the newly emerged male and female moths and let them mate under moth covers for at least 4 h at room temperature (Fig. 4). Transfer the mated moths to 5 °C for 1–2 days, separate the female moths, followed by incubation at 25 °C on starch glue coated paper and let them lay eggs for 2 h separately under moth covers. Harvest the eggs from the paper by immersing in water.
2. Arrange the freshly laid eggs on a glass slide with the dorsal side pointing to the right. Fix them with cyanoacrylic glue.

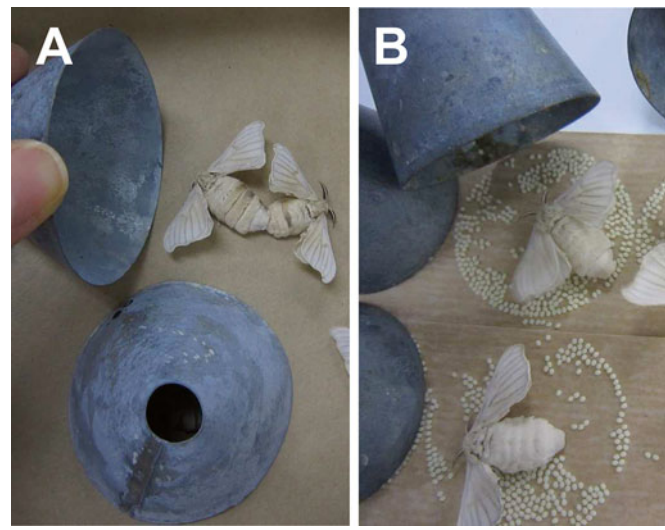


Fig. 4 Preparation of eggs for microinjection. A pair of newly emerged male and female moths are placed under a moth cover (**a**), allowed to mate at room temperature for 4 h and stored at 5 °C for 1 or 2 days. Female moths start to lay eggs immediately after transfer to room temperature. The eggs are harvested 2 h later (**b**)

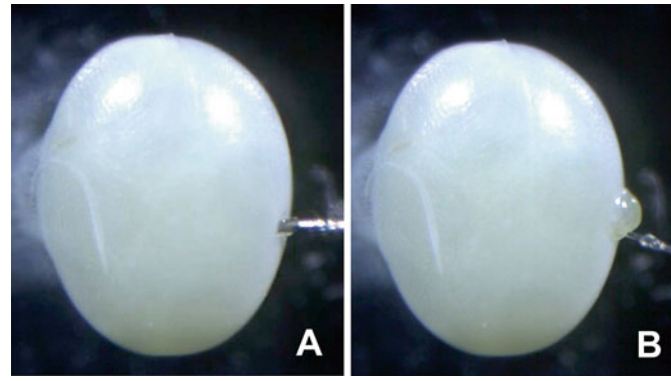


Fig. 5 RNA microinjection into *B. mori* eggs. A silkworm egg has the shape of the letter D. We positioned the convex dorsal part of the egg to the right. (**a**) In the first step a small hole is created in the egg by a tungsten needle. (**b**) A glass capillary is inserted into the hole and the RNA solution is injected into the egg (a tiny droplet of cytoplasm exudes from the egg). The eggs are microinjected 2–6 h after oviposition

3. Solubilize the in vitro prepared RNA (from Subheading 3.2) encoding the appropriate TALEN pair and adjust the concentration to 0.2–1.0 μg/μl with injection buffer.
4. Make a small hole in the egg chorion on the dorsal side using a tungsten needle (Fig. 5). Insert the tip of a glass capillary in the hole and inject 1–5 nl of RNA solution.
5. Seal the hole with cyanoacrylic glue and sterilize the egg using formaldehyde vapors for 5 min.
6. Incubate the eggs at 25 °C in humidified atmosphere until hatching (*see* **Note 10**).

3.4 TALEN Activity Assays

It is important to verify the function of newly designed TALENs (*see* **Note 11**). There are two variants of this method, differing in **steps 4** and **5** below.

1. Use 30–50 eggs microinjected as described in Subheading 3.3 and incubate for 3–4 days at 25 °C.
2. Pool the eggs together and extract the genomic DNA.
3. Amplify a targeted DNA region by PCR.
4. To obtain the first basic information on the TALEN efficiency, perform a restriction reaction directly on the PCR products and roughly estimate the TALEN efficiency by comparing the density of the digested and undigested bands.
5. Alternatively, ligate the mix of PCR products from **step 3** above into a TA plasmid vector, transform competent *E. coli* and confirm positive clones by colony PCR. Use restriction digestion or DNA sequencing of the target DNA to obtain a more reliable quantitative estimate of TALEN activity.

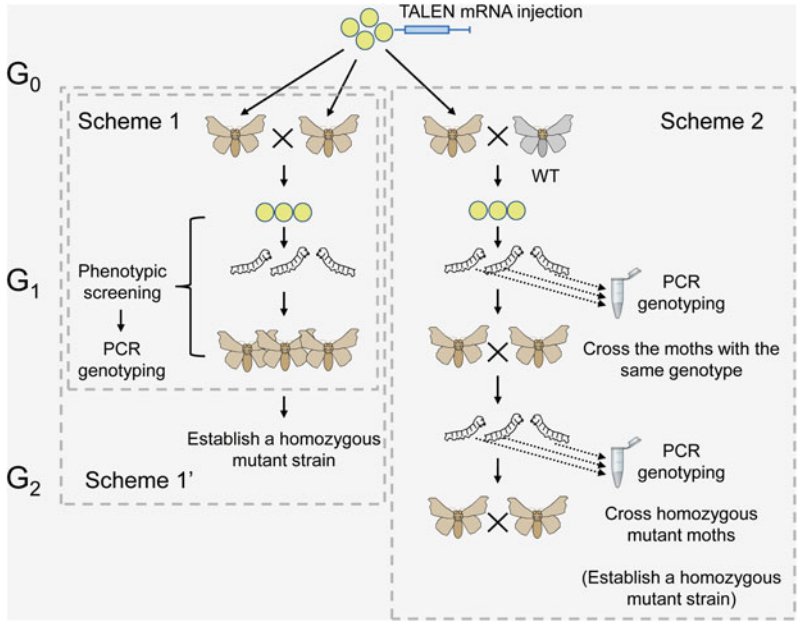


Fig. 6 Crossing schemes used for the detection of *B. mori* NHEJ mutants. Scheme 1 (*left*) contains a cross between the G_0 moths; it often produces heteroallelic mutants and is usually used for fast mutant detection and producing new alleles of genes with known phenotypes. Scheme 2 (*right*) involves backcrossing of G_0 moths with WT and produces homoallelic mutants

3.5 Germline Mutagenesis Protocols

There are two variants of crossing schemes that we use for TALEN knockout experiments (Fig. 6). They differ mainly in the first step, where we either cross the G_0 individuals with their siblings (scheme 1) or backcross them with WT (scheme 2) (*see* **Notes 12–14**).

3.5.1 Crossing Scheme 1

1. Let the injected G_0 moths mate with each other to obtain fertile eggs.
2. Incubate the eggs at 25 °C in a humidified atmosphere separately for each brood. Use 25–50 eggs of each brood for the PCR pilot test using the same procedure as the TALEN activity assay (*see* Subheading 3.4) (Fig. 7). For further work use several broods which show high frequency of mutations.
3. Let the remaining eggs from the selected broods hatch, rear the G_1 silkworms and examine their phenotypes when they reach suitable developmental stage.
4. Perform (if necessary, noninvasive) genotyping. Extract genomic DNA from each G_1 individual, amplify the target region using PCR, and directly sequence the PCR fragment using the PCR primers designed to amplify the target region.
5. Analyze the phenotype and the target sequences for each G_1 individual.

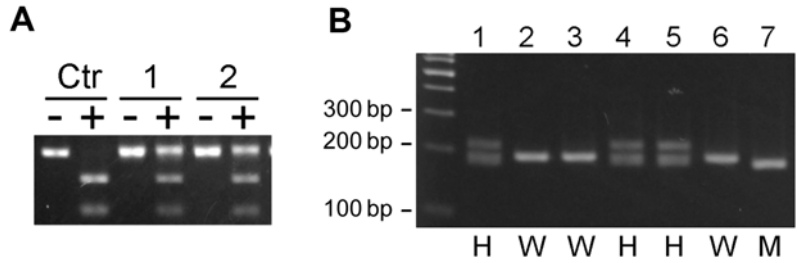


Fig. 7 Examples of *B. mori* genotyping. (**a**) Sampling of G_1 broods by embryo analysis. Genomic DNA was extracted from 20 to 25 G_1 embryos obtained by the crossing of G_0 and WT moths. The PCR products without (–) and with (+) restriction digestion were analyzed by agarose gel (all embryos in the control, Ctr, contained the restriction site, whereas in samples 1 and 2 the uncut DNA suggested the presence of mutations in the mix.) (**b**) Analysis of G_2 progeny individuals. The DNA was isolated from one adult middle leg. The target region was amplified by PCR and the difference between mutants, WT and heterozygous individuals was detected by size fractionation on an agarose gel. *Lanes 1*, *4*, and *5* are heterozygotes, *2*, *3*, and *6* are WT and *lane 7* is a mutant. Product sizes: 191 bp (WT), 184 bp (mutant)

6. If construction of a mutant strain is required, let the G_1 moths mate with each other or with WT moths and repeat genotyping and sibling crosses to obtain the mutant line of a desired genotype (Fig. 6, scheme 1′) (*see* **Note 15**).

3.5.2 Crossing Scheme 2

1. Let the injected G_0 moths mate with WT moths to obtain fertile eggs.
2. Incubate the eggs at 25 °C in a humidified atmosphere. Use 25–50 eggs of each brood for a PCR pilot test (the same procedure as for the TALEN activity assay—*see* Subheading 3.4). Select several broods which show a high frequency of mutations.
3. Use the remaining eggs from the selected broods for silkworm rearing to obtain G_1 moths. Extract DNA from each G_1 individual using a droplet of hemolymph from the final instar larva or a leg from a newly emerged moth.
4. Amplify the target region by PCR, and analyze the products using agarose gel electrophoresis. If possible, directly sequence the mutant alleles.
5. Choose male and female moths (heterozygotes) with the same desirable mutant allele and let them mate to obtain fertile eggs. Keep the homozygotes as a mutant line.
6. If homozygous mutant moths do not survive, keep the heterozygotes as a mutant line. Some mutations are not completely lethal and rare survivors can be found in larger scale crosses (*see* **Notes 16–18**).

3.6 Homology-Directed Repair

1. DNA repair process may be influenced by the co-injection of donor DNA together with TALEN RNA (*see* **Note 19**). Single stranded oligonucleotides are able to induce specific small mutations within a locus. Design the oligonucleotide donor (not more than 100 nucleotides long) with sequences homologous to the two 20–30-bp regions across the double strand break and introduce a novel sequence (for example, restriction site and/or stop codons) between them. It is desirable that at least one side is designed to be homologous to the adjacent sequence of the double strand break.
2. Follow the protocols in Subheadings 3.1 and 3.2.
3. Mix poly(A)-tailed mRNAs for the appropriate TALEN pair at concentration of 0.2–0.4 μg/μl together with the donor oligonucleotide at a concentration of 0.2–0.4 μg/μl in injection buffer and keep the solution on ice.
4. Inject the nucleic acid solution into the eggs, seal the small hole in the egg with cyanoacrylic glue, and surface-sterilize with formaldehyde vapors for 5 min. Incubate the embryos at 25 °C in a humidified chamber following the protocol in Subheading 3.3.
5. Perform the G_0 and G_1 sampling as described in Subheading 3.5.1. Test for insertion of the novel restriction site (*see* **Notes 20** and **21**).

4 Notes

1. We receive the best resolution of small PCR fragments on agarose gels containing mix of 1 % LO3 (Takara-bio) and 3 % NuSieve GTG (Lonza).
2. We use mMESSAGE mMACHINE T7 kit (Ambion, Carlsbad, CA, USA). Alternatively, a more expensive mMESSAGE mMACHINE T7 Ultra kit can be used for RNA synthesis. It includes a novel cap analog which functions twice as well as the one in the previous kit, but we did not observe a significant difference in the efficiency of mutagenesis.
3. In most cases TALENs are used for functional analysis by knocking out a gene of interest and observing the resultant mutant phenotype. Mutagenesis is usually performed with a single TALEN pair and will produce a small deletion or insertion causing a frameshift and/or insertion of different amino acids or premature truncation of the encoded protein. For such a simple NHEJ mutagenesis experiment it is usually enough to microinject 200 eggs. It is theoretically possible to obtain several mutations from just a few fertile moths but we prefer to inject larger groups of eggs to increase the chance of success.

4. Some experiments require generation of larger deletions, such as the removal of an entire gene or a gene cluster. In such cases, the desirable mutants are detected by the presence of PCR products of the predicted size. As shown by Ma et al. [29], it is possible to use two pairs of TALENs to delete DNA fragments of 800 bp or longer. Such large mutations generated with two TALEN pairs occurred with a similar frequency to single TALEN pair deletions performed in parallel. Experiments with pairs of engineered nucleases simultaneously targeting two desired sites were performed in other model systems and showed that it is possible to delete much larger fragments (e.g., deletion of genomic segments of up to 15 Mb between two distal target sites were observed in cultured human or other mammalian cells) [36, 37].
5. In some cases we need precise modifications of the encoded proteins, for example a change in just one or two amino acids (e.g., in studies of the enzyme active center) or to make exact point mutations. It is possible to combine the TALEN RNA microinjection with a co-injected DNA donor and to perform homologous recombination with an oligonucleotide as discussed in Subheading 3.6. The efficiency of DNA insertions using oligonucleotides is usually lower than the frequency of mutations in simple NHEJ, and the detection of mutations is made easier by using a mutant-specific primer (*see* Subheading 3.6).
6. We prefer TALENs with 15–21 RVDs and a spacer length of around 15 nucleotides. The kit allows combining maximally ten RVDs in one cloning step (plasmid pFusA). The terminal RVD (which is shorter) has to be cloned separately (plasmid pLR) and the remaining 4–10 RVDs are subcloned into an appropriate pFusB4-10 vector.
7. Longer TALENs containing more than 21 RVDs are constructed with a modified protocol using plasmids pFusA30A and pFusA30B, which are also components of the Golden Gate Cloning Kit (Addgene).
8. If we have problems with false positives in Golden Gate assembly, we use fresh chemicals—DNA ligase buffer, high quality ligase (NEB 400 units/μl or equivalent), and fresh plasmids isolated using standard plasmid midiprep protocol.
9. The pBlue-TAL expression vector (GenBank: KF724948.1) contains 136 N-terminal and 63 C-terminal amino acid residues from the TAL effector. It also contains a T7 promoter for in vitro transcription and encodes the *Xba* I-ending poly(A) tail, which is sufficient for RNA stability [31].
10. If problems occur with embryo survival, check buffer sterility and compare the microinjection results between RNA and negative control physiological solutions.

11. Since about 10 % of the newly designed TALENs might have a very low or no activity and since it is relatively easy to obtain a sufficient number of *B. mori* eggs, we recommend pre-examining TALEN function by one of the pilot tests performed on the embryonic cells (mix of somatic and germline cells). Since we select our TALEN targets so they will surround naturally occurring cleavage sites for some restriction endonucleases, we examine the modification of the target area in the embryo DNA by the loss of a restriction site using PCR and restriction digestion.
12. By using crossing scheme 1 (Fig. 6) in the germline mutagenesis protocol, we can obtain homozygous mutants in G_1 and use their phenotypes for screening, but they may represent heteroallelic combinations. The other scheme uses DNA-based screening and produces homoallelic combinations of mutations but takes longer.
13. Some of the genes responsible for distinct phenotypes such as pigmentation of eggs, adult eyes and larval integument, cocoon color, etc. are mapped to a limited region of a chromosome and the sequence data for identifying candidate genes can be obtained from *B. mori* genome databases. TALEN mutagenesis is instrumental in determining the exact gene responsible for a particular phenotype. In such experiments we recommend to combine a TALEN knockout experiment with a G_0 sibling cross (scheme 1), which will provide a quick assessment. Since the knockout efficiency in the silkworm is often more than 50 %, a significant number of homozygous mutants can be obtained in G_1.
14. The *B. mori* genome sequencing project identified more than 14,000 genes, many of which have orthologs in other organisms. A number of interesting candidate genes possibly involved in key physiological and developmental processes can be identified and their role examined using the reverse genetic approach. When we try to knockout one of such genes, we usually use scheme 2 (Fig. 6) to establish mutant lines first, since the gene functions are not known and lethality might occur in homozygous condition.
15. By using crossing scheme 1 we usually analyze more allelic combinations. This may be informative if we study a gene with unknown phenotype, since we can detect more phenotypic variants (we can even mix silkworms from different broods to increase the variability). When using crossing scheme 2 we usually analyze only a limited number of mutants.
16. While both crossing schemes are useful to observe lethal phenotypes, we have to be careful when trying to establish mutant lines using crossing scheme 1. If the efficiency of a TALEN is nearly 100 %, we may lose most of the mutant alleles in G_1 because of lethality.

17. If we suspect that mutations cause decreased viability or lethality we may consider keeping part of the mutants in a diapause state by crossing some of the G_0 moths with a diapausing strain. We can do the experiments on a portion/subset of the genotyped individuals while others might be kept for an extended period of time as a backup.
18. If we expect that mutations would cause lethality we may also consider generating a *Bombyx* strain overexpressing a modified copy of the gene to be mutagenized (lacking the target site). We can mutagenize the gene in transgenic strain and segregate the transgene or block its expression in later generations.
19. One of the major goals of gene targeting is gene replacement by homologous recombination. This would allow to knock-in a marker gene (e.g., GFP, DsRed) inside a mutation, so that an induced mutation can be identified by the presence of a recombinant marker phenotype. The availability of knock-in targets is crucial for the maintenance of lethal mutations. We performed the first experiments with the *BmBLOS2* gene and donor DNA carrying a marker gene flanked by 2 kb of homologous region. We observed fluorescent protein mosaics in 9.4–50 % of the G_0 larvae and a single individual among the 11770 G_1 larvae carrying the knock-in GFP allele [9].
20. We tested co-injection of oligonucleotides (44–86 nucleotides long) with TALEN RNA and observed a relatively high efficiency of homology-directed repair (12–79 % of G0 moths yielded HDR progeny). The addition of a donor oligonucleotide did not affect the total number of mutations.
21. We usually use crossing scheme 1 for knock-in experiments because the frequency of the precise mutation is much lower than that of usual NHEJ mutagenesis. Further, the desired mutants are easily detected by PCR using specific primers or by restriction digestion of PCR products of the target region.

Acknowledgements

This research was supported by KAKENHI Grant Number 23580083 from Japan Society for the Promotion of Science and grant P305/10/2406 from the Grant Agency of the Czech Republic. We acknowledge the use of research infrastructure that has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 316304. We would like to thank Dr. František Sehnal for critical reading of the manuscript.

References

1. Marsh RE, Corey RB, Pauling L (1955) An investigation of the structure of silk fibroin. Biochim Biophys Acta 16(1):1–34
2. Butenandt A, Karlson P (1954) Über die Isolierung eines Metamorphose-Hormone der Insekten in Kristallisierten Form. Z Naturf 9B:389–391
3. Suzuki Y, Brown DD (1972) Isolation and identification of the messenger RNA for silk fibroin from Bombyx mori. J Mol Biol 63(3): 409–429
4. Suzuki Y, Tsuda M, Takiya S, Hirose S, Suzuki E, Kameda M, Ninaki O (1986) Tissue-specific transcription enhancement of the fibroin gene characterized by cell-free systems. Proc Natl Acad Sci U S A 83(24):9522–9526
5. Suzuki Y (1977) Differentiation of the silk gland. A model system for the study of differential gene action. Results Probl Cell Differ 8:1–44
6. Kafatos FC (1972) The cocoonase zymogen cells of silk moths: a model of terminal cell differentiation for specific protein synthesis. Curr Top Dev Biol 7:125–191
7. Kafatos FC, Tzertzinis G, Spoerel NA, Nguyen HT (1995) Chorion genes: an overview of their structure, function and transcriptional regulation. In: Wilkins AS, Goldsmith MR (eds) Molecular model system in the lepidoptera. Cambridge University Press, Cambridge, pp 181–216
8. International Silkworm Genome Consortium (2008) The genome of a lepidopteran model insect, the silkworm Bombyx mori. Insect Biochem Mol Biol 38(12):1036–1045
9. Daimon T, Kiuchi T, Takasu Y (2014) Recent progress in genome engineering techniques in the silkworm, Bombyx mori. Dev Growth Differ 56(1):14–25
10. Xia Q, Li S, Feng Q (2014) Advances in silkworm studies accelerated by the genome sequencing of bombyx mori *. Annu Rev Entomol 59:513–536
11. Banno Y, Shimada T, Kajiura Z, Sezutsu H (2010) The silkworm-an attractive BioResource supplied by Japan. Exp Anim 59(2):139–146
12. Yasukochi Y (1998) A dense genetic map of the silkworm, Bombyx mori, covering all chromosomes based on 1018 molecular markers. Genetics 150(4):1513–1525
13. Miao XX, Xub SJ, Li MH, Li MW, Huang JH, Dai FY, Marino SW, Mills DR, Zeng P, Mita K, Jia SH, Zhang Y, Liu WB, Xiang H, Guo QH, Xu AY, Kong XY, Lin HX, Shi YZ, Lu G, Zhang X, Huang W, Yasukochi Y, Sugasaki T, Shimada T, Nagaraju J, Xiang ZH, Wang SY, Goldsmith MR, Lu C, Zhao GP, Huang YP (2005) Simple sequence repeat-based consensus linkage map of Bombyx mori. Proc Natl Acad Sci U S A 102(45):16303–16308
14. Yamamoto K, Narukawa J, Kadono-Okuda K, Nohata J, Sasanuma M, Suetsugu Y, Banno Y, Fujii H, Goldsmith MR, Mita K (2006) Construction of a single nucleotide polymorphism linkage map for the silkworm, Bombyx mori, based on bacterial artificial chromosome end sequences. Genetics 173(1):151–161
15. Yamamoto K, Nohata J, Kadono-Okuda K, Narukawa J, Sasanuma M, Sasanuma S, Minami H, Shimomura M, Suetsugu Y, Banno Y, Osoegawa K, de Jong PJ, Goldsmith MR, Mita K (2008) A BAC-based integrated linkage map of the silkworm Bombyx mori. Genome Biol 9(1):R21
16. Mita K, Morimyo M, Okano K, Koike Y, Nohata J, Kawasaki H, Kadono-Okuda K, Yamamoto K, Suzuki MG, Shimada T, Goldsmith MR, Maeda S (2003) The construction of an EST database for Bombyx mori and its application. Proc Natl Acad Sci U S A 100(24):14121–14126
17. Suetsugu Y, Futahashi R, Kanamori H, Kadono-Okuda K, Sasanuma S, Narukawa J, Ajimura M, Jouraku A, Namiki N, Shimomura M, Sezutsu H, Osanai-Futahashi M, Suzuki MG, Daimon T, Shinoda T, Taniai K, Asaoka K, Niwa R, Kawaoka S, Katsuma S, Tamura T, Noda H, Kasahara M, Sugano S, Suzuki Y, Fujiwara H, Kataoka H, Arunkumar KP, Tomar A, Nagaraju J, Goldsmith MR, Feng Q, Xia Q, Yamamoto K, Shimada T, Mita K (2013) Large scale full-length cDNA sequencing reveals a unique genomic landscape in a lepidopteran model insect, Bombyx mori. G3 (Bethesda) 3(9):1481–1492
18. Suetsugu Y, Minami H, Shimomura M, Sasanuma S, Narukawa J, Mita K, Yamamoto K (2007) End-sequencing and characterization of silkworm (Bombyx mori) bacterial artificial chromosome libraries. BMC Genomics 8:314
19. Atsumi S, Miyamoto K, Yamamoto K, Narukawa J, Kawai S, Sezutsu H, Kobayashi I, Uchino K, Tamura T, Mita K, Kadono-Okuda K, Wada S, Kanda K, Goldsmith MR, Noda H (2012) Single amino acid mutation in an ATP-binding cassette transporter gene causes resistance to Bt toxin Cry1Ab in the silkworm, Bombyx mori. Proc Natl Acad Sci U S A 109(25):E1591–E1598

20. Futahashi R, Sato J, Meng Y, Okamoto S, Daimon T, Yamamoto K, Suetsugu Y, Narukawa J, Takahashi H, Banno Y, Katsuma S, Shimada T, Mita K, Fujiwara H (2008) yellow and ebony are the responsible genes for the larval color mutants of the silkworm Bombyx mori. Genetics 180(4):1995–2005
21. Ito K, Kidokoro K, Sezutsu H, Nohata J, Yamamoto K, Kobayashi I, Uchino K, Kalyebi A, Eguchi R, Hara W, Tamura T, Katsuma S, Shimada T, Mita K, Kadono-Okuda K (2008) Deletion of a gene encoding an amino acid transporter in the midgut membrane causes resistance to a Bombyx parvo-like virus. Proc Natl Acad Sci U S A 105(21):7523–7527
22. Liu C, Yamamoto K, Cheng TC, Kadono-Okuda K, Narukawa J, Liu SP, Han Y, Futahashi R, Kidokoro K, Noda H, Kobayashi I, Tamura T, Ohnuma A, Banno Y, Dai FY, Xiang ZH, Goldsmith MR, Mita K, Xia QY (2010) Repression of tyrosine hydroxylase is responsible for the sex-linked chocolate mutation of the silkworm, Bombyx mori. Proc Natl Acad Sci U S A 107(29):12980–12985
23. Qiao L, Xiong G, Wang RX, He SZ, Chen J, Tong XL, Hu H, Li CL, Gai TT, Xin YQ, Liu XF, Chen B, Xiang ZH, Lu C, Dai FY (2014) Mutation of a cuticular protein, BmorCPR2, alters larval body shape and adaptability in silkworm Bombyx mori. Genetics 196:1103. doi:10.1534/genetics.113.158766
24. Tamura T, Thibert C, Royer C, Kanda T, Abraham E, Kamba M, Komoto N, Thomas JL, Mauchamp B, Chavancy G, Shirk P, Fraser M, Prudhomme JC, Couble P (2000) Germline transformation of the silkworm Bombyx mori L. using a piggyBac transposon-derived vector. Nat Biotechnol 18(1):81–84
25. Uchino K, Imamura M, Shimizu K, Kanda T, Tamura T (2007) Germ line transformation of the silkworm, Bombyx mori, using the transposable element Minos. Mol Genet Genomics 277(3):213–220
26. Uchino K, Sezutsu H, Imamura M, Kobayashi I, Tatematsu K, Iizuka T, Yonemura N, Mita K, Tamura T (2008) Construction of a piggyBac-based enhancer trap system for the analysis of gene function in silkworm Bombyx mori. Insect Biochem Mol Biol 38(12):1165–1173
27. Imamura M, Nakai J, Inoue S, Quan GX, Kanda T, Tamura T (2003) Targeted gene expression using the GAL4/UAS system in the silkworm Bombyx mori. Genetics 165(3):1329–1340
28. Takasu Y, Kobayashi I, Beumer K, Uchino K, Sezutsu H, Sajwan S, Carroll D, Tamura T, Zurovec M (2010) Targeted mutagenesis in the silkworm Bombyx mori using zinc finger nuclease mRNA injection. Insect Biochem Mol Biol 40(10):759–765
29. Ma S, Zhang S, Wang F, Liu Y, Xu H, Liu C, Lin Y, Zhao P, Xia Q (2012) Highly efficient and specific genome editing in silkworm using custom TALENs. PLoS One 7(9):e45035
30. Sajwan S, Takasu Y, Tamura T, Uchino K, Sezutsu H, Zurovec M (2013) Efficient disruption of endogenous Bombyx gene by TAL effector nucleases. Insect Biochem Mol Biol 43(1):17–23
31. Takasu Y, Sajwan S, Daimon T, Osanai-Futahashi M, Uchino K, Sezutsu H, Tamura T, Zurovec M (2013) Efficient TALEN construction for Bombyx mori gene targeting. PLoS One 8(9):e73458
32. Takasu Y, Tamura T, Sajwan S, Kobayashi I, Zurovec M (2014) The use of TALENs for nonhomologous end joining mutagenesis in silkworm and fruitfly. Methods 69(1):46–57
33. Wang Y, Li Z, Xu J, Zeng B, Ling L, You L, Chen Y, Huang Y, Tan A (2013) The CRISPR/Cas system mediates efficient genome engineering in Bombyx mori. Cell Res 23(12):1414–1416
34. Kim C, Kim K, Park D, Seol Y, Hahn J, Park S, Kang P (2010) An integrated database for the enhanced identification of silkworm gene resources. Bioinformation 4(10):436–437
35. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39(12):e82
36. Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, Voytas DF, Long CR, Whitelaw CB, Fahrenkrug SC (2012) Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci U S A 109(43):17382–17387
37. Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, Kim S, Lee C, Jeong E, Chung E, Kim D, Lee MS, Go EM, Song HJ, Kim H, Cho N, Bang D, Kim JS (2013) A library of TAL effector nucleases spanning the human genome. Nat Biotechnol 31(3):251–258

Chapter 12

GeneKnockout by Targeted Mutagenesis in a Hemimetabolous Insect, the Two-Spotted Cricket *Gryllus bimaculatus*, using TALENs

Takahito Watanabe, Sumihare Noji, and Taro Mito

Abstract

Hemimetabolous, or incompletely metamorphosing, insects are phylogenetically basal. These insects include many deleterious species. The cricket, *Gryllus bimaculatus*, is an emerging model for hemimetabolous insects, based on the success of RNA interference (RNAi)-based gene-functional analyses and transgenic technology. Taking advantage of genome-editing technologies in this species would greatly promote functional genomics studies. Genome editing using transcription activator-like effector nucleases (TALENs) has proven to be an effective method for site-specific genome manipulation in various species. TALENs are artificial nucleases that are capable of inducing DNA double-strand breaks into specified target sequences. Here, we describe a protocol for TALEN-based gene knockout in *G. bimaculatus*, including a mutant selection scheme via mutation detection assays, for generating homozygous knockout organisms.

Key words Targeted mutagenesis, Cricket, TALEN, Homozygous knockout, Hemimetabolous insect

1 Introduction

The two-spotted cricket, *Gryllus bimaculatus* (Orthoptera: Gryllidae), is one of the most abundant cricket species. *G. bimaculatus* inhabits tropical and subtropical regions of Asia, Africa, and Europe. It can be easily bred in the laboratory and has been widely used to study insect physiology and neurobiology [1]. This species was recently established as a model system for studies on molecular mechanisms of development and regeneration. To analyze gene functions in *G. bimaculatus*, the RNA interference (RNAi) technique can be used. For genes involved in embryonic development, double-stranded RNA (dsRNA) is injected into the body cavity of adult female crickets [2]. RNAi effects are observed through the subsequently laid eggs. To study regeneration mechanisms of the leg, dsRNA can be injected into the body cavity of

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_12,

the nymph [3], followed by amputation of the metathoracic tibia and observation of the RNAi effects during regeneration of the lost part of the leg.

We recently established a technique to generate transgenic crickets with the *piggyBac* transposase [4, 5]. Using *eGFP-expressing* transgenic lines, we performed live imaging analysis of fluorescently labeled embryonic cells and nuclei. Blastoderm cells were found to move dynamically, retaining their positional information to form the posteriorly localized germ anlage [4]. The generation of transgenic crickets would also be useful for analyzing the functions of genes and *cis*-regulatory elements.

Although the above RNAi and transgenic systems are effective for analyzing gene function, each system has some shortcomings. For example, gene functions can be easily and efficiently inhibited by RNAi, but they cannot be inhibited completely because some RNA can remain un-degraded. In the case of *piggyBac*-based transgenesis, we cannot control both the copy number and the genomic locus of a transgene because, by nature, transgenes are randomly integrated into a genome. Thus, a technique is needed for modifying the cricket genome at a specific site, in order to conduct further sophisticated gene-functional analyses.

Transcription activator-like (TAL) effector nucleases (TALENs) can be used to induce targeted DNA double-strand breaks (DSBs) into specific regions of the genome [6]. TALENs consist of an engineered array of TAL effector repeats fused to the *FokI* cleavage domain. DNA-binding domains of TALENs have tandem repeat units of 33–35 amino acids in length. The nucleotide binding preference is determined by two adjacent amino acids, known as the repeat variable di-residue [7]. *FokI* requires dimerization to cleave DNA. Dimerization is achieved through the binding of two TALEN molecules to contiguous target sequences in each DNA strand, separated by a spacer sequence. Induction of DSBs by TALENs activates the DNA damage response [8]. A DSB can be repaired by nonhomologous end joining, in which short insertions or deletions are generated at the cleavage site [9], or by homologous recombination with a DNA template, in which gene knock-in results in a perfect repair or a sequence replacement (if a modified template is used) [10].

TALEN-mediated mutations have been induced in several organisms, including fruit flies [11], zebrafish [12], rats [13], and mice [14]. In these animals, both somatic and germline mutations were shown [11–14]. However, obtaining constitutional homozygous mutants is laborious in cases without specific genetic markers to breed with, or without morphologically detectable phenotypes to identify individuals with a mutated allele.

To facilitate the production of homozygous knockout animals via genome editing, we examined an effective strategy for selecting animals with a mutant allele, independent of their genetic tractability or their phenotypic characteristics. We reported the first research in hemimetabolous insects to show the effectiveness of TALENs for generating

knockout animals [15]. We designed TALENs targeted to the *G. bimaculatus laccase2* (*Gb'lac2*) locus. This gene was selected based on previous studies, in which RNAi targeting of *Gb'lac2* in the nymphal stages of the cricket resulted in loss of cuticle tanning after molting.

After in vitro transcription, TALEN mRNAs were injected into cricket eggs just after fertilization. Approximately 17 % of the fifth instar G_0 larvae generated were somatic mutants with a mosaic cuticle pattern, with or without tanning. To generate knockout crickets, we designed a two-step screening method based on mutant detection assays using SURVEYOR nuclease. Mutations transmitted from the founder crickets were detected in ~17 % of G_0 adults in the first round of screening. Heterozygous offspring were selected with mutation detection assays in the second round of screening. Subsequent sibling crosses created homozygous knockout crickets exhibiting the cuticle tanning defect phenotype (Fig. 1).

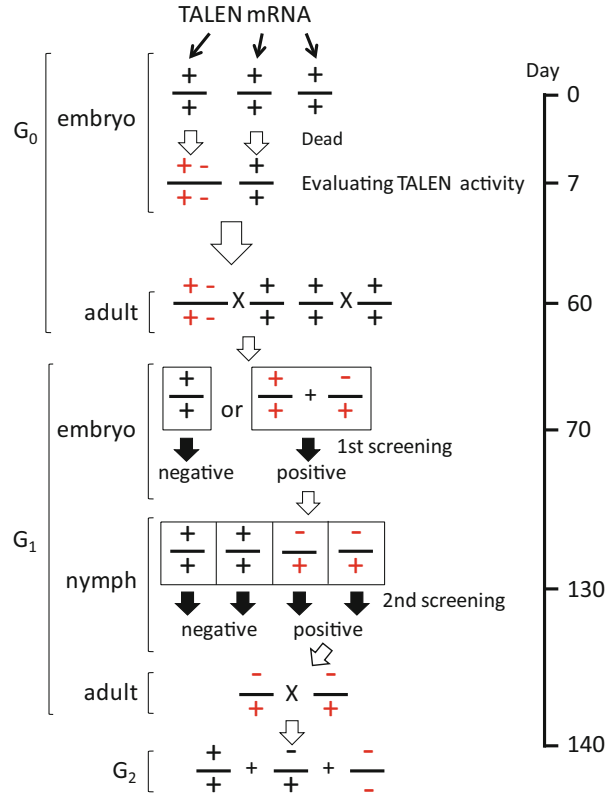


Fig. 1 Illustration of the scheme to isolate homozygous mutations in a gene of interest. TALEN activity is evaluated by the SURVEYOR nuclease assay in ten mutagenized G_0 embryos (day 7). Mutagenized G_0 adults are crossed to wild-type adults (day 60). G_1 embryos are checked for heterozygous mutations in a first round of screening with the SURVEYOR nuclease assay (day 70). Positive G_1 individuals are developed into final instar nymphs and subjected to a second round of screening with the SURVEYOR nuclease assay (day 130). Positive G_1 adults are crossed by strain to obtain homozygous mutants in the G_2 generation (day 140). Reprinted from Ref. 20 with permission from Elsevier

In this chapter, we provide protocols for generating homozygous knockouts of an endogenous gene in *G. bimaculatus* via TALEN-based targeted mutagenesis. This approach can be used to generate knockout animals, independent of the specific genotype or phenotype. In principle, these protocols can be adapted to any organism for which TALEN mRNA can be injected into eggs and the organism reared through a few generations.

2 Materials

2.1 Breeding Crickets

1. Artificial fish food (e.g., TetraFin and TetraGold (Tetra)).
2. Paper towels for dishes to collect eggs and for water tubes.
3. Insect cages.

2.2 Injection into Cricket Eggs

2.2.1 Injection

1. Glass capillary (ID: 0.6 mm; OD: 1 mm).
2. Micropipette puller (P-1000, Sutter Instruments).
3. Glass slide.
4. Double-sided tape.
5. Inverted microscope, 10× eyepiece lens and 10× objective lens (Fig. 2a).
6. 50-mL glass syringe (Fig. 2a).
7. Micromanipulator (Leica Micromanipulator M) (Fig. 2a).
8. Mineral oil.

2.2.2 Preparation of Injection Solution

1. Plasmid purification kit.
2. TE buffer: 10 mM Tris–HCl, pH 8.0, 1 mM EDTA.
3. Phenol:chloroform:isoamyl alcohol (25:24:1, v/v).
4. Restriction enzyme.
5. mMessage mMachine T7 Ultra Kit (AM1345, Ambion).

2.3 Analysis of Gene Disruption in the Cricket Genome

1. Genomic DNA purification kit.
2. Taq DNA polymerase.
3. Primer pairs (10 μM).
4. Thermal cycler.
5. SURVEYOR Mutation Detection Kit (706020, Integrated DNA Technologies).
6. Agar for electrophoresis.

3 Methods

3.1 Maintenance of Wild-Type Crickets

Rear nymphs and adults at 29 °C and 50 % humidity under a 10-h light:14-h dark photoperiod. Feed nymphs and adults with artificial fish food (*see* **Note 1**). Under these conditions, first to third

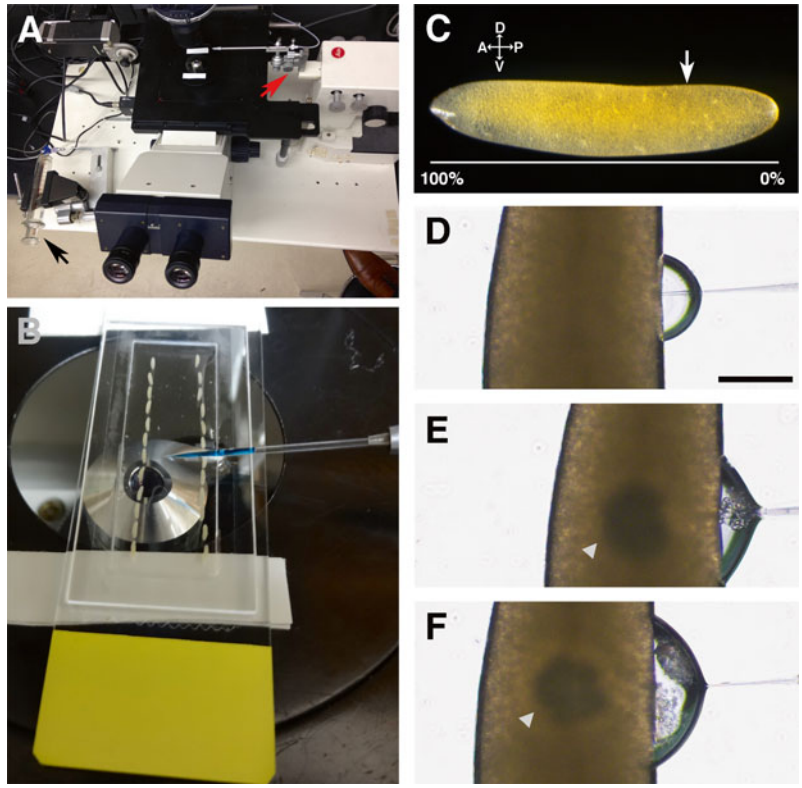


Fig. 2 Injection system for cricket eggs. (**a**) Injection apparatus for cricket eggs. This step uses an inverted microscope, a 50-mL glass syringe (*black arrow*), and a micromanipulator (*red arrow*). (**b**) Egg-lined slide and injection needle with solution (*blue colored*; visualized with Fast Green FCF for this figure) are shown. The line of eggs is as straight as possible in the well of the plastic frame on the double-sided tape. (**c**) Picture of a cricket egg. Anterior end is left, dorsal side is up. The injection needle is inserted from the dorsal side of the egg near the posterior end (20–25 % of the egg length from the posterior end; *arrow*). (**d**) Picture of the injection needle stuck vertically in the egg surface. Scare bar: 0.2 mm. (**e**) Solution is injected into the middle region of the egg. Injected solution can be observed as a small bubble, which will be as large as a third part of the diameter in the egg; *arrowhead*. (**f**) When the injection needle is removed from the egg, the injected solution remains (*arrowhead*). Adapted from Ref. 20 with permission from Elsevier

instar nymphs will molt every 2 or 3 days, fourth to sixth instar nymphs will molt every 4 or 5 days, seventh and eighth instar nymphs will molt every week, and adults will emerge after the eighth molt. The generation time of the cricket is about 2 months.

Maintain first to third instar nymphs in plastic insect cases, separated by hatching day (*see* **Note 2**). Maintain fourth to eighth instar nymphs in plastic containers (e.g., cloth cases) separated by week. Maintain adults in plastic containers, separated by the week in which the cricket reaches adulthood. Each plastic insect case should contain one 50-mL plastic tube plugged with a paper towel

for supplying water, as well as several crumpled paper towels as a shelter. Each plastic container for rearing nymphs and adults should contain a pile of moist and folded paper towels (~1.5 cm in height) in a plastic dish for supplying water (*see* **Note 3**) and for adult females to lay eggs during the dark period.

Replace the paper towels containing laid eggs with new towels each day, collecting the freshly laid eggs with forceps. Place the collected eggs separately on moist paper towels. Incubate the eggs under the same conditions as used for rearing adult crickets. After the first instar nymphs hatch from the eggs (~13th day of incubation), transfer the nymphs to the rearing cases.

3.2 Preparation of TALEN Left and Right (TALEN-L/R) mRNAs

There are several choices available for obtaining plasmids for the in vitro transcription of TALEN-L/R mRNAs. We previously used the Custom TALEN™ Service offered by Cellectis Bioresearch for these plasmids [15]. Currently, Thermo Fisher Scientific (Life Technologies) and Transposagen Biopharmaceuticals also provide TALEN plasmid construction services. For plasmid construction in the laboratory, Golden Gate Assembly [16], Platinum Gate Assembly (modified Golden Gate Assembly) [17], and FLASH Assembly [18] are available (*see* **Note 4**). In *Gryllus*, TALEN pairs constructed by Transposagen Biopharmaceuticals and Platinum Gate Assembly induced mutations into the target genomic locus with high efficiency (Watanabe et al., unpublished data).

1. Purify plasmids for the TALEN-L/R mRNAs using a plasmid purification kit.
2. Suspend plasmids in 50 μL of TE buffer.
3. Linearize the plasmids for mRNA synthesis with a restriction enzyme (*Sma*I, in the case of our plasmids [17]).
4. Extract linearized DNAs using phenol:chloroform:isoamyl alcohol.
5. Perform ethanol precipitation, and resuspend DNA in TE buffer at 0.5 μg/μL.
6. Synthesize TALEN-L/R mRNAs with the mMessage mMachine T7 Ultra Kit, according to the manufacturer's protocol.
7. Resuspend RNA pellets in nuclease-free water, and adjust to a final concentration of 2 μg/μL.
8. Aliquot the solution into a 1.5-mL tube, and store at −80 °C (*see* **Note 5**).

3.3 Injection of TALEN mRNAs into Cricket Eggs

3.3.1 Collection and Preparation of Cricket Eggs

1. Prepare the injection room: The room should be warm (25–28 °C) and the humidity should be about 40–60 %, with no blowing air to avoid drying the eggs.
2. Deprive crickets of any moist laying material overnight (at least 8–10 h) before collecting eggs.

3. Make and place an egg dish to collect eggs in the cage with the crickets: Cut two to three layers of paper towels into small squares, and place them in a Petri dish. Moisten the towels generously with water (*see* **Note 6**). Make a cover of two paper towel squares. Tuck the edges of the squares under the dish, and anchor the towels under the Petri dish lid.
4. Allow females to oviposit eggs for 1 h.
5. Remove the egg dish and place in a 28 °C incubator for 1 h (*see* **Note 7**).
6. Prepare injection capillaries and slides during the incubation periods in the above steps.
   (a) Use a glass capillary and a micropipette puller. For the micropipette puller, use the following program: HEAT 743, PULL 0, VEL 20, TIME 250; HEAT 743, PULL 0, VEL 20, TIME 250; HEAT 743, PULL 0, VEL 20, TIME 250; HEAT 743, PULL 70, VEL 25, TIME 150.
   (b) Put a piece of double-sided tape on the glass slide.
   (c) Paste a custom-ordered plastic frame on the double-sided tape.
7. After incubation for 1 h, collect eggs from the paper towels in the egg dish by swishing the towel squares in a Tupperware of water. Collect the eggs from the water with a tea strainer.
8. Soak the eggs in 70 % ethanol for 3–5 s by placing the strainer in a Petri dish containing ethanol and swirling.
9. Dip the strainer in fresh water to rinse the eggs.
10. Moisten and gently wring out a paper towel. Spread the eggs on the paper towel, and separate individuals by using a wet paintbrush.
11. Pick up individual eggs with an egg spatula (yellow tip with double-sided tape). Line the eggs up along the long edge of the injection well. The dorsal side of the embryo (concave side of the egg) should face the middle of the slide. It is important that the line of eggs be as straight as possible (Fig. 2b).
12. Gently push the eggs into the tape to ensure that they will not roll when injected.
13. Cover eggs with 500–600 μL of mineral oil (*see* **Note 8**).
14. Place eggs in a moist incubation chamber (*see* **Note 9**).
15. Set up the injection needle and injection solution.
   (a) Break a capillary tube in half, and place the halves on double-sided tape on a glass slide. Secure from the top with another piece of tape.
   (b) Under the microscope, gently tap the sharp end of pulled capillary into the broken end of the capillary tube.

This process will break the fine end of the needle, to be approximately 5–7 μm in diameter.

(c) Mix 1 μg/μL each of the TALEN L and R mRNAs for the injection solution (*see* **Note 10**).

(d) Load the injection solution into the blunt end of the broken capillary tube that is stuck to the glass slide with double-sided tape.

(e) Under the microscope, insert the pulled capillary into the capillary tube and into the solution.

(f) Pull the syringe to back-load the solution into the needle.

3.3.2 Injection of TALENs into Eggs, and Rearing of Injected Eggs

1. Put the egg-lined slide onto the stage of an inverted microscope.
2. Check the anterior and posterior poles of an egg under the microscope. The anterior end is pointed, and the posterior end is round (Fig. 2c).
3. Determine the region for injection at the posterior end of the egg (20–25 % of egg length from the posterior end; Fig. 2c).
4. Insert the injection needle into the egg with weak pressure, to avoid back-loading of the egg yolk (an unclogging technique). It is important that the injection needle be as vertical to the egg surface as possible because an improper angle can cause the needle to break (Fig. 2d).
5. Forcefully inject 2–4 nl of TALEN-L/R mRNAs into the middle region of the egg, using a syringe. It should be possible to observe the injected solution as a small bubble, which will be as large as a third part of the diameter in the egg (Fig. 2e).
6. Quickly remove the injection needle from the egg with weak pressure, to avoid back-loading of the egg yolk (Fig. 2f).
7. Move to the next egg, and repeat **steps 2–6**. One injection needle should be sufficient for injection into several hundreds of eggs.
8. After all of the eggs on the slide have been injected, place the slide in a moist incubation chamber for 2 days at 28 °C.
9. On day 2, record the total number of surviving embryos.
10. Prepare a humid Petri dish to incubate the eggs: Cut paper towels, place them into the Petri dish, and moisten with double-distilled H_2O. Press out most of the excess water from the towels. The towels should be moist, not soaking wet.
11. Transfer the eggs to the humid Petri dish: Slowly slip fine forceps under the egg, moving gently back and forth to loosen the egg from the tape. Use scientific cleaning wipes to remove as much mineral oil as possible (*see* **Note 11**). Place the eggs in a humid Petri dish, with ~40 eggs every few layers. Cover the dish with a lid, and incubate in a cricket room until the eggs hatch.

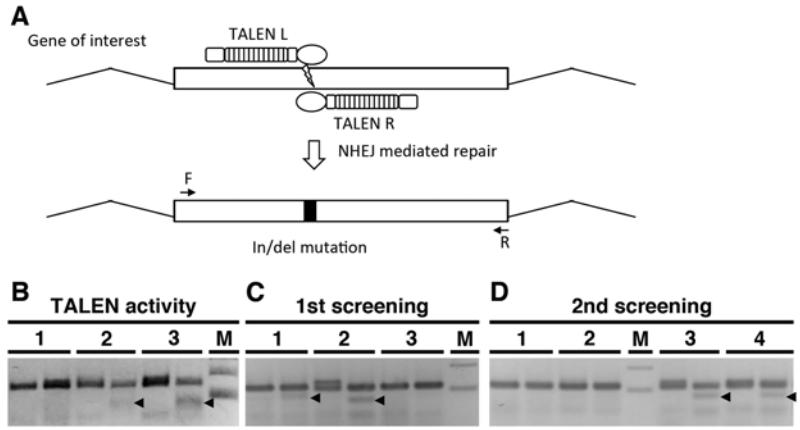


Fig. 3 Isolation of homozygous mutants after two screening stages**.** (**a**) Illustration of the binding and function of TALENs. After the introduction of double-strand breaks (DSBs) in the DNA between targeted binding sites, nonhomologous end joining can occur, resulting in deletions or insertions. The gene function of the targeted gene is disrupted, and subsequent binding of TALENs does not result in *FokI* dimerization or DSB generation. Primers "F" and "R" are designed to bind genomic DNA for amplifying 150–300 bp of the PCR product. (**b**–**d**) Agarose gel pictures of the SURVEYOR nuclease assays. *Left lane* of each sample is untreated DNA, and *right lane* is DNA treated with SURVEYOR nuclease. (**b**) SURVEYOR nuclease assay for evaluating TALEN activity on day 7. The expected band cleaved by SURVEYOR nuclease was detected in eggs no. 2 (weak) and no. 3 (strong), but not no. 1 (*arrowhead*). (**c**) First round of screening using the SURVEYOR nuclease assay on day 70. The expected band was detected in G_1 lines no. 1 (weak) and no. 2 (strong), but not no. 3 (*arrowhead*). (**d**) Second round of screening with the SURVEYOR nuclease assay on day 130. The expected band was detected in G_1 nymphs no. 3 and no. 4, but not no. 1 or no. 2 (*arrowhead*). *M* marker DNA. Reprinted from Ref. 20 with permission from Elsevier

3.4 Generation of Homozygous Knockout Cricket Lines via Two Selection Stages (Fig. 1)

1. On day 0, inject TALEN-L/R mRNAs into the posterior end of ~150 eggs (*see* **Note 12**) (*see* Subheading 3.3).
2. On day 7, evaluate the TALEN activity using ten eggs.
   (a) Separately extract genomic DNA from ten TALEN-treated eggs, using a genomic DNA purification kit.
   (b) Amplify the 150 to 300 bp fragment, including the TALEN target site, by PCR in 20 μL. Divide the PCR products into halves for SURVEYOR nuclease treatment or non-treatment (*see* **Note 13**) (Fig. 3a).
   (c) To allow complementary but mismatched strands to anneal, incubate 10 μL of the PCR products at 95 °C for 5 min, ramp the temperature from 95 to 85 °C at −2 °C/s, and ramp the temperature from 85 to 25 °C at −0.1 °C/s.
   (d) Add 0.5 μL Nuclease S and 0.5 μL Enhancer S to the reannealed PCR products.
   (e) Incubate the mixture at 42 °C for 45 min (*see* **Note 14**).

(f) Visualize the SURVEYOR nuclease-treated and -untreated products by agarose gel electrophoresis immediately after incubation (*see* **Note 15**) (Fig. 3b).

(g) Confirm the number of eggs for which mutations have been introduced into a TALEN target site.

3. Approximately 70–80 % of nymphs will hatch on day 13. This hatching rate will be lower if genes related to embryonic development have been targeted. If the knockout phenotype is visible (as is the case in *laccase2*), then nymphs will exhibit a partial knockout phenotype because G_0 nymphs are chimeric insects consisting of TALEN-induced mutant cells and wild-type cells.
4. From days 14–50, keep the G_0 nymphs in the insect case.
5. To avoid crossing to the mutant line, divide the G_0 nymphs into males and females at the time that they become eighth instar nymphs (about day 50).
6. At 3–7 days after the nymphs become adults (about day 60), mate the G_0 adults to wild-type adults, and collect 200–300 G_1 eggs.
7. One week after the eggs have been laid (about day 70), perform the first round of screening with SURVEYOR nuclease, to identify cricket lines containing TALEN-induced mutants.
   (a) Extract genomic DNA from batches of 25 embryos in each line (*see* **Note 16**), using a genomic DNA purification kit.
   (b) Perform SURVEYOR nuclease treatment (*see* Subheading 3.4, 2(b)–(f)).
   (c) Isolate SURVEYOR nuclease-sensitive (positive) lines, which contain the mutants (Fig. 3c).
8. Keep the G_1 eggs and nymphs from the positive lines of the first screening until the eighth instar stage (day 70–120).
9. A few days after the eggs/nymphs from **step 8** become eighth instar nymphs (about day 120–130), perform the second round of screening with SURVEYOR nuclease to identify the G_1 heterozygous mutants.
   (a) Separately extract genomic DNA from the T3 leg tips of 24 nymphs (12 males and 12 females) in each line, using a genomic DNA purification kit.
   (b) Perform SURVEYOR nuclease treatment (*see* Subheading 3.4, 2(b)–(f)).
   (c) Isolate the SURVEYOR nuclease-sensitive (positive) crickets, which correspond to the G_1 heterozygous mutants (Fig. 3d).
10. Keep the G_1 heterozygous mutants until the adult stage.
11. At 3–7 days after the G_1 heterozygous mutants become adults (about day 140), mate the G_1 heterozygous mutant adults individually, and collect the G_2 eggs.

12. G_2 eggs include homozygous mutants, heterozygous mutants, and wild types as Mendelian characteristics.
13. Genotype the eggs by the SUVEYOR nuclease system or by direct sequencing of the target region.

4 Notes

1. Provide flesh feed every 3 or 4 days because crickets prefer to eat flesh feed. For the first instar nymphs, crushed feed was provided using a coffee mill.
2. The maximum population per one insect case ($15 \times 20 \times 15$ cm^3) was approximately 300 nymphs. Too dense of an insect population promotes cannibalization.
3. Fresh water in a plastic dish should be provided daily, and the paper towels should be replaced every 3 days.
4. For design of the TALEN target site, the Web-based tool, TAL Effector Nucleotide Targeter 2.0 (https://tale-nt.cac.cornell.edu/) [19], was used according to the instructions.
5. To avoid repeated freeze-thaw cycles, the solution was divided into 1 μL or 1.5 μL aliquots.
6. Crickets will produce few eggs in the paper towels if there is too much water. To reduce the amount of water, drop the extra water by flipping the plastic water dish upside down.
7. The mortality rate of the eggs can be reduced experimentally, although it is not necessary.
8. Use fresh oil. Using old oil can be lethal to the injected eggs. Generally, we divide the purchased fresh oil into 30 mL aliquots.
9. Line the bottom of a square plastic dish with a moistened paper towel.
10. An injection solution of 2 μL, including 1 μL of each mRNA, is sufficient to inject approximately 200 eggs.
11. Performed rapidly before the eggs dry out.
12. In general, approximately 20 % of the eggs that are injected will develop to fertile adults. Thus, 150 injected eggs should develop to 30 fertile adult crickets. Five or six mutant lines can be generated if the efficiency of germline mutation is 20 %, as shown in our previous study [15].
13. In this step, it is recommended to check the products of amplification by agarose gel electrophoresis.
14. If the digested bands appear smeared during electrophoresis, reduce the reaction time to 15 or 30 min.

15. The digested bands are often broad because the size of the insertions/deletions can vary among the mutagenized sequences. By using gels that have 6 or 7 mm wells, a clearer band pattern may be achieved.
16. By using a larger number of embryos, the sensitivity of mutation detection may be reduced.

Acknowledgement

This work was supported by JSPS KAKENHI (23687033/25650080) to T.M. and 22124003 to T.M. and S.N.

References

1. Niwa N, Inoue Y, Nozawa A, Saito M, Misumi Y, Ohuchi H, Yoshioka H, Noji S (2000) Correlation of diversity of leg morphology in *Gryllus bimaculatus* (cricket) with divergence in dpp expression pattern during leg development. Development 127:4373–4381
2. Mito T, Sarashina I, Zhang H, Iwahashi A, Okamoto H, Miyawaki K, Shinmyo Y, Ohuchi H, Noji S (2005) Non-canonical functions of *hunchback* in segment patterning of the intermediate germ cricket *Gryllus bimaculatus*. Development 132:2069–2079
3. Nakamura T, Mito T, Bando T, Ohuchi H, Noji S (2008) Dissecting insect leg regeneration through RNA interference. Cell Mol Life Sci 65:64–72
4. Shinmyo Y, Mito T, Matsushita T, Sarashina I, Miyawaki K, Ohuchi H, Noji S (2004) *piggyBac*-mediated somatic transformation of the two-spotted cricket, *Gryllus bimaculatus*. Dev Growth Differ 46:343–34
5. Nakamura T, Yoshizaki M, Ogawa S, Okamoto H, Shinmyo Y, Bando T, Ohuchi H, Noji S, Mito T (2010) Imaging of transgenic cricket embryos reveals cell movements consistent with a syncytial patterning mechanism. Curr Biol 20:1641–1647
6. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holms MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148
7. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501
8. Rémy S, Tesson L, Ménoret S, Usal C, Scharenberg AM, Anegon I (2010) Zinc-finger nucleases: a powerful tool for genetic engineering of animals. Transgenic Res 19:363–371
9. Lieber MR, Ma Y, Pannicke U, Schwarz K (2003) Mechanism and regulation of human non-homologous DNA end-joining. Nat Rev Mol Cell Biol 4:712–720
10. Weiner A, Zauberman N, Minsky A (2009) Recombinational DNA repair in a cellular context: a search for the homology search. Nat Rev Microbiol 7:748–755
11. Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM, Jiao R (2012) Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. J Genet Genomics 39:209–215
12. Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JR (2011) Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29:697–698
13. Tesson L, Usal C, Ménoret S, Leung E, Niles BJ, Remy S, Santiago Y, Vincent AI, Meng X, Zhang L, Gregory PD, Anegon I, Cost GJ (2011) Knockout rats generated by embryo microinjection of TALENs. Nat Biotechnol 29:695–696
14. Kato T, Miyata K, Sonobe M, Yamashita S, Tamano M, Miura K, Kanai Y, Miyamoto S, Sakuma T, Yamamoto T, Inui M, Kikusui T, Asahara H, Takada S (2013) Production of *Sry* knockout mouse using TALEN via oocyte injection. Sci Rep 3:3136
15. Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N, Nakamura T, Bando T, Ohuchi H, Yamamoto T, Noji S, Mito T (2012) Non-transgenic genome modifications in

a hemimetabolous insect using zinc-finger and TAL effector nucleases. Nat Commun 3:1017

16. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39:e82
17. Sakuma T, Hosoi S, Woltjen K, Suzuki K, Kashiwagi K, Wada H, Ochiai H, Miyamoto T, Kawai N, Sasakura Y, Matsuura S, Okada Y, Kawahara A, Hayashi S, Yamamoto T (2013) Efficient TALEN construction and evaluation methods for human cell and animal applications. Genes Cells 4:315–326
18. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30:460–465
19. Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, VanDyk JK, Bogdanove AJ (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40:117–122
20. Watanabe T, Noji S, Mito T (2014) Gene knockout by targeted mutagenesis in a hemimetabolous insect, the two-spotted cricket *Gryllus bimaculatus*, using TALENs. Methods 69:17–21

Chapter 13

Methods for TALEN Evaluation, Use, and Mutation Detection in the Mosquito *Aedes aegypti*

Sanjay Basu, Azadeh Aryan, Mary Etna Haac, Kevin M. Myles, and Zach N. Adelman

Abstract

The generation and study of transgenic *Aedes aegypti* mosquitoes provides an essential tool for elucidating the complex molecular biology of this important vector. Within the field, genetic manipulation has surpassed the proof of principle stage and is now utilized in both applied and theoretical vector control strategies. The application of new instruments, technologies and techniques allows ever more controlled experiments to be conducted. In this text we describe microinjection of *Ae. aegypti* embryos in the context of evaluating and performing genomic editing with transcription activator-like effector nucleases (TALENs).

Key words Mosquito, Aedes, TALEN, Mutagenesis, HRMA, Nuclease, Gene editing

1 Introduction

Aedes aegypti mosquitoes are known vectors of a number of medically significant viruses and parasites including dengue viruses and yellow fever. Increased global travel, amongst other factors, has lead to the expansion of dengue to previously non-endemic regions of Europe, the USA, and China [1]. It is therefore imperative to understand the molecular biology of this complex vector, and over the last 20 years molecular techniques available to investigate this organism have advanced significantly. At present, the technology and techniques have advanced sufficiently to generate transgenic lines or follow a reverse-engineered genetic approach.

One of the most exciting breakthroughs in the last 2 years has been the application of gene-editing techniques that can truly generate genetic knock-out mosquito lines. Transcription activator-like effector nucleases (TALENs) have been shown to work effectively in *Ae. aegypti* [2] and there is also evidence for successful

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_13,

homologous recombination [3, 4]. The adaptation of other technologies, such as high-resolution melt analysis (HRMA), is very useful in helping to characterize and track gene-editing events. To the uninitiated these experiments may appear time consuming and daunting but are actually straightforward in their work-flow. In this chapter we have worked to create structured and robust protocols for the development of standard procedures building upon previous methods [5] to introduce and evaluate TALEN-induced genetic lesions in *Aedes aegypti*.

Assessment of gene editing activity using germline experiments is time consuming and labor intensive, requiring at least 2 months for *Ae. aegypti*. To this end we describe two rapid assays to make initial assessments of TALEN activity. Both assays have their own particular advantages and disadvantages. The single strand annealing (SSA) assay is the more sensitive of the two assays reviewed here. In this assay the TALEN target site is placed between two direct repeat sequences. The sequence upstream of the target site is the first 300 bp of the firefly luciferase gene and the full length of firefly luciferase open reading frame (ORF) is located downstream of the target site. Successful double stranded DNA break induction can be repaired using the single strand annealing pathway, resulting in the collapse of the direct homolog sequences and expression of the full length firefly luciferase ORF [2]. The disadvantage of this method is that it measures activity against an extrachromosomal (plasmid) target, not the endogenous target. Also, each unique target must first be cloned into a plasmid, limiting the ability of this method to scale. The second assay uses high resolution melt analysis (HRMA) [6] to determine differences in the melting temperature of PCR amplicons generated by primers flanking the targeted region of interest. Insertion/deletion events (indels) will alter the melt profile of the product in comparison to the wild-type control. The advantages of this method are that no cloning step is required, and that multiple TALENs or other gene-editing agents (CRISPRs) can be tested within the same injection mix as the assay allows for the same template to be employed in a number of different reactions. Two or three pairs of TALENs can be injected and assessed together, rather than individually (as in the case of the SSA assay). Disadvantages are that this method is not as sensitive as the SSA-based assay and is susceptible to false positive results if the amplified region contains natural sequence variants; this is easily avoided through careful prescreening of potential target sites [7].

Once an active nuclease is identified in one or both transient assays, the described injection protocol can then be used for germline modification experiments. The final section describes our screening protocol to identify heteritable mutations in the absence of a phenotypic marker using HRMA.

2 Materials

All DNA plasmids for injection must be prepared using an endotoxin-free kit. All insectary rearing should be conducted in standard *Ae. aegypti* laboratory conditions (28 °C, 60–80 % relative humidity, photoperiod of 16 h light–8 h dark) with all water required for mosquito rearing either reverse-osmosis treated or double-distilled.

2.1 Microinjection of Ae. aegypti Embryos

1. 50 ml conical tube.
2. X-large cotton balls non-sterile.
3. 1.5″ hole puncher.
4. Chromatography paper.
5. Plastic coverslips.
6. Double-sided adhesive nontoxic tape.
7. 100 mm × 15 mm petri dish.
8. Whatman paper Grade #40
9. Whatman paper Grade #50.
10. 40 ml beaker.
11. Parafilm.
12. 150 mm × 15 mm petri dish.
13. Adhesive putty.
14. Borosilicate microcapillaries, inner diameter 0.64 mm, outer diameter 1.0 mm, length 100 mm.
15. Micropipette Puller.
16. 3 % H_2O_2: 3 % solution in DEPC treated-water.
17. Microelectrode Beveler with fine abrasive plate.
18. 10× Injection Buffer: 50 mM KCl, 1 mM NaH_2PO_4 in nuclease free water. Filter-sterilize using 0.22 μM syringe filter (adapted [8]).
19. Chilled centrifuge.
20. Microloaders.
21. Flashlight aspirator.
22. Stereomicroscope.
23. Tweezers: Dumont #5 Inox 11 cm.
24. Halocarbon oil 27 (Sigma-Aldrich, St. Louis, MO).
25. Microscope and manipulator to position and operate the microinjector needle holder.
26. Femtojet, programmable microinjector (Eppendorf, Hauppauge, NY).
27. Fine-spray water bottle.

2.2 Single Strand Annealing Assay

1. Dual-Luciferase Reporter Assay System (Promega, Madison, WI).
2. Glomax-Multi+ Microplate Multimode Reader (Promega, Madison, WI).
3. Fine paintbrush size 3/0.
4. Long forceps 12″.
5. Microcentrifuge pestle for 1.5 ml microcentrifuge tube.
6. Costar White Solid 96-well plate.
7. SSA test plasmid [7].

2.3 High Resolution Melt Analysis Transient Assay

1. LightScanner mastermix (Biofire Defense, Salt Lake City, UT).
2. Hard-shell black–white 96-well plate.
3. Mineral Oil.
4. Optical film.
5. Gradient thermocycler.
6. 1× TAE buffer: add together 4.84 g Tris, 1.1 ml Glacial Acetic Acid, 2 ml 0.5 M EDTA (pH 8.0). Make upto 1 l with reverse-osmosis treated water then autoclave.
7. TAE-agarose gel: 2 % agarose in 1× TAE buffer. Dissolve 2 g agarose in 100 ml 1× TAE solution (described in previous step).
8. LightScanner (catalog number LSCN-ASY-0011) (Biofire Defense, Salt Lake City, UT).
9. Genomic DNA from Tissue Kit.
10. Pellet pestle cordless motor.
11. Spectrophotometer.

2.4 Germline Experiment

1. Phire Animal Tissue Direct PCR kit (catalog number #F-140) (Thermo Scientific, Pittsburgh, PA).
2. 70 % ethanol: 70 % ethanol solution in water.
3. LC Green Plus + (catalog number BCHM-ASY-0005) (Biofire Defense, Salt Lake City, UT).
4. Fish food.
5. Soup cup 16 oz.
6. Mesh fabric (Nylon Tulle).
7. 15 ml disposable beaker.
8. 5 ml disposable polystyrene beakers.
9. Cardboard popcorn container 46 oz.
10. Glue-gun.
11. Staples and stapler.
12. 4″ cotton stockinette.

13. 14″ cable ties.
14. Wide polystyrene drosophila tube (28.5 mm × 95 mm).
15. Flug (Genesee, San Diego, CA).
16. Mosquito rearing pan—28 quart storage box and lid (23″ × 16 ¼″ × 6″).
17. Disposable plastic Pasteur pipette.
18. Pyrex glass petri dish, 150 mm × 20 mm.
19. Small tub and lid (500 ml volume).
20. Plastic sheeting for use with a 2 mm Impulse Heat Sealer.

3 Methods

3.1 Microinjection of Aedes aegypti Embryos

3.1.1 Preparation

1. A cage of mosquitoes should be blood fed at least 72 h before the first day of embryo injections (*see* **Note 1**).
2. Prepare the laying chambers/tubes by placing a ball of wet cotton in the cone of a 50 ml conical tube. Place a damp 1.5″ diameter circular chromatography paper gently on top of the cotton. The dampness of the cotton and paper should not be in excess; another cotton ball is used as the stopper at the open end (*see* **Note 2**).
3. Prepare 10–12 coverslips with double-sided adhesive tape along one edge. Trim the tape back to the edge of the coverslip using scissors.
4. Prepare a petri dish (100 mm diameter) with a circle of Whatman paper-grade #40. The upturned petri dish lid makes manipulation under the microscope easier as the field of view can quickly be adjusted by moving the dish. Saturate the petri-dish mounted paper with water then transfer the laying chamber-tube paper onto the petri-dish mounted paper. The level of water saturation is critical for transferring the embryos as too little moisture will promote desiccation of embryos, whereas too much will make manipulation of the embryos problematic.
5. Prepare pieces of transfer embryo paper. The line of embryos for transfer is generated upon a different piece of Whatmann paper-grade #50. The different grade ensures that fewer fibers are transferred to the adhesive tape upon the coverslip.
6. Prepare embryo recovery beakers. After injecting, the embryos are moved to a labeled 40 ml recovery beaker containing a ball of damp cotton and sealed with Parafilm.
7. Needle manufacture (detailed in [7]).
   (a) Prepare a covered container for the needles to be stored in; a 100 mm petri dish or tip box with two lines of adhesive putty is ideal.

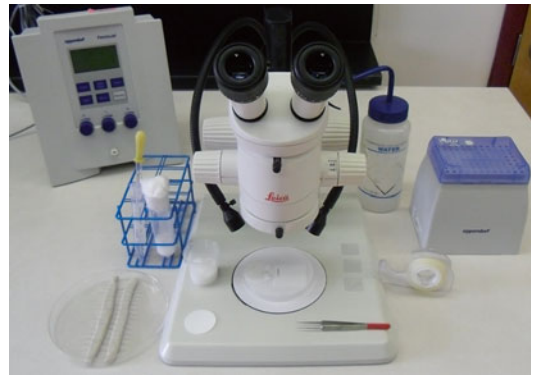


Fig. 1 A prepared work station ready for lining up embryos for microinjection. Upon the desk from the left then proceeding clockwise: needles within their container, a laying chamber tube and a 15 ml tube of Halocarbon Oil with a glass Pasteur pipette ready for covering dessicated embryos, the Eppendorf Femtojet, a water bottle with a fine jet spray nozzle, Eppendorf microloaders (within their tip-rack) and double-sided adhesive tape. Upon the microscope stage (from *left* to *right*) are a *small circle* of paper for the laying chamber, a recovery beaker, a Whatman paper *circle* upon a petri dish lid showing a *line* of embryos ready for transfer, a pair of tweezers and three prepared coverslips

(b) Generate needles for injection (*see* **Notes 3** and **4**).

(c) If the injection mix includes RNA components then use treated needles to minimize RNase contamination (*see* **Note 5**).

8. Bevel needle. Our preferred angle of bevel is approximately: 22.5° (*see* **Note 6**).
9. Assemble the injection mix components in 10× Injection Buffer (make 1× Injection Buffer by dilution with water), if the mix includes RNA components then prepare needles and injection mix on that day. If the injection mix only includes DNA components then needles and injection mix can be prepared further ahead of time (*see* **Notes 7–10**).
10. Figure 1 shows a prepared work station for microinjection with many of the described items.

3.1.2 Method

1. Within a prechilled centrifuge, spin the injection mix at >16,000 × *g* for >20 min at 4 °C then maintain the injection mix chilled (on ice) while injecting (*see* **Note 10**).
2. Use Microloader tips to backfill the needle with injection mix (*see* **Note 10**).
3. Gently transfer 20–30 gravid females into the laying chamber using an aspirator. Leave the tube in darkness in insectary conditions for approximately 30 min (*see* **Notes 11** and **12**). A small cardboard box works well.

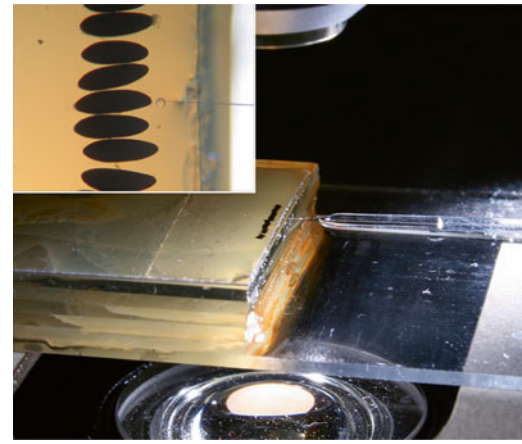


Fig. 2 The microinjection setup. A coverslip with a line of embryos covered in Halocarbon oil placed upon the slide rostrum. The elevation allows the needle to easily access the posterior pole of the embryo without touching the stage. Inset image shows injection microscope view of embryos and needle (magnification ×50)

4. Use tweezers (Dumont #5 Inox 11 cm or similar) to line up the embryos next to each other orientating the posterior pole in one direction. Line up 100–130 embryos in a single line that will span most of the width of the coverslip. When selecting embryos for lining up, observe and select embryos of a similar color. No matter how many embryos are lined up the embryos should be between 90 and 120 min old when injecting.
5. Dry the transfer paper by pressing down upon a folded kimwipe tissue before transfer to the coverslip. Be aware that the embryos will start to dessicate once the water is removed from around them (*see* **Note 13**).
6. Transfer the lined-up embryos to the prepared coverslip. Firmly but gently press the adhesive tape strip of the coverslip against the line of embryos. The edge of the coverslip is ideally only 2–3 mm away from the posterior poles of the lined up embryos (*see* **Note 14**).
7. Dessicate the embryos upon the coverslip. When they have achieved the desired desiccation level (first appearance of dimpling), cover in Halocarbon oil 27.
8. Using tweezers, remove any fibers that have also become stuck to the adhesive tape. This prevents the fibers obstructing or damaging the needle while injecting.
9. Transfer the coverslip to the injection microscope stage. Line up the edge of the coverslip closely with the edge of the injection microscope slide rostrum (*see* Fig. 2). Bring the bevel of the needle into focus using the micromanipulator then withdraw the needle from view and then focus the first embryo

with the stage controls. Slowly bring the needle back into the microscope view until it is in the center of the image.

10. Using the Femtojet Injector, inject the embryos back to their full turgidity using the stage controls to pierce the embryos upon the fixed needle (*see* Notes **15**–**21**).
11. Drain the halocarbon oil from the injected embryos and using forceps transfer the embryos carefully to a damp paper surface for their development time to complete. Uninjected embryos are removed or squashed at this stage (*see* **Note 22**).
12. Embryos recover on a small piece of filter paper, Whatman paper grade #40 placed within a prepared recovery beaker. The number of transferred embryos are counted and noted then the recovery beaker is sealed with a stretched piece of Parafilm over the top. The recovery beaker is placed back in the insectary for development under regular conditions.
13. If the injection mix includes RNA components then after injecting run a sample of the mix using TAE-agarose gel electrophoresis to observe if any RNA degradation has occurred.

3.2 Single Strand Annealing Assay

3.2.1 Preparation

Clone target site (including both TALE binding regions and spacer) into the spacer region of the SSA test vector [7]. The activity of Firefly luciferase and Renilla luciferase can be measured by using the Dual-Luciferase Reporter Assay System with the Glomax-Multi+ Microplate Multimode Reader according to the manufacturer's protocol. Different target sites can be assembled into a single SSA vector using gene synthesis technology. Numerous TALEN constructs can be tested within a single SSA test vector by using a multiple cloning site to allow insertion of the synthetic fragment.

3.2.2 Method

1. Assemble injection mix consisting of the SSA construct, a pair of TALEN expression plasmids and a control plasmid for normalization of the dual luciferase assay (*see* **Note 23**). Use 100–200 ng/μl of each TALEN expression construct, 200 ng/μl of SSA reporter plasmid and 200 ng/μl of the control plasmid. Add DEPC water to make a total volume of 50 μl.
2. Spin the injection mix down at the maximum speed for at least 10 min at 4 °C.
3. Follow the injection protocol as stated above, ideally inject embryos in 3–6 replicates of 100–120 embryos each. Allow the embryos to recover for 24 h after injection, then transfer them into a labeled 1.5 ml microcentrifuge tube using a fine paintbrush. Snap-freeze the tubes in liquid nitrogen (*see* **Note 24**).
4. Remove the tubes from the liquid nitrogen one at the time using long forceps (*see* **Note 25**).
5. Dilute 5× Passive Lysis Buffer to 1× with nuclease-free molecular biology grade water (can be stored at 4 °C for up to 1 month).

6. Prepare sufficient 1× Passive Lysis Buffer (from the Promega kit) for 100 µl per tube of embryos.
7. Remove one tube at a time and crush within 50 µl of 1× Passive Lysis Buffer using a microcentrifuge pestle. Add an additional 50 µl of 1× Passive Lysis Buffer to the tube and place in a rack until all samples have been crushed (*see* **Note 26**).
8. Snap-freeze all samples again in liquid nitrogen then transfer them back to a rack to thaw.
9. Once thawed, spin down samples at ~3000 × *g* for 10 s at room temperature.

3.2.3 Luciferase Assay

1. Prepare the Luciferase Assay Reagent II (LARII) by resuspending the lyophilized luciferase assay substrate in 10 ml of the supplied Luciferase Assay Buffer II (*see* **Note 27**).
2. Prepare the Stop & Glow substrate. Substrate is supplied as a 50× concentrate; dilute 1:50 into Stop &Glow buffer, each sample will require 80 µl per sample (*see* **Note 28**).
3. Equilibrate samples and reagent to ambient temperature.
4. Load 20 µl of each lysate into a Costar white solid 96-well plate.
5. Follow the manufacturer's protocol Glomax-Multi+ Microplate Multimode Reader to inject 80 µl LARII for each sample using one injector, read the luminescence (delay 10 s, integration 2 s) then inject 80 µl Stop & Glo substrate using the second injector (delay 10 s, integration 2 s).
6. Create a ratio by normalizing the firefly RLU value using the Renilla RLU value.

3.3 High Resolution Melt Analysis Transient Assay

3.3.1 Preparation

1. Use the Lightscanner Primer design program to design a pair of primers for an assay. A good starting point for HRMA primer pairs design is a product length of 80–120 bp, primer Tm°C of 50–72 °C, primer lengths of 19–30 nt and both 5′ and 3′ exclusion buffers set at 0 (*see* **Note 29**).
2. Set up the following reaction with the Lightscanner mastermix within a black–white 96-well plate.
   (a) Add 20 µl of mineral oil to each sample and 10 µl of the following mastermix reaction: 4 µl 2.5× Mastermix, 3 µl DEPC water, 1 µl of each primers (each to a final concentration of 250 µM).
   (b) Add the specific sample DNA (1 µl) to each single well; 10–20 ng of genomic DNA per 10 µl reaction is recommended (*see* **Note 30**).
   (c) Once the plate is loaded, cover with optical film and ensure an airtight seal over each well and around the edge of the plate.

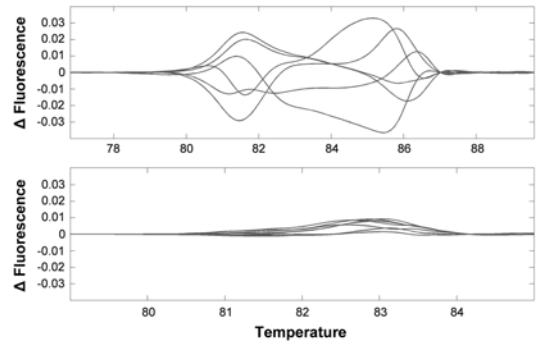


Fig. 3 Variability within the uninjected control samples reduces the sensitivity of HRMA analysis of gene editing events. The HRMA melt profiles (*top*) show excessive variability between replicate control DNAs ($n=6$), compared to those (*bottom*) that show relatively little variation ($n=7$)

(d) Centrifuge the plate in a swinging bucket rotor at 400 × *g* for 2 min at room temperature.

(e) Perform gradient PCR using a standard thermocycler (95 °C 2 min, 40 cycles of [94 °C 30 s, 50–72 °C 30 s], 94 °C 30 s, 25 °C 30 s).

(f) Load 7 μl of each sample into a TAE-agarose gel (2 %) for analysis. Check the PCR product is the predicted size.

(g) Thermal melt profiles can be generated using a Lightscanner (60–95 °C, hold 57 °C).

3. Using the Lightscanner Call-IT 2.0 software, examine the melt profiles within the annealing temperature determined by the gel image analysis. Find the temperature that has the greatest initial fluorescence tempered with the greatest difference in fluorescence change over the melt profile. This is the optimal annealing temperature for the reaction (*see* **Note 31**).

4. Perform the PCR reaction again (95 °C 2 min, 40 cycles of [94 °C 30 s, *x*°C 30 s], 94 °C 30 s, 25 °C 30 s), where "x" is the optimized annealing temperature identified in the preparation steps, using DNA from six to nine groups of 24 h old embryos. Generate thermal melt profiles using a Lightscanner (50–95 °C, hold 47 °C). If excessive variation between samples is observed, primers should be redesigned and validated prior to proceeding (Fig. 3).

3.3.2 Method

1. Assemble injection mix consisting of the 300 ng/μl of each TALEN expression construct and 1× injection buffer. Add DEPC water to make a total volume of 50 μl.

2. Follow the injection protocol as stated above. Inject embryos in three replicates of 100–120 embryos each. Allow the embryos to recover for 24 h after injection then transfer the

injected embryos into a 1.5 ml microcentrifuge tube using a fine paintbrush. Snap-freeze labeled tubes in liquid nitrogen (*see* **Note 25**).

3. Also collect 6–9 samples of uninjected embryos of the same developmental time as these will serve as controls for the HRMA analysis.
4. Extract DNA from the experimental and control embryo samples using Macherey-Nagel Tissue NucleoSpin Kit. Follow the stated protocol with the following considerations.
   (a) Homogenize the embryos (100–300) within a microcentrifuge tube using a microcentrifuge pestle in 180 µl of T1 buffer (*see* **Note 26**).
   (b) Incubate the proteinase K-T1 buffer homogenate 16 h.
   (c) Before loading the homogenate into the spin column, spin the sample >14,000 × *g* for 2 min at room temperature to pellet the embryo debris then transfer the solution to the column.
   (d) Elute in 100 µl Elution buffer and measure the DNA content using a NanoDrop spectrophotometer.
5. Follow the previous directions detailed in the transient assay preparation steps to set up the Lightscanner mastermix reactions within a hard-shell 96-well plate. Include 6–9 control samples of uninjected embryos of the same developmental time period and at least one no-template control.
6. Perform the PCR reaction (95 °C 2 min, 40 cycles of [94 °C 30 s, *x*°C 30 s], 94 °C 30 s, 25 °C 30 s), where "x" is the optimized annealing temperature identified in the preparation steps. Thermal melt profiles are generated using a Lightscanner (50–95 °C, hold 47 °C).
7. Using the Lightscanner Call-IT 2.0 software, examine the melt profiles to determine whether indel events are occurring within the target region. After normalizing the data, set the uninjected controls as standards. Any indel events will cause a shift in melt profile of the PCR product from the uninjected controls. A difference curve of 0.1 or greater (from the set standards) is typically sufficient to justify a germline experiment; smaller differences may also be acceptable depending on how much variation is present amongst the control samples.

3.4 Germline Experiment

3.4.1 Preparation

1. Optimize PCR for your target sequence using the Phire Animal Tissue Direct PCR kit.
   (a) Set up a tube with 20 µl Dilution Buffer + 0.5 µl DNA Release Additive.
   (b) Place one leg (or piece of leg) from a wild-type adult mosquito into the tube (*see* **Note 32**) (see setup featured in Fig. 4).

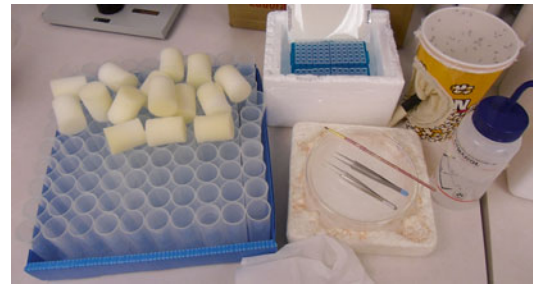


Fig. 4 Preparation of materials for setting up a 96-well plate HRMA tracking experiment. Clockwise from the *left*, the drosophila tubes ready for the adult mosquitoes with flugs to plug the tube, an ice-bucket with a 96-well plate already prepared with Phire mastermix in each sample well, a 46 oz popcorn-cage with adult mosquitoes, an ethanol bottle to wipe the tweezer legs between samples, an ice-chilled glass petri dish with tweezers and a paintbrush to move and transfer the mosquitoes

(c) Incubate at room temperature for 2–5 min then 98 °C (within a thermocycler) for 2 min. Allow to cool to room temperature while preparing PCR mix.

2. Using a multichannel pipette add 20 μl mineral oil to intended sample wells of a hard-shell 96-well plate.
3. Prepare the following: 19 μl Phire mastermix for each well across the gradient range; 10 μl 2× Phire Animal Tissue PCR Buffer, 1 μl primer mix (final concentration ~500 nM), 0.4 μl Phire II polymerase, 2 μl LCGreen, 5.6 μl DEPC water. Pipette Phire master mix into the sample wells.
4. Transfer 1 μl of Phire-dilution sample to each well of the gradient range.
5. Once the plate is loaded, cover with optical film and ensure an airtight seal over each well and around the edge of the plate.
6. Centrifuge the plate in a swing bucket rotor for 1 min at 400 × *g* at room temperature.
7. Run Phire gradient PCR program (including a heteroduplex formation step) using a thermocycler (98 °C 5 min, 40 cycles of [98 °C 5 s, 60–72 °C 5 s, 72 °C 20 s], 72 °C 1 min, 94 °C 30 s, 25 °C 30 s, 4 °C ∞).
8. Generate thermal melt profiles using a Lightscanner (60–95 °C, hold 55 °C).
9. Using the Lightscanner Call-IT 2.0 software, examine the melt profiles within the annealing temperature. Find the temperature that has the greatest initial fluorescence tempered with the greatest difference in fluorescence change over the melt profile. This will be the optimal annealing temperature for the assay.

10. Hatch wild type strain mosquitoes 4–7 days prior to hatching the injected embryos. This ensures that when G_0 male and female have emerged the wild type females/males are ready to cross with them. Add a small amount of fish food in the pan.
11. Prepare between 80 and 100 soup-cup cages to house sibling-matings for the progeny of each G_0 survivor using a 16 oz soup cup and the corresponding lid.
    (a) The cardboard inner top is punched out from the lid ring and discarded.
    (b) A mesh fabric (Nylon Tulle) is then placed over the cup opening and trapped down by replacing the ring of the adapted lid. The tension in the mesh needs to be sufficient to secure the mosquitoes inside and to support raisins and a small ball of damp cotton.
    (c) For each soup-cup cage prepare an oviposition beaker (15 ml disposable beaker) with a piece of chromatography paper curled around the inside.
    (d) After sex-separating the wild type pupae, place 30–35 individuals in 5 ml disposable polystyrene beakers into soup cup-cages. Provide a damp piece of the cotton wool and a number of raisins as sources of water and sugar.
12. Prepare 10–15 disposable popcorn-cages using a 46 oz cardboard containers:
    (a) Using a glue-gun, stick two 46 oz cups together, one inside the other.
    (b) A 10 cm diameter hole is cut through one side.
    (c) Cut an entry sock of ~25 cm of 4″ cotton stockinette.
    (d) Using a stapler secure the entry sock stockinette to the inside lip of the side hole.
    (e) Two cable ties (14″) are secured together to form a loop, these then trap a double layer of mesh (Nylon Tulle) to the top opening. The mesh is trapped to provide tension across the top of the opening allowing support of a blood feeder without excessive sagging.
    (f) Seal the mesh and cable ties in place using a hot glue gun.
13. Prepare laying tubes for single blood-fed female G_0 mosquitoes.
    (a) Place a damp ball of cotton into the bottom of a wide polystrene drosophila tube.
    (b) Place a damp 1.5″ diameter circular chromatography paper gently on top of the cotton.
    (c) Use a tissue or paper towel to wipe dry the inside edge of the conical tube to discourage mosquitoes ovipositing eggs anywhere other than the circular paper. The dampness of the laying surface is crucial in that if the surface is

too wet the mosquitoes will be trapped and prevent optimal laying. If too dry the mosquitoes will not lay.

(d) Plug the opening with cotton or a flug.

3.4.2 Method

1. Assemble injection mix with 300 ng/μl of each TALEN expression plasmid in 1× injection buffer.
2. Follow the injection protocol as stated above; inject approximately 1000 embryos.
3. At 5 days post-injection, prepare a pan containing 4 l of reverse osmosis-purified water with an appropriate amount of ground fish food (*see* **Note 33**).
4. Hatch the injected embryos within the labeled pan (*see* **Notes 34** and **35**).
   (a) Using a pair of tweezers, carefully transfer the paper containing injected eggs to the surface of water with the injected eggs face down.
5. Monitor the hatching pan daily for pupae, pick any developing pupae and separate the males and females into popcorn cages (*see* **Notes 36** and **37**).

G_0 Females

1. Set up popcorn-cages with 20–30 wild type males each and a equal number of G_0 females, anesthetize the adult mosquitoes using CO_2 and then transfer them to 150 mm × 20 mm glass petri dishes where they can be examined and manipulated with ease (*see* **Note 38**).
2. Feed the G_0 female cages 2–3 days after establishing the mating cross.
3. One day after blood feeding, anesthetize the mosquitoes, cut away the top mesh using a razor blade then transfer the mosquitoes (mated G_0 females and wild type males) to a prechilled glass petri dish. Using a pair of tweezers, place a single blood-fed G_0 female mosquito into a polystyrene drosophila laying tube then plug the tube (*see* **Note 39**).
4. Once the females have recovered, place the tubes gently in a horizontal orientation within a small cardboard box in the insectary. Keep 3–4 and record the number of tubes that show eggs have been laid. At this point the adult female mosquito can be discarded.
5. To hatch, gently fill the drosophila laying tube with water ensuring that the eggs are completely covered, along with a drop of very diluted fish-food solution.
6. After 2–3 days any viable G_1 larvae will have hatched. At this point, larvae should be moved to an appropriately sized rearing pan.

7. Transfer G_1 pupae from each founder female into a individual soup cup-cage allowing the sibling-adults to cross. At this time also add an oviposition beaker (*see* **Note 40**).
8. Blood feed the soup cup-cages 2–3 days after establishing the mating cross.
9. Two days after blood feeding, half-fill the oviposition beakers in the G_1 inter-cross soup cup-cages with water by spraying/trickling water through the mesh.
10. After 3 or 4 days examine and record the number of soup cup-cages that show eggs have been laid.
11. When the G_1 females have laid the eggs, anesthetize them using CO_2 and transfer all the individuals to a labeled micro-centrifuge tube. Snap-freeze and store at –80 °C until ready for genomic DNA extraction.
12. Store the labeled G_2 oviposition papers (*see* **Note 41**).
13. Extract genomic DNA from the G_1 adult bodies. Use Macherey-Nagel Tissue NucleoSpin kit. Elute DNA from spin column in 100 μl.
14. Use HRMA to analyze sequence at TALEN recognition site as detailed previously in HRMA Transient Assay section.
15. If indel mutation is identified by HRMA, hatch the corresponding G_2 oviposition paper.
16. Place G_2 male and female pupae into separate cages prior to genotyping by HRMA.
17. To identify individual adults with mutant alleles use Phire Animal Tissue Direct PCR kit:
    (a) Set up 96-well plate with 20 μl Dilution Buffer + 0.5 μl DNA Release Additive in each well.
    (b) Place one leg (or piece of leg) from a wild-type adult mosquito within 1 well (*see* **Note 32**).
    (c) Seal plate with optical PCR film, incubate at room temperature for 2–5 min then 98 °C (within a thermocycler) for 2 min. Allow to cool to room temperature while preparing PCR mix.
18. Follow the previous directions detailed in the preparation steps (Germline experiment, Preparation, **Steps 2** and **3**) to prepare the 96-well plate in which the Lightscanner reaction will be performed.
19. Transfer 1 μl of Phire-dilution sample to the 96-well plate, we recommend using 96-well plate plan to track samples.
20. Run Phire PCR at optimized temperature (including a heteroduplex formation step): (98 °C 5 min, 40 cycles of [98 °C 5 s, *x*°C 5 s, 72 °C 20 s], 72 °C 1 min, 94 °C 30 s, 25 °C 30 s, 4 °C ∞).

21. Generate thermal melt profiles using a Lightscanner (60–95 °C, hold 55 °C) and analyze using the Lightscanner Call-IT 2.0 software.
22. Identify individual adult mosquitoes with mutant alleles.

G_0 Males

1. When crossing G_0 males, first visually check the soup cups of virgin wild type females to see if there are any males in the group; these should be discarded. Anesthetize the adult mosquitoes using CO_2 and then transfer them to 150 mm × 20 mm prechilled glass petri dishes. From the dishes the adults can be transferred using tweezers into the prepared soup cup-cages.
2. Put a single G_0 male with five virgin females into each soup cup-cage. At this time also add an oviposition beaker into the soup cup cage (*see* **Note 40**).
3. Feed the soup cup-cages 2–3 days after establishing the mating cross.
4. Two days after blood feeding fill the oviposition cups in the male soup cups with water by spraying/trickling water through the mesh.
5. After 3 or 4 days examine and record the number of soup cup-cages that show eggs have been laid.
6. Using CO_2, anesthetize the adults within those cages that have laid eggs. Those that have not can either be refed or discarded.
7. Using tweezers, transfer the G_1 oviposition paper into a small tub of water that has already been prepared with some fish-food. Alternatively the oviposition paper can be stored for screening later.
8. Continue the protocol as stated above for G_1 progeny (of G_0 females) by inter-crossing the adults within a soup cup-cage (*see* **step 13**), blood-feeding and collecting the G_2 oviposition paper.

4 Notes

1. Preferentially the mosquitoes should be fed 96 h in advance of the injection day. The number of females blood fed is dependent upon the number of injections needed for the experiment. For approximately 1000 injections, a minimum of 200 females is recommended.
2. Circular papers are manufactured using a hole puncher (1.5″) from chromatography paper. Use a tissue paper to wipe dry the inside edge of the conical tube to discourage mosquitoes ovipositing eggs anywhere other than the circular paper. The dampness of the laying surface is crucial in that if the surface is too wet the mosquitoes will be trapped and prevent optimal laying. If too dry the mosquitoes will not lay.

3. Handle the microcapillaries at the ends of the tubes to avoid oils or other contaminants being present at the heating point.
4. We recommend using the Sutter P2000 Laser puller with borosilicate microcapillaries of inner diameter: 0.64 mm, outer diameter: 1.0 mm, length 100 mm and program: Heat = 270, FIL = 3, VEL = 37, DEL = 250, PUL = 140. Micropipette-pullers are also available from other companies, but the program must be determined empirically through embryo injection survival rate. When pulling the needles, monitor the Heat-on-time value generated after every pull to ensure that the needles are consistent. Needles pulled outside a narrow variance should be discarded at this stage.
5. Handle microcapillaries with gloves. To treat needles for RNA work: soak all microcapillaries in 3 % H_2O_2 solution overnight. Thoroughly rinse the microcapillaries numerous times in reverse osmosis treated water. Bake dry in an oven until all liquid has evaporated. Pull needles on the day of injections using standard embryo injection program.
6. To bevel microcapillaries, the 104D fine abrasive plate is recommended. Before use, blow any dust or fine particles away from the beveling surface using compressed air. The bevel of each needle can be checked using a microscope (magnification 100×–200× is ideal). The bore of the needle should also checked as if much greater than 1 μm then the increased trauma to the embryo maybe lethal.
7. If the injection mix includes RNA components then normal RNA procedures should be followed (gloves, barrier tips, RNA ase-free needles).
8. Each component of the injection mix should be centrifuged at maximum (>16,000 × *g*) for a minimum of 10 min at 4 °C prior to pipetting to minimize debris carryover into the mix.
9. When assembling the injection mix, it is recommended use a maximum concentration of 800 ng/μl total nucleic acid as a higher concentration may block the needles.
10. Numerous steps can be taken to minimize clogging of the needle. The injection mix volume should be 30 μl or greater, preferably in a 0.5 ml tube. This volume allows that the needle-backfiller can draw fluid from the uppermost level of the mix. When backfilling the needle, the end of the microloader should be as close to the needle point as possible. Dispense the injection mix gently while withdrawing the microloader to minimize the occurrence of bubbles within the needle. The injection mix can be left spinning at max (~16,000 × *g*) at 4 °C in a benchtop centrifuge until it is needed for injection. While injecting, the injection mix should be left on ice and additionally should be spun down every 2 h for at least 10 min using previously stated parameters.

11. Ideally the aspirator can be adapted to have a narrow piece of tubing at the end (~5 mm inner diameter) that facilitates both the selection of gravid females and depositing them into the laying chamber, down the side of the stopper. Shake and blow the mosquitoes gently down the aspirator tube into the laying tube.
12. Check the laying chamber after 30 min for the appearance of embryos. If none can be observed then the oviposition time can be increased to 45–60 min. Multiple laying chambers can be setup simultaneously. The promptness of mosquitoes to oviposit varies based with each strain and time post-bloodmeal. We have found that most strains will lay eagerly after 4–6 days post blood meal but any further delay may result in reabsorption of the embryos or reduction in longevity.
13. Turn on the Femtojet Injector so that the injection pressure has already been achieved while the embryos are desiccating. This allows the operator to begin injecting shortly after moving the coverslip to the injection microscope rostrum.
14. Tamp down the embryos of the coverslip using an arm of the forceps to push the embryo more firmly into the adhesive tape before injection. Be aware that once the embryos are transferred they will begin to desiccate.
15. Place the coverslip upon the microscope stage or the microscope rostrum ready for injecting. This is manufactured by gluing four or five microscope slides together with the lower one offset to enable the microscope stage to grasp the edges while the light is not obscured by the specimen holder. Using a rostrum also allows greater variability in the angle of injection.
16. The angle of the injection can be accomplished between 0° and 20° but we prefer approximately 5°.
17. The injection mix should be flowing freely from the tip of the needle. This allows the operator to inject the embryo without actually using the inject function of the Femtojet Injector. The pressure compensation value (pc) should be varied dependent upon the bore of the needle and the viscosity of the mix. The flow should be continuous but not too fast as the disruption to the embryo's internal environment will increase mortality rates. If the yolk from the embryo is excessively spilling out then could indicate that the desiccation stage was too short or that the needle is too blunt.
18. Occasionally, it can be observed that embryos are being displaced from their adhesive tape position rather than pierced by the needle. This can occur when excess water from the transfer step causes the adhesive tape to soften, the pressure to stick them to the coverslip was not sufficient, when the needle is too blunt or when the embryos have darkened too much.

19. If the embryos are invaginating excessively before the needle is piercing the embryo then the needle maybe too blunt or the desiccation was too long.
20. If the needle becomes clogged with debris then the "Clean" function may be used to unblock the aperture. If this does not resolve the problem (very slow or no injection mix exiting the needle) then replace the needle.
21. After repeated use the needle will blunt and should be replaced. Replacement is dependent upon the actual injections but ideally after two or three lines when using borosilicate glass.
22. Removal of halocarbon oil can be achieved in different ways. We have found the easiest to be gently spraying the coverslip with a water spray bottle. The fine jet can be directed just above the line of embryos as the coverslip is held perpendicularly to some paper towels. This encourages the oil to drain off the embryos and also softens the adhesion of the tape easing their removal with forceps.
23. It is advised that the experimental and control plasmids use different promoters to express luciferase to minimize any promoter cross-talk between them. As a normalization control plasmid, either *pkhsp82-Renilla* or *pIE1-Renilla* can be used with good effect to express Renilla luciferase [7].
24. It is advisable to also collect up the non-transferred uninjected embryos to act as a negative control; this hatch rate will give a baseline for the survival rate. An additional more stringent control for the injection process may include the recovery of lined-up but uninjected embryos to assess the potential mortality caused by the physical lining up and recovery stages.
25. Samples can be stored at –80 °C at this stage. If samples have been stored then transfer them back into liquid nitrogen ready for homogenizing.
26. When using a pestle motor, add an initial volume of 30 μl and then crush as this will minimize liquid spraying out of the sample tube. Once the sample has been homogenized add the additional volume to make up the desired total.
27. LARII is stable for 1 month at –20 °C and 1 year at –80 °C. Make 1 ml aliquots of LARII and use 80 μl per sample.
28. Make up the 1× Stop & Glow buffer in a glass vial or siliconized polypropylene tube (a culture tube works well).
29. One obvious source of variation within the HRMA output can originate from SNPs within the sequence. SNPs can be accounted for if known samples are considered as standards; therefore sequence data of the target region for a number of individuals is advised. Where possible, the assay primers and

the gene-editing elements should be designed around an invariable region, therefore any indel events within the target region are more easily recognized.

30. One of the most crucial preparation steps is to identify an optimal annealing temperature for the PCR reaction that yields the product for analysis. It is therefore imperative that the template preparation is standardized to minimize variability.
31. Check the gel image of the gradient PCR products to ensure that the optimal temperature identified by the melt profile results in a specific PCR product of the predicted size.
32. Ideally the entire leg is immersed in the reaction volume. Be sure to wipe the forcep tips and legs with 70 % ethanol between individual samples to limit contamination.
33. To reduce chance of cross-contamination, maintain separate labeled disposable transfer Pasteur pipettes and small beakers (for food solutions) for each line.
34. If some eggs float on the surface of water, then trap them under a small piece of tissue. This prevents them from sticking to the side of the pan where a drop in water level through evaporation might cause them to become non-viable. It also helps ensure any remaining oil is siphoned away from the embryo surface.
35. Add a small amount of fish food in the pan. The pan with the injected-recovered egg papers should be maintained for at least 2 weeks to ensure all viable eggs will hatch. During this time the water should not be changed. If too much food is present then the water may turn cloudy. Only add more food when all has been consumed.
36. We recommend to have 25–30 G_0 females in each medium-sized cage.
37. If there is uncertainty about the size of male and female pupae (as the G_0 pupae sometimes differentiate less obviously than the wild type), place each pupa in a drosophila tube (28.5 mm × 95 mm) in a small amount of water and stopper the tube with a flug or cotton. Check the tubes every day and as soon as they emerge, release them in the male and females popcorn-cages.
38. Anesthetize adult mosquitoes using CO_2 then transfer them to a prechilled glass petri dish (150 mm × 20 mm) on ice. Be aware of the amount of time the mosquitoes are knocked out, as extended period of times anesthetized may have an effect on their longevity. Check for any contaminating adults of the other sex then transfer adults into the appropriate cage.
39. Place the female mosquito gently inside the tube as it is laid flat (rather than standing). This will allow the female to recover on a dry surface rather than against the damp paper.

40. Placing the dry oviposition beaker in the soup cup-cage at this same time saves repeatedly anesthetizing the adult mosquitoes after the mated females have blood fed. Also punch a hole in the side of the container to allow the addition of more pupae if necessary.
41. To store *Aedes aegypti* oviposition papers first gently but firmly blot the oviposition paper upon a paper towel or tissue. The amount of moisture within the paper is crucial, too much will encourage mold growth but too little will result in very few or no eggs hatching. Place the dried (but not desiccated) paper within a plastic pouch then seal the edges around the enclosed paper using a 2 mm Impulse Heat Sealer. In this manner the oviposition papers can be stored for up to 4 months at room temperature.

Acknowledgments

This protocol was developed with the support of grants [AI085091 and AI099843] from the NIAID. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIAID.

References

1. WHO (2014) Dengue and severe dengue: fact sheet No. 117. World Health Organization, Geneva
2. Aryan A, Anderson MA, Myles KM, Adelman ZN (2013) TALEN-based gene disruption in the dengue vector *Aedes aegypti*. PLoS One 8(3): e60082. doi:10.1371/journal.pone.0060082
3. Liesch J, Bellani LL, Vosshall LB (2013) Functional and genetic characterization of neuropeptide Y-like receptors in *Aedes aegypti*. PLoS Negl Trop Dis 7(10):e2486. doi:10.1371/journal.pntd.0002486
4. McMeniman CJ, Corfas RA, Matthews BJ, Ritchie SA, Vosshall LB (2014) Multimodal integration of carbon dioxide and other sensory cues drives mosquito attraction to humans. Cell 156(5):1060–1071. doi:10.1016/j.cell.2013.12.044
5. Jasinskiene N, Juhn J, James A (2007) Microinjection of *A. aegypti* embryos to obtain transgenic mosquitoes. J Visual Exp 5:219
6. Bassett Andrew R, Tibbit C, Ponting Chris P, Liu J-L (2013) Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. Cell Reports 4(1):220–228. doi:10.1016/j.celrep.2013.06.020, http://dx.doi.org
7. Aryan A, Myles KM, Adelman ZN (2014) Targeted genome editing in *Aedes aegypti* using TALENs. Methods 69:38–45. doi:10.1016/j.ymeth.2014.02.008
8. Coates CJ, Jasinskiene N, Miyashiro L, James AA (1998) *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. Proc Natl Acad Sci U S A 95(7): 3748–3751

Chapter 14

Methods for TALEN-Mediated Genomic Manipulations in *Drosophila*

Jiyong Liu, Yawen Guo, Changqing Li, Yixu Chen, and Renjie Jiao

Abstract

TALEN (transcription activator-like effector nuclease) is a powerful tool for gene disruption and other genomic modifications. In the past 3 years or so, it has attracted eyes from every corner of the biological world, due to its characteristics of simplicity, high efficiency, low toxicity, and applicability across almost all species. In our lab, we first reported the TALEN-mediated gene disruption in *Drosophila*, and recently employed this technique in precisely modifying the *Drosophila* genome, such as in vivo tagging and gene correction. Here, we describe in detail the protocols and experiences in TALEN-mediated genomic modifications to share with the Drosophilists all over the world.

Key words TALEN, *Drosophila*, Genomic modifications, NHEJ, HR

1 Introduction

TALEs (transcription activator-like effectors) are a family of transcriptional factors which are generated in the plant pathogen *Xanthomonas*, and translocated into the nuclei of its host cells to activate specific host genes [1]. Natural TALEs recognize the specific target sequence in a one module to one nucleotide manner. The TALE modules are identical in amino acid sequence except for the 12th and 13th aa of each repeat, namely the repeat-variable diresidues (RVDs) [2]. It is the RVDs that determine DNA recognizing specificity, which has been decoded as follows: HD to C, NG to T, NI to A, NN to A or G, and NS to A, C, G, or T [3].

TALENs are designed to work in a way similar to ZFNs (zinc-finger nucleases) [4] (Fig. 1a). Once manipulated with the expression of TALENs, the cells use two pathways to repair the DSBs (double-strand breaks) resulting from the TALEN activity: NHEJ (nonhomologous end joining) and HR (homologous recombination). NHEJ is more straightforward, but error prone, leading to small insertions and/or deletions (indels) at the DSB sites, which is ideal for disrupting genes. HR precisely repairs the DSB using a

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_14, © Springer Science+Business Media New York 2016

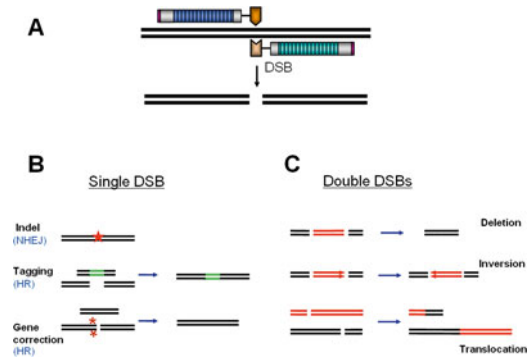


Fig. 1 A schematic illustration of TALEN-mediated genomic modifications. (**a**) The left and right TALEN bind to their adjacent target sequences leading to the formation of a heterodimer of the *Fok I* nucleases which cleaves the DNA in the spacer region to generate a DSB. (**b**) Cells repair the single DSBs through NHEJ or HR pathways, leading to three different outcomes as indicated. (**c**) Upon the generation of two different DSBs, large scale of chromosomal alterations such as fragmental deletion, chromosomal inversion, and translocation can be produced. Modified from Liu et al. [9]

homologous sequence as the template, which makes it possible to generate modifications such as nucleotides replacement, and in vivo tagging [5] (Fig. 1b).

In addition to small modifications as described above, TALEN can also be used to induce large-scale modifications of the genome, such as large fragmental deletion, inversion, and translocation upon the production of DSBs by two pairs of TALENs (Fig. 1c). The maximum deletion reported thus far is about 5.5 Mb in zebrafish [6], and in *Drosophila* this remains elusive.

In this chapter, we describe in detail the protocols of TALEN-induced mutagenesis (by NHEJ) (*see* Subheading 3.1) and other genomic modifications (by HR) (*see* Subheading 3.2) in *Drosophila*. NHEJ-based mutagenesis is mainly achieved within the following steps: TALEN construction and mRNA preparation, microinjection, and genetic and/or molecular screening of mutants. The entire process can be accomplished for an ordinary fly lab in about 1 month [7]. HR-based genomic modifications are similar to that of NHEJ except that donor DNA constructs have to be co-injected with the TALEN mRNAs.

2 Materials

2.1 Fly Strains

1. *yw*: for injection.
2. w^{1118}: for injection.
3. w^{1118}, $lig4^{169}$ (BL28877): for injection.

4. *yw*; *BCG/CyO*: for genetic screen and stock establishment.
5. *yw*; *TM3,Sb/TM6B,Tb*: for genetic screen and stock establishment.

2.2 TALEN Construction and Other Molecular Manipulations

1. Mono-/double-/triple- and tetra-units: to be used for Unit Assembly. These are not required if other strategies are employed for TALEN construction [8].
2. pCS2-PEAS and pCS2-PERR: for the construction of TALENs via Unit Assembly [8]. These are not required if other strategies are used for TALEN construction.
3. pBluescript KS: for the construction of homologous donors. Other kinds of vectors can be used as alternatives.
4. Plasmid Midi Kit (QIAGEN GmbH, Hilden, Germany): extraction of donor DNAs with high quality. Alternative kits can also be used.
5. Squishing Buffer (SB): 10 mM Tris–HCl, pH 8.2, 1 mM EDTA, 25 mM NaCl and 0.2 % Triton X-100. Used for genomic DNA preparation. Prepare on ice prior to use.

2.3 In Vitro Transcription of TALENs

1. SP6 mMESSAGE mMACHINE Kit (Ambion, Austin, TX, USA): for in vitro transcription of TALENs.

3 Methods

3.1 TALEN-Mediated Mutagenesis Through NHEJ

3.1.1 Selection of TALEN Targets

The selection of TALEN target sites can be carried out with the help of a series of programs such as TAL Effector-Nucleotide Targeter (TALE-NT) 2.0 (https://boglab.plp.iastate.edu/) (*see* **Note 1**). The specification of particular parameters is explained as follows (for more details, *see* Liu et al. [9]).

1. The repeat number of the designed TALENs. The typical number of TALE repeats for the natural TALEs is 17.5, with a range from 1.5 to 33.5 [1]. The reported minimum number of the repeats for a functional TALE is 10.5 [3]. We recommend the length of TALEN-binding sequence to be between 12 and 17 bp, based on the balance of binding specificity and practical construction of the TALEN pairs using Unit Assembly (*see* Subheading 3.1.3, **step 1**).
2. The length of the spacer sequence (the sequence between the left and right TALEN-binding sequence). TALENs function in a form of *Fok I* heterodimer. Improper DNA distance between the two TALENs may obstruct the formation of a heterodimer, and consequently lower the efficiency of TALENs. The reported spacer sequence for functional TALEN pairs normally range from 13 to 30 bp. In our experiments, the number is between 14 and 18 bp (*see* **Note 2**).

3. Following the characteristic of most naturally occurred TALEs, the TALEN-binding sequence usually starts after a T [10]. As an alternative, a preceding C is permissive, although it may compromise the TALE activity [11].
4. Evenly distribute the weak (NI, NG, or NK) and strong (HD or NN) RVDs, avoiding weak or strong RVD clusters. It was reported that more than 3–4 evenly distributed strong RVDs were important for the TALE activity [12]. Although NH or NK target "G" with more specificity [13, 14], NN can also to be used to recognize "G" to meet the requirement of incorporating of strong RVDs (*see* **Note 3**).

Following the above guidelines, functional TALENs can be designed to produce DSBs in the fly genome. To achieve the goals of site specific mutagenesis or other genomic modifications, two other points should be borne in mind:

1. If the purpose is to disrupt the function of a gene by NHEJ, the TALEN site is recommended to be located in the coding region, especially essential domains or the region that is not far after ATG start codon. If the purpose is to carry out precise genomic modifications by HR (*see* Subheading 3.2), please select a target site within or near the sequence to be modified.
2. To avoid possible off-target(s), a BLAST search is recommended to make sure that the selected target site is unique.

3.1.2 Confirmation of the Selected TALEN Target Sequences

To exclude potential single-nucleotide polymorphisms (SNPs), we should molecularly confirm the accuracy and authenticity of selected TALEN target sequence. Extract the genomic DNA of the TALEN recipient flies:

1. Amplify the DNA fragment of about 500 bp that includes the selected pair of TALEN-binding sites (*see* **Note 4**).
2. Sequence the amplified DNA to confirm the accuracy of the selected target sites. Attention should also be paid to the restriction enzyme site. If applicable, digest the amplified DNA fragments with the selected restriction enzyme (*see* **Note 2**) to make sure for complete digestion (*see* **Note 5**).

3.1.3 Construction of TALENs

The construction of a TALEN can be divided into two steps: assembly of TALE repeats and construction of a TALEN expression plasmid. The main challenge of TALEN construction is the assembly of the DNA binding arrays due to the high sequence homology of individual TALE unit. Several kinds of methods have been developed, including Unit Assembly [8], *R*estriction *E*nzyme *A*nd *L*igation (REAL) [15], and Golden Gate-based methods [16], which are practical for the majority of ordinary fly labs, and solid-phase-based strategies such as *F*ast *L*igation-based *A*utomatable *S*olid-phase *H*igh-throughput (FLASH) system [17],

and *I*terative *C*apped *A*ssembly (ICA) [18], which are developed for high-throughput production of TALENs. Another choice of TALEN construction is commercial service from the companies including Cellectis Bioresearch (http://www.cellectis-bioresearch.com/), Life Technologies (http://www.lifetechnologies.com/), and System Biosciences (http://www.systembio.com/).

1. Assembly of TALE repeats using Unit Assembly. This method starts with the existing library of mono-, double-, triple-, and tetra- TALE units. Using this strategy, one can finish the construction of a pair of TALEs (left and right TALEs) which recognize 12–17 bp DNA, respectively, through two rounds of standard digestion-ligations within 4 days (*see* **Notes 6** and **7**).
2. Construction of TALEN expression plasmids. Cut the complete TALE repeats and put corresponding fragments into relevant *Fok* I-containing vectors to get TALEN expression plasmids (*see* **Note 8**).

For more details of TALEN construction with Unit Assembly, please see our previously published protocols [9].

3.1.4 Preparation of TALEN mRNAs and Microinjection Solution

The TALEN mRNAs transcribed in vitro are directly injected into the embryos of the recipient flies.

1. Fully linearize the circular TALEN plasmids, purify the linear DNA using commercial gel extraction kit and dissolve the DNA in RNase-free water (*see* **Note 9**).
2. In vitro transcription of TALEN mRNAs. The left and right TALENs are transcribed in vitro, respectively, using the purified DNA as templates (*see* **Note 10**).
3. Recover the transcribed mRNAs and dissolve into RNase-free water (*see* **Note 11**).
4. Mix the left and right TALEN mRNAs for the microinjection solution with a final concentration of about 250–500 ng/μL each. Aliquot the injection solution and store at −70 °C (*see* **Note 12**).

3.1.5 Microinjection

The capped mRNAs of TALEN pairs are injected into the embryos in which they are translated into functional TALENs to cleave the target DNA.

1. Select appropriate flies for embryo injection. The genotype of embryos used for injection depends on the purpose of the experiment. Essentially, for NHEJ-based mutagenesis, it is recommended to use either *yw* or w^{1118} flies; for HR-based genomic modifications, it is recommended to use *lig4* flies (BL28877) (*see* **Note 13**). You can also directly inject the embryos that carry one or more other genetic elements/mutations such as FRT (*see* **Note 14**).

2. Microinjection of the TALEN mRNAs. Briefly, collect the embryos of the recipient flies for 30 min before dechorionating the embryos with sodium hypochlorite solution. After a thorough wash with water, the dechorionated embryos are carefully aligned on a slide before putting in a desiccator for drying about 30 min. The embryos are then covered with halocarbon oil and used for microinjection with the prepared injection solution (*see* Subheading 3.1.4, **step 4**).

For more details of fly embryo injection, please see previously published protocols by Bachmann and Knust [19] (*see* **Note 15**).

3.1.6 Screen and Establishment of Genetically Modified Fly Stocks

After injection of TALENs in F_0 flies, the heritable mutations can be molecularly identified in F_1 flies and the stable mutant stocks can be established in F_2. The crossing scheme is shown in Fig. 2a.

1. Singly cross the enclosed virgin F_0 flies with the opposite sex of double-balancer flies (*see* **Note 16**).
2. When sufficient crawling larvae are seen in the tube, sacrifice the F_0 fly, extract the single genome, PCR, carry out enzyme digestions to identify mosaic (somatic mutant) F_0 (*see* **Notes 17** and **18**). We detected 43.8 and 58.3 % mosaic male F_0s in *yellow* and *CG9797* mutagenesis, respectively [7] (Fig. 2b).
3. Collect F_1 virgin flies and singly cross with the double-balancer flies to establish stocks. For the mosaic F_0 flies, we suggest to pick 20 F_1 flies (ten males and ten virgin females) for the stock establishment, due to the higher probability of obtaining heritable mutation (*see* **Note 19**). For other F_0s, 10 F_1 flies are picked for stock establishment.

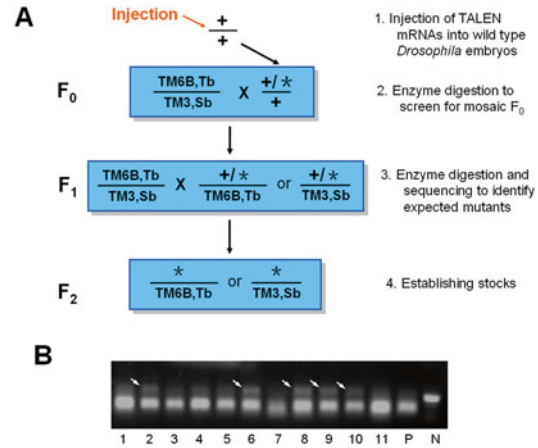


Fig. 2 Genetic and molecular work-flow-chart of TALEN-mediated mutagenesis through NHEJ. (**a**) Genetics and timing steps. (**b**) Molecular identification of *CG9797* mutations in F_0 flies. The undigestable DNA was detected in the lanes of 2, 6, 8, 9, 10, indicating alterations of wild-type genomic DNA in F_0 somatic cells. *P* positive control, *N* negative control. Modified from Liu et al. [7]

4. The molecular identification of heritable mutants. Similar to **step 2** of this section, sacrifice the F_1 flies and extract their single genomes for PCR and enzyme digestions after sufficient F_2 offspring have been produced. The F_1s that carry half of undigestable DNAs are candidate mutants (*see* **Note 20**).
5. Extract the undigestable DNAs for sequencing to confirm the mutations (*see* **Note 21**).

3.2 TALEN-Mediated Genomic Modifications by HR

TALEN-mediated genomic modifications by HR depend on the presence of an exogenous donor during the process of DSB repair. The main difference between HR and the TALEN-mediated NHEJ is that for the former one needs to construct a donor plasmid, co-inject the donor with relevant TALENs, and the screening strategy may also need to be modified accordingly.

3.2.1 Preparation of TALENs

For preparing TALENs, from target sites selection to generation of capped TALEN mRNAs, one needs to follow what have been described in the above sections (*see* Subheadings 3.1.1–3.1.4, **step 3**).

3.2.2 Preparation of Donor DNAs

1. The donor DNAs containing the designed modifications such as nucleotides replacement, deletion, and tag insertion and the flanking left and right homolog arms are cloned into a donor plasmid (we used the vector of pBluescript KS, other vectors would also be OK). In our experiments, the homologous arms ranged from 1.0 to 4.2 kb, with a typical length of 1.5 kb [5] (*see* **Notes 22** and **23**).
2. In the TALEN-mediated HR experiments, we suggest to use a circular DNA as the donor. It has been shown that the circular donor is more efficient than linear DNAs in the experiment of ZFN-mediated HR [20]. However, it is unclear whether this is also the case for TALEN-mediated HR.

3.2.3 Microinjection

Different from the TALEN-mediated NHEJ, here donor DNA has to be co-injected with the TALENs (*see* **Note 24**).

1. Extract the donor DNAs using Plasmid Midi Kit (*see* Subheading 2.2, **item 4**) to get good-quality DNA, especially to be completely free of RNase. Dissolve the DNA in RNase-free water.
2. Mix the donor DNAs with the TALEN mRNAs to get a final concentration of about 500–750 ng/μL for the donor DNA, and about 250–500 ng/μL for each TALEN mRNA (*see* **Note 24**). Aliquot the mixed solution and store at −70 °C.
3. Microinjection of the prepared solution according to the standard procedure as described by Bachmann and Knust [19] (*see* **Note 15**). The genotype of the recipient embryos should be *lig4* (*see* **Note 13**).

3.2.4 Screen for and Identification of the Modified Flies

The genetic screen, identification, and stock establishment methods are similar to as described above (*see* Subheading 3.1.6) with modified identification strategies depending on different modifications.

1. For HR-mediated deletion, one can identify the deletion by detecting a shortened PCR product, or loss of a priming sequence.
2. For HR-mediated insertion, one can identify the insertion by detection of an exogenous restriction enzyme site or other markers, such as w^+, eGFP, and Myc.

3.2.5 Distinguish Ends-Out from the Ends-In Events

The purpose of HR-based genomic modifications is to precisely modify the target sequence through ends-out recombinations. However, it is possible for ends-in events to occur as well during the process of DSB repair since the co-injected donor DNA is circular, leading to unwanted modifications [5, 21]. Therefore, it is important to bear in mind that sufficient mutant candidates should be examined to make sure you will get the designed ends-out HR recombination events. It is not clear for now what determines that ratio of ends-out versus ends-in events, or whether this ratio depends on the design of the donor constructs such as the lengths of the homologous arms, or the relative distance between the DSB site and the homologous arms.

4 Notes

1. Other Web-based programs for TALE/TALEN target designing include E-TALEN (http://www.e-talen.org); idTALE (http://idtale.kaust.edu.sa/index.html); Mojo Hand (http://www.talendesign.org); TALEN Hit (http://talen-hit.cellectis-bioresearch.com/search); and ZiFit Targeter (http://zifit.partners.org/ZiFiT).
2. For the convenience of mutation identification, we strongly recommend to choose a pair of TALEN-binding sequences spanning a spacer that contains an efficient restriction enzyme site. The disruption of the DNA sequence in the spacer may also affect the restriction site, which can be detected by the following enzyme digestions (*see* Subheading 3.1.6, **steps 2** and **4**).
3. NN is an efficient RVD to target "G," but also recognizes "A" [3].
4. Please make sure that the PCR is efficient enough for the following molecular screens (*see* Subheading 3.1.6, **steps 2** and **4**).
5. Unless you will not take the PCR-enzyme digestion strategy to screen for mutations, this step is compulsory.

   Highly efficient enzyme digestion is recommended for the mutant screening (*see* Subheading 3.1.6, **steps 2** and **4**).

   It is important that one can easily distinguish the digested DNA from the undigested.

6. We use colony PCR to identify the positive clones with the following primers: M13-47 (5′-CGCCAGGGTTTTCCC AGTCACGAC-3′) and RV-M (5′-AGCGGATAACAATTT CACACAGGA-3′). Please note that each TALE unit is 102 bp and the amplified backbone sequence is 156 bp [8].
7. It is recommended to do sequencing to check the constructed TALE repeats. The mentioned primer pairs of M13-47 and RV-M are also used for sequencing.
8. In the TALEN-mediated DNA cleavage, *Fok* I works as a heterodimer. Thus the left and right TALE repeats are cloned into pCS2-PEAS and pCS2-PERR, respectively [8].
9. From purification of the linearized DNA to storing of the prepared injection solution, all steps should be strictly RNase free.

   The final concentration of DNA templates depends on the reaction system of in vitro transcription kit. Take SP6 mMESSAGE mMACHINE Kit as an example: a concentration of more than 1 μg/6 μL is required for a 20 μL reaction setup.
10. Set up a proper reaction system depending on your experiment. Generally, half of the standard reaction setup of the Ambion SP6 mMESSAGE mMACHINE Kit (10 μL) is sufficient for a NHEJ mutagenesis or other kinds of HR-mediated genomic modifications.

    Be aware that mature TALEN mRNAs are required for injection; thus it is recommended to Kits that are designed to synthesize capped and poly-adenylated mRNAs.
11. We use lithium chloride precipitation to recover the transcribed mRNAs following the instruction of the Ambion kit.

    It is advisable to check the quality of the transcribed products.

    Quantify to determine the concentration of the mRNA solution with a photometer. It is recommended to get a final concentration of no less than 500 ng/μL.
12. Please mix the solution thoroughly. 30 μL of mixed solution is sufficient for one mutagenesis experiment.
13. DNA ligase IV (*lig4*) is an essential component of NHEJ-mediated DSB repair. It is reported that deficiency of *lig4* suppresses NHEJ and promotes HR [20].
14. It is important that the flies to be used for injection are healthy.
15. Bear in mind to avoid possible RNase contamination. We suggest relatively low amount (about 2 μL) loading of the injection solution each time.

    Centrifuge the injection solution at the maximum speed (about 13523 × *g*) for 2 min before it is loaded into the injection needle, avoiding possible blockage of the needle.

    The number of injected embryos (F_0) needs to be estimated properly. In our experiments of *yellow* and *CG9797* mutagenesis,

only about a quarter of the hatched larvae developed into fertile flies. About 17.2–66.7 % of the fertile F_0 yielded mutant offspring varying in different cases [7]. We suggest at least 50 single crosses of F_0 for a NHEJ mutagenesis, and about 100–200 single crosses for the HR-based genomic modifications.

16. We used a double-balancer fly stock for single crosses with the F_0 to balance any possible mutant chromosomes. For the mutagenesis of second chromosome genes, *BCG/CyO* flies can be used. For the third chromosome genes, *TM3,Sb/TM6B,Tb* flies are a good choice.
17. This step is optional. But the identification of somatic mutant F_0 may help to increase the chance of getting heritable F_1 mutations with the same amount of screening work (*see* Subheading 3.1.6, **step 3**).
18. Extraction of a single genome: Mix the squishing buffer (SB) (*see* Subheading 2.2, **item 5**) and 1 % (V/V) proteinase K (20 mg/mL) to get the extraction solution on ice; homogenize single F_0 flies each in the tube with 5 μL extraction solution and then add 35 μL extraction mixture; mix well and incubate at 37 °C for 30 min and at 95 °C for 5 min.

    PCR: Take 5 μL of the single-genome extract as the template for a 50 μL PCR. The primers used in Subheading 3.1.2, **step 1** are utilized here.

    Enzyme digestions: Take about 500 ng amplified DNA (if the enzyme is efficient enough, the PCR product can be directly used for digestion, or it is necessary to do an ethanol-precipitation of the amplified DNA) for enzyme digestion (*see* Subheading 3.1.2, **step 2**). Because of the mosaicism of the mutation in F_0 somatic cells, any remaining of uncut DNA after complete digestion is an indication of mutation event.
19. In the TALEN-mediated mutagenic experiments for *yellow* and *CG9797*, 65.4 and 85.7 % mosaic F_0 males, respectively, yielded expected mutant F_1 flies [7].
20. Because the candidate F_1 flies are heterozygous for the expected mutation, only one copy of the interested gene is disrupted which may not be digestable any more by the enzyme that cuts only the wild-type genome.

    One can also perform the molecular identification work with the homozygous F_3 flies if you wish and time permits.
21. It is worth to know these facts: First, in some cases, it is impossible to get heritable mutation by TALENs, due to severe lethality of the injected embryos for unknown reasons. Secondly, in some cases, the designed TALENs do not work efficiently. Therefore, it may be wise to simultaneously design more than one pair of TALENs that target the same gene.

22. It is reported that unlimited extension of the homolog arm does not apparently improve the HR efficiency [20].

    It is recommended to introduce acceptable modifications in the donor to avoid the binding of TALENs to the donor DNA.

23. Please use high-fidelity DNA polymerases to amplify DNA to avoid possible introduction of un-designed alterations into the PCR products. Sequencing is recommended to exclude the possibility of introduction of extra mutations, especially the intolerable mutations.

24. Please be aware of RNase-free protections that are similar to what is required in Subheadings 3.1.4 and 3.1.5.

Acknowledgement

We thank Dr. Bo Zhang for providing reagents and stimulating discussions. Thank Ms. Xuehong Liang for technical assistance. This work was supported by the grants from NSFC (Nos. 31201007 and 31100889) and the 973 grant 2012CB825504.

References

1. Boch J, Bonas U (2010) *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. Annu Rev Phytopathol 48:419–436
2. Bogdanove AJ, Voytas DF (2011) TAL effectors: customizable proteins for DNA targeting. Science 333:1843–1846
3. Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326: 1509–1512
4. Bibikova M, Golic M, Golic KG et al (2002) Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. Genetics 161:1169–1175
5. Yu Z, Chen H, Liu J et al (2014) Various applications of TALEN- and CRISPR/Cas9-mediated homologous recombination to modify the *Drosophila* genome. Biol Open 3:271–280
6. Gupta A, Hall VL, Kok FO et al (2013) Targeted chromosomal deletions and inversions in zebrafish. Genome Res 23:1008–1017
7. Liu J, Li C, Yu Z et al (2012) Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. J Genet Genomics 39:209–215
8. Huang P, Xiao A, Zhou M et al (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29:699–700
9. Liu J, Chen Y, Jiao R (2014) TALEN-mediated *Drosophila* genome editing: protocols and applications. Methods 69:22–31
10. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501
11. Doyle EL, Booher NJ, Standage DS et al (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. Nucleic Acids Res 40:W117–W122
12. Streubel J, Blucher C, Landgraf A et al (2012) TAL effector RVD specificities and efficiencies. Nat Biotechnol 30:593–595
13. Cong L, Zhou R, Kuo YC et al (2012) Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. Nat Commun 3:968
14. Christian ML, Demorest ZL, Starker CG et al (2012) Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS One 7, e45383
15. Reyon, D., Khayter, C., Regan, M.R., et al. (2012) Engineering designer transcription activator-like effector nucleases (TALENs) by REAL or REAL-Fast assembly. In: Frederick, M. et al. (ed) Current protocols in molecular biology 100:12.15:12.15.1–12.15.14

16. Engler C, Gruetzner R, Kandzia R et al (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. PLoS One 4:e5553
17. Reyon D, Tsai SQ, Khayter C et al (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30:460–465
18. Briggs AW, Rios X, Chari R et al (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40:e117
19. Bachmann A, Knust E (2008) The use of P-element transposons to generate transgenic flies. Methods Mol Biol 420:61–77
20. Beumer KJ, Trautman JK, Bozas A et al (2008) Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. Proc Natl Acad Sci U S A 105:19821–19826
21. Rong YS, Golic KG (2000) Gene targeting by homologous recombination in *Drosophila*. Science 288:2013–2018

Chapter 15

Targeted Mutagenesis in Zebrafish by TALENs

Peng Huang, An Xiao, Xiangjun Tong, Shuo Lin, and Bo Zhang

Abstract

Zebrafish is a valuable model organism to study vertebrate development, organ regeneration and to generate human disease models. As an important member of the arsenal of genome editing, TALE nucleases (TALENs) have implicated in broad applications in zebrafish reverse genetic studies. In this chapter, we describe the detailed protocols of TALEN-mediated genome manipulations in zebrafish, including targeted gene disruption by indel mutations, deletion of large genomic regions by using two pairs of TALENs, and precise genome modification by homologous recombination (HR).

Key words Zebrafish, TALE nuclease, TALEN, Gene targeting, Reverse genetics, Genome manipulation, Gene disruption, Targeted mutagenesis, Chromosomal deletion, Homologous recombination

1 Introduction

As a vertebrate model organism, zebrafish (*Danio rerio*) has showed a unique power for studying gene functions of embryonic development and organ regeneration, as well as in modeling human diseases, due to the transparency of its embryos, rapid and in vitro development, as well as its tremendous reproductivity. However, efficient reverse genetics techniques have not been established until recently, thanks to the fast development of the revolutionary gene targeting technology based on engineered endonucleases (EENs), including zinc finger nuclease (ZFN), TALE nuclease (TALEN), and very recently, the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) system [1–4].

Our and others' studies demonstrated that TALENs held great promise to achieve targeted genome modifications for a variety of applications in zebrafish. TALENs can create DNA double strand breaks (DSBs) at their target sites, which can be repaired through homologous recombination (HR) pathway in the presence of a homologous donor template, or through an error-prone nonhomologous end-joining (NHEJ) pathway [5, 6]. Based on these

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_15,

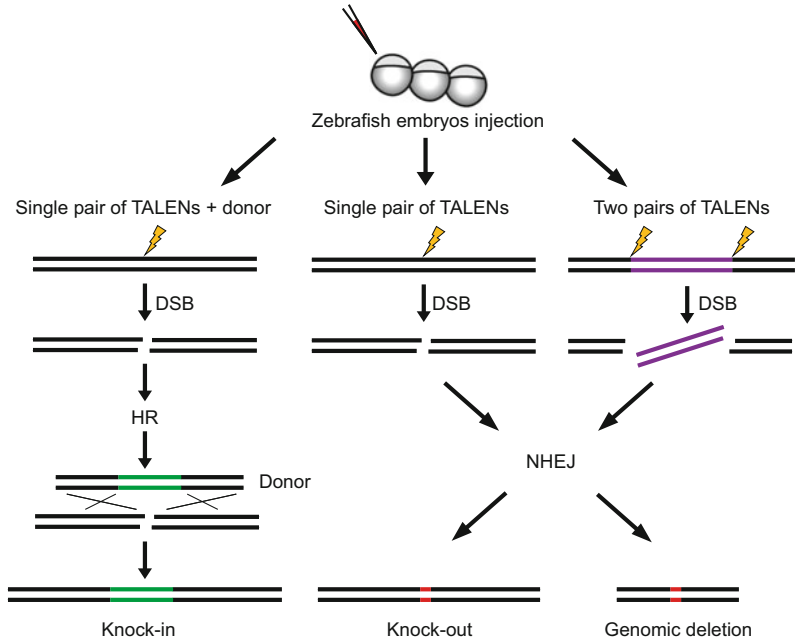


Fig. 1 TALEN-mediated genome manipulations in zebrafish. After being introduced as mRNAs into zebrafish embryos by microinjection, TALEN proteins will cleave the genome at the target sites in pairs and generate DSBs. In the presence of a homologous donor template, the chromosome can be repaired through HR pathway, by which knock-in mutations could be achieved. Without homologous templates, the DSBs are repaired by the error-prone NHEJ pathway, which could lead to indel mutations at the TALENs target site or even large genomic deletions between two TALENs target sites

mechanisms, different TALEN-mediated genome manipulation approaches have been reported in zebrafish (Fig. 1) [7, 8]. Since DSB can increase HR efficiency, co-injection of TALENs mRNA and a homologous donor template can theoretically achieve site-specific knock-in of any desired DNA sequence, including big reporter genes such as GFP or short DNA fragments such as *loxP*, or other types of precise genome modification [8]. Genome editing through HR can precisely modify genomes and is a powerful tool for gene function study or disease modeling. More complicated genome manipulations, such as conditional gene knockout in a tissue or stage specific manner, also rely on efficient HR. Nevertheless, the simplest way to disrupt a gene through TALEN is to introduce indel mutations at the target site by using one single pair of TALENs. Indel mutations could cause frameshift and abolish gene function when the mutations happened in the exons. However, sometimes a simple indel mutation might not be enough to fully disrupt the function of the target gene, especially when the target genes are huge in size and/or complicated in expression, such as having multiple transcripts, multiple splice variants, or alternative translation initiation sites. In some other cases, one might need to disrupt multiple adjacent genes or gene clusters,

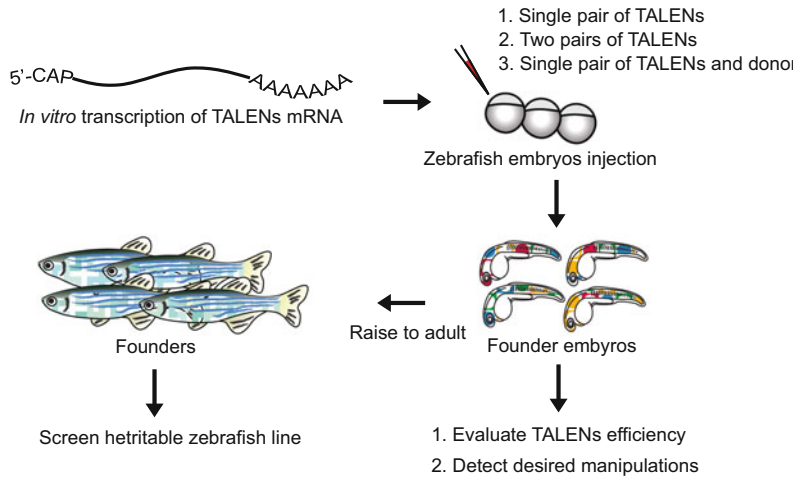


Fig. 2 Overview of genome manipulations using TALENs in zebrafish. TALEN mRNAs are synthesized by in vitro transcription. To evaluate the cleavage efficiency of TALEN pairs, a single pair of TALEN mRNAs is injected into zebrafish embryos at one-cell stage. At 2 dpf, the injected embryos are used to evaluate TALENs efficiency. After assessment, different TALEN-based mutagenesis approaches can be carried out. The injected embryos are then tested for targeted manipulations in somatic cells from founder embryos at 2 dpf. Stable mutant zebrafish lines are identified after the injected founder embryos are raised to adults

where indel mutations are also not helpful. As a complement, these problems can be solved by co-injection of two pairs of TALENs, which can delete the whole genomic region between the two target sites and therefore ensure the complete disrupt of the target gene(s). This method is also useful for the study of the functions of noncoding genes as well as gene regulatory elements [9–12].

In this chapter, we provide the detailed protocols for the generation of and screening for targeted zebrafish mutants generated by microinjection of TALEN mRNAs into zebrafish embryos. We will mainly focus on the instruction for three types of zebrafish genome manipulations, including targeted gene disruption by indel mutations, large genomic deletion by two pairs of TALENs, and gene knock-in by HR, which should have covered most of the important applications of gene targeting practice in zebrafish (Figs. 1 and 2). We believe that these methods will be a valuable resource for the zebrafish community.

2 Materials

2.1 Reagents and Solutions for Molecular Biology Experiments

1. NotI (or other appropriate restriction enzymes) and reaction buffer: For linearization of TALEN expression vector to prepare the template for in vitro transcription.
2. Other restriction enzymes (REs): For the evaluation of TALEN efficiency. Variable according to different target sites.

3. SP6 (or other appropriate promoters) mMessage mMachine kit (Ambion, USA): For making TALEN mRNA through in vitro transcription.
4. PCR purification kit.
5. Gel extraction kit.
6. Primers: To amplify target regions. Variable according to different target sites.
7. PCR DIG Probe Synthesis Kit (Roche, USA).
8. 50 mM NaOH.
9. 1 M Tris–HCl (pH 8.0).
10. Lysis buffer: 10 mM Tris–HCl pH 8.2, 10 mM EDTA, 200 mM NaCl, 0.5 % SDS, 200 μg/mL proteinase K.
11. Solution I (Takara, Japan) or T4 DNA ligase.
12. RNase-free water, tips, and Eppendorf tubes.
13. pMD18-T simple vector or other equivalents: For TA-cloning.
14. RNaseZap (Ambion, USA) (*optional*).
15. Hybond-N+ membrane (Amersham, USA).

2.2 Reagents, Consumables, and Equipment for Zebrafish Husbandry and Microinjection

1. Wild-type zebrafish or other desired zebrafish strains.
2. Mating tanks (for zebrafish).
3. E3 embryo buffer: 5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$.
4. Stereomicroscope.
5. Microinjection molds or other equivalents: To hold zebrafish embryos for microinjection.
6. 0.5 % phenol red.
7. Dumont #5 Tweezer (Inox, 11 cm) or other equivalents: To break injection needles.
8. 1 μL disposable capillaries (R: 0.25 %, CV: 0.6 %): For quantification of injection volumes (CAMAG, Switzerland).
9. Glass capillaries: For making injection needles (e.g., O.D. 1.0 mm, I.D. 0.58 mm; Harvard Apparatus, USA).
10. PN-30 Puller (Narishige, Japan) or other equivalents.
11. Nitrogen gas and tank.
12. MPPI-2 Pressure Injector (Applied Scientific Instrumentation, USA), or PLI-90 Pico-Injector (Harvard Apparatus, USA), or other equivalents.
13. Microloader tips (Eppendorf, USA): For filling the microinjection capillaries/needles (*optional*).

3 Methods

3.1 Preparation of TALEN mRNAs by In Vitro Transcription

1. We use "Unit Assembly" method to construct TALEN expression cassettes into pCS2 vector [5], which has a SP6 promoter upstream to the coding region and can be linearized for making mRNA through in vitro transcription. For linearization, 20 μg TALEN expression plasmids (called pCS2-TALEN) are digested by NotI or other unique restriction enzyme after the TALEN coding region overnight at 37 °C (*see* **Note 1**). To monitor the extent of linearization, load 1 μL of the reaction mixture to 0.8 % agarose gel and examine by electrophoresis (Fig. 3a) (*see* **Note 2**).
2. When the digestion is complete, purify the linearized template by using a DNA purification kit and elute with 20 μL nuclease-free water (*see* **Note 3**). Determine the concentration of the linearized template by a spectrophotometer (e.g., NanoDrop) (*optional*: *The linearized plasmid can be stored at −20°C and used later*).
3. We use the SP6 mMessage mMachine kit to synthesize capped mRNA (*see* **Note 4**). To set up an in vitro transcription reaction, mix 10 μL 2× NTP/CAP with 2 μL 10×reaction

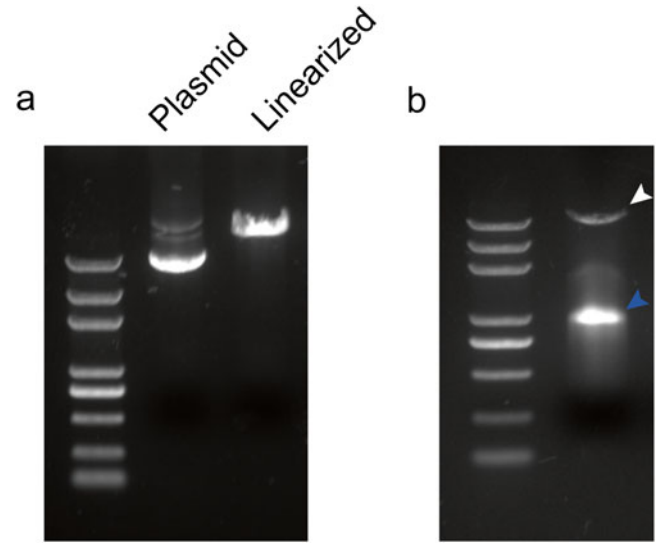


Fig. 3 Linearization of DNA template and in vitro transcription of TALENs mRNA. (**a**) Linearization of the template DNA (plasmid) was monitored by electrophoresis after NotI digestion. The original undigested plasmid was used as a control. No band of circular plasmid is visible after the template being completely linearized. (**b**) After in vitro transcription, TALEN mRNA can be visualized by electrophoresis. The majority of mRNA is shown as one band, which is indicated by the *blue arrow*. The *white arrow* indicates the linearized DNA template

buffer, 1 μg linearized DNA template from the above step, and 2 μL SP6 enzyme mix, and then supplement the volume to 20 μL with nuclease-free water, mix well by pipetting. Incubate the mixture at 37 °C for 2 h. To monitor the mRNA synthesis, load 0.5 μL reaction mixture to 1 % agarose gel and examine by electrophoresis (Fig. 3b) (*see* **Note 5**).

4. If the transcription is successful, add 1 μL TURBO DNase supplied by the kit and incubate at 37 °C for 15 min to remove the DNA template.
5. Stop reaction by adding 30 μL LiCl and 30 μL nuclease-free water provided by the kit. Mix well and store at −20 °C for at least 30 min. Then centrifuge at 4 °C for 15 min at top speed. The RNA pellet should be visible at the bottom of the tube. Remove the supernatant and wash with 1 mL cold 70 % ethanol. Centrifuge at 4 °C for 10 min at top speed. Remove the 70 % ethanol, air-dry the pellet, and add 50 μL nuclease-free water to dissolve the pellet (*see* **Note 6**). Determine the concentration of mRNA by a spectrophotometer and aliquot them into small volumes (e.g., 5 μL). Store the aliquots at −80 °C for later use and long-term storage (*see* **Note 7**).

3.2 Microinjection of TALEN mRNA Pairs into Zebrafish Embryos

1. The adult zebrafish pairs intended to be used for gene targeting should be genotyped to confirm the sequence of the TALEN target site (*see* **Note 8**). In the afternoon before the injection day, set up several mating tanks. Put one pair of zebrafish in each tank and separate the male from the female by a divider. In the morning of injection day, remove the divider from one tank each time and collect embryos as soon as possible.
2. We use 1 μL disposable capillaries to calibrate the injection needles. Press the foot switch and count how many drops are delivered into the capillary. The total volume can be measured by the length of the liquid in the capillary by using a ruler. Then the volume of each drop can be calculated by dividing the total volume by the number of drops (*see* **Note 9**).
3. TALEN mRNAs usually need to be injected in pairs (one target site corresponds to two TALEN binding sites). They can be injected into the cytoplasm or the yolk of one-cell stage zebrafish embryos. Usually we inject 100–300 pg mRNA for each pair of TALENs into the cytoplasm (*see* **Note 10**). When using different batches of mRNA, we recommend reevaluating the optimal injection dosage (*see* **Note 11**).
4. After injection, incubate the embryos in E3 embryo buffer at 28.5 °C. Save some uninjected sibling embryos as a control and process them the same as the injected ones for the determination of the efficiency of TALEN pairs. The dead and deformed embryos are counted and removed at 5–6 h post-fertilization (hpf) and 1 day post-fertilization (dpf).

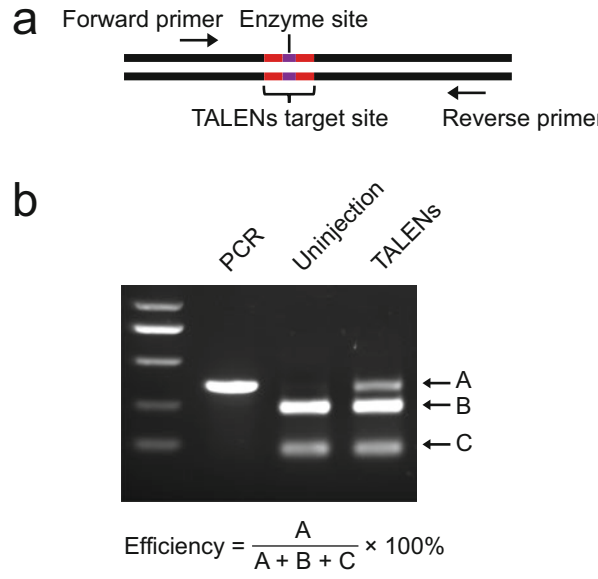


Fig. 4 Restriction enzyme digestion assay (or called RFLP assay) to evaluate TALENs efficiency. (**a**) Schematic diagram of primers used to amplify the genomic fragment spanning the TALEN target site by PCR. (**b**) In this case, the PCR fragments were digested with BamHI. The uninjected embryos were used as a control in which the PCR fragment was completely digested, resulting in two smaller bands (bands B and C) comparing with the original PCR product. In the injected embryos (TALENs), certain fraction of the PCR fragment (band A) was resistant to the enzyme digestion due to the disruption of the enzyme site (BamHI in this case). The efficiency of TALENs can be estimated by calculating the percentage of the uncuttable fragment (band A) out of the total PCR product (sum of bands A, B, and C) by measuring the intensity of each band. In this case, the efficiency of the TALENs is about 20 %

3.3 Evaluation of TALEN Efficiency in Founder Embryos

1. We generally use restriction enzyme-resistant assay to evaluate the efficiency of TALEN pairs and also for the screening of heritable mutations (Fig.4a) (*see* **Note 12**). This method is simple and cost-effective, and can obtain both the efficiency data as well as the sequence of the indel mutations from the same sample. If no proper enzyme site(s) are available at the target site (ideally to be in the spacer region between the two TALEN binding sites), other methods can also be adopted, such as CEL-I or T7E1 enzyme assay, melting curve assay and sequencing.
2. The genomic DNA of control (uninjection siblings) and injected embryos is extracted using NaOH lysis method [13]. Five normally developed 2–4 dpf embryos are pooled into one PCR tube. Remove extra E3 buffer and add 50 μL 50 mM NaOH. The embryos are lysed by heating for 10 min at 95 °C using a PCR machine, then cool down to 4 °C. Vortex the tube briefly to break up the embryos, then add 5 μL Tris–HCl

(pH 8.0) to neutralize NaOH and centrifuge for 5 min at 12,000 rpm (~13,500×*g*). The supernatant contains crude genomic DNA and is ready to be used as PCR templates.

3. Take 1 μL crude genomic DNA extract as template and assemble a 10 μL reaction, then perform PCR (*see* **Note 13**). After amplification, take 2 μL PCR products and digest with the proper restriction enzyme for more than 2 h. Run the DNA on a 2 % agarose gel. If the DNA from control embryos is digested completely, the efficiency of the TALEN pair can be estimated by measuring the intensity of each band and calculate the percentage of the resistant band (=uncuttable band, due to mutations in the restriction site) (Fig. 4b).
4. To further confirm the genomic lesions, the uncuttable (resistant) band can be cut out and extracted from the gel and verified by sequencing after cloning into TA-cloning vectors.

3.4 Gene Disruption by Indel Mutations

1. If the TALEN pair showed a measurable targeting efficiency (e.g., >1 %; >10 % is better) and the injected embryos showed a reasonable survival rate (e.g., >70 %) during efficiency evaluation, enough amount of the same batch of embryos are raised to adult as founders (F_0) for the screening of heritable mutations (*see* **Note 14**).
2. The mosaic F_0 fish are outcrossed with wild-type zebrafish. After breeding, each F_0 fish is placed and raised separately in a single tank until the F_1 embryos are evaluated. The F_1 embryos collected from each individual F_0 are labeled correspondingly and used to screen for stably inherited genomic lesions. We use the same strategy as described in Subheading 3.3 to identify stable mutations in F_1 embryos. We normally screen 40 (at least 20) F_1 embryos for each F_0 fish. If the TALEN efficiency is higher than 3 %, usually more than one positive F_0 fish (i.e., the fish that carry heritable mutations in its germ cells) can be identified from ten founders (*see* **Note 15**).
3. Once a positive F_0 fish has been identified, the indel mutations are further verified by sequencing (Fig. 5). If the indel mutations are satisfactory and will most likely disrupt the expression and/or function of the target gene (e.g., those mutations that will lead to frameshift and/or generate premature stop codons), enough amount of the rest of the same batch of F_1 embryos or the F_1 embryos collected from further breeding of the same F_0 fish are raised to adulthood (*see* **Note 16**).
4. F_1 zebrafish heterozygous for the mutation of interest (i.e., the F_1 carrier) are identified by genotyping of the genomic DNA from fin clips from each individual. The genomic DNA is isolated and evaluated similarly as for the embryos described in Subheading 3.3. Once the indel mutations are verified, allele PCR strategy can be used to identify the desired F_1 carriers

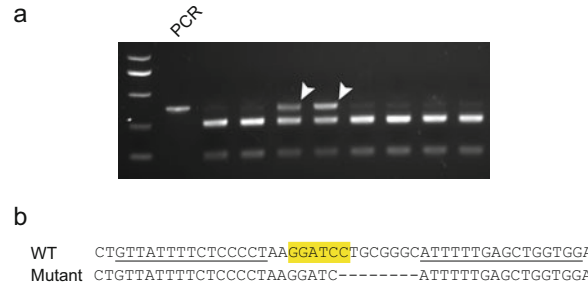


Fig. 5 Screening for germ-line transmission of indel mutations from founders. (**a**) The offspring (mixed F_1 embryos) of each founder by outcross were screened using restriction enzyme digestion assay. The white arrows indicate the PCR fragments carrying indel mutations which are resistant to enzyme digestion. (**b**) The PCR fragments from (**a**) were gel extracted and the mutant allele was verified by sequencing. TALENs binding sites are shown in underline. BamHI site is highlighted in *yellow*

using a specific primer against the mutation allele. Ideally, at least one pair of male and female adult heterozygotes should be recovered and maintained for further analyses for each particular indel mutation. If not in pairs, one need to outcross the heterozygous F_1 and screen for paired heterozygous F_2.

5. Homozygous zebrafish mutants can be obtained by in-cross of pairs of heterozygous carriers (best to be siblings). Individual embryos can be genotyped by genomic DNA extraction followed by PCR and RE digestion or sequencing, even after fixation and processing for in situ hybridization.
6. For long-term storage of zebrafish mutations, sperms of the heterozygous or homozygous (if available) mutants can be frozen through cryopreservation and kept in liquid nitrogen [14].

3.5 Disruption of Gene/Gene Clusters or Noncoding Regions by Genomic Deletions

1. To completely disrupt a gene, multiple adjacent genes/gene clusters, or other large genomic regions by deletion mutations, a pair of TALEN target sites (i.e., four TALEN binding sites) are selected flanking the genomic region to be deleted (*see* **Note 17**). Individual TALEN expression vectors are constructed and transcribed into mRNA as described in Subheading 3.1. The mutagenesis efficiency of each TALEN pair is evaluated as described in Subheadings 3.2 and 3.3. TALEN pairs showing high targeting efficiency (e.g., >30 % for each target site) can be used for co-injection and to screen for genomic deletion mutations (*see* **Note 18**).
2. Two pairs of TALEN mRNAs are co-injected into zebrafish embryos. To examine the success of deletion, the genomic regions containing the target sites are amplified by PCR (Fig. 6a). The shorter amplicons with appropriate length usually represent the alleles with deletions (Fig. 6b).

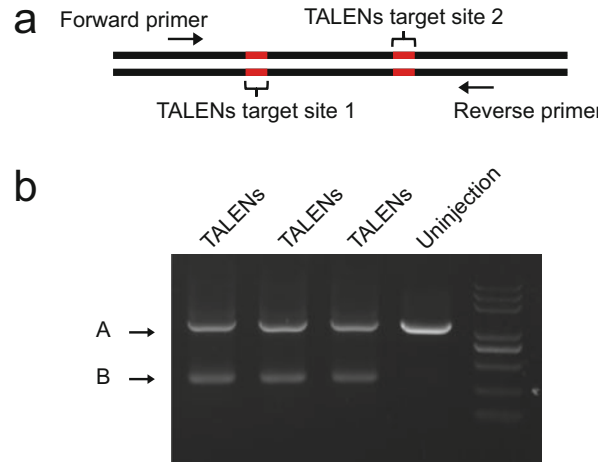


Fig. 6 Evaluation of genomic deletions induced by two pairs of TALENs. (**a**) Schematic diagram of primers used to amplify genomic fragment. (**b**) In this case, PCR fragments amplified from uninjected embryos were used as a control and only one longer PCR product (band A) was obtained. In the injected embryos, an additional and shorter PCR product (band B), which corresponds to the deletion allele, was clearly visible

3. We use the same strategy as described in Subheading 3.4 to raise and outcross the mosaic F_0 fish, screen F_1 embryos, as well as identify and verify the deletion alleles in F_1 adults and obtain homozygous mutants, except that no RE digestion is necessary for the identification of large deletion mutations.

3.6 Precise Genome Modification Through Homologous Recombination Induced by TALENs

1. The selection of TALEN target sites, preparation and microinjection of pairs of TALEN mRNAs, and evaluation of efficiency of inducing indel mutations are as described. TALEN pairs showing high indel-inducing efficiency (>50 %) are selected for the following homologous recombination experiments.
2. Selection of proper parental fish for HR: Isolate the genomic DNA individually from the fin clips of at least ten pairs of healthy adult fish (better to be ten pairs of siblings derived from the same inbred pair of fish) and sequence the flanking regions of the TALEN target site. Choose the fish pairs with homogeneous and absolute identical sequence in the region to be used as homologous arms for the following experiments (*see* **Note 19**).
3. Construction of the donor vector for HR: Amplify the homologous arms from the genome of the selected parental zebrafish by PCR with high-fidelity DNA polymerase. The length of homologous arms at each side could be around 1 kb (*see* **Note 20**). Ligate the PCR products into appropriate vectors and transform into *E. coli* to obtain the donor plasmid. The vector should contain the sequence/elements used for precise genome

modifications (e.g., GFP coding sequence or *loxP* sequence, to replace or insert into the TALEN target site in the zebrafish genome by HR), and the homologous arms should be ligated into the flanking regions of these modification sequence/elements in proper orientations. Prepare the potential donor plasmids from each colony and choose the one containing the correct homologous arms by sequencing individually.

4. Cut out and purify the donor DNA fragment containing the homologous arms and the specific modification sequence/elements from the donor vector by proper restriction enzymes (*see* **Note 20**). Inject 100–300 pg of each TALEN mRNA and 50–100 pg linear donor DNA fragment into one-cell stage zebrafish embryos. The embryos only injected with TALEN mRNAs are prepared simultaneously and used for negative controls.
5. To identify successful events of HR in the injected founder embryos, isolate the genomic DNA from groups of embryos and amplify each homologous arm together with the modification sequence by PCR with one primer located within the specific modification sequence and the other flanking the corresponding homologous arm (*see* **Note 21**). The HR events can be further evaluated and confirmed by amplifying the entire homologous region (including both the left and right homologous arms) from a single PCR reaction by using a primer pair located outside but flanking the entire homologous region followed by sequencing of the PCR product.
6. Similar strategies as described in Subheading 3.4 can be used to raise and outcross the mosaic F_0 fish, screen F_1 embryos to identify germ-line transmission events, and further verify the HR alleles in F_1 adults (*see* **Note 22**). The same primer pairs from the above step can be used here.
7. The precise HR integration of the donor fragment in the identified positive F_1 fish can be further confirmed by Southern blot (Fig. 7).

3.7 Southern Blot Analysis in Zebrafish to Confirm HR Events

1. Mix dozens of 4–5 dpf zebrafish larvae (*see* **Note 23**) in a 1.5 mL Eppendorf tube and digest the embryos with the lysis buffer overnight at 55 °C.
2. Isolate the genomic DNA from the digestion mixture through the conventional phenol–chloroform extraction and ethanol precipitation method, dissolve in TE buffer.
3. Digest approximately 20 μg genomic DNA with selected restriction enzyme(s) completely (usually need overnight digestion) (*see* **Note 24**).
4. To monitor the completeness of digestion, pick up an aliquot from the digestion mixture and check by electrophoresis with 1 % agarose (*see* **Note 25**). Sometimes fresh restriction enzyme(s) could be supplemented to the reaction mixture after a long period of incubation to ensure complete digestion.

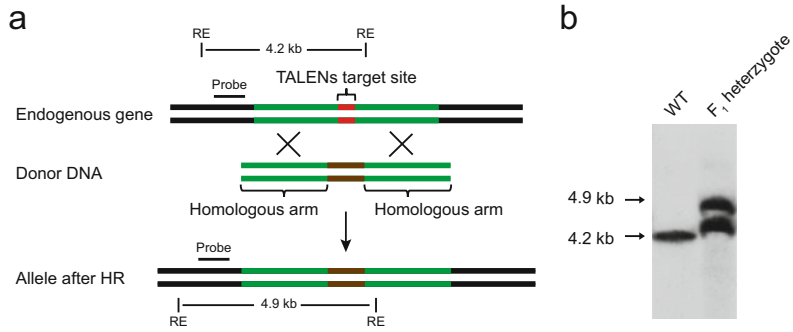


Fig. 7 Evaluation of HR induced by TALENs. (**a**) Schematic diagram of the structures of endogenous gene and target site, the donor template, as well as the targeted allele. The restriction enzymes (RE) and the probe used for Southern blot are also indicated. (**b**) A representative Southern blot result. In this case, the genomic fragments from the wild-type embryos were used as the control, which only shows a single hybridization band at 4.2 kb. For the genomic DNA from the embryos obtained by outcross of F_1 heterozygous for HR allele, an additional band of 4.9 kb, which represents the HR targeted allele, is clearly visible, as expected

5. The completely digested genomic DNA is separated on 1.0 % agarose gel and transferred to nylon membrane A (e.g., Amersham Hybond-N^+) by conventional blotting method.
6. The probes for Southern blot are labeled with DIG-dUTP or ^{32}P-dNTP and the molecular hybridization is carried out following the corresponding manufacturer's instructions (*see* **Note 26**).
7. The hybridization signal can be detected by using chemiluminescent CDP-Star or by exposure to X-ray film directly.

4 Notes

1. NotI is a unique enzyme site in the pCS2-TALEN plasmids constructed through our "Unit Assembly" method. Other expression vectors can also be used as long as there are proper promoters available for in vitro transcription (i.e., T3, T7 or SP6). If using other vectors, inspect the restriction enzyme sites and use the unique ones downstream of the TALEN coding sequence for linearization.
2. Make sure all the circular plasmids are digested completely. Use the original (undigested) plasmids as a control in electrophoresis and there should be no visible bands at the same position as the undigested original plasmid in the lane loaded with the linearized plasmid on the agarose gel. Insufficient linearization will lead to low yield of mRNA from in vitro transcription.
3. All the reagents and consumables used for in vitro transcription should be free of RNase. Use RNaseZap to make an RNase-free environment if necessary.

4. There are three kinds of mMessage mMachine kit from Ambion, only different in promoters (i.e., T3, T7 and SP6). Choose the correct one according to the promoter in the TALEN expression vectors.
5. The size of mRNA may not be accurately revealed in the agarose gel and may show up as multiple bands, possibly due to the formation of secondary structures of the mRNA. As long as the majority of mRNA is shown as one single band, it should be fine for the following experiments.
6. Do not let the RNA pellet dry completely, otherwise it will be hard to resuspend.
7. Use a new aliquot of mRNA each time for the microinjection, since it will degrade if thawing and freezing frequently.
8. Since TALENs are highly sequence specific, mismatches in the target sites will dramatically decrease their activity. To avoid low efficiency or failure in targeting due to the strong sequence polymorphisms of the zebrafish genome, we strongly recommend confirming the real sequence of each predicted target site in individual fish by PCR and direct sequencing and use the confirmed zebrafish or their offspring to perform the following mutagenesis experiments.
9. Usually we adjust the parameters of pressure and duration time of the microinjection machine to calibrate the volume of each drop to be 1–2 nL. The injection needle should be recalibrated after any changes in the pressure, duration time, balance, or re-breaking of the tips of the needles after getting plugged.
10. The mRNA injected into the embryonic cytoplasm is theoretically more active than into the yolk. Thus higher dosage might be needed to get similar targeting efficiency when injecting the TALEN mRNAs into the yolk. The effective injection dosage can be optimized by serial dilutions of TALEN mRNA pairs.
11. For different batches of TALEN mRNA, even though the concentrations of mRNA might be the same, the activity may still be different due to variations of the capping efficiency during in vitro transcription.
12. TALEN target sites can be easily identified with quite a few choices in a given sequence. Therefore, we generally prefer to pick up the sites that have unique restriction enzyme sites in their spacer region for the ease of determining targeting efficiency.
13. The crude lysate obtained from this fast genomic DNA extraction method is sufficient for the amplification of small fragment (<500 bp). Accordingly, we normally design the primers to amplify 200–500 bp fragments. Try other methods for the genomic DNA extraction if the amplification is not successful or the product is longer.

14. Based on our experience with more than 100 TALEN target sites in zebrafish, about 74 % TALEN pairs could show a measurable targeting efficiency, ranging from 1 % to nearly 100 %. The higher the efficiency, the easier to get heritable mutations through germ-line transmission.
15. According to our experience, male fish seem to give better efficiency of germ-line transmission. On the other hand, male fish develop faster towards sex maturation. Under optimal feeding conditions (e.g., optimal temperature, proper population density, sufficient food and extra feeding, etc.), male fish can start to fertilize eggs after 2 months, or even as short as 1 month and a half. So we recommend screening the male founder fish first for detecting mutations in F_1 embryos.
16. We recommend keeping two different indel mutation alleles (better covering two different aberrant open reading frames) for each target site.
17. Again, similar to the situation in generating indel mutations, single nucleotide polymorphisms (SNPs) and other genomic variations must be carefully excluded by PCR and sequencing in the TALEN target sites of the parental fish used for microinjection, especially when the sites are located in intronic or intergenic regions.
18. The efficiency of large genomic deletions depends on both the activity of TALEN pairs and the size of deletion. Theoretically, the higher the TALEN efficiency and/or smaller the deletion, the easier to identify germ-line transmission of large deletion mutations. However, it is difficult to precisely evaluate the efficiency of large genomic deletions in founder embryos, since usually the deletion alleles can be amplified more efficiently by PCR, whereas the wild-type allele is usually much longer than the deletion allele and therefore poorly amplified, or sometimes cannot be amplified at all if the sequence is too long. Nevertheless, we have successfully identified heritable deletion mutations spanning up to 122 kb genomic regions.
19. Polymorphisms between the target site and the donor template will markedly decrease the efficiency of HR. Therefore, it is crucial to make sure to use the parental fish whose sequences in both of the two homologous chromosomes are identical to the TALEN target sites and the homologous arms in the donor template.
20. About the design of homologous arms: According to our experience with injected founder embryos, more HR events could be detected with longer and linear homologous arms [13]. We practically choose ~1 kb homologous arms since they both are amenable to PCR amplification with primer pairs

spanning both left and right homologous regions (when the insertion element is <1 kb) and can produce satisfactory HR events. The position of the homologous arms should be outside of but as close as possible to the TALEN target sites.

21. We recommend mixing every 4–5 embryos as a group for PCR evaluation, and evaluating at least ten groups for each batch of injection.
22. The efficiency of HR is generally much lower than that of indel mutations or genomic deletions, so at least 100 founder fish are usually necessary to be used for screening in order to identify successful HR events.
23. We usually collect zebrafish larvae (offspring of the adult zebrafish to be evaluated) to extract enough genomic DNA for Southern blot. In our experience, fin clips from individual adult fish are hardly enough to extract sufficient genomic DNA for this purpose.
24. In our hands, HindIII, PstI, and PvuII worked well for the digestion of zebrafish genomic DNA prepared by our method, whereas BamHI, EcoRV, KpnI, NcoI, and SacII cannot completely digest the genomic DNA even after overnight incubation, for unknown reason(s).
25. The digestion can be considered complete if the lane looks smeary and no bands can be distinguished. If one can still see some discontinuous dominant bands, it means the digestion is not yet completed and longer incubation time is needed.
26. Compared to the radioactive isotope ^{32}P, digoxigenin is safer and easier to manipulate. In addition, the digoxigenin-labeled probes can be stored for a longer time (up to 1 year), but the sensitivity and signal to noise ratio may be lower than the radioactive probes.

Acknowledgments

We thank Yan Shen and Yuying Gao for the management of TALEN constructs; Yingdi Jia and Jingliang Chen for maintenance of our zebrafish facility; and all the members in our lab for their efforts on optimizing the TALEN applications. This work was partially supported by the National Natural Science Foundation of China (NSFC) (31110103904) and the National Key Basic Research Program of China (2012CB945101, 2011CBA01000, 2015CB942803). This work was also supported by the NSFC grant 81371264.

References

1. Blackburn PR, Campbell JM, Clark KJ, Ekker SC (2013) The CRISPR system: keeping zebrafish gene targeting fresh. Zebrafish 10(1):116–118
2. Huang P, Zhu Z, Lin S, Zhang B (2012) Reverse genetic approaches in zebrafish. J Genet Genomics 39(9):421–433
3. Wei C, Liu J, Yu Z, Zhang B, Gao G, Jiao R (2013) TALEN or Cas9: rapid, efficient and specific choices for genome modifications. J Genet Genomics 40(6):281–289
4. Xiao A, Wu Y, Yang Z, Hu Y, Wang W, Zhang Y, Kong L, Gao G, Zhu Z, Lin S, Zhang B (2013) EENdb: a database and knowledge base of ZFNs and TALENs for endonuclease engineering. Nucleic Acids Res 41(Database issue):D415–D422
5. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29(8):699–700
6. Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JR (2011) Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29(8):697–698
7. Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, Krug RG 2nd, Tan W, Penheiter SG, Ma AC, Leung AY, Fahrenkrug SC, Carlson DF, Voytas DF, Clark KJ, Essner JJ, Ekker SC (2012) In vivo genome editing using a high-efficiency TALEN system. Nature 491(7422):114–118
8. Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, Hu Y, Luo Z, Huang P, Wu Q, Zhu Z, Zhang B, Lin S (2013) TALEN-mediated precise genome modification by homologous recombination in zebrafish. Nat Methods 10(4):329–331
9. Gupta A, Hall VL, Kok FO, Shin M, McNulty JC, Lawson ND, Wolfe SA (2013) Targeted chromosomal deletions and inversions in zebrafish. Genome Res 23(6):1008–1017
10. Lim S, Wang Y, Yu X, Huang Y, Featherstone MS, Sampath K (2013) A simple strategy for heritable chromosomal deletions in zebrafish via the combinatorial action of targeting nucleases. Genome Biol 14(7):R69
11. Ma AC, Lee HB, Clark KJ, Ekker SC (2013) High efficiency in vivo genome engineering with a simplified 15-RVD GoldyTALEN design. PLoS One 8(5), e65259
12. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, Zu Y, Li W, Huang P, Tong X, Zhu Z, Lin S, Zhang B (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res 41(14), e141
13. Meeker ND, Hutchinson SA, Ho L, Trede NS (2007) Method for isolation of PCR-ready genomic DNA from zebrafish tissues. Biotechniques 43(5):610, 612, 614
14. Wang D, Jao LE, Zheng N, Dolan K, Ivey J, Zonies S, Wu X, Wu K, Yang H, Meng Q, Zhu Z, Zhang B, Lin S, Burgess SM (2007) Efficient genome-wide mutagenesis of zebrafish genes by retroviral insertions. Proc Natl Acad Sci U S A 104(30):12428–12433

Chapter 16

Mutagenesis in Xenopus and Zebrafish using TALENs

Yun Liu, Hui Zhao, and Christopher H.K. Cheng

Abstract

Transcription activator-like effector nucleases (TALENs) have been proven to be effective for gene specific targeting across species. Here we describe the validated protocol for TALEN assembly as well as methods for generating gene knockout animals of *Xenopus* and zebrafish. This protocol covers selection of TALEN targeting sites, TALEN assembly with a modified Golden Gate method, injection of TALEN mRNAs into *Xenopus* and zebrafish embryos as well as the detection of somatic and germ-line transmitted mutations. Finally, the establishment of knockout *Xenopus* and zebrafish lines is also described. This protocol will facilitate broader applications of TALENs in developmental biology.

Key words TALENs, Gene knockout, Gene disruption, Mutagenesis, *Xenopus*, Zebrafish

1 Introduction

Xenopus tropicalis and *Danio rerio* are excellent models used in studies of embryonic development and disease. However, it is very difficult to induce gene specific mutagenesis using conventional gene disruption approaches. A major breakthrough has been made in recent years by using engineered nucleases, e.g. the zinc finger nucleases (ZFNs) [1], the transcription activator-like effector nucleases (TALENs) [2, 3], and more recently, the RNA mediated CRISPR/Cas9 [4–6].

These engineered nucleases induce double-stranded DNA breaks (DSB) at their target sites. Repair of the DSB by error-prone nonhomologous end joining (NHEJ) can introduce insertions or deletions (indels) at the cleavage site. Among these three nucleases, TALENs have been proven to be highly effective for gene disruption and produce less off-target effects.

1.1 Transcription Activator-Like Effector Nucleases

TALENs have been successfully applied in various species including plant [7], *Drosophila* [8], *C. elegans* [9], zebrafish [10, 11], *Xenopus* [2, 3], mouse [12, 13], rat [14], livestock [15], human somatic cells [16], and human pluripotent stem cells [17].

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_16,

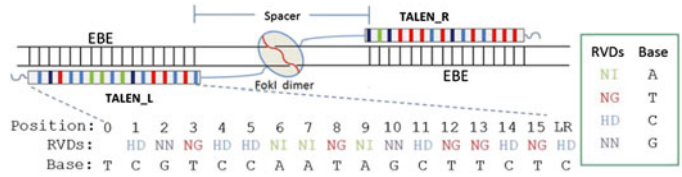


Fig. 1 A diagram of target recognition by engineered TALENs. A full TALENs target site is composed of two half-sites separated by a spacer. FokI dimer is formed in the spacer region where DSB is introduced. The binding specificity of TALENs is determined by the RVDs of each TALE repeat. The codons for RVD-base interaction are shown on the right. *LR* last repeat, *EBE* effector binding elements

The architecture of TALENs consists of two distinct domains, an engineered DNA-binding domain derived from the transcription activator-like effector (TALE) that can bind to the target DNA sequence [18, 19], and the restriction endonuclease Fok I domain fused to the C-terminal of TALE that confers the nuclease activity of TALENs [20] (Fig. 1).

The TALE proteins were originally discovered in *Xanthomonas* plant bacteria [18, 19]. Their DNA binding specificity relies on the central repeat domain. Each repeat unit consists of 34 amino acids, and their sequences are almost identical except for two amino acids at the 12th and 13th positions. These two amino acids determine the binding specificity to DNA bases, and are called "repeat-variable di-residues" (RVDs). It is known that NI, NG, HD, and NN preferentially recognize adenine (A), thymine (T), cytosine (C), and guanine (G)/adenine (A), respectively (Fig. 1). Following this code, a given DNA sequence can theoretically be recognized by a combined central repeats of TALE proteins.

Each TALEN monomer has its own DNA recognition domain, and binds to its targeting site individually. Upon dimerization of FokI domains after two monomers bind to adjacent predetermined sites (effector binding elements, EBEs), the Fok I will induce DNA double-strand breaks at the spacer region between two EBEs (Fig. 1).

1.2 TALEN EBE Selection and Golden Gate Assembly

Currently, a number of TALEN assembly platforms have been established such as Golden Gate assembly [16], PCR-based modular assembly [21], FLASH assembly [22] as well as some commercially available approaches including GeneArt® Precision TALs (Life Technologies) and Cellectis bioresearch. Among them, the Golden Gate assembly has been proven to be easy, fast, and cost-effective, and therefore suitable for most laboratories. The Golden Gate strategy for TALEN assembly was chosen in our studies. It is based on type IIS restriction enzymes that can cleave DNA base pairs out of their binding sites [23]. In addition to the original report [16], we have made two constructs, pCS2-TALEN-ELD/KKR, which are feasible for microinjection into both zebrafish and *Xenopus* embryos with high gene disruption efficiency [3, 24].

1.3 Choosing the TALEN Binding Site

First, position 0 ought to be T, as this is always observed in the target sites of the natural TALEs. Second, the half-site length is recommended to be 16–18 bp. In our hands, longer half-site length decreases TALENs activity. Third, the ideal spacer length is around 16–18 bp. The spacer length is critical since an ideal spatial space is needed to form the FokI dimer. Fourth, less than five G in a half-site is recommended since the RVD NN which bind to G could also bind to A. The RVDs of HD, NI, and NG have a high preference for C, A, and T, respectively; but the RVD NN can recognize both G and A [25]. We try to avoid G to reduce the potential off-target effects caused by the ambiguity of NN. There are also two web-based programs publicly available (http://taleffectors.com/tools, and https://tale-nt.cac.cornell.edu/node/add/talen) for TALEN EBE selection.

1.4 DNA Repair After Double Strand Break Induced by TALENs

After TALE target selection, we assemble the TALE repeats into the modified pCS2-TALEN-ELD/KKR vectors. TALENs with wild type Fok I can form homodimers that may cause off-target effects. ELD/KKR are modified Fok I nuclease domains carrying three point mutations in each monomer, and form a heterodimer. TALENs harboring ELD and KKR were proven to have less off-target effects and higher efficiency because of the obligatory heterodimer formation [26]. Upon dimerization of ELD and KKR, a pair of TALENs that binds to adjacent EBEs at a chosen locus is expected to produce DNA double-strand cleavage at the spacer, generating a DNA double-strand break (DSB). Repair of the DSB by NHEJ results in small insertions or deletions (indels) at the cleaved site, leading to disruption of gene function.

1.5 Preparation of TALEN mRNAs for Micro injection

The protocol for TALEN assembly will be described in a subsequent section on experimental procedure. After assembly of the TALEN constructs, the two TALEN plasmids will be linearized by NotI and the mRNAs for microinjection will be transcribed in vitro using the linearized DNA as templates. After purification, the quality and integrity of the synthesized mRNAs will be checked on a denaturing agarose gel, and concentrations will be determined by spectrophotometry.

1.6 Assessment of Mutagenesis Rate in Somatic Cells

The somatic mutations in the injected embryos are examined at 48 h after injection. Five to eight injected embryos are randomly pooled and their genomic DNA extracted for further PCR assay [3]. To assess the frequency of NHEJ-induced mutagenesis at the target sites, three primers were employed for this assay (Fig. 2). Forward Primer (F) and reverse primer (R) can bridge the entire two EBE regions, and the forward primer 1 (F1) and reverse primer can link the spacer region and the downstream EBE region. PCR is performed with F and R using the extracted genomic DNA as template. The amplicons will be cloned into the vector pMD18-T

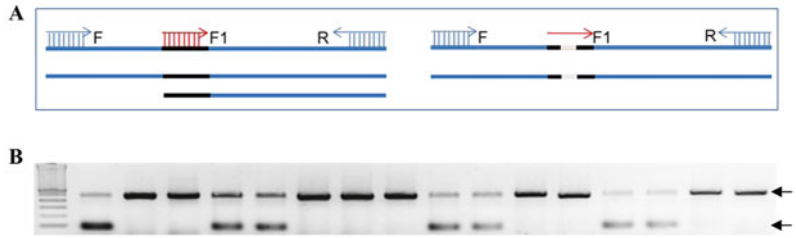


Fig. 2 Rapid detection of mutated clones by competitive PCR. (**a**) Schematic diagram of PCR-based mutation detection. Three primers were designed and used in the PCR reactions. F1 is located at the spacer region. In single clones containing the wild-type sequence, two bands would be amplified. Whereas in clones containing the mutated sequence, one single band only would be amplified because mutations in the spacer sequence destroy the binding site of the F1 primer. (**b**) A gel picture representing the PCR-based detection of mutations. The two expected bands in the wild-type clones are indicated by *arrows*. The larger band is brighter in the mutated clones because of loss of competition by the F1 primer in these clones

by TA cloning, and colonies will subsequently be examined by colony PCR using primer F, F1 and R. In single clones containing the wild-type sequence, two bands would be amplified. Whereas in clones containing the mutated sequence, one single band only would be amplified because mutations in the spacer sequence destroy the binding site of the F1 primer.

1.7 Evaluation of Germ-Line Transmission

Germ-line transmission is essential for establishment of a gene knockout animal line. In addition to evaluating the rate of somatic mutations, the rest of the injected embryos will be raised to adulthood for examining germ-line transmission of mutations. The F0 frogs or fish are mosaic, carrying various mutations in their somatic cells. However, this situation will not normally affect germ-line transmission as individual sperm or oocyte only carries one mutation. To carry out germ-line transmission of mutations, individual F0 animal will be mated with wild type animal. The fertilized embryos (F1 generation) will be collected and cultured for 2 days. Genomic DNA will be extracted from individual embryos and analyzed by the PCR assay described above for assessing the mutagenesis rate. If a F1 embryo carries a mutation, it should be heterozygous, and theoretically approximately 50 % of its genomic DNA will harbor the mutation. The mutagenesis positive colonies will be sequenced directly to confirm sequence alternations. The ratio for germ-line transmission is defined as the number of mutated F1 embryos to the total examined F1 embryos. Siblings of the F1 embryos will be raised to sexual maturity. Genotyping for individual F1 animals will be performed by PCR and subsequently sequencing using genomic DNA extracted from its tail. We often find that several F1 embryos carry the same mutation.

In that case, these two founders will be mated and the fertilized embryos (F2) will be collected. In theory, 25 % of the embryos are homozygous at the target locus. The F2 progenies will be raised and genotyped for population expansion of this knockout line.

2 Materials

2.1 TALEN Assembly Vectors

1. Golden Gate vectors (Addgene Cat #1000000016).
2. TALENs backbone vectors (pCS2-TALEN-ELD and pCS2-TALEN-KKR).

2.2 Primer List

1. pCR8_F: TTGATGCCTGGCAGTTCCCT.
2. pCR8_R: CGAACCGAACAGGCTTATGT.
3. NTalF: GATGACAAGGGTACCGTG.
4. CTalR: CTAGTTGGGATCCGGCAAC.

2.3 Reagents

1. T4 DNA ligase.
2. NotI.
3. BsaI.
4. BsmbI.
5. Plasmid-Safe™ ATP-Dependent DNase.
6. mMessage mMachine SP6 Kit.
7. RNeasy MiniElute Cleanup Kit.
8. PCR Purification Kit.
9. DNA Miniprep Kit.
10. DNeasy blood & tissue Kit.
11. Ampicillin.
12. Spectinomycin.
13. Tetracycline.
14. DH5α.
15. Microinjector.
16. Needle Puller.
17. hCG.
18. MS222.
19. L15 medium.
20. Testis medium: 50 % L15 medium, 10 % calf serum.
21. MMR buffer: 0.1 M NaCl, 2.0 mM KCl, 1 mM $MgSO_4$, 2 mM $CaCl_2$, 5 mM HEPES (pH 7.8), 0.1 mM EDTA.

3 Methods

3.1 Construction of Highly Effective TALENs

Select TALEN binding site using the following rules. First, position 0 ought to be T. Second, the half-site length is recommended to be 16–18 bp. Third, the ideal spacer length is around 16–18 bp. Fourth, less than five G in a half-site (*see* **Note 1**).

1. Obtain the TALEN Golden Gate toolkit from Addgene (http://www.addgene.org/TALeffector). The TALEN expression plasmids (pCS2-TALEN-ELD and pCS2-TALEN-KKR) could be obtained from our laboratory (Please forward such request to Prof. Christopher H.K. Cheng). For each obtained clone, culture the bacteria in 5-ml LB with appropriate antibiotics (tetracycline, spectinomycin or ampicillin) overnight (Fig. 3). Extract plasmid DNA using a standard Plasmid Miniprep Kit. Adjust the concentration of plasmids to 100 ng/μl. Label plasmids pNI1–pNI10 as A1–A10, pNN1–pNN10 as G1–G10, pNG1–pNG10 as T1–T10, and pHD1–pHD10 as C1–C10. Store the plasmids at 4 °C. Repeated thaw–freeze cycles should be avoided.
2. For each target, a pair of TALEN monomers should be constructed to bind the left (TALEN-L) and the right (TALEN-R) half-site. Labeling the 5′ T as position 0 and the remaining nucleotides with serial numbers (e.g., 1, 2, 3, and *X*). Pick up ten RVDs recognizing positions 1–10 (the A at position 1 is

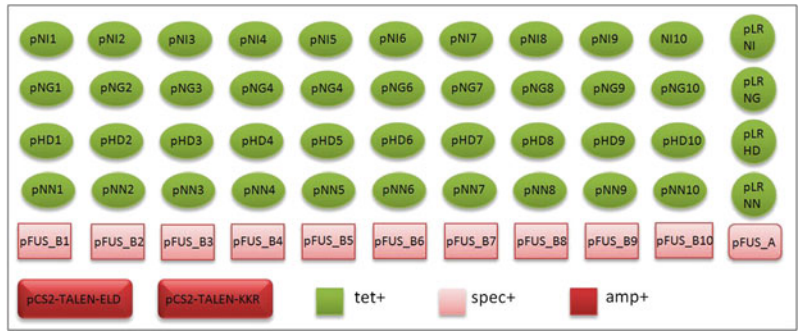


Fig. 3 Contents of the Golden Gate TALEN kit. The toolkit includes 40 RVD modular plasmids (*shown in ellipse*), 11 middle array plasmids (*shown in rectangle*), 4 last repeat plasmids (*shown in circle*) and 2 TALEN expression plasmids (pCS2-TALEN-ELD and pCS2-TALEN-KKR). The 40 modular plasmids recognize 4 genetic bases at 10 different positions. The middle array plasmid pFUS_A serves as the vector backbone for insertion of the RVDs recognizing nucleotides at positions 1–10. The middle array plasmids pFUS_B1 to pFUS_B10 serve as the vector backbones for insertion of RVDs recognizing nucleotides at positions 11–20. The last-repeat plasmids recognize the nucleotide in the last position. The two TALEN expression plasmids serve as the vector backbones for insertion of the assembled RVDs and express functional TALENs. The antibiotic property of each plasmid is coded by a different color. *Tet+* tetracycline resistance, *spec+* spectinomycin resistance, *amp+* ampicillin resistance

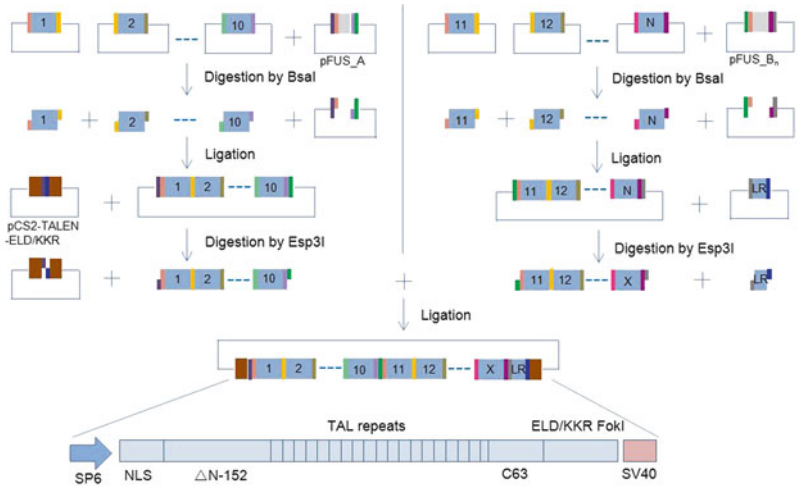


Fig. 4 Golden Gate assembly of customized TALENs in two digestion–ligation rounds. Each plasmid could generate 4-bp complementary overhangs (color-coded) and ligate together in a serial order. The first round Golden Gate assembly is to digest and ligate the TAL repeats recognizing each target base (except the last base) into the two array plasmids (pFUS_A and pFUS_B). The ten modular plasmids recognizing the first ten target DNA base are cloned into the middle array vector pFUS_A. The modular plasmids recognizing the 11th–20th target DNA base are cloned into the corresponding array vector pFUS_B. The second round Golden Gate assembly is to digest and ligate the assembled array pFUS_A, pFUS_B and the last repeat into the TALENs expression vector pCS2-TALEN-ELD or pCS2-TALEN-KKR. SP6, SP6 RNA polymerase binding site; NLS, nuclear localization signal; SV40, SV40 polyadenylation signal sequence; LR, last repeat; ΔN152, truncation of the N-terminal 152-aa; C63, retention of C-terminal 63-aa in the TALENs scaffold

recognized by pNI1, the C at position 2 is recognized by pHD2, etc.) and the middle array backbone plasmid pFUS_A. Pick up RVDs recognizing positions 11-(X-1) (the A at position 11 is recognized by pNI1, the C at position 12 is recognized by pHD2, etc.) and the corresponding middle array backbone plasmid pFUS_B (e.g., if pFUS_B is to contain five RVDs, pick up pFUS_B5 as backbone). The plasmid recognizing the last nucleotide will be assembled later.

3. The customized TALENs are assembled in two digestion–ligation rounds (Fig. 4). Set up the first digestion–ligation reaction as tabulated in Tables 1 and 2 (*see* **Note 2**) and run the reaction on a PCR machine as described in Table 3.
4. Add 1 μl Plasmid-Safe nuclease (Epicentre) to the reaction mix and incubate at 37 °C for 30 min.
5. Gently add all the reaction mix into 100 μl DH5α competent cells. Incubate on ice for 10 min. Heat-shock the competent cells at 42 °C for 2 min. Place the competent cells on ice for 2 min. Add 200 μl LB medium and culture the competent cells at 37 °C with shaking for 40 min. Centrifuge the culture at 4000 × *g* for 2 min. Remove 180 μl supernatant and reconstitute

Table 1
Components of Golden Gate cloning round 1 (position 1–10)

Components	Amount (μl)
10× NEB Buffer 4	1
25 mM ATP	0.4
RVD Plasmids 1–10 (100 ng/μl)	0.6×10
pFUS_A (100 ng/μl)	1
BsaI (10 U/μl)	0.6
T4 ligase (2000 U/μl)	0.6
dH_2O	To 10 μl

Table 2
Components of Golden Gate cloning round 1 (position 11–19)

Components	Amount (μl)
10× NEB Buffer 4	1
25 mM ATP	0.4
RVD Plasmids 1-N (100 ng/μl)[a]	0.6×N
pFUS_Bn (100 ng/μl)	1
BsaI (10 U/μl)	0.6
T4 ligase (2000 U/μl)	0.6
dH_2O	To 10 μl

[a] *N* = 1–9

Table 3
Digestion–ligation program for Golden Gate cloning round 1

Cycle number	Temperature (°C)	Time (min)
1–6	37	15
	16	15
7	37	15
8	80	5

the competent cells. Smear the competent cells on an LB plate supplemented with 50 μg/ml spectinomycin (spec+). Blue/white screening of the correct clones is optional. Culture the plate overnight at 37 °C.

6. Pick up five pFUS_A clones and three pFUS_B clones. Culture the single clones in 60 μl LB spec+ medium for 3–4 h with shaking at 37 °C.
7. Perform colony PCR as described in Tables 4, 5, and 6. Run the PCR products on a 1.5 % agarose gel. Smearing ladder bands will appear in the correct clones (Fig. 5) (*see* **Note 3**).
8. Culture one correct pFUS_A/pFUS_B clone with 1.5 ml spec$^+$ LB medium overnight. Extract plasmid DNA with Miniprep Kit according to the manufacturer's instructions. Elute the plasmid DNA with 40 μl dH_2O. Measure the plasmid DNA concentrations on a spectrophotometer.
9. Set up the second digestion–ligation reaction as described in Table 7. Clone TALE repeats recognizing the left half-site into the pCS2-TALEN-ELD backbone and those recognizing the right half-site into the pCS2-TALEN-KKR backbone.

Table 4
Primers for Golden Gate assembly and sequencing

Primer name	Primer functions
pCR8_F1 pCR8_R1	Amplification and sequencing of the assembled TALE repeats in the array plasmids
NTalF CTalR	Amplification and sequencing of the assembled final TALEN expression plasmids

Table 5
Components of colony PCR in Golden Gate cloning round 1

Components	Amount (μl)
10× Taq DNA polymerase buffer	1
10 mM dNTP	0.2
pCR8_F1 (10 μm)	0.5
pCR8_R1 (10 μm)	0.5
Bacterial culture template	0.5
Taq DNA polymerase	0.05
dH_2O	To 10 μl

Table 6
Colony PCR program for Golden Gate cloning round 1

Cycle number	Denature	Anneal	Extend
1	95 °C, 5 min		
2–34	95 °C, 20 s	52 °C, 20 s	72 °C, 60 s
35			72 °C, 5 min

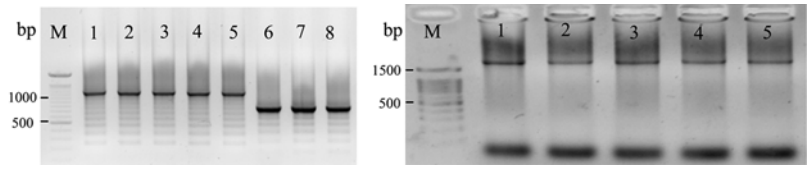


Fig. 5 Representative gels of the assembled TALE arrays. *Left panel*, *Lane 1–5*: the correct pFUS_A array containing ten RVDs; *Lane 6–8*: the correct pFUS_B array containing five RVDs. *Right panel*, *Lane 1–5*: the correct final TALEN expression array containing 15.5-RVDs

Table 7
Components of Golden Gate cloning round 2

Components	Amount (μl)
10× NEB Buffer 3	1
25 mM ATP	0.4
pFUS_LA or pFUS_RA (100 ng/μl)	1
pFUS_LB or pFUS_RB (100 ng/μl)	1
pCS2_TALEN_ELD or pCS2_TALEN_KKR (100 ng/μl)	1
Last repeat (100 ng/μl)	1
Esp3I (10 U/μl)	0.4
T4 ligase (2000 U/μl)	0.4
dH_2O	To 10 μl

10. Run program on PCR machine as described in Table 8.
11. Transform the ligation mix into competent cells. Plate the transformants on an LB plate supplemented with 50 μg/ml ampicillin (amp+) and incubate overnight at 37 °C. Pick up five single clones from each plate and culture the clones in 60 μl LB amp+ medium for 3–4 h with shaking at 37 °C.
12. Perform colony PCR as described in Tables 9 and 10. Run the PCR products on a 1.5 % agarose gel. Smearing ladder bands will appear in the correct clones (Fig. 5) (*see* **Note 4**).

Table 8
Digestion–ligation program for Golden Gate cloning round 2

Cycle number	Temperature (°C)	Time (min)
1–6	37	15
	16	15
7	37	15
8	80	5

13. Overnight culture of one correct clone in 4 ml amp+ LB medium. Isolate the plasmid using Miniprep Kit. Sequence the final TALEN plasmid with primers NtalF and CtalR (Table 4). To analyze the sequencing result, translate the DNA sequence into protein using translate tool (http://web.expasy.org/translate/). Find the RVDs to determine whether the order of RVDs is correct.

3.2 TALEN mRNA Preparations

1. Linearize each TALEN plasmid with NotI as described in Table 11. Purify the digested plasmid using the PCR Purification Kit. Elute the plasmid DNA with 30 μl RNase-free water. UV exposure of the DNA template for in vitro transcription should be avoided.
2. Prepare TALEN mRNAs using the mMESSAGE mMACHINE SP6 kit. Set up the in vitro transcription reaction as shown in Table 12. Incubate the mix at 37 °C for 2 h. Add 1 μl TURBO DNase and incubate the mix at 37 °C for 15 min.
3. Purify the mRNA using RNeasy Mini Kit according to the manufacturer's instructions. Elute the mRNA with 30 μl RNase-free water. Otherwise purify the mRNA using reagent provided in the mMESSAGE mMACHINE SP6 Kit.
4. Check the quality and integrity of the mRNA on a denaturing agarose gel and measure the RNA concentrations on a spectrophotometer. The total yield of the RNA is about 10 μg. RNA should be aliquoted and stored at −80 °C. All procedures should be carefully done to avoid possible RNase contamination.

3.3 Microinjection of TALEN mRNA into *X. tropicalis* Embryos

1. In vitro fertilization (IVF). Prime female frogs by injecting 20 IU human chorionic gonadotropin (hCG) into the dorsal lymph sac at night before IVF. Prime male frogs with 100 IU hCG at the same time. Inject another dose of 100 IU hCG into female frogs in the morning of the second day. Place the primed female frogs into a water tank of temperature not below 23 °C. The female frogs normally lay eggs within a period of 2–6 h after the booster injection.

Table 9
Colony PCR reaction mix for Golden Gate cloning round 2

Components	Amount (μl)
10× Taq DNA polymerase buffer 10 mM dNTP	1 0.2
NtalF (10 μM)	0.5
CtalR (10 μM)	0.5
Bacterial culture template	0.5
Taq DNA polymerase	0.05
dH_2O	To 10 μl

Table 10
Colony PCR program for Golden Gate cloning round 2

Cycle number	Denature	Anneal	Extend
1	95 °C, 5 min		
2–34	95 °C, 20 s	52 °C, 20 s	72 °C, 90 s
35			72 °C, 5 min

Table 11
Components of TALEN plasmid linearization

Components	Amount (μl)
10× NEB buffer 3	2
NotI	1
TALEN_L/TALEN_R (10 μg)	Variable
dH_2O	To 20 μl

Table 12
Components of in vitro transcription

Components	Amount (μl)
NotI-linearized plasmid DNA	3
2× NTP/ATCA	5
10× SP6 reaction buffer	1
SP6 enzyme mix	1

2. Place one primed male frog into a water tank containing 0.2 % ethyl 3-aminobenzoate methanesulfonate salt (MS222) to anesthetize this male frog. Sacrifice the male frog after it becomes immobile. Dissect the testes out and keep them in the testis medium and perform IVF right away. Unused testes can be stored in L15+ 10 % calf serum at 4 °C for 3 h. Collect eggs by squeezing a primed female frog into a petri dish. Cut the testis into very small pieces, add 1 ml testis medium, and mix them well. Add sperm mix into the collected eggs, and mix eggs with sperm gently but thoroughly. Wait for 3–5 min, and then add 1/9× MMR buffer into the dish to cover eggs till completion of fertilization. Wait at least 20 min and keep eggs at about 25 °C during the fertilization process. Remove jelly coat from the eggs using 3 % cysteine hydrochloride (dissolved in 1/9× MMR pH 7.8–8.0). Normally the dejelly procedure takes 8–10 min.
3. Microinjection can be performed at room temperature (about 22 °C) using standard procedure (http://tropicalis.berkeley.edu/home/manipulate_embryos/microinjection/microinjection.html). We normally inject 500 pg mRNA/embryo (TALEN-L and TALEN-R mRNA, 250 pg each) into one-cell stage embryos. After all embryos are injected, transfer the injected embryos into a new petri dish containing 3 % Ficoll in 1/9× MMR, maintain embryos in Ficoll solution for 1 h, and then transfer the injected embryos into a petri dish containing 1/9× MMR with gentamicin (50 μg/ml). Culture the injected embryos at 25 °C (*see* **Note 5**).

3.4 Microinjection of TALEN mRNA into Zebrafish Embryos

1. On the night before microinjection, set up 6–10 pairs of fish into individual breeding tanks. Keep the male and female fish separated. On the injection morning, remove the divider to let male and female fish mate. Collect eggs at the bottom of the breeding tank.
2. Prepare the injection solution as tabulated in Tables 13 and 14. For single locus targeting, a pair of TALEN mRNAs should be mixed. For dual loci targeting (fragment deletion), two pairs of TALEN mRNAs should be mixed. Injection solution should be freshly prepared and should not be reused.
3. Microinjection should be performed as described previously [27]. Adjust the injection volume as 3 nl per injection. The TALEN mRNA mix should be directly injected into the cytoplasm of one-cell stage embryos.
4. Calculate the number of the dead/deformed/normal embryos at 48 h post fertilization. Remove the dead and severely deformed embryos (*see* **Note 6**).

Table 13
Components of microinjection solution for single locus

Components	Amount (μl)	Amount (ng)
TALEN_L mRNA	Variable	500
TALEN_R mRNA	Variable	500
RNase-free water	To 3 μl	

Table 14
Components of microinjection solution for single locus

Components	Amount (μl)	Amount (ng)
TALEN_L1 mRNA	Variable	250
TALEN_L2 mRNA	Variable	250
TALEN_R1 mRNA	Variable	250
TALEN_R2 mRNA	Variable	250
RNase-free water	To 3 μl	

3.5 Detection of TALEN-Induced Mutations and Deletions in Embryos

3.5.1 For Indel-Mutation Detection (Single Locus Targeting)

1. At 48 h post fertilization, randomly collect and pool 5–7 TALEN-injected *X. tropicalis* embryos or 8–12 TALEN-injected zebrafish embryos for genomic DNA isolation.
2. Set up genomic PCR reaction as described in Table 15 and run the genomic PCR as described in Table 16. Gel purify the target bands. Elute DNA with 30 μl dH_2O.
3. Ligate the purified PCR product into TOPO-4.1 vector as tabulated in Table 17. Transform the ligation product. Pick up 32 single clones and culture these clones in a 60 μl amp+ medium on a 96-well plate with shaking at 37 °C for 3–4 h.
4. Set up competitive PCR as described in Table 18 and run the PCR program in Table 19. Gel purify the larger bands and sequence the purified samples using T7 primer.
5. Blast the sequencing results with the genomic DNA sequence at Ensembl database to identify the exact mutated genotypes.

3.5.2 For Fragment Deletions (Dual Loci Targeting, Fig. 6)

1. Pool 8–12 TALEN-injected zebrafish embryos for genomic DNA isolation. Set up PCR reactions as shown in Table 15 (*see* **Note 7**). Run the genomic PCR using the program in Table 20. Gel purify PCR products and elute DNA with 30 μl dH_2O.

Table 15
Components of genomic PCR

Components	Amount (μl)
10× Taq DNA polymerase buffer	1
10 mM dNTP	0.2
Forward primer (10 μM)	0.5
Reverse primer (10 μM)	0.5
Genomic DNA	0.5
Taq DNA polymerase	0.05
dH_2O	To 10 μl

Table 16
Program for PCR amplification of genomic DNA isolated from pooled embryos

Cycle number	Denature	Anneal	Extend
1	95 °C, 5 min		
2–30	95 °C, 20 s	52 °C, 20 s	72 °C, 60 s
35			72 °C, 10 min

Table 17
Components of TA cloning

Components	Amount (μl)
TOPO-4.1 vector	1
Insert DNA (~0.3 pmol)	4
Solution I	5
RNase-free water	To 10 μl

2. Ligate the purified PCR product into the TOPO-4.1 vector and transform the competent cells. Pick up 12 single clones and culture each clone in 60 μl amp + medium on a 96-well plate at 37 °C for 3–4 h.
3. Run colony PCR as shown in Tables 18 and 19. Purify the target bands. Sequence the purified samples using proper primer (The primer binding site should be at least 100 bp upstream or downstream from the expected mutation). Compare the sequencing result with the zebrafish genome database at Ensembl to identify the exact mutated genotypes.

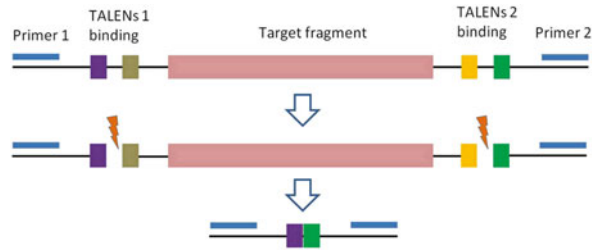


Fig. 6 Schematic diagram of large genomic deletion using two pairs of TALENs. To delete a large DNA fragment, two DSBs were introduced simultaneously on each side of the targeted genomic fragment using two pairs of TALENs. Repair of the DSB by ligation of the broken ends will lead to deletion of the flanked genomic fragment. The genomic deletion could be detected by PCR using properly designed primers

Table 18
Components of competitive colony PCR

Components	Amount (μl)
10× Taq DNA polymerase buffer	1
10 mM dNTP	0.2
Forward primer 1 (10 μM)[a]	0.25
Forward primer 2 (10 μM)[a]	0.25
Reverse primer (10 μM)	0.5
Bacterial cultures	0.5
Taq DNA polymerase	0.05
dH_2O	To 10 μl

[a]The ratio of the two forward primers could be adjusted

Table 19
Program for colony PCR

Cycle number	Denature	Anneal	Extend
1	95 °C, 5 min		
2–34	95 °C, 20 s	52 °C, 20 s	72 °C, 60 s
35			72 °C, 10 min

Table 20
PCR amplification of genomic DNA with fragment deletions

Cycle number	Denature	Anneal	Extend
1	95 °C, 5 min		
2–40	95 °C, 20 s	52 °C, 20 s	72 °C, 60 s
35			72 °C, 10 min

3.6 Establishment of Mutant Animal Lines

3.6.1 Establishment of *X. tropicalis* Mutant Lines

1. Raise the sibling TALEN injected *X. tropicalis* embryos till sexual maturity according to (http://tropicalis.berkeley.edu/home/husbandry/raisetads.html). This normally takes about 5–6 months, and longer period is also expected depending on the feeding conditions. A reliable indicator for maturation is the presence of the dark stripes in males on the interior of the forearm.
2. When TALEN injected tadpoles reached stage 45 or older, tail tissue can be collected to identify the tadpoles carrying mutation. Anesthetize tadpoles in a petri dish containing 0.025 % tricaine in distilled water. When tadpoles become immobile, use scalpel to cut away the posterior 1/4 to 1/3 of the tail.
3. Extracting genomic DNA from the dissected tail. Somatic mutation can be detected with PCR-based method (described in Subheading 3.5).
4. When TALENs targeted frog are ready for germ-line transmission, outcross a male or female P0 frog with a wild-type frog. Prime frogs using the protocol described in Subheading 3.3. After the booster injection, the animals should be divided into male/female pairs and placed in 1/9× MMR with gentamycin. The mating tanks should be kept in a quiet dark place. Amplexus normally begins 1–3 h after boosting and can last up to 6 h. Usually, female frogs lay eggs from 2 h to 6 h after the second hCG injection. Keep frogs at about 25 °C throughout the priming and natural mating periods.
5. Collect embryos and remove jelly coat from the eggs using 3 % cysteine hydrochloride dissolved in 1/9× MMR (pH 7.8–8.0). Culture the embryos in 1/9× MMR at 25 °C until stage 41. Randomly collect 20 embryos, and extract genomic DNA from each embryo, detect the mutation rate after germ-line transmission using PCR-based method described in Subheading 3.4 or as described previously [28]. Leave the sibling embryos to develop to adult frogs.
6. If the results are positive, collect tail tissue when TALEN injected tadpoles reach stage 45 or older for founder selection.

7. Genotype each heterozygous F1 larva and select the larvae carrying ORF-shift mutations and raise them to adulthood for mating.
8. After population expansion of selected F1 adult frogs, intercross the F1 frogs carrying the same mutation. Approximately 25 % offspring from intercross of F1 adult frogs are expected to be homozygous at the targeted locus according to Mendel's law. Expand the frog population with this genome type, and maintain this mutant line.

3.6.2 Establishment of Zebrafish Mutant Lines

1. Pick out 3–4 mature P0 fish and cross each P0 fish with wild-type fish individually.
2. To detect germ-line transmitted indel-mutations, pool 10–12 F1 zebrafish embryos from each cross for genomic DNA isolation, genomic PCR, gel purification, colony assay, and DNA sequencing as described in Subheading 3.5.
3. To detect germ-line transmission of large deletions, collect 24–32 F1 embryos from each cross for genotyping. Isolate genomic DNA from single embryos as described previously [29]. Genomic DNA contamination should be carefully avoided. Perform genomic PCR using the program in Table 20. Purify the target band and sequence the purified samples to identify the exact mutated DNA sequence.
4. If desired mutations or deletions are detected in the F1 embryos, raise the embryos to adulthood.
5. To genotype the F1 adults with indel-mutations, cut a small piece of the tail fin for genomic DNA isolation. Set up the genomic PCR reaction as shown in Table 15 and perform genomic PCR using the program in Table 16. Purify the target band and sequence the purified samples to identify the exact mutated DNA sequence. Two peaks will appear in the sequencing result of the F1 heterozygotes. One peak represents the wild-type allele and the other peak represents the mutated allele. The exact mutated genotype should be determined by reading each peak manually. To get homozygous mutants in F2 generation, F1 heterozygotes harboring the same mutation need to be obtained (*see* **Note 8**).
6. To genotype the F1 adults with large genomic deletions, cut a small piece of the tail fin for genomic DNA isolation and perform genomic PCR and DNA sequencing as described in Subheading 3.5.2.
7. Cross the F1 harboring the same mutations or deletions. About 25 % of F2 fish will be homozygous mutants. The genotype of F2 could be analyzed by genomic PCR and sequencing.

4 Notes

1. To avoid off-target effects, the TALEN binding sites should not be chosen in genomic region containing highly repetitive sequences. Based on the need, one can choose to prepare one pair of TALENs to induce small indel mutations of the targeted region or to prepare two pairs of TALENs to delete a large genomic region flanked by the two TALEN targeted loci. Often, the deletion of a large genomic region is a more favorable approach to eliminate a gene function because the genotyping of large genomic deletions is much easier than genotyping of indel mutations and the indel mutation may not completely eliminate the gene function in some cases (e.g. about 1/3 indels mutations may not cause ORF shift and alternative splicing of the disrupted exon). In our experience, the germ-line mutation frequency of a large genomic deletion is lower than that of indel mutations. However, we can still obtain founders with large genomic deletions by screening a limited number of P0 animals.
2. This reaction buffer system is crucial for successful TAL repeats assembly. When supplied with ATP, this buffer is compatible with T4 ligase.
3. In our experience, about 80 % of the clones contain the correct RVDs.
4. In our experience, about 90 % of the clones are correct.
5. Adding gentamicin into the 1/9× MMR could increase survival rate of the injected embryos.
6. Occasionally, TALENs induce embryonic cell toxicity. If most of the injected embryos are deformed or dead, lower the injection amount of TALEN mRNA.
7. The PCR primer should be carefully designed. Often, we design two pairs of primers to ensure that the large deletion could be readily detected.
8. Intercross of F1 heterozygotes with different loss-of-function alleles could also generate gene knockout animal in F2. To quickly identify the phenotypic consequences of target gene knockout, one can even intercross the P0 animals. However, the phenotypes should be interpreted with caution and some phenotypic changes could be caused by possibly off-target effects of TALENs. Outcross the founder could reduce off-target effects.

Acknowledgments

This work is supported by grants from the Research Grants Council of Hong Kong (N_CUHK413/12 to H Z and 463013 to CHKC), the National Natural Science Foundation of China (31172394 to CHKC) and the Shenzhen Municipal R&D Fund of Science and Technology (GJHS20120702105523308 to CHKC). We thank colleagues in our laboratories for helpful discussion on this project.

References

1. Young JJ, Cherone JM, Doyon Y, Ankoudinova I, Faraji FM, Lee AH, Ngo C, Guschin DY, Paschon DE, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Harland RM, Zeitler B (2011) Efficient targeted gene disruption in the soma and germ line of the frog Xenopus tropicalis using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A 108(17): 7052–7057. doi:10.1073/pnas.1102030108, 1102030108 [pii]
2. Ishibashi S, Cliffe R, Amaya E (2012) Highly efficient bi-allelic mutation rates using TALENs in Xenopus tropicalis. Biol Open 1(12): 1273–1276. doi:10.1242/bio.20123228, BIO20123228 [pii]
3. Lei Y, Guo X, Liu Y, Cao Y, Deng Y, Chen X, Cheng CH, Dawid IB, Chen Y, Zhao H (2012) Efficient targeted gene disruption in Xenopus embryos using engineered transcription activator-like effector nucleases (TALENs). Proc Natl Acad Sci U S A 109(43):17484–17489. doi:10.1073/pnas.1215421109, 1215421109 [pii]
4. Guo X, Zhang T, Hu Z, Zhang Y, Shi Z, Wang Q, Cui Y, Wang F, Zhao H, Chen Y (2014) Efficient RNA/Cas9-mediated genome editing in Xenopus tropicalis. Development 141(3):707–714. doi:10.1242/dev.099853
5. Chang N, Sun C, Gao L, Zhu D, Xu X, Zhu X, Xiong JW, Xi JJ (2013) Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Res 23(4):465–472. doi:10.1038/cr.2013.45
6. Hruscha A, Krawitz P, Rechenberg A, Heinrich V, Hecht J, Haass C, Schmid B (2013) Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. Development 140(24):4982–4987. doi:10.1242/dev.099085
7. Zhang Y, Zhang F, Li X, Baller JA, Qi Y, Starker CG, Bogdanove AJ, Voytas DF (2012) TALENs enable efficient plant genome engineering. Plant Physiol 161:20–27. doi:10.1104/pp.112.205179, pp.112.205179 [pii]
8. Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, Zhu N, Shen Y, Chen Y, Zhang B, Deng WM, Jiao R (2012) Efficient and specific modifications of the Drosophila genome by means of an easy TALEN strategy. J Genet Genomics 39(5):209–215. doi:10.1016/j.jgg.2012.04.003, S1673-8527(12)00077-X [pii]
9. Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, Amora R, Miller JC, Leung E, Meng X, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Meyer BJ (2011) Targeted genome editing across species using ZFNs and TALENs. Science 333(6040):307. doi:10.1126/science.1207773, science.1207773 [pii]
10. Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29(8):699–700. doi:10.1038/nbt.1939, nbt.1939 [pii]
11. Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JR (2011) Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 29(8):697–698. doi:10.1038/nbt.1934, nbt.1934 [pii]
12. Sung YH, Baek IJ, Kim DH, Jeon J, Lee J, Lee K, Jeong D, Kim JS, Lee HW (2013) Knockout mice created by TALEN-mediated gene targeting. Nat Biotechnol 31(1):23–24. doi:10.1038/nbt.2477
13. Wefers B, Meyer M, Ortiz O, Hrabe de Angelis M, Hansen J, Wurst W, Kuhn R (2013) Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. Proc Natl Acad Sci U S A 110(10):3782–3787. doi:10.1073/pnas.1218721110
14. Tesson L, Usal C, Menoret S, Leung E, Niles BJ, Remy S, Santiago Y, Vincent AI, Meng X, Zhang L, Gregory PD, Anegon I, Cost GJ (2011) Knockout rats generated by embryo

microinjection of TALENs. Nat Biotechnol 29(8):695–696. doi:10.1038/nbt.1940, nbt.1940 [pii]

15. Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, Voytas DF, Long CR, Whitelaw CB, Fahrenkrug SC (2012) Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci U S A 109(43):17382–17387. doi:10.1073/pnas.1211446109, 1211446109 [pii]
16. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res 39(12), e82. doi:10.1093/nar/gkr218, gkr218 [pii]
17. Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R (2011) Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29(8):731–734. doi:10.1038/nbt.1927, nbt.1927 [pii]
18. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326(5959):1509–1512. doi:10.1126/science.1178811, 1178811 [pii]
19. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326(5959):1501. doi:10.1126/science.1178817, 1178817 [pii]
20. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11(9):636–646. doi:10.1038/nrg2842, nrg2842 [pii]
21. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29(2):149–153. doi:10.1038/nbt.1775, nbt.1775 [pii]
22. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK (2012) FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 30(5):460–465. doi:10.1038/nbt.2170, nbt.2170 [pii]
23. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. PLoS One 3(11):e3647. doi:10.1371/journal.pone.0003647
24. Liu Y, Luo D, Zhao H, Zhu Z, Hu W, Cheng CH (2013) Inheritable and precise large genomic deletions of non-coding RNA genes in zebrafish using TALENs. PLoS One 8(10):e76387. doi:10.1371/journal.pone.0076387
25. Christian ML, Demorest ZL, Starker CG, Osborn MJ, Nyquist MD, Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF (2012) Targeting G with TAL effectors: a comparison of activities of TALENs constructed with NN and NK repeat variable di-residues. PLoS One 7(9), e45383. doi:10.1371/journal.pone.0045383, PONE-D-12-13580 [pii]
26. Doyon Y, Vo TD, Mendel MC, Greenberg SG, Wang J, Xia DF, Miller JC, Urnov FD, Gregory PD, Holmes MC (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods 8(1):74–79. doi:10.1038/nmeth.1539, nmeth.1539 [pii]
27. Rosen JN, Sweeney MF, Mably JD (2009) Microinjection of zebrafish embryos to analyze gene function. J Vis Exp 25:1115. doi:10.3791/1115
28. Lei Y, Guo X, Deng Y, Chen Y, Zhao H (2013) Generation of gene disruptions by transcription activator-like effector nucleases (TALENs) in Xenopus tropicalis embryos. Cell Biosci 3(1):21. doi:10.1186/2045-3701-3-21
29. Foley JE, Maeder ML, Pearlberg J, Joung JK, Peterson RT, Yeh JR (2009) Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. Nat Protoc 4(12):1855–1867. doi:10.1038/nprot.2009.209

Chapter 17

Genome Editing in Mice Using TALE Nucleases

Benedikt Wefers, Christina Brandl, Oskar Ortiz, Wolfgang Wurst, and Ralf Kühn

Abstract

Gene engineering for generating targeted mouse mutants is a key technology for biomedical research. Using TALENs as sequence-specific nucleases to induce targeted double-strand breaks, the mouse genome can be directly modified in zygotes in a single step without the need for embryonic stem cells. By embryo microinjection of TALEN mRNAs and targeting vectors, knockout and knock-in alleles can be generated fast and efficiently. In this chapter we provide protocols for the application of TALENs in mouse zygotes.

Key words Knock-in, Knockout, One-cell embryo, Gene targeting, TALEN, Mouse

1 Introduction

Engineering of the mouse genome to create targeted mutants is a key technology for biomedical research. The ascent of the mouse as genetic model organism is based on spontaneously occurring homologous recombination (HR) of gene targeting vectors encoding knockout or knock-in mutations in embryonic stem (ES) cells and the subsequent generation of chimaeric mice [1]. Alternatively, HR can be strongly stimulated at target sites by the creation of double-strand breaks (DSB) using sequence-specific nucleases. The first generation of such reagents, zinc-finger nucleases, are tedious to construct but demonstrated the power of this technology in a variety of biological systems, including the germ line in one-cell mouse embryos, thus making the creation of mouse mutants independent of ES cells [2, 3]. The application of nucleases for genome engineering was further stimulated by the discovery of the four letter DNA recognition code of the transcription activator-like (TAL) proteins [4, 5]. Based on the experience with ZFNs, the TAL-derived system could be readily adapted for the use as sequence-specific TAL nucleases (TALEN) [6, 7]. TALEN proteins bind to one strand of the DNA helix in a 5′-3′ direction. By combination of TAL elements following the four letters

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_17,

recognition code it is possible to design TALENs against virtually any target sequence preceded by a 5′-thymidine [4]. Functional TALENs are commonly built using a target sequence length of 15–18 bp, which corresponds to 14.5–17.5 TAL repeats. Due to the obligate dimerization of the FokI nuclease domains, a complete TALEN target site includes two TALEN binding sites. In order to position the two C-terminal FokI domains in the correct position, the two binding sites must be located on the opposite DNA strands. Moreover, the distance between the TALEN binding sites, the spacer region, is a crucial factor for the dimerization of the FokI domains. The ideal length of the spacer to enable dimerization is described as 14–16 bp [7]. Within the last years TAL proteins and TALENs were further characterized and applied for gene editing in mice and other species. These efforts showed that TALENs represent a versatile tool for gene modification in various species equal or superior to ZFNs. In particular, the majority of in silico designed TALENs exhibit considerable specific nuclease activity [7]. Taking advantage of the modular nature of the four basic TAL components a growing number of cloning protocols that enable to construct TALEN expression vectors within short time were reported. Following the ZFN paradigm TALENs were readily adopted to achieve genome modifications in mouse embryos. For delivery into the embryo´s cytoplasm or pronucleus the TALEN coding regions are first in vitro transcribed into mRNA. Upon the microinjection of mRNA into zygotes, the nucleases are translated, imported into nuclei, and cause DSBs in the target gene (Fig. 1).

DSBs can be repaired by the HR pathway [8] using gene targeting vectors that provide sequence homology regions flanking a desired genetic modification. Alternatively, DSBs can be closed by the nonhomologous end joining (NHEJ) pathway that religates the open DNA ends without repair template [9]. The DNA ends are frequently edited by the removal of nucleotides and frameshift (knockout) mutations can be obtained by targeting coding regions. In the presence of a coinjected targeting vector or oligonucleotide as templates, DSBs can be repaired by HR, enabling the introduction of precisely targeted mutations such as codon replacements or reporter gene insertions. Since the modified alleles are also present in the founder´s germ cells heterozygous mutant progeny can be obtained by breeding of founder mutants (Fig. 1). TAL nucleases were successfully applied for the targeting of 25 genes in zygotes derived from inbred or hybrid mice, suggesting that TALENs provide a routine tool for the manipulation of the mouse germ line (Table 1). Knockout alleles can be obtained at high frequency upon cytoplasmic TALEN microinjections such that up to 70 % of the pups represent founder mutants [10]. A single day of microinjection and embryo transfer is therefore sufficient to obtain a variety of knockout

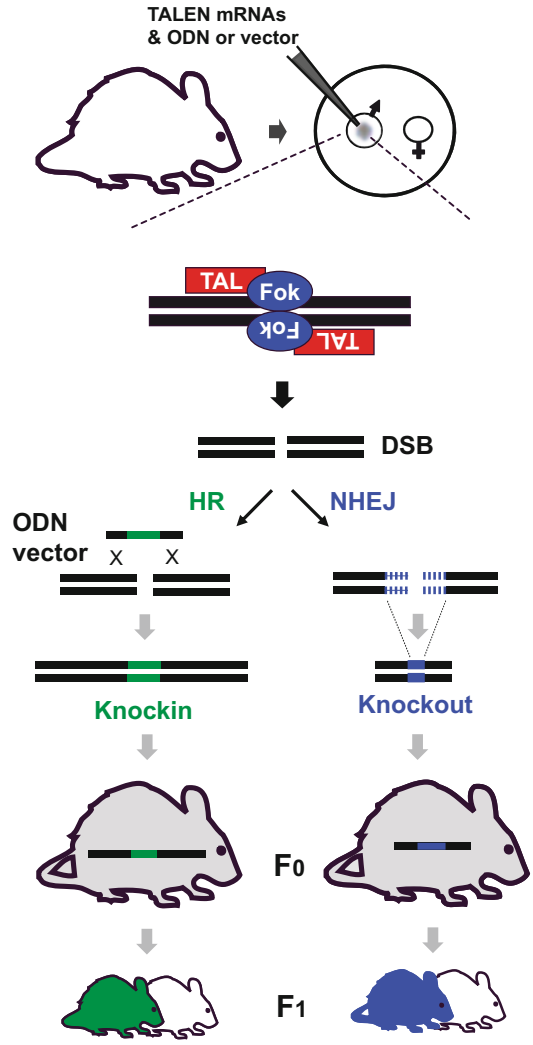


Fig. 1 Gene targeting in mouse zygotes using TALENs. For gene editing one-cell embryos are collected from wild-type female mice. TALEN mRNAs together with a targeting vector or a single-stranded oligodeoxynucleotide (ODN) are microinjected into the larger (male) pronucleus. The embryos are then transferred into foster females and represent the F0 generation of founder mutants. Upon mRNA translation, TAL-FokI fusion proteins (TALENs) recognize their binding sites on opposing strands of the target site and the dimerization of the two Fok domains generate a targeted double-strand break (DSB). DSBs can be repaired by homologous recombination (HR) or nonhomologous end-joining (NHEJ). In the presence of a repair template harboring homologous sequences to the target site and the desired mutation, HR mediates sequence replacement or insertion, leading to a knock-in (KI) mutant. Alternatively, the error-prone NHEJ repair closes the DSB, leading to random nucleotide deletions (or insertions) causing frameshift knockout (KO) mutations. Founder animals (F0), carrying the mutant allele in their germ line, are mated to wild-type mice to transfer mutant alleles to their offspring (F1)

Table 1
Gene editing in mice using TALEN

Target gene	Genotype	Delivery	Founder mutants Knockout (%)	Founder mutants Knock-in (%)	Reference
Pibf1, Sepw1	C57BL/6	Cytoplasm	56–77	n.a.	[10]
Rab38	FVB×C57BL/6	Pronucleus and cytoplasm	6	2	[14]
Zic2	CD1, C3H, C57BL/6	Cytoplasm	10–46	n.a.	[17]
Lepr and nine others	FVB, C57BL/6	Cytoplasm	13–67	n.a.	[18]
Fus, C9orf72	FVB×(C57BL/6×DBA)	Pronucleus	41	6.8	[15]
Fats	C57BL/6	Pronucleus	62	n.a.	[19]
miRNAs	C57BL/6×DBA	Cytoplasm	3–30	n.a.	[20]
Sry	C57BL/6×DBA	Cytoplasm	13	n.a.	[11]
Rosa26	C57BL/6	Pronucleus and cytoplasm	64	9	[13]
Eif2s3y	C57BL/6×DBA	Cytoplasm	not specified	n.a.	[12]
Agouti	FVB×C57BL/6	Cytoplasm	67–82	n.a.	[21]
Fgf10	C57BL/6×DBA	Cytoplasm	27–48	n.a.	[22]
Satb1	C57BL/6×DBA	Pronucleus	25	3	[23]
Scn8a	C57BL/6×SJL	Pronucleus	37	7	[24]
Rab38	FVB×C57BL/6	Pronucleus	6	n.a.	[16]

n.a. not applied

alleles. The production of knock-in alleles upon the pronuclear microinjection of TALEN mRNAs and targeting molecules relies on HR that occurs at lower frequency (2–9 %). Notably, TALENs enabled for the first time editing of genes located on the Y chromosome [11, 12] and can also be used for the insertion of for example reporter genes into the *Rosa26* locus, a frequently used docking site for the ubiquitous expression of transgene constructs [13]. The overall mutagenesis rate using TALENs in zygotes certainly depends on critical parameters such as the concentration and volume of the injected RNA/DNA solution, the TALEN nuclease activity, and possibly the differential susceptibility of various target genes for nucleases and repair proteins. The protocol presented in this chapter is based on our work on the pronuclear injection of TALEN mRNAs and targeting molecules

resulting into knockout or knock-in alleles at single target sites [14, 15]. In addition, we recently showed that two pairs of TALENs can be used to achieve genomic deletions [16].

2 Materials

2.1 Targeting Constructs

1. TALE nuclease expression vectors (*see* **Note 1**).
2. Targeting molecules for HDR (cloned gene targeting vector or single-stranded oligodeoxynucleotide (ssODN)) (*see* **Note 2**).
3. Plasmid DNA Maxiprep kit (*see* **Note 3**).

2.2 In Vitro Validation of TALEN Activity

1. Murine cell line (e.g., Neuro-2a) (*see* **Note 4**).
2. Complete growth medium.
3. 0.25 % trypsin–EDTA.
4. 1× phosphate buffered saline (PBS).
5. Mammalian cell transfection kit.
6. Genomic DNA extraction kit.
7. Gene-specific primers flanking the TALEN target region (*see* **Note 5**).
8. PCR purification kit.
9. T7 endonuclease I (T7EI).

2.2.1 Mismatch-Sensitive T7 Endonuclease I Assay

2.2.2 Cellular TALEN Reporter Assay

1. Synthetic complementary oligonucleotides containing the TALEN target site.
2. Generic TALEN reporter vector (Addgene #45964) (*see* **Note 6**).
3. Mammalian cell line (e.g., HEK293).
4. Complete growth medium.
5. 0.25 % trypsin–EDTA.
6. 1× phosphate buffered saline (PBS).
7. Mammalian cell transfection kit.
8. β-Gal reporter gene assay.
9. Plate luminometer.

2.3 TALEN mRNA Production

1. Restriction enzyme for linearization of TALEN expression plasmid.
2. Plasmid DNA purification kit.
3. In vitro transcription kit (*see* **Note 7**).
4. RNA purification kit (*see* **Note 8**).
5. Embryo-tested water.
6. $T_{10}E_{0.1}$ microinjection buffer: 10 mM Tris–HCl, 0.1 mM EDTA pH 7.4 (*see* **Note 9**).

2.4 Preparation of Microinjection Aliquots

1. Dialysis membranes.
2. Centrifugal filter.

2.5 Superovulation and Fertilized Embryo Collection

1. C57BL/6N female donor mice (3–6 weeks old) (*see* **Note 10**).
2. C57BL/6N male stud mice (8 weeks old) for mating with the females (*see* **Note 11**).
3. Hyaluronidase stock solution: 10 mg/ml hyaluronidase in KSOM medium. Aliquot in single-use aliquots of 50 µl and keep at –20 °C no longer than 6 months.
4. PMSG (Pregnant Mare's Serum Gonadotrophin) stock solution: 50 IU/ml PMSG dissolved in sterile water. Keep as single-use aliquots at –20 °C no longer than 2 months.
5. hCG (human chorionic gonadotrophin) stock solution: 50 IU/ml hCG dissolved in sterile water. Keep as single-use aliquots at –20 °C no longer than 2 months.
6. Complete KSOM embryo culture medium: KSOM, 100 IU/ml penicillin, 100 mg/ml streptomycin.
7. Embryo-tested mineral oil.
8. Mouth-operated transfer pipettes.
9. Stereomicroscope.
10. Humidified incubator at 37 °C and 5 % CO_2.

2.6 Pronuclear Microinjection of One-Cell Embryos

1. Inverted stereomicroscope microscope equipped with 40× objective, phase-contrast, polarizer, and differential interference contrast for pronucleus microinjection.
2. Micromanipulators for holding and microinjection capillaries.
3. Microinjector for pronucleus injection (*see* **Note 12**).
4. Holding capillaries (inner diameter 20–25 µm).
5. Microinjection capillaries (inner diameter 1.6 µm).
6. Microloader pipette tips.

2.7 Embryo Transfer into Recipient Females

1. Mouth-operated transfer pipettes.
2. Anesthetic working solution (*see* **Note 13**): 1.0 ml 2 % xylazine hydrochloride, 0.25 ml 50 mg/ml Ketamine hydrochloride, 4 ml saline. Store at 4 °C up to 2 weeks.
3. CD-1 pseudopregnant females.
4. CD-1 vasectomized males.
5. Surgical instruments and stereomicroscope.

2.8 Isolation of Genomic DNA

1. Tissue lysis stock solution: 50 mM Tris pH 8.0, 1 mM EDTA pH 8.0, 0.03 % SDS, 20 mM NaCl.
2. Tissue lysis buffer: 2.5 ml tissue lysis stock solution, 125 µl 20 mg/ml proteinase K.
3. Genomic DNA purification kit.

2.9 Genotyping

1. High-fidelity DNA polymerase.
2. Site-specific PCR primers.
3. PCR purification kit.
4. PCR cloning kit.
5. T7 endonuclease I.
6. HRMA device.
7. HRMA master mix kit.
8. Mineral oil.

3 Methods

3.1 In Vitro Validation of TALEN Activity

3.1.1 Mismatch-Sensitive T7 Endonuclease I Assay

1. Transfect a murine cell line with a pair of TALEN expression vectors following the transfection reagent's manufacturer's instructions.
2. Cultivate cells for 48 h.
3. Isolate genomic DNA from the transfected cells using a DNA extraction kit.
4. PCR-amplify the respective genomic locus and purify the PCR product using a PCR purification kit.
5. Determine the concentration using a UV spectrophotometer.
6. In a 0.2 ml tube mix 200 ng of PCR Product with 2 μl of T7EI buffer and add H_2O to a final volume of 19 μl.
7. Incubate in a PCR thermocycler: 95 °C for 5 min, cool to 85 °C (cooling rate −2 °C/min), cool to 25 °C (cooling rate −0.1 °C/min).
8. Add 10 U T7 endonuclease I.
9. Incubate at 37 °C for 15 min.
10. Add 2 μl of 0.25 M EDTA to stop the reaction.
11. Analyze the reaction products by agarose or acrylamide gel electrophoresis. TALEN activity is represented by smaller fragments resulted by the digestion at the mismatch site.

3.1.2 Cellular TALEN Reporter Assay

1. Design and order two complementary oligonucleotides that contain the respective TALEN target site and suitable overhangs for the generic TALEN reporter vector.
2. Anneal the two oligonucleotides by incubation at 95 °C for 5 min and slow cool-down to RT.
3. Clone the annealed oligonucleotides into the linearized generic TALEN reporter vector.
4. Transfect a murine cell line with a pair of TALEN expression vectors together with the cloned TALEN reporter vector following the transfection reagent's manufacturer's instructions.

5. Cultivate cells for 48 h.
6. Lyse cells and measure β-galactosidase activity following the manufacturer's instruction.

3.2 Preparation of TALE Nucleases mRNA

1. Linearize 10 μg of TALEN expression vector with a restriction enzyme cutting downstream of the poly-A signal.
2. Purify linearized plasmid using a DNA purification kit (*see* **Note 14**).
3. Load 1 μl on an agarose gel to check for complete digestion and to determine DNA concentration.
4. Prepare the in vitro transcription reaction (protocol for the mMessage mMachine T7 Ultra kit; *see* **Note 7**): add 10 μl T7 2× NTP/ARCA, 2 μl 10× T7 reaction buffer, 6 μl linearized plasmid (>1 μg), and 2 μl T7 enzyme mix in a 1.5 ml tube.
5. Mix reaction by pipetting and incubate for 2 h at 37 °C.
6. Prepare the polyadenylation reaction: to the in vitro transcription reaction add 36 μl nuclease-free water, 20 μl 5× E-PAP buffer, 10 μl 25 mM $MnCl_2$, 10 μl 10 mM ATP solution, and 4 μl E-PAP enzyme.
7. Mix reaction by pipetting and incubate for 1 h at 37 °C.
8. Purify the mRNA (protocol for the MEGAclear Kit; *see* **Note 8**): to the 100 μl RNA reaction add 350 μl binding buffer and 250 μl 100 % EtOH.
9. Apply sample to the column, spin for 1 min at 10,000 × *g*, and discard flow-through.
10. Wash column twice with 500 μl wash solution
11. Add 50 μl elution solution and incubate column for 5 min at 65 °C.
12. Spin for 1 min at 10,000 × *g*.
13. Repeat elution with 50 μl fresh elution solution.
14. Precipitate mRNA: to 100 μl eluate add 10 μl 5 M NH_4Ac and 275 μl 100 % EtOH
15. Incubate for >30 min at −20 °C.
16. Spin for 15 min at 16,000 × *g* and discard supernatant.
17. Wash mRNA pellet with 500 μl 70 % EtOH (*see* **Note 15**).
18. Spin for 1 min at 16,000 × *g* and discard supernatant.
19. Let the mRNA pellet air-dry for ~2 min and then resuspend it in 40 μl $T_{10}E_{0.1}$ microinjection buffer at 37 °C for 5 min (*see* **Note 16**).
20. Determine mRNA concentration using a UV spectrophotometer and store the mRNA solution at −80 °C until preparation of injection aliquots (*see* **Note 17**).

3.3 Preparation of Microinjection Aliquots

1. Preparation of an ssODN: dissolve the lyophilized ssODN in embryo-tested water to a concentration of 1 μg/μl (*see* **Note 18**).
2. Preparation of a gene targeting vector: precipitate 15 μg of plasmid DNA by adding 0.1 volumes of 3 M NaAc and 2.5 volumes of 100 % EtOH. Incubate for 30 min at RT, then pellet the DNA by centrifugation at 16,000 × *g* for 10 min. Discard supernatant and wash pellet with 500 μl 70 % EtOH. Spin for 1 min at 16,000 × *g* for 1 min, discard the supernatant and air-dry the pellet for 2 min. Resuspend the DNA pellet in 30 μl of $T_{10}E_{0.1}$ microinjection buffer for 15 min at 37 °C (*see* **Note 19**).
3. Determine DNA concentration of ssODN or gene targeting vector by UV spectrophotometry.
4. Mix the TALEN mRNA and targeting molecule stock solutions to a final concentration of 45 ng/μl of each mRNA and 15 ng/μl of targeting molecule in $T_{10}E_{0.1}$ microinjection buffer (*see* **Note 20**).
5. Filter the injection mix using an Ultrafree centrifugal filter column by centrifugation at 12,000 × *g* for 1 min (*see* **Note 21**).
6. Aliquot the filtrate into clean tubes and store the microinjection samples at −80 °C until the day of embryo injection.

3.4 Superovulation and Fertilized Embryo Collection

1. At day -3, the donor females are injected with 150 μl (7.5 UI) of freshly thawed PMSG aliquot at 1 p.m.-2 p.m.
2. 48 h later (at day -1), the donor females are injected with 150 μl (7.5 UI) of a freshly thawed aliquot of hCG at 12 noon-1 p.m.
3. After hCG injection, each donor female is mated overnight with a stud male and on the microinjection day (day 0) the copulation plugs are checked.
4. On the microinjection day (day 0), the foster females are identified (plug positive) and separated for later embryo transfers. The donor females (hCG-treated and mated) are sacrificed using a method approved by your institution. The oviducts are dissected out and each placed into a drop of 50 μl at 37 °C of KSOM medium (freshly supplemented with hyaluronidase stock solution, working concentration of 0.3 mg/ml) in a clean 6-cm Petri dish.
5. To dissect the cumulus complex with the fertilized oocytes, tear the ampullae with a pair of fine forceps and incubate it for 3–5 min until the cumulus cells fall off.
6. To wash off the hyaluronidase solution and remove residual cumulus cells, transfer the embryos to a prewarmed 50 μl drop of KSOM and pool all embryos in a prewarmed 200 μl drop of KSOM.

7. Transfer again the embryos to a new prewarmed 200 µl drop of KSOM and keep them at 37 °C until they are injected (no longer than 2 h) (*see* **Note 22**).

3.5 Pronuclear Microinjection of One-Cell Embryos

1. Load the pronucleus injection capillary with 2 µl of injection solution by using a microloader pipette tip. Connect the capillary to the microinjector and attach the injector to the micromanipulator. Set the injection conditions to 240 hPa as injection pressure and to 140 hPa as compensation pressure (*see* **Note 12**).
2. Mount the injection chamber, pipet 100 µl KSOM in the middle, and cover the drop with 200 µl of mineral oil.
3. Transfer the embryos to be injected into the upper region of the 100 µl KSOM drop within the injection chamber.
4. Fix the first fertilized embryo using the holding pipette and place it into the center of the KSOM drop.
5. Push the injection capillary to penetrate the zona pellucida and guide the capillary into the larger (male) pronucleus. Remain capillary in the pronucleus for 0.5–2 s until a swelling of the pronucleus becomes visible. Carefully withdraw the capillary and place the injected embryo in the lower part of the KSOM drop.
6. Repeat the procedure until all embryos are injected.
7. Transfer the surviving injected embryos into a 200 µl KSOM drop in a 6-cm petri dish covered with mineral oil. Incubate the plate at 37 °C, 5 % CO_2 until the implantation.

3.6 Embryo Transfer into Recipient Females

1. The day before the embryo transfer, the foster females are mated with vasectomized males.
2. For the embryo transfer, anesthetize a foster female with 0.25 ml of Anesthetic and make a 5-mm skin incision parallel to the dorsal midline above the position of the left or right oviduct. The ovarian fat pad should become visible.
3. Using a vessel clamp pull out the ovary and fix the fat pad. Locate the infundibulum and tear the bursa using two fine forceps.
4. Transfer the injected embryos to a fresh 200 µl aliquot KSOM without mineral oil to avoid coating of the transfer capillary. Load the capillary with 10 embryos (20 embryos are transferred per foster mother, 10 per oviduct).
5. Insert the tip of the transfer capillary into the infundibulum and carefully release the embryos. Relocate the ovary and fat pad into the abdomen and sew the peritoneum and skin.
6. Repeat the surgery for the second ovary and the second group of ten embryos. Repeat the same procedure until all embryos are transferred.

3.7 Isolation of Genomic DNA from Either ear-Punch Biopsies or Tail-Tip Biopsies of Founder Mice

3.7.1 Ear-Punch Biopsies

1. Ear-mark pups at the age of 3–4 weeks and collect biopsy in 0.2 ml tube.
2. Add 22 µl of tissue lysis buffer.
3. Dissolve tissue at 55 °C for 15 min. Then vortex sample and incubate for another 15 min.
4. Add 170 µl water to each tube and heat it to 100 °C.
5. Incubate for 10 min and store DNA sample at 4 °C no longer than 8 weeks.

3.7.2 Tail-Tip Biopsies

1. Collect tail-tip biopsies from pups and extract genomic DNA using a genomic DNA purification kit. Follow the manufacturer's instructions.

3.8 Genotyping of F_0 Founders and Their Progeny

3.8.1 Genotyping by Means of PCR and Direct Sequencing

1. Design site-specific primers to amplify the targeted region from the genomic DNA template. The resulting amplicon should be 200–1000 bp long.
2. Optimize PCR conditions and perform amplification using a high-fidelity polymerase.
3. Visualize PCR product by means of gel electrophoresis. Only one specific amplicon should be obtained.
4. Purify the PCR product using a PCR purification kit. Follow the manufacturer´s instructions.
5. Use forward or reverse primer from **step 1** to sequence the purified PCR product.
6. Screen for mutations by comparing founder sequence to wild-type sequence and checking for mixed sequencing peak at the desired mutation site (*see* **Note 23**).
7. In presence of mixed sequencing peaks subcloning should be performed using a PCR cloning kit to identify all actual mutations. Sequence 5–10 clones.

3.8.2 Genotyping by Means of PCR and T7 Endonuclease I Analysis

1. In a total volume of 19 µl mix 200 ng of purified PCR product (from Subheading 3.8.1, **step 4**) with 2 µl T7EI buffer.
2. Allow the formation of heteroduplexes using the following thermocycler conditions: 95 °C for 5 min, cool to 85 °C (cooling rate –2 °C/s), cool to 25 °C (cooling rate –0.1 °C/s).
3. Add 10 U of T7 endonuclease I and incubate at 37 °C for 15 min.
4. Stop the reaction by adding 2 µl 0.25 M EDTA and purify using a PCR purification kit.
5. Perform agarose gel electrophoresis and compare samples to a wild-type control. Mutant genotypes can be identified due to the presence of additional shorter bands.
6. Perform subcloning and sequencing of PCR products derived from mutant mice using a PCR cloning kit.

3.8.3 Genotyping by Means of Restriction Fragment Length Polymorphism (RFLP) Analysis

This genotyping strategy can be applied in case of depletion or introduction of an enzymatic restriction site caused by either NHEJ-mediated deletion or a HR-mediated knock-in at the target site.

1. In a total volume of 19 μl mix 500 ng of purified PCR product (from Subheading 3.8.1, **step 4**) with 5 U restriction enzyme and the respective digestion buffer.
2. Incubate restriction digest for 1 h at the enzyme's recommended temperature.
3. Visualize the digested samples by running a gel electrophoresis and check for the expected fragment sizes. Compare to wild-type samples.
4. Subclone and sequence the PCR product from animals with the expected band size patterns.

3.8.4 Genotyping by Means of High-Resolution Melt (HRM) Analysis

1. Design specific PCR primers for HRM analysis with an annealing temperature of 60–65 °C. The resulting amplicon should be 100–300 bp of length.
2. In a 96-well black/white measurement plate, set up a PCR reaction including 40 ng of genomic DNA of founder mice, 400 nM of each gene-specific primer, and 1× HRMA master mix kit in a total volume of 10 μl (*see* **Note 24**). Include a water control and wild-type control in your experiment.
3. Add 15 μl mineral oil to each well and seal plate with adhesive foil. Centrifuge briefly and perform PCR reaction using the following conditions: 98 °C, 30 s; 40 cycles (98 °C, 5 s; Tm, 5 s; 72 °C, 5 s); 72 °C, 1 min; denaturation at 98 °C, 1 min; rapid cool down to 25 °C for heteroduplex formation.
4. Analyze samples by using a HRMA device according to the manufacturer's instructions. Sequence alterations can be identified through different melting curves when compared to a wild-type control.
5. Subclone and sequence promising PCR products using a PCR cloning kit.

4 Notes

1. TALEN expression vectors can be either created using one of the publicly available TALEN construction kits (distributed by the noncommercial repository Addgene, www.addgene.org/TALEN) or ordered from commercial sources. Ensure that the vector plasmid contains a eukaryotic promoter for cellular TALEN activity assays and a T7 or U6 promoter for in vitro transcription.

2. The selection of the type of targeting molecule is dependent of the type of the intended mutation: synthesized ssODNs, which are limited in length and contain short homology regions of ~60 bp, are usually used to introduce subtle mutations like substitutions or small insertions or deletions of some base pairs. Cloned gene targeting vectors are used for larger modifications like insertions of cDNA, reporters, etc. To prevent unwanted mutations, use ssODNs of highest synthesis grade and check the sequence integrity of targeting vectors by Sanger sequencing.
3. Plasmid DNA quality is critical for in vitro transcription and microinjection. Use high quality Plasmid DNA Maxiprep and Plasmid DNA purification kits.
4. Any murine cell line can be used for the T7 endonuclease I assay. We chose Neuro-2a cells because they grow fast in standard growth medium and are easy to transfect.
5. The resulting amplicon should be 200–1000 bp in length. Use primers with high melting temperature to prevent unspecific amplicons.
6. The pCMV-TALEN-Rep contains a BstBI and an NruI restriction site to generate a 5′-GC overhang and a 3′ blunt end for cloning of the complementary oligonucleotides.
7. We recommend the mMessage mMachine T7 Ultra kit (Life Technologies) which contains an improved cap analog for higher yield of capped mRNA and a polyadenylation step, which is mandatory for gene targeting in mouse zygotes. However, in vitro transcription kits from other companies can be used, too. Ensure that the promoter of the transcription kit fits to the TALEN expression vector.
8. We recommend the MEGAclear kit (Life Technologies); however, RNA purification kits from other companies can be used, too.
9. For the preparation of microinjection buffer use highest quality reagents and embryo-tested water to prevent embryotoxicity upon microinjection.
10. The genetic background is dependent on your experimental needs. If another background is used, the protocol for superovulation has to be adopted empirically.
11. Stud males should be mated only once per week and replaced every 3–4 months.
12. We recommend the Eppendorf Femtojet and the given parameters in the protocol refer to the Femtojet. However, other mechanical or electronic microinjectors will work, too.
13. Use an anesthetic protocol that fulfills your institutional and governmental regulations.

14. Elute DNA in 30–40 μl elution buffer. The resulting concentration should be >166 ng/μl.
15. Use embryo-tested water for the preparation of 70 % EtOH.
16. Do not overdry the mRNA pellet because otherwise it becomes hard to resuspend. Add buffer as soon as pellet becomes transparent.
17. It is recommended to check the integrity of the mRNA sample after purification. This can either be done by standard denaturing gel electrophoresis or by RNA analysis chips, like the Agilent RNA 6000 Nano Kit.
18. To prevent embryo-toxicity due to impurities of the ssODN synthesis, we recommend dialyzing the dissolved ssODN against embryo-tested water using membrane filters.
19. We use supercoiled gene targeting vectors for microinjection to prevent unwanted, random integration of linearized DNA fragments into the genome.
20. We prepare for each day of microinjection one single-use aliquot of 10 μl total volume, i.e., for an experiment of 3 injection days 30 μl of injection mix is required.
21. Filtering the injection mix removes dust particles that could clog the injection capillary.
22. Approximately 300 embryos should be obtained from 15 C57BL/6N donor females.
23. Mixed sequences can be deconvoluted using appropriate online tools (e.g., Indelligent) or software packages (e.g., Mutation surveyor).
24. To avoid inaccuracies measure all samples in duplicates.

Acknowledgements

This work was supported by the European Union within the EUCOMM project (LSHG-CT-2005-018931, to W.W.), by the German Ministry of Education and Research within the TAL-CUT (03 V0261) and DIGTOP project (01GS0858) (to W.W. and R.K.) of the VIP and NGFN-Plus program.

References

1. Capecchi MR (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet 6:507–12
2. Meyer M, de Angelis MH, Wurst W et al (2010) Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proc Natl Acad Sci U S A 107:15022–6
3. Carbery ID, Ji D, Harrington A et al (2010) Targeted genome modification in mice using zinc-finger nucleases. Genetics 186:451–9

4. Boch J, Scholze H, Schornack S et al (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512
5. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501
6. Christian M, Cermak T, Doyle EL et al (2010) Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186:757–761
7. Miller JC, Tan S, Qiao G et al (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148
8. Heyer W-D, Ehmsen KT, Liu J (2010) Regulation of homologous recombination in eukaryotes. Annu Rev Genet 44:113–139
9. Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79:181–211
10. Sung YH, Baek I-J, Kim DH et al (2013) Knockout mice created by TALEN-mediated gene targeting. Nat Biotechnol 31:23–24
11. Kato T, Miyata K, Sonobe M et al (2013) Production of Sry knockout mouse using TALEN via oocyte injection. Sci Rep 3:3136
12. Matsubara Y, Kato T, Kashimada K et al (2015) TALEN-mediated gene disruption on Y Chromosome reveals Critical Role of EIF2S3Y in Mouse Spermatogenesis. Stem Cells Dev 24(10):1164–70
13. Kasparek P, Krausova M, Haneckova R et al (2014) Efficient gene targeting of the Rosa26 locus in mouse zygotes using TALE nucleases. FEBS Lett 588:3982–3988
14. Wefers B, Meyer M, Ortiz O et al (2013) Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. Proc Natl Acad Sci U S A 110:3782–3787
15. Panda SK, Wefers B, Ortiz O et al (2013) Highly efficient targeted mutagenesis in mice using TALENs. Genetics 195:703–713
16. Brandl C, Ortiz O, Röttig B et al (2014) Creation of targeted genomic deletions using TALEN or CRISPR/Cas nuclease pairs in one-cell mouse embryos. FEBS Open Bio 5:26–35
17. Davies B, Davies G, Preece C et al (2013) Site specific mutation of the Zic2 locus by microinjection of TALEN mRNA in mouse CD1, C3H and C57BL/6J oocytes. PLoS One 8, e60216
18. Qiu Z, Liu M, Chen Z et al (2013) High-efficiency and heritable gene targeting in mouse by transcription activator-like effector nucleases. Nucleic Acids Res 41, e120
19. Ma K, Wang J, Shen B et al (2013) Efficient targeting of FATS at a common fragile site in mice through TALEN-mediated double-hit genome modification. Biotechnol Lett 36(3):471–9
20. Takada S, Sato T, Ito Y et al (2013) Targeted Gene Deletion of miRNAs in Mice by TALEN System. PLoS One 8, e76004
21. Li C, Qi R, Singleterry R et al (2014) Simultaneous gene editing by injection of mRNAs encoding transcription activator-like effector nucleases into mouse zygotes. Mol Cell Biol 34:1649–1658
22. Yasue A, Mitsui SN, Watanabe T et al (2014) Highly efficient targeted mutagenesis in one-cell mouse embryos mediated by the TALEN and CRISPR/Cas systems. Sci Rep 4:5705
23. Sommer D, Peters A, Wirtz T et al (2014) Efficient genome engineering by targeted homologous recombination in mouse embryos using transcription activator-like effector nucleases. Nat Commun 5:3045
24. Jones JM, Meisler MH (2014) Modeling human epilepsy by TALEN targeting of mouse sodium channel Scn8a. Genesis 52:141–148

Chapter 18

Genome Editing in Rats Using TALE Nucleases

Laurent Tesson, Séverine Remy, Séverine Ménoret, Claire Usal, Reynald Thinard, Chloé Savignard, Anne De Cian, Carine Giovannangeli, Jean-Paul Concordet, and Ignacio Anegon

Abstract

The rat is an important animal model to understand gene function and model human diseases. Since recent years, the development of gene-specific nucleases has become important for generating new rat models of human diseases, to analyze the role of genes and to generate human antibodies. Transcription activator-like (TALE) nucleases efficiently create gene-specific knockout rats and lead to the possibility of gene targeting by homology-directed recombination (HDR) and generating knock-in rats. We describe a detailed protocol for generating knockout and knock-in rats via microinjection of TALE nucleases into fertilized eggs. This technology is an efficient, cost- and time-effective method for creating new rat models.

Key words TALE nucleases, TALEN, Gene editing, Targeted integration, Rat model, Homology-directed repair, Nonhomologous end joining, Knockout, Knock-in, Genomics, Genetic engineering, *Rosa26*, *Hprt*

1 Introduction

The rat is widely used as an experimental model [1], and during the last 20 years, a number of technologies have been developed to modify the rat genome including pronuclear microinjection of DNA [2], lentiviral transgenesis [3], ENU mutagenesis [4], and transposon mutagenesis [5]. But the lack of knockout or knock-in mutants has been an obstacle to fully develop the potential of rat model.

Since 2009, the emergence of engineered nucleases (ZFNs, TALE nucleases, meganucleases, and CRISPR/Cas9) permitted easy generation of many different knockout rats [6–9]. The application of gene-specific nucleases results in double strand DNA breaks (DSB) at desired loci that can be repaired by error-prone nonhomologous end-joining (NHEJ) [10]. Deletions and/or insertions (indels) generated by NHEJ lead to consequent rupture

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_18,

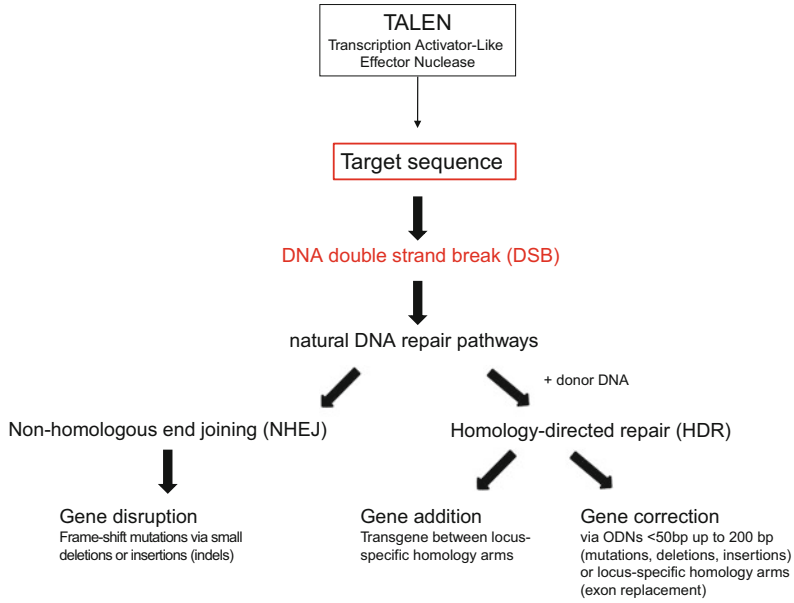


Fig. 1 A basic diagram of TALE nucleases function to create double strand breaks and stimulate NHEJ or HDR

of the coding frame, early stop codon generation and subsequent mRNA degradation [10]. In addition, genetic engineering by HDR is important to create knock-in rats expressing transgenes in a controlled way and to introduce precise gene modifications. The incorporation of exogenous sequences either placed between the homologous arms of a donor DNA sequence or as single-stranded oligonucleotides (ssODNs) is increased by the process of HDR after nuclease-induced DSB [11] (Fig. 1). Derivation of germ-line-competent rat ES cells [12] to generate knockout and knock-in rats [13, 14] by HDR was done, but they are still less robust than mouse ES cells [15].

In rats, ZFNs and TALE nucleases were used to generate HDR-modified rats [16, 17], in the case of rats with short insertions. More recently, we achieved the generation of HDR with TALE nucleases in rats in two different loci (*Hprt and Rosa26*) using three different pairs of nucleases and a donor DNA including an expression cassette between two homology arms designed as shown in Fig. 2a. In a third targeted locus (*IgM*) an exon exchange of rat exon 2 for the human one was performed successfully [18].

In this chapter, we describe the protocol used to generate knockout and knock-in rats by combining in vitro transcription of TALE nucleases-encoding plasmids, testing mRNA activity in cultured rat cells, preparing donor DNA and delivery to the one-cell embryo via standard transgenic microinjection techniques.

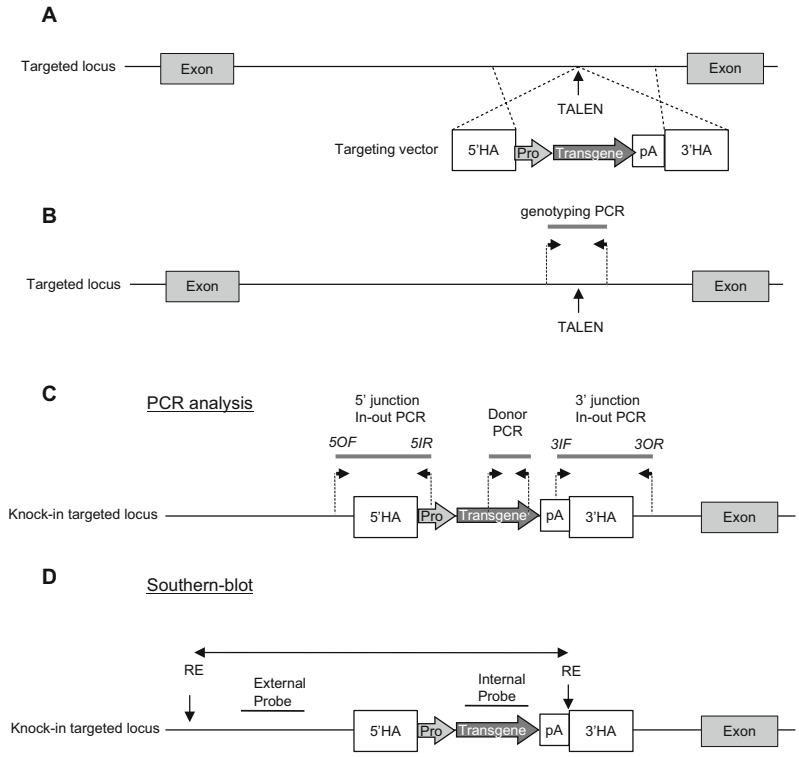


Fig. 2 Founder identification. (**a**) TALEN indicates the target site of the TALE nucleases to induce DSBs. Targeting vector with homology arms upstream (5′HA) and downstream (3′HA) of the TALEN site with expression cassette in between. Pro = promoter; pA = polyadenylation signal. (**b**) Primers position for genotyping analysis. (**c**) Primers position for targeted integration of donor and junction analysis. Primers could be chosen along the expression cassette to prove the integration of donor DNA but not the targeting integration. For that, junction in-out PCR must be done for each side of the donor DNA: at the 5′ side, with a primer upstream to the homology arm (5OF) and a primer in the expression cassette (5IR). At the 3′ side, with a primer inside the expression cassette (3IF) and a primer downstream to the homology arm (3OR). (**d**) Southern blot could be performed using a probe inside the donor DNA (internal probe) and a probe outside the donor DNA (external probe). RE = restriction enzyme site

2 Materials

2.1 TALE Nucleases

The generation of TALE nucleases is described in other chapters of this issue.

Usually, 2–3 TALE nucleases pairs are generated to target a given sequence and the most effective in vitro transfection of rat cells is then used in transgenesis.

2.1.1 TALE Nucleases mRNA Production

1. 10 μg of each plasmid encoding the monomers of TALE nucleases.
2. In vitro transcription kit.
3. Poly(A) tailing kit if not included in the transcription kit used.
4. Purification for transcription reactions kit.
5. MOPS-formamide agarose gel.

6. 1× MOPS buffer: 20 mM MOPS, 5 mM Na acetate, 2 mM EDTA.
7. Formamide.
8. 37 % formaldehyde.
9. Formamide loading buffer: 0.05 % xylene cyanol, 0.05 % bromophenol blue in formamide.
10. TE 10/1: 10 mM Tris–HCl pH 8.5, 1 mM EDTA.

2.1.2 Screening of TALE Nucleases and Gene Targeting in Rat Cells

1. C6 cells (ATCC; cat no. CCL-107).
2. F-12 medium supplemented with 5 % fetal bovine serum and 15 % horse serum.
3. 0.25 % Trypsin–EDTA.
4. 1× phosphate buffered saline (PBS).
5. Mammalian cell transfection kit.
6. Genomic DNA extraction kit.

2.2 HR Targeting Vectors

Depending of the sequence conversion by HDR needed, ODNs or plasmid-targeting vectors could be used.

1. Gene synthesis grade ODNs or isogenic genomic template PCR-amplified or gene synthetized.
2. Homology arms specific primers.
3. High-fidelity Taq DNA polymerase.
4. TA cloning kit.
5. Endo-free plasmid kit.
6. DNA purification kit.
7. Microinjection buffer: 5 mM Tris, 0.1 mM EDTA, pH 7.4, DNase–RNase-free water.

2.3 Gene Targeting in Rats

All experiments were compliant with Animal Protection Law of the French Republic (Article R214-89), which is in compliance with the European Community Council recommendations for the use of laboratory animals 86/609/ECC and EU directive 2010/63/EU for animal experiments. Experiments were approved by the CEEA Pays de la Loire committee (ref CEEA-2011-45).

2.3.1 Superovulation and Fertilized Embryo Collection

1. Immature Sprague–Dawley females (4–5 weeks old and 75–100 g in weight).
2. Fertile Sprague–Dawley males (2 months old) for mating with the immature females. They should be replaced every 8–12 months.
3. Pregnant Mare's Serum Gonadotrophin (PMSG). Working solution: 125 IU/mL made up with 0.9 % (w/v) NaCl. Store frozen at –20 °C in 1 mL aliquots.

4. Human chorionic gonadotrophin (hCG). Working solution: 150 IU/mL made up with 0.9 % (w/v) NaCl. Store at 4 °C (up to 2 weeks) in 1 mL aliquots.
5. Hyaluronidase. Stock solution: 10 mg/mL in PBS. Store at –20 °C (stable for several months) in 50 μL aliquots.
6. Embryo culture medium: M16 medium supplemented with 10 % fetal bovine serum, 2 mM L-glutamine, 100 IU/mL penicillin, and 100 mg/mL streptomycin.
7. Embryo-tested mineral oil.
8. 35-mm petri dishes.
9. Egg transfer Pasteur pipette assembled into a mouth-operated system made up of a mouthpiece, rubber tube (~40 cm), and a pipette holder.
10. Stereomicroscope with under-stage illumination.
11. Humidified incubator at 37 °C and 5 % CO_2.

2.3.2 Microinjection of One-Cell Embryos

1. Inverted microscope equipped with a 10× lens for low-magnification work and a 40× objective for microinjection.
2. Two micromanipulators for both holding and microinjection pipettes.
3. Microinjector.
4. Micropipette puller.
5. Holding pipette (Glass capillary).
6. Microinjection pipette (Glass capillary with filament).
7. Tips for loading capillaries with mRNAs or DNA solution.
8. Injection chamber.

2.3.3 Transfer of Microinjected Embryos into Recipient Females

1. Recipient Sprague–Dawley females (8–16 weeks of age) that have successfully had at least 1 litter (virgins often eat pups).
2. Vasectomized Sprague–Dawley males are needed to engender pseudopregnancy in the recipient females. They should be replaced every 8–12 months.
3. Surgical microscope and fiber optic illumination.
4. 1 mg/mL Adrenalin.
5. Embryo transfer pipette assembled into a mouthpiece.

2.4 Analysis of Gene Targeting in the 8 to 10-Day-Old Rats

2.4.1 Preparation of DNA from Tail Pups Biopsies

1. Tissue digestion buffer: 100 mM Tris–HCl pH 8.5, 10 mM EDTA, 0.2 % (w/v) SDS, 200 mM NaCl. Add Proteinase K 200 μg/mL immediately before use from a freshly made stock solution.
2. Phenol, Tris saturated pH 7.9.
3. Chloroform.
4. 100 % ethanol.

5. 70 % ethanol.
6. TE 10/1: 10 mM Tris–HCl pH 8.5, 1 mM EDTA.

2.4.2 Genotyping of Modified Rats by PCR and Sequencing

1. High-fidelity Taq DNA polymerase.
2. Genotyping specific primers.
3. T7 endonuclease I (M0302L; New England Biolabs).
4. Specific primers for transgene, 5′ and 3′ junction (in-out) PCR.
5. TA cloning kit.
6. Sequencing primers specific of the targeted gene.

3 Methods

3.1 Preparation of TALE Nucleases mRNA

The TALE nucleases plasmids may have a T7 promoter upstream of the TALE nucleases, whereas others may have a T3 or SP6 promoter.

3.1.1 In Vitro Transcription of TALE Nucleases mRNA

1. Linearize 10 μg of TALE nucleases expression plasmid DNA using an enzyme situated 3′ of the TALE nuclease expressing sequences and 5′ of the polyA sequence in 100 μL reaction containing 1× buffer, 1× BSA, and 80 U of enzyme restriction, at 37 °C for 2 h (*see* **Note 1**).
2. Extract the reactions with 100 μL of phenol/chloroform (1 vol/1 vol), pH 8.0, and centrifuge at 13,000 × *g* for 10 min.
3. Transfer the aqueous phase to a clean tube and precipitate with 4 μL 5 M NaCl and 250 μL 100 % ethanol. Place 1 h at –20 °C and centrifuge at 13,000 × *g* for 15 min at +4 °C.
4. Decant supernatant and wash pellet with 300 μL 70 % ethanol and centrifugation at 13,000 × *g* for 3 min
5. Air dry the pellet for 5 min and resuspend in 20 μL of TE 10/1.

3.1.2 In Vitro Transcription

Ambion's mMachine mMessage T7 Ultra kit uses the anti-reverse cap analog (ARCA) and improves the yield of capped mRNA. This kit provides also polyadenylation-tailing reaction. Other kits could be used with T3/SP6 promoters. If you use another kit without polyA-tailing step, immediately following in vitro transcription, use a polyadenosine (polyA)-tailing kit such as the Poly(A) tailing kit (*see* **Note 2**).

3.1.3 Purification and Gel Visualization of mRNA

1. Immediately following the polyA-tailing reaction, use a kit such as the Ambion MegaClear kit to purify the mRNA and elute in an RNAse-free solution.
2. Measure RNA concentration and OD 260/280 ratio (*see* **Note 3**).
3. Mix 1 μL of the eluted TALE nucleases mRNA and 1 μL of formamide loading buffer.
4. Heat to 70 °C for 3 min and place on ice.

5. Load to 1× MOPS-formamide agarose gel and run in 1× MOPS buffer at 12 V/cm for 20 min (*see* **Note 4**).
6. Dilute mRNA to 10 ng/μL final concentrations of the combination of each TALE nuclease in the mixture and store at −80 °C until use. Try microinjection first with a high mRNA concentration (10 ng/ul) and if you have embryos toxicity then decrease the mRNA concentration to 5 or 2 ng/μL.

3.2 Validating TALE Nucleases Activity Assay in Cultured Rat Cells

1. Two transfections should be performed: one with the TALE nucleases plasmids (each containing a monomer), the other with a plasmid donor containing an eGFP expression cassette to estimate transfection efficiency and serve as a negative control for TALE nucleases activity. Transfect ~0.3 million C6 cells with 1.5 μg each TALE nuclease plasmid or eGFP plasmid. We used the Amaxa nucleofection kit V for Amaxa shuttle.
2. Two days post-transfection, harvest cells via trypsinization.
3. Genomic DNA is prepared with DNA preparation kit following manufacturer's instructions.
4. PCRs were performed with a high-fidelity Taq DNA polymerase, run on a gel to verify size and specific amplification (without nonspecific products).
5. Analysis of genomic DNA for mutations induced by TALE nucleases with T7 endonuclease I assay which is a mismatch-specific digestion assay. The following PCR program is prone to make heteroduplexes formation presenting mismatches (one strand wild-type and one strand mutated) which should be digested by T7 endonuclease I. PCR products (typically 10 μL of 25 μL) were denatured in NEBuffer 2 (1× final in a total volume of 20 μL) and reannealed according to the following thermocycler conditions: 95 °C for 5 min, 95 to 25 °C at −0.5 °C/s, and 4 °C for 15 min. 5 μL of 0.5 U of T7 endonuclease I diluted in NEBuffer 2 (1×) were added, the mixture was incubated at 37 °C for 20 min.
6. Perform an electrophoresis with the mixture on a 2.4 % agarose gel containing ethidium bromide or freeze at −20 °C for later analysis.
7. Visualize gel for evidence of nuclease cleavage, which will be clear as new shorter bands not present in the control lane.

3.3 Preparation of Donor DNA for Microinjection (See Note 5)

1. ssODNs: resuspend in water at 1 mg/mL.
2. Supercoiled plasmid vectors: production of plasmid vectors needs to be done with endofree-plasmid kit.
3. Linear donor DNA:
   - Previous endofree-plasmid is digested with appropriate restriction enzymes to excise the donor DNA for HR.

Table 1
Results of donor DNA + TALE nucleases microinjection

Target locus	Dose mRNA/DNA (ng/μL)	No. injected eggs (% viable eggs)	No. E15 (e) or pups (p) (%)[a]	No. of GFP+ animals (%)[a]	No. of RI-positive animals (%)[a]	No. of HR-positive animals (%)[a]	No. of indels positive animals (%)[a]
Hprt1	50+50/2	113 (73.5)	8e (13.3)	0 (0)	0 (0)	0 (0)	1 (1.67)
	20+20/2	209 (78.5)	17e (10.7)	0 (0)	0 (0)	0 (0)	2 (1.26)
	10+10/2	164 (72.6)	8e (8.2)	1 (1.02)	0 (0)	1 (1.02)	2 (2.04)
Hprt2	20+20/2	69 (78.3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	10+10/2	258 (72.1)	32e (20.3)	1 (0.63)	0 (0)	1 (0.63)	15 (9.49)
	5+5/2	148 (85.8)	33e (26.2)	4 (3.17)	1 (0.79)	3 (2.38)	17 (13.5)
	5+5/2[b]	159 (71.1)	25e (23.8)	1 (0.95)	1 (0.95)	0 (0)	4 (3.81)
Rosa26	20+20/2	134 (52.2)	5e (10.6)	0 (0)	0 (0)	0 (0)	1 (2.13)
	10+10/2	180 (76.7)	8e (8.8)	3 (3.30)	1 (1.10)	2 (2.20)	3 (3.30)
	5+5/2	161 (80.8)	11p (9.2)	2 (1.67)	0 (0)	2 (1.67)	5 (4.17)
IgM	10+10/10	1063 (72.7)	5e–72p (15.8)	/	8 (1.64)	3 (0.62)	54 (11.1)
	10+10/10[b]	410 (75.1)	64p (22.9)	/	2 (0.72)	0 (0)	35 (12.5)

TALE nucleases, as mRNA, were injected at different concentration in combination with of donor DNA, either in its linear form or in its circular form, both into the cytoplasm and into the male pronucleus. Potential toxicity was assessed by the number of pups obtained following the transfer of injected eggs. The number of E15 or alive pups, which have integrated the donor DNA sequence either by random integration (RI) or by HDR integration (PCR positive both at the 5′ and 3′ ends), or which have presented indels mutations, are reported in the three last columns of the table

[a]Percentages indicated between brackets correspond to percentage of transferred embryos

[b]Circular form of donor DNA. For *Hprt* and *Rosa* locus, homology arms are 800 bp in length, whereas for *IgM* locus they are 750 bp and 1460 bp

- Perform an electrophoresis on a 0.5× TAE agarose gel.
- Cut the digested linear donor DNA from agarose gel.
- Electroeluted and purified linear donor DNA.
- Linear DNA is quantified using a NanoDrop-1000 and stored at −20 °C until use (*see* **Note 6**).

4. ssODNs (working concentration is 15 ng/μL), plasmid vectors or linear donor DNA (working concentration is 2–10 ng/μL) are mixed with different TALENs mRNA concentration in microinjection buffer (Table 1) and stored at −80 °C until use.

3.4 Gene Targeting in Rat Embryos

3.4.1 Superovulation and Fertilized Embryos Collection

Sprague–Dawley female rats are administrated gonadotrophins prior to mating to increase the number of released eggs (superovulation). This technique can yield around 30–40 eggs per donor female.

1. Inject prepubescent Sprague–Dawley female rats intraperitoneally with 25 IU of PMSG between 12 pm and 1 pm on day -2, followed by 30 IU of hCG between 3 and 4 pm on day 0.
2. Individually mate each hormone-treated female with one fertile male overnight. On the morning of day 1, check the females for copulation plugs.
3. Sacrifice the isoflurane-anesthetized female rats by cervical dislocation around 10 am on the morning of day 1.
4. Excise the oviducts and transfer them to a dish containing PBS or M2 medium at room temperature. Embryos, enclosed by cumulus mass cells, can be released from the swollen ampullae (the upper portion of the oviduct) by gently tugging and opening the walls of the ampullae with fine forceps.
5. Transfer the embryos using an egg transfer pipette to a dish containing a pre-warmed hyaluronidase solution (500 μg/mL in PBS), which enzymatically digests the cumulus cells thus releasing the embryos. A few minutes of treatment are sufficient (longer incubation can be toxic for embryos); gentle up and down pipetting can facilitate the process.
6. Transfer the embryos to PBS to wash off the hyaluronidase solution and preserve their viability.
7. Finally, transfer the embryos in fresh pre-warmed M16 embryo culture medium in a humidified 37 °C incubator under 5 % CO_2 until microinjection.

3.4.2 Delivery of TALE Nucleases mRNA with or Without Donor DNA to One-Cell Embryos

The TALE nucleases mRNAs alone are delivered into the cytoplasm of fertilized single-cell embryos. The mixture TALE nucleases mRNAs and donor DNA is injected in two steps, first into the pronucleus and then into the cytoplasm during the withdrawal of the injection pipette.

1. Thaw a newly frozen aliquot (~10 μL) of mRNA ± DNA solution at room temperature and briefly spin at 10,000 rpm. Keep the solution on ice during the microinjection and discard after 1-day session of microinjection.
2. Load approximately 2–3 μL of nucleic acid solution into the microinjection pipette with a microloader tip.
3. The loaded micropipette is held in place onto a micromanipulator connected to the N_2 gas-operated pressure injector.
4. Transfer one-cell fertilized embryos (in batches of 30–40) in a microdrop of embryo culture medium in the injection chamber and cover with mineral oil to prevent evaporation and maintain osmolarity.

5. Mount the chamber on the stage of an inverted microscope and monitor the injection procedure under 400× magnification.
6. Hold fertilized embryos (pronuclei visible) in place against the holding pipette using gentle negative pressure. Hold the micropipette loaded in place onto a micromanipulator.
7. Using the micromanipulator to guide the pipette, push the tip through the zona pellucida into the cytoplasm or into the pronucleus.
8. Using gentle positive pressure, usually 40 psi for 1 s, the solution flows continuously from the pipette (*see* **Note 7**). If it is delivered into the cytoplasm, the injected solution spreads to form a drop having a diameter approximately similar to the pronuclei's one. The delivery into the pronucleus is successful when the pronucleus swells to double its original volume.
9. The microinjection pipette should be changed every 30–40 embryos.
10. After the injection, transfer the surviving embryos in embryo culture medium equilibrated at 37 °C, 5 % CO_2, and then keep them in a 37 °C humidified incubator under 5 % CO_2 until implantation.

3.4.3 Transfer of Embryos into Recipient Females

1. Obtain pseudopregnant Sprague–Dawley females needed to host the microinjected embryos by mating sexually mature Sprague–Dawley females in estrus with vasectomized mature males, the night before the day of microinjection. When transfer is performed the same day of microinjection higher yields of newborn animals are obtained. Confirm mating the next morning by checking for a plug. We mate 16 vasectomized males individually with females and usually obtain between 1 and 3 matings. The non-mated females are used for another mating 10 days after the first attempt.
2. Transfer injected embryos into the oviduct of host females by using an embryo transfer pipette, preferably the same day as the microinjection to increase the rate of implantation. In general, no more than 24 embryos are transferred bilaterally into the oviducts.

3.5 Analysis of DNA Mutations from Rat Tissues

1. Harvest short tail biopsies (~6 mm) from 8 to 10-days-old rats.
2. Add 500 μL tissue digestion buffer to the tail biopsies in a 1.5 mL tube.

3.5.1 Analysis of DNA from Rat Tissues from Rat Pups

3. Incubate at 56 °C for overnight.
4. Extract the reactions with 100 μL of phenol/chloroform (1 vol/1 vol), pH 8.0, and centrifuge at 13,000 × *g* for 10 min.
5. Transfer the aqueous phase to a clean tube and precipitate with 4 μL 5 M NaCl and 250 μL 100 % ethanol.

6. Invert tubes 20–30 times until nucleic acids precipitate.
7. Centrifuge at 13,000 × *g* for 10 min.
8. Wash nucleic acid pellet with 1 mL 70 % ethanol.
9. Allow pellet to dry for 5 min. Do not dry the pellet to completion.
10. Dissolve the pellet in 100 μL TE for >15 min.

3.5.2 Analysis of Gene Modifications by PCR (Fig. 2b)

1. Design primers that will amplify PCR products comprised between 300 and 600 bp of the region of interest from rat genome.
2. Optimize PCR conditions so that only one specific product is obtained using high-fidelity Taq DNA polymerase. PCR amplify the region of interest from the crude nucleic acid preparation.
3. Performed a sample electrophoresis on a 2 % agarose gel in 0.5× TAE or freeze at −20 °C for later analysis.
4. If the deletions are large, the visualization of PCR products on gel may allow identifying some mutated founders but all animals should be analyzed using the T7 nuclease assay and sequencing (*see* **Note 8**).

3.5.3 Enzymatic Mismatch Cleavage with T7 Endonuclease I Assay and Sequencing

1. Perform enzymatic mismatch cleavage with T7 endonuclease I assay as described in Subheading 3.2 **step 5**.
2. The products are separated by electrophoresis to visualize evidence of T7 nuclease cleavage. The presence of mismatches is visualized by the generation of lower molecular weight bands**.**

3.5.4 Confirmation of Mutations by Sequencing of PCR Products

1. PCR products could be directly sequenced and the presence of mutations is visualized by comparison with wild-type sequences (*see* **Note 9**).
2. If more than one mutation is present in the same founder animal, PCR products should be cloned via TA cloning kit (e.g., TOPO PCR cloning system and 10–20 bacteria colonies sequenced independently) (*see* **Note 10**).

3.5.5 Analysis of Knock-In Mutants for DNA Donor Including Expression Cassette

1. Four distinct amplifications need to be done:
   (a) Primer set amplified the target region for TALE nucleases activity and/or positive control for genomic DNA quality (Fig. 2b) (*see* **Note 11**).
   (b) Primer set amplified sequences in the introduced DNA sequences between the homology arms to detect either knock-in or random integration (Fig. 2c).
   (c) Primer sets, called in-out primers, amplified the 5′ and 3′ junctions of targeted insertion and sign up the targeted integration events (Fig. 2c).

2. Analysis of concatemers presence.
   - Primers "in" (5IR-3IF) should be used to amplify integrated concatemers (head-to-tail) of donor DNA (*see* **Note 12**).
3. Southern blot
   Southern blot analysis confirmed integration of the donor DNA to the target loci. Both external and internal probes have to be used (Fig. 2d) (*see* **Note 13**).
4. Off-target
   Potential off-target sites could be predicted, for example, on Prognos Web site (http://baolab.bme.gatech.edu/Research/BioinformaticTools/prognos.html) which gives several ranking sites with PCR primers for each site. The five highest ranked off-target sites should be tested by PCR and should be sequenced to show if mutations happened in these off-target sites.

3.6 Results

TALE nucleases are a very useful tool for the generation of knockout and/or knock-in rats. In our hands, the mutation rate with TALE nucleases mRNA microinjection in one-cell rat embryos is 2.28–11.1 % of transferred embryos [19], and the frequencies of HDR obtained were between 0.63 and 2.38 % [18]. Nevertheless, it is important to improve efficiencies, particularly for HDR, and ongoing research on the different steps involved in these techniques will likely increase efficiency and specificity.

Linear or supercoiled donor DNA sequence could be used in HDR strategies. For our HDR insertions, we obtained HDR events only with linear forms, whereas others in rats using ZFNs generated HDR events both with linear and supercoiled forms [16, 18] (Table 1).

Off-target activity of TALE nucleases is quite low [7, 20]. Off effects were detected at high concentrations of TALE nucleases, such as when using mRNA and not DNA encoding sequences [7].

Despite potential low levels of mosaicism, most of the mutated founders transmitted the mutation to the offspring [7, 18].

In summary, two major advantages of this approach are the ability to apply ZFNs/TALE nucleases or CRISPR/Cas9 to any rat strain and to accelerate the generation of knockout or knock-in animals (~4 months) vs. the use of embryonic stem cells (>12 months) [21].

4 Notes

1. To get a high quality mRNA, it is very important to have a plasmid prepared with an endofree-plasmid kit. Plasmids prepared with minipreps column often lead to poor mRNA quality.
2. We have never used TALE nucleases mRNA without polyadenylation but this has been done for microinjection in livestock [22].

3. Our normal yield is about 30 μg/reaction, using Ambion's transcription and tailing kits and Ambion's purification kit. OD ratio 260/280 is always above 2.0.
4. New running buffer, high voltage and a short running time are critical for maintaining mRNA integrity during electrophoresis with a proper size marker. TALE nucleases mRNAs runs at ~2.5–3 kb as one defined band with minor smearing below the major band.
5. ssODNs could be used for small gene correction. The range size is between <50 up to 200 nucleotides which is the synthesis limit. It is important to keep the maximum amount of homology in the sequence [23]. For gene addition with the use of homology arms, we successfully obtained HDR with 800 bp either side of an expression cassette. The size of the homology arm can vary in length, but are typically longer than 500 bp [24]. The distance between the homology arms and the DSB could be a parameter influencing HDR. The efficiency decreases rapidly with distance (100 bp) from the nuclease cleavage point. Our results on IgM locus (exon exchange with homology at 150 bp from the break point) were comparable to those using homology arms contiguous to the DSB.
6. Both linear and circular donor DNA could be used. In the genome, linear DNA is more prone to random integration than the circular DNA and subjected to often form concatemers. Even though the mechanism of concatemer formation is not clearly understood, it might be limited by in situ excision of donor sequences from plasmids by flanking donor sequences by nuclease target sites [25].
7. The pressure and duration of each injection had to be adjusted. If the pressure is set too high, the embryo will be lysed. The volume injected is dependent on the pressure, the injection time, and the size of the hole in the injection pipette.
8. If large deletion occurs, no amplification could happen, and it is recommended to perform a control PCR to eliminate the possibility of inhibition. If there are no PCR inhibitors, then use primers amplifying larger product (1 Kb or more). PCR artifacts could also happen. For example, we had a PCR amplification of only 247 bp band and not the wild type one (395 bp). This founder rat seems to be a homozygous mutated rat with a 148 bp deletion but when the germ-line transmission was analyzed, we found heterozygous mutated rats and, surprisingly, wild-type rats. In the founder, the smallest band was preferentially amplified over the wild type one, maybe due to competitive PCR and/or mosaicism. This could happen only for large deletion.
9. If mixed sequencing peaks are present, they can be deconvoluted with software (Mutation Surveyor) or free Web-based tools (e.g., Indelligent).

10. If in the mosaic founder rat, one deletion results in a mutation of interest, you must make F1 offspring to look at mutation transmission.
11. This primer set could be used for ssODNs analysis insertion (restriction site for example). When HDR is done with a targeting vector (expression cassette), the use of this primer set could result in any amplification in the case of homozygous knock-in rat.
12. Primer 5IR could be used alone to show head-to-head concatemers and primer 3IF alone for tail-to-tail concatemers. In our hands, with or without concatemers, GFP expression is equivalent in rats with knock-in of an expression cassette in *Rosa26* locus. We, usually, used linear DNA as donor DNA but circular DNA might be used to reduce concatemer formation.
13. This technique is completely described in Southern E, Nat Protoc, 2006 [26]. Despite a correct gene targeting, as identified by in-out PCRs and Southern blot, HDR could resulted in ectopic integration of the targeting vector including flanking homology arms of variable size [27].

Acknowledgments

This work was in part funded by Region Pays de la Loire, Biogenouest, IBiSA, and TEFOR (Investissements d'Avenir, French government).

References

1. Aitman TJ, Critser JK, Cuppen E et al (2008) Progress and prospects in rat genetics: a community view. Nat Genet 40:516–522
2. Ménoret S, Remy S, Usal C et al (2010) Generation of transgenic rats by microinjection of short DNA fragments. Methods Mol Biol 597:81–92
3. Remy S, Tesson L, Usal C et al (2010) New lines of GFP transgenic rats relevant for generative medicine and gene therapy. Transgenic Res 19:745–763
4. Zan Y, Haag JD, Chen KS et al (2003) Production of knockout rats using ENU mutagenesis and a yeast-based screening assay. Nat Biotechnol 21:645–651
5. Izsvák Z, Fröhlich J, Grabundzija I et al (2010) Generating knockout rats by transposon mutagenesis in spermatogonial stem cells. Nat Methods 7:443–445
6. Geurts AM, Cost GJ, Freyvert Y et al (2009) Knockout rats via embryo microinjection of zinc-finger nucleases. Science 325:433
7. Tesson L, Usal C, Menoret S et al (2011) Knockout rats generated by embryo microinjection of TALENs. Nat Biotechnol 29:695–696
8. Menoret S, Fontaniere S, Jantz D et al (2013) Generation of Rag1-knockout immunodeficient rats and mice using engineered meganucleases. FASEB J 27:703–711
9. Li D, Qiu Z, Shao Y et al (2013) Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. Nat Biotechnol 31:681–683
10. Gaj T, Gersbach CA, Barbas CF 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 31:397–405
11. Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. Science 300:763
12. Buehr M, Meek S, Blair K et al (2008) Capture of authentic embryonic stem cells from rat blastocysts. Cell 135:1287–1298

13. Yamamoto S, Nakata M, Sasada R et al (2012) Derivation of rat embryonic stem cells and generation of protease-activated receptor-2 knockout rats. Transgenic Res 21:743–755
14. Meek S, Buehr M, Sutherland L et al (2010) Efficient gene targeting by homologous recombination in rat embryonic stem cells. PLoS One 5, e14225
15. Zheng S, Geghman K, Shenoy S et al (2012) Retake the center stage–new development of rat genetics. J Genet Genomics 39:261–268
16. Cui X, Ji D, Fisher DA et al (2011) Targeted integration in rat and mouse embryos with zinc-finger nucleases. Nat Biotechnol 29:64–67
17. Ponce de Leon V, Merillat AM, Tesson L et al (2014) Generation of TALEN-mediated GRdim knock-in rats by homologous recombination. PLoS One 9, e88146
18. Remy S, Tesson L, Ménoret S et al (2014) Efficient gene targeting by homology-directed repair in rat zygotes using TALE nucleases. Genome Res 24:1371–1383
19. Ménoret S, Tesson L, Remy S et al (2014) Gene targeting in rats using transcription activator-like effector nucleases. Methods 69:102–107
20. Huang P, Xiao A, Zhou M et al (2011) Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 29:699–700
21. Dwinell MR, Lazar J, Geurts AM (2011) The emerging role for rat models in gene discovery. Mamm Genome 22:466–475
22. Carlson DF, Tan W, Lillico SG et al (2012) Efficient TALE-mediated knockout in livestock. Proc Natl Acad Sci U S A 109:17382–17387
23. Bedell VM, Wang Y, Campbell JM et al (2012) In vivo genome editing using a high-efficiency TALEN system. Nature 491:114–118
24. Beumer KJ, Trautman JK, Christian M et al (2013) Comparing zinc finger nucleases and transcription activator-like effector nucleases for gene targeting in Drosophila. G3 (Bethesda) 3:1717–1725
25. Ochiai H, Sakamoto N, Fujita K et al (2012) Zinc-finger nuclease-mediated targeted insertion of reporter genes for quantitative imaging of gene expression in sea urchin embryos. Proc Natl Acad Sci U S A 109:10915–10920
26. Southern E (2006) Southern blotting. Nat Protoc 1:518–525
27. Mangerich A, Scherthan H, Diefenbach J et al (2009) A caveat in mouse genetic engineering: ectopic gene targeting in ES cells by bidirectional extension of the homology arms of a gene replacement vector carrying human PARP-1. Transgenic Res 18:261–279

Chapter 19

Designer Nuclease-Mediated Generation of Knockout THP1 Cells

Tobias Schmidt*, Jonathan L. Schmid-Burgk*, Thomas S. Ebert, Moritz M. Gaidt, and Veit Hornung

Abstract

Recent developments in the field of designer nucleases allow the efficient and specific manipulation of genomic architectures in eukaryotic cell lines. To this end, it has become possible to introduce DNA double strand breaks (DSBs) at user-defined genomic loci. If located in critical coding regions of genes, thus induced DSBs can lead to insertions or deletions (indels) that result in frameshift mutations and thereby the knockout of the target gene. In this chapter, we describe a step-by-step workflow for establishing knockout cell clones of the difficult-to-transfect suspension cell line THP1. The here described protocol encompasses electroporation, cell cloning, and a deep sequencing-based genotyping step that allows the in-parallel analysis of 96 cell clones per gene of interest. Furthermore, we describe the use of the analysis tool *OutKnocker* that allows rapid identification of cell clones with all-allelic frameshift mutations.

Key words CRISPR/Cas9, Genome editing, Knockout cell lines, Barcoded deep sequencing, Genotyping

1 Introduction

Targeted genome editing technologies based on designer nucleases such as ZFNs [1], TALENs [2], or RGENs [3–5] enable highly efficient and flexible induction of DNA double-strand breaks (DSB) in eukaryotic genomes. DSBs can initiate two different repair pathways that can be exploited to change the genomic architecture at user-defined genomic loci: Homologous recombination (HR) specifically repairs DSB sites using a homologous DNA template like a sister chromatid or exogenous DNA [6], whereas non-homolougous end joining (NHEJ) is an error-prone mismatch repair pathway that can give rise to insertions or deletions [7, 8]. NHEJ-induced indels can result in frameshift

* Tobias Schmidt and Jonathan L. Schmid-Burgk contributed equally to this work

Ralf Kühn et al. (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0_19,

mutations; however, loss of functional protein expression in cells can only be achieved by targeting critical exons of all alleles present. Therefore, upon the introduction of indel mutations into a polyclonal cell population, it is necessary to perform single cell cloning and genotyping of the targeted genetic locus to obtain cell lines with all-allelic frameshift mutations.

In this chapter, we describe a step-by-step protocol for efficient targeting of the difficult-to-transfect suspension cell line THP1 using the CRISPR/Cas9 system (Fig. 1). For this purpose, we decided to use the CRISPR/Cas9 system due to its superior genome editing activity. However, this protocol is also suitable for a pair of TALENs or ZFNs. Upon electroporation of THP1 cells with expression plasmids encoding a sgRNA (single guide RNA) of interest and a codon-optimized version of the Cas9 protein, THP1 cells expressing high levels of Cas9 are enriched by fluorescence-activated cell sorting (FACS). Subsequently monoclonal cell lines are generated by limiting dilution cloning. Thus obtained cell clones are then duplicated and screened for the desired genetic lesion of interest. To this end, genomic DNA of one replicate of each cell clone is obtained and subjected to a two-step genotyping PCR that introduces barcodes for multiplexing. The genotyping amplicons are then subjected to a deep sequencing reaction, whereas the resulting sequencing raw data are specifically evaluated for out-of-frame indel mutations using the software tool *OutKnocker* (www.outknocker.org) [9]. Altogether our protocol provides easy access to the generation of cellular loss-of-function models in the human system.

2 Materials

In the following we describe the generation of knockout THP1 cells using the TLR2 gene as an example [9].

2.1 Oligonucleotide Primer for Sequencing

2.1.1 Locus-Specific Primer for the First Level PCR

hsTLR2 genotyping FWD:

5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-TGGAATATGCAGCCTCCGGATTG -3′

hsTLR2 genotyping REV:

5′-TGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-TCAGAACTATCCACTGGTGAAAC -3′

Underlined is the gene-specific part of the primer. The 5′ part is the adapter sequence for the second level PCR

2.1.2 Barcode Primer for the Second Level PCR

Name	Sequence
J501	AATGATACGGCGACCACCGAGATCTACAC GCGAAGAT ACACTCTTTCCCTACACGACGCT
J502	AATGATACGGCGACCACCGAGATCTACAC CCCTTGTG ACACTCTTTCCCTACACGACGCT

(continued)

Name	Sequence
J503	AATGATACGGCGACCACCGAGATCTACAC TCGCATCA ACACTCTTTCCCTACACGACGCT
J504	AATGATACGGCGACCACCGAGATCTACAC AGAATCAA ACACTCTTTCCCTACACGACGCT
J505	AATGATACGGCGACCACCGAGATCTACAC GTTTGAGC ACACTCTTTCCCTACACGACGCT
J506	AATGATACGGCGACCACCGAGATCTACAC GACTAACG ACACTCTTTCCCTACACGACGCT
J507	AATGATACGGCGACCACCGAGATCTACAC CTCACATA ACACTCTTTCCCTACACGACGCT
J508	AATGATACGGCGACCACCGAGATCTACAC CCTGAAAC ACACTCTTTCCCTACACGACGCT
J701	CAAGCAGAAGACGGCATACGAGAT TGTAGTGC GTGACTGGAGTTCAGACGTGTGCT
J702	CAAGCAGAAGACGGCATACGAGAT GCGTCGAC GTGACTGGAGTTCAGACGTGTGCT
J703	CAAGCAGAAGACGGCATACGAGAT GGTCTTCT GTGACTGGAGTTCAGACGTGTGCT
J704	CAAGCAGAAGACGGCATACGAGAT AAATGTCC GTGACTGGAGTTCAGACGTGTGCT
J705	CAAGCAGAAGACGGCATACGAGAT GTTGAAAC GTGACTGGAGTTCAGACGTGTGCT
J706	CAAGCAGAAGACGGCATACGAGAT TCTTTACG GTGACTGGAGTTCAGACGTGTGCT
J707	CAAGCAGAAGACGGCATACGAGAT ATGCCTGG GTGACTGGAGTTCAGACGTGTGCT
J708	CAAGCAGAAGACGGCATACGAGAT CAATAAGG GTGACTGGAGTTCAGACGTGTGCT
J709	CAAGCAGAAGACGGCATACGAGAT CGCCGTAA GTGACTGGAGTTCAGACGTGTGCT
J710	CAAGCAGAAGACGGCATACGAGAT TAAGGCTT GTGACTGGAGTTCAGACGTGTGCT
J711	CAAGCAGAAGACGGCATACGAGAT TTGCTGCC GTGACTGGAGTTCAGACGTGTGCT
J712	CAAGCAGAAGACGGCATACGAGAT CTCAATGT GTGACTGGAGTTCAGACGTGTGCT

2.2 Enzymes

1. Proteinase K.
2. High fidelity DNA Polymerase for PCR.

2.3 Buffers and Solutions

1. LB medium.
2. Ampicillin stock solution 100 mg/ml in ethanol.

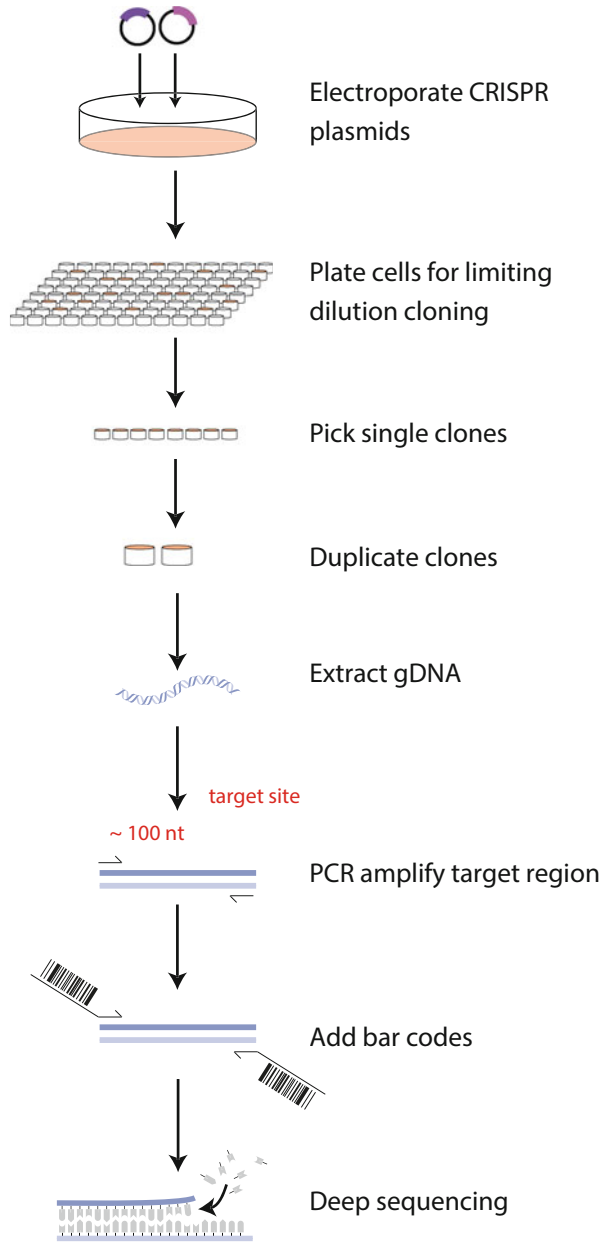


Fig. 1 Workflow of THP1 knockout cell line generation. A Cas9 expression plasmid is co-electroporated with a gene-specific sgRNA expression plasmid. After optional enrichment of positively electroporated cells by FACS sorting, cells are seeded at limiting dilution conditions into 96-well plates. Clones are grown for 2 weeks. Wells containing growing cells are identified and split into two duplicate 96-well plates. After lysing one of the duplicates, a two-step PCR strategy is used to amplify the targeted genomic locus and to add adaptor and barcode sequences for subsequent deep sequencing on the MiSeq platform

3. RPMI medium.
4. OptiMEM.
5. 100 mM sodium pyruvate.
6. 2 mg/ml ciprofloxacin solution.
7. Fetal calf serum (FCS).
8. THP1 culture medium: 1× RPMI, 10 % FCS, 1 mM sodium pyruvate, 10 μM ciprofloxacin.
9. Direct lysis buffer: 0.2 mg/ml proteinase K, 1 mM $CaCl_2$, 3 mM $MgCl_2$, 1 mM EDTA, 1 % Triton X-100, 10 mM Tris pH 7.5.

2.4 Plasmids

1. pCMV-mCherry-T2A-Cas9 (*see* **Note 1**).
2. pLKO.1-sgRNA-CMV-GFP (*see* **Note 1**), e.g., pLKO.1-sgRNA_hsTLR2-CMV-GFP.

2.5 Cells

1. THP1 monocytic leukemia cells (ATCC® TIB-202™).

2.6 Kits

1. Plasmid Maxi Prep kit.
2. Gel Extraction kit.
3. MiSeq 300 cycle v2 sequencing kit.

2.7 Devices

1. Bacterial shaker.
2. Photometer to quantitate DNA concentrations.
3. GenePulser electroporation device (Bio-Rad) (*see* **Note 2**).
4. Cell sorter.
5. PCR thermocycler.
6. MiSeq bench top sequencing system (Illumina).

3 Methods

3.1 Target Site Selection and sgRNA Construction

1. To generate a functional knockout cell line, we propose locating a CRISPR target site within the first critical coding exons of the gene of interest. To knockout human TLR2, we would chose the following target site located in exon 3 of TLR2: GACTGTACCCTTAATGGAGTTGG.
2. The sgRNA design and synthesis is performed according to published criteria [10, 11].

3.2 Large-Scale DNA Amplification of CRISPR Plasmids

1. 300 ml of LB medium supplemented with 100 μg/ml ampicillin are prepared per plasmid in a 1 L Erlenmeyer flask.
2. 300 ml cultures each are inoculated from glycerol stock of the plasmids pCMV-mCherry-T2A-Cas9 and pLKO.1-sgRNA_hsTLR2-CMV-GFP.

3. The cultures are shaken at 37 °C for 18 h.
4. The plasmid DNA is purified using a standard plasmid DNA Maxi preparation kit according to the manufacturer's instructions.

3.3 Electroporation of THP1 Cells

1. THP1 cells are seeded at a density of 2×10^5/ml and are incubated in a 37 °C incubator with 5 % CO_2 for 24 h.
2. 2.5×10^6 cells are resuspended in 250 µl Opti-MEM, mixed with 2.5 µg pCMV-mCherry-T2A-Cas9 plasmid DNA and 2.5 µg pLKO.1-sgRNA_hsTLR2-CMV-GFP plasmid DNA in a 4 mm cuvette and are electroporated using an exponential pulse at 250 V and 950 µF utilizing a Bio-Rad electroporation device (*see* **Note 2**).
3. Cells are then allowed to recover for 2 days in 6-well plates filled with 4 ml medium per well (Fig. 1).

3.4 FACS Sorting

1. FACS sorting of at least 12,000 mCherry positive cells is performed (*see* **Notes 1** and **3**).
2. Cells are kept on ice and used for limiting dilution cloning within 2 h.

3.5 Limiting Dilution Cloning

1. The efficiency of electroporation may vary between different experiments. To this end we make use of three different cell densities for limiting dilution (200 cells/ml, 100 cells/ml and 50 cells/ml) (*see* **Note 4**).
2. 12,000 sorted cells are diluted in a total of 60 ml medium and plated to three round-bottom 96-well plates at a density of 200 cells/ml (100 µl per well).
3. 30 ml medium are added to the remaining 30 ml of cell suspension and cells are plated to three round-bottom 96-well plates at a density of 100 cells/ml (100 µl per well).
4. 30 ml medium are added to the remaining 30 ml of cell suspension and cells are plated to three round-bottom 96-well plates at a density of 50 cells/ml (100 µl per well).
5. Clones are grown at 37 °C, 5 % CO_2 until they become clearly detectable by eye (usually for 2 weeks).
6. Plates are scanned in a plate reader for absorption at 600 nm or by phase-contrast microscopy. Next, grown clones are picked and transferred in duplicates to two fresh 96-well plates
7. One duplicate is cultivated further while the other one is used for lysis to obtain gDNA for subsequent genotyping.

3.6 Cell Lysis

1. From the clone duplicate that is used for genotyping the medium is removed and the cells are resuspended in 30 µl of direct lysis buffer per well. Of this reaction 10 µl are transferred to a 384-well PCR plate (*see* **Note 5**).

2. The lysis samples are incubated at 65 °C for 10 min and at 95 °C for 15 min using a standard PCR thermocycler. The lysates can directly be used for a subsequent genomic PCR reaction.

3.7 Primer Design for the First Level PCR Reaction

1. A genomic amplicon of 220–270 nt length is chosen with the nuclease target site located roughly in the center of this amplicon (*see* **Note 6**).
2. Primer of 23 nt each with a GC content around 50 % and a GC content of exactly 50 % in the two 3′-terminal primer base positions are chosen (N_{23}).
3. The primer sequences are elongated by the following 5′ sequences that are required as constant regions for the second level PCR reactions:
   (a) FWD: ACACTCTTTCCCTACACGACGCTCTTCCGATCT - N_{23}.
   (b) REV: TGACTGGAGTTCAGACGTGTGCTCTTCCGATCT - N_{23}.
4. For TLR2 specific primer pairs refer to Subheading 2.1.1.
5. The primer are ordered in 96-well plate format at 25 nmol synthesis scale, desalted and resuspended in TE pH 8.0 at 50 μM.

3.8 Dual PCR Barcoding

1. The first level PCR reactions are performed in 384-well plates in a volume of 6.25 μl per well using the following protocol (*see* **Note 7**):
   - 1.25 μl 5× Phusion buffer.
   - 0.125 μl dNTP mix 10 mM.
   - 0.0625 μl forward primer 50 μM.
   - 0.0625 μl reverse primer 50 μM.
   - 1 μl PCR-compatible lysate (*see* Subheading 3.6) as a template.
   - 0.0625 μl Phusion polymerase.
   - 3.6875 μl water.
   - Annealing temperature: 60 °C.
   - Elongation time: 15 s.
   - Cycle number: 19 cycles.
2. A universal 96-well MiSeq barcode primer dilution plate is prepared. To this end eight forward primer (J501–J508, row-wise) are combined with 12 reverse primer (J701–J712, column-wise) to obtain 96 different primer pair combinations in a 96-well plate at a final concentration of 5 μM per primer (Fig. 2, *see* **Note 8**).
3. Of the first level PCR reactions, 1 μl are transferred to a new 384-well plate to serve as a template in the second level PCR reaction

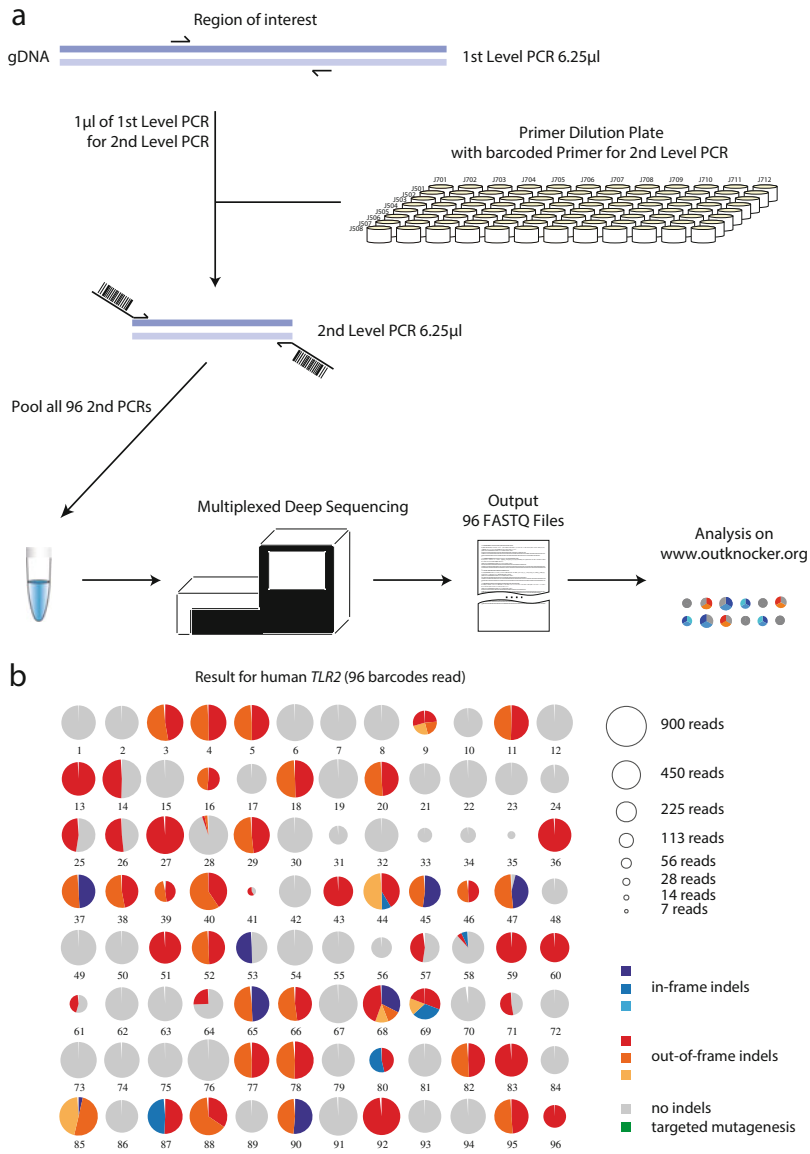


Fig. 2 Barcoded deep sequencing. (**a**) Subsequent to the lysis procedure, the first level PCR is performed on the genomic DNA (gDNA) using a pair of locus-specific primer. Then 1 µl is transferred to a second level PCR in which sequencing adapters and barcodes are introduced using 96 unique primer combinations premixed on a 96-well plate. To this end, eight forward primer (J501–J512, row-wise) and 12 reverse primer (J701–712, column-wise) are distributed over the 96-well plate to generate 96 unique combinations. All second level PCRs of all gene-targeting experiments are then pooled, precipitated, and subjected to the deep sequencing procedure. Upon sequencing, 96 barcode-deconvoluted FASTQ files are acquired. Sequencing files can be analyzed and checked for genotypes with the publicly available software OutKnocker (www.outknocker.org). (**b**) Analysis of 96 sequencing FASTQ files by the outknocker software for human TLR2 in THP1 cells (data from ref. [9]). Each pie chart represents a FASTQ file and as such a single cell clone. The size of the chart is proportional to the number of sequencing reads analyzed. *Colored pie pieces* indicate mutated alleles (*blue* = inframe indels, *red* = out of frame indels). *Grey pieces* represent wt alleles

using the position-dependent combination of barcode primer and the following reaction and cycling conditions (*see* **Note 7**):

- 1.25 µl 5× Phusion buffer.
- 0.125 µl dNTP mix 10 mM.
- 0.625 µl barcode primer mix (5 µM each primer).
- 1 µl PCR template from the first level PCR reaction.
- 0.0625 µl Phusion polymerase.
- 3.1875 µl water.
- Annealing temperature: 60 °C.
- Elongation time: 15 s.
- Cycle number: 19 cycles.

3.9 Deep Sequencing

1. Crude PCR products from the second-level PCR are pooled by centrifuging the inverted PCR plates for 10 s into empty reservoirs (e.g., clean pipet tip boxes).
2. The mixed PCR products are size-separated using a 1.5 % agarose gel run at 100 V.
3. After visualization with ethidium bromide under UV light, DNA bands from 300 bp to 450 bp are cut out and are purified using a gel extraction kit according to the manufacturer's protocol.
4. Eluted DNA is precipitated by adding 0.1 volumes of 3 M NaOAc (pH 5.2) and 1.1 volumes of isopropanol. After centrifugation for 15 min at 4 °C, the resulting pellets are washed in 70 % EtOH, air-dried, and resuspended in 30 µl water.
5. Non-soluble fractions are spun down for 3 min and 25 µl of the supernatant are transferred to a new tube.
6. The DNA concentration is quantified using a spectrophotometer. Typical DNA concentrations are 50–150 ng/µl.
7. Deep sequencing is performed according to the manufacturer's protocol using the MiSeq bench top sequencing system. Data are obtained as 96 barcode-deconvoluted FASTQ files. Approximately 5,000–10,000 cell clones can be analyzed in a single MiSeq sequencing run (*see* **Note 9**).

3.10 Deep Sequencing Data Evaluation Using OutKnocker

1. In a web browser, the evaluation software OutKnocker is loaded (www.OutKnocker.org).
2. The gene name, e.g., human TLR2 is entered.
3. The reference sequence is entered in capital letters, spanning the locus-specific portions of the first PCR primer set. The reference sequence must start from the forward primer sequence. The respective amplicon for human TLR2 would be:
TGGAATATGCAGCCTCCGGATTGTTAACGTTTC
CACTTTACCTGGATCTATAACTCTGTCATTATCAG

ATGCTCTAAAATTACCAACTCCATTAAGGGTACAG
TCATCAAACTCTAATTCTAACAATCCAGAAATCTG
ATTCAAAAGTTTCATAACCTGAAACAAACTTTCAT
CGGTGATTTTCACATTTCTAAATGTAAACTTTTTA
ATCAATGAATTTGTTTCACCAGTGGATAGTTCTGA

4. The nuclease target site is entered in capital letters in the same orientation as the reference sequence. The target site should include the PAM motif. The respective target site for human TLR2 would be: GACTGTACCCTTAATGGAGTTGG.
5. All 96 MiSeq FASTQ files are selected using the file browser button. Files must be sorted according to their file names in ascending order before selecting them.
6. The "START" button is clicked. The evaluation process can take several minutes.
7. In the result page, every cell clone is depicted by a pie chart ordered according to the 96-well plate format of the second PCR primer. Pie chart pieces indicate the relative abundance of indel alleles with a frameshift (red color), in-frame indels (blue color), or alleles without indel events (grey color).
8. Knockout clones are selected by clicking on pie charts that are fully colored in red.
9. To verify the genotyping result, individual alleles of the selected clone are depicted as alignments to the reference sequence on the bottom of the results page. If a base is sequenced with reduced confidence, it is depicted smaller than the other bases, with the letter size linearly correlating with the relative abundance of that specific base. Red bases indicate mismatches to the reference sequence.

4 Notes

1. pCMV-mCherry-T2A-Cas9 expresses the Cas9 protein separated from mCherry by a T2A peptide. pLKO.1-sgRNA_hsTLR2-CMV-GFP expresses the sgRNA driven by the Pol III promoter U6 and it encodes for GFP driven by a CMV promoter. In our experience it suffices to enrich Cas9 expressing cells by sorting for mCherry expression, since high Cas9 expression is crucial for efficient genome editing activity.
2. We recommend optimizing the electroporation/transfection procedure when using a different electroporation device or cell line.
3. Enrichment of transfected cells is an optional step and can be omitted for cell lines with higher transfection efficiencies (e.g., HEK293 cells).

4. Limiting dilution cloning is amenable to most cell lines; however, we observed that cell lines behave differently under limiting dilution conditions. Moreover, we found that cell viability is reduced after FACS sorting and that this underlies considerable variation. Therefore, counting of 12,000 cells by the FACS sorting device allows only an estimation of the number of monoclones that will grow at a given dilution. As such, we found that plating THP1 cells at 5, 10, and 20 cells per well provides sufficient variation to select 96 monoclones in the majority of cases while limiting invested material and efforts to a reasonable amount. However, when working with a new cell line we recommend titrating the cell density over a broad range of cells per well (e.g., 0.5–16 cells per well) to determine the best limiting dilution conditions.
5. Optionally, the cell lysis can be performed in 96-well PCR plates or PCR strips by transferring 30 μl to the PCR vessel.
6. In order to detect very large deletions we designed the amplicon for the genomic PCR in a way that the CRISPR target site is located roughly in the center of the amplicon. When placing the target site asymmetrically it is possible that large deletions result in loss of one primer-binding site. Consequently, the PCR would fail and the deletion would not be detected by sequencing.
7. Usually we use one quadrant of a 384-well plate for the PCRs of 96 clones to keep the PCR reaction volume as small as possible. The other three quadrants can be used for different gene-targeting experiments. When using 96-well plates or strips for PCR we suggest doubling the volume of the PCR reaction (12.5 μl) and the lysate respectively (2 μl).
8. Optionally, less than 96 barcodes can be used. However, we recommend to genotype 96 monoclones per gene targeting project when using difficult-to-target cell lines such as THP1 cells. Moreover, we recommend preparing the 96 primer pair combinations for the second level PCR at bulk (e.g., 100 μl). This plate can be stored at –20 °C and conveniently used at later time points.
9. A typical MiSeq 300 cycle v2 run yields 10–15 million reads. We found that 1000–2000 reads are sufficient to genotype a single cell clone. Therefore, 5000–10,000 clones can be analyzed in a single sequencing run. If less than 15 amplicons are sequenced in a single run, please refer to Illumina technical literature about low-complexity sequencing runs.

Acknowledgement

This work is financially supported by SFB 670, SFB 704, ERC starting grant, the *ImmunoSensation* cluster of excellence to V.H., and by the *Studienstiftung des deutschen Volkes* (J.L.S.-B.)

References

1. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. Nat Rev Genet 11(9):636–646. doi:10.1038/nrg2842
2. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29(2):143–148. doi:10.1038/nbt.1755
3. Hsu PD, Lander ES, Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. Cell 157(6):1262–1278. doi:10.1016/j.cell.2014.05.010
4. Kim H, Kim JS (2014) A guide to genome engineering with programmable nucleases. Nat Rev Genet 15(5):321–334. doi:10.1038/nrg3686
5. Gaj T, Gersbach CA, Barbas CF 3rd (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 31(7):397–405. doi:10.1016/j.tibtech.2013.04.004
6. Jasin M, Rothstein R (2013) Repair of strand breaks by homologous recombination. Cold Spring Harb Perspect Biol 5(11):a012740. doi:10.1101/cshperspect.a012740
7. Chiruvella KK, Liang Z, Wilson TE (2013) Repair of double-strand breaks by end joining. Cold Spring Harb Perspect Biol 5(5):a012757. doi:10.1101/cshperspect.a012757
8. Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79:181–211. doi:10.1146/annurev.biochem.052308.093131
9. Schmid-Burgk J, Schmidt T, Gaidt M, Pelka K, Latz E, Ebert T, Hornung V (2014) OutKnocker: a web tool for rapid and simple genotyping of designer nuclease edited cell lines. Genome Res 24(10):1719–1723. doi:10.1101/gr.176701.114
10. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 31(9):822–826. doi:10.1038/nbt.2623
11. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096):816–821. doi:10.1126/science.1225829

INDEX

Ralf Kühn et al. (eds.), (eds.), *TALENs: Methods and Protocols*, Methods in Molecular Biology, vol. 1338,
DOI 10.1007/978-1-4939-2932-0, © Springer Science+Business Media New York 2016

N

O

P

R

S

T

U

X

Z